Journal of Multimedia

ISSN 1796-2048
Volume 9, Number 7, July 2014

Contents

REGULAR PAPERS

Study of Chinese Text Similarity Based on Difference Factor in Word-Number

Yan Niu, Qing Zou, and Yaqing Han
Hubei University of Technology, Wuhan, China
Email: ny@mail.hbut.edu.cn, 309921125@qq.com, 7137100@qq.com

***Abstract*—Text similarity calculation is the basic work in the application of Chinese information processing. A high-quality text similarity calculation method must be accurate and efficient, that is, it can be able to compare texts from the level of text natural language meaning, and arrive at the similarity distinction similar to artificial reading based on a full understanding of the author or text source semantic. At the same time, it should also be an efficient algorithm to save the processing time in facing large amount of text information to be processed. Through the research of many domestic and foreign literature, analysis and further research on current situation of similarity calculation, this paper intended to present a new method to improve the performance of similarity calculation, namely a Chinese text similarity algorithm based on word-number difference, which combined the traditional based on statistics and the narrow semantic method that meant the combination of the statistical efficiency and semantic accuracy. Combining the advantages of statistics and semantic category also means the necessity to face and overcome disadvantages of the two kinds of methods. This paper attempted to take the difference in word-number as the breakthrough point, took advantage of the diversity of Chinese word-number, combining with the word frequency, number and meaning, in order to successfully extend the word similarity calculation to the text similarity calculation. Finally, introduced the self built small text set as test object, compared similarity calculation of different methods in the laboratory environment. It shows that the similarity calculation method based on difference in word-number performances better than the traditional methods based on statistical and semantic. Through artificial comparison of the test results of research on this topic in accuracy and speed of segmentation, provide a new approach for Chinese text similarity calculation**

***Index Terms*—Similarity; Factor in Word-Number; Segmentation**

I. INTRODUCTION

As the basic work in organization and management of mass information, the calculation of text similarity has a decisive impact on text classification and information retrieval. It also has a wide application prospect in the fields of search engine, document check, artificial intelligence, digital library and so on.

Text similarity calculation means calculating the similarity between documents in a certain way. In similarity calculation of English texts, English owns the features that each word has its own meaning, and words are separated by a space in any sentence. However, different from English, Chinese characters have different meanings in different combinations. Therefore, in Chinese text similarity calculation, the particularity of Chinese should be put into full consideration. Do Chinese word segmentation to the text, and with the help of Chinese segmentation results, can make the computer capable of identifying the contents of natural text. On the basis of different kinds of Chinese word segmentation, both of the calculation method and calculation results Chinese text similarity would be different [1].

Calculation of text similarity is the basic work in the application of Chinese information processing. A high-quality calculation method of text similarity must be accurate and efficient, that is, comparison can be done from the level of natural language meaning of texts and similarity distinction can approach actual reading by full understanding of author or semantics of text source [2]. Meanwhile, the method must use an efficient algorithm to save processing time during processing large amounts of text information.

To improve the utilization efficiency of information and enhance the precision of information retrieval, high-quality text similarity calculation method is a necessary measure. The significance of the study on text similarity is how to use the most accurate and efficient similarity algorithm to calculate the text similarity between texts of different word-number and belong to different categories quickly and accurately [3]. As a basic work of application like text classification and information retrieval, the efficiency of similarity calculation plays a decisive role in them.Study of this topics is of realistic significance in the following aspects:

(1) As information disseminator and audience, humans express thoughts and meanings through text in the information exchange process, so the natural language is rich in semantics. But the human communication, is based on human biology cognitive understanding, for the computer, biological understanding is more difficult. In natural language, phenomena of polysemy and synonym are very common. At the same time, other complex phenomena abound like synonyms antonyms, and parts of speech, grammar etc. Improvement of information retrieval quality and efficiency, need to be verified with semantics, the calculation of text similarity, is to try to overcome these difficulties, make the computer can compare text from natural language meaning perspective,

doi:10.4304/jmm.9.7.865-872

doing the basic work in other applications for Chinese information. The calculation of text similarity is the foundation for stability, rapid development of information. High efficiency in the operation of computer, can replace primitive data collection and record, and with the realization of the similarity calculation, can make a computer identify text approximately, then help human free from complex, repetitive text work to a greater extent.

(2) Accelerate the speed of query target information. Due to the wide spread of computer network, the whole world owns hundreds of billions of Webpage, and when we carry on the information retrieval, there will be lots of text with the retrieval condition. Through the calculation of text similarity, can compare text with retrieval conditions, sort the resulting text according to the retrieval of different conditions. The rapid development of Baidu and Google, shows huge range of application and broad market space of retrieval. Of course the search engine also involves a variety of computer technology, but the text similarity, is one of the key [4].

(3) Improve the comprehensive of information query relevant information. Similarly, as the large data scale of the network space, when the user query data, a lot of material will be retrieved. However, due to a variety of language expression, if there is no perfect calculation of text similarity, data of less relevant would be filtered by search conditions. Besides, text similarity can expand the traditional keyword dominant retrieval method, especially in similar document query, can query to more comprehensive relevant information, according to the document category keyword combined with the text similarity.

The semantic problem is the most difficult to define and inevitable problems in text information processing, which may be the relationship between a linguistic sign and it's coreference, or the concept of associated human perception and language symbols, or response to a series of external stimuli. No matter merplotactic language English, or semotactic language Chinese, both of whose multiple attribute can be discussed from different angles.

(1) Semantics is the product of the interaction of objective and subjective. A semantic unit can be referring to objective terminology and common name, virtual name generated from possible world or imaginary world, verbs express actions, adjectives and adverbs express quality and properties, or prepositions and conjunctions express relationship etc.

(2) Semantics is the product of the combination of language form and content. The content and form are integrated into word, phrase, sentence of shared pronunciation or morphology or known as symbolic units by cognitive thinking, and these units become words in application or participation in communication, then form the external language. The functional description and the complicated relationship set (synonymy, hyponymy, antonymy, part and whole etc.) in use of symbolic units can reveal its semantic properties [5].

(3) Semantics is the product of the combination of individual and general. The semantic comestible from the cognitive subject- the awareness and understanding of human towards the objective world. The generality reflects the static meaning and the static properties of words, while the Individuality reflects the dynamic meaning and dynamic characteristics of words.

Only on the basis of fully familiar with the semantic features and composition, can the computer be used to identify language approximately, thus calculate text similarity with the way meeting human cognitive habits.

II. Difference Factor in Word-Number

This case often occurs in the calculation of the text similarity: in three texts A, B, and C with 50, 50, and 500 Chinese characters respectively. Since text A has fewer counts of Chinese characters, the words in text A tends to be included in text C after sentence splitting and word segmentation, which makes the calculation result be a very high similarity, that is, the phenomenon that a long corpus covers a short one occurs. However, in the viewpoint of text content, the texts A and B both with 50 words are closer semantically. In order to solve this problem, we introduce the concept of difference factor in word-number.

Word is the smallest unit in Chinese information processing, so rapid and effective segmentation to text is the precondition of text similarity calculation. In Chinese, words can be divided into three kinds of a total of eleven classes according to the part of speech, including symbols of special use, such as character punctuation, unit, digital, various western letters and other non Chinese. The three kinds include component words, special element words and non component words, while the eleven classes include nouns, verbs, adjectives, numerals, measure words, pronouns, adverbs, prepositions, conjunctions, auxiliary, onomatopoeia and interjections. Among these, nouns, verbs, adjectives, numerals, measure words and pronoun word class, Is the main weight part of text semantic expression, namely the key to similarity calculation, while other function words only take part in last cumulative part in the similarity calculation [6].

A. Concept of Difference Factor in Word-Number

Since the effect of difference in word-number on text similarity calculation exists objectively, this paper puts forward the concept of difference factor in word-number, in order to pertinently show the effect of difference in word-number on text similarity calculation, especially the effect of semantic level. Participation of the difference factor in word-number would reduce the variation caused by the difference in word-number in the other calculation of similarity.

The difference factor in word-number is formed by the total count of text characters, count of notional words, count of function words, and special identifiers. The notional words include nouns, verbs, adjectives, numerals, quantifiers, and pronouns. The function words include adverbs, prepositions, conjunctions, auxiliaries, onomatopoeias, interjections and 398 special identifiers. The total count of text characters is the number of all Chinese characters in the text. After the Chinese segmentation, total number of notional words, function

words, and special identifiers can be counted respectively according to their characteristics.

Let A be the set of notional words in the text, and m be their count, then $A=(a_1,a_2,...,a_m)$. Let h be the frequency of each notional word occurs in the text, then $H=(h_1,h_2,...,h_m)$, $\{A,H\}=[(a_1,h_1),(a_2,h_2),...,(a_m,h_m)]$. Let B be the set of function words, and n be their count, then $B=(b_1,b_2,...,b_n)$. Let C be the set of special identifiers, and k be their count, then $C=(c_1,c_2,...,c_k)$. Due to the particularity of Chinese, the count of characters contained in the words is unfixed. Let L be the count of characters contained in notional words, since $A=(a_1,a_2,...,a_m)$, then $L=(l_1,l_2,...,l_m)$, $\{A,H\}=[(a_1,h_1),(a_2,h_2),...,(a_m,h_m)]$, then $\{L,H\}=[(l_1,h_1),(l_2,h_1),...,(l_m,h_m)]$, which indicates a two-dimensional array made by the count of characters contained in the word and the occurrence number of words in the text, as shown in the figure. Since function words and special identifiers have no independent semantics in the text, only their total count is used to calculate the difference factor in word-number.

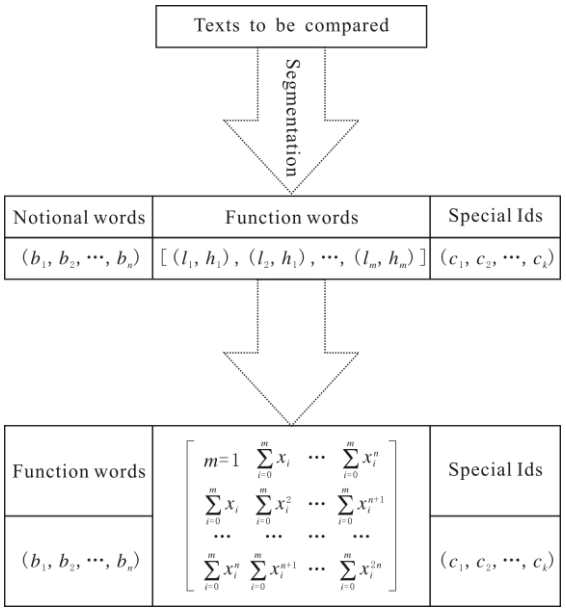


Figure 1. Statistics and reorganization of segmentation results

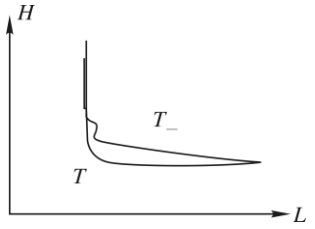


Figure 2. Curves of test text

B. Calculating the Difference Factor in Word-Number

Take some Chinese text for example. Consider the text as T_1, the contrast text (created based on the T_1 after Chinese segmentation) as T_2, and the source text, which T_1 is excerpted from, as T_3. Count the total number of notional words, function words, and special identifiers respectively through Chinese words segmentation.

Assume that in the text T_1, there are 25 notional words, which are distributed as $\{L,H_1\}$,while in the text T_2, there are 33 notional words, which are distributed as $\{L,H_2\}$.

Take the words as specified data points (x_i,y_i) $(i=0,1,...,m)$, a function can be constructed according to their occurrence number in the text:

$$\begin{cases} h_n(x)=\sum_{k=0}^{n} a_k x^k \in \phi \\ I=\sum_{i=0}^{m}[p_n(x_i)-y_i]^2=\sum_{i=0}^{m}(\sum_{k=0}^{n} a_k x_i^k - y_i)^2 = \min \end{cases} \tag{1}$$

where $I=\sum_{i=0}^{m}(\sum_{k=o}^{n} a_k x_i^k - y_i)^2$ when n=1. Based on multivariate function of $(a_1,a_2,...,a_n)$, there is:

$$\begin{cases} \dfrac{\partial I}{\partial a_j}=2\sum_{i=0}^{m}(\sum_{k=0}^{n} a_k x_i^k - y_i)x_i^j \in \phi \\ \sum_{k=0}^{m}(\sum_{i=0}^{n} x_i^{j+k})a_k = \sum_{i=0}^{m} x_i^j y_i \end{cases} \tag{2}$$

The linear equations of $(a_1,a_2,...,a_n)$ can be represented by the following matrix:

$$\begin{bmatrix} m=1 & \sum_{i=0}^{m} x_i & \cdots & \sum_{i=0}^{m} x_i^n \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 & \cdots & \sum_{i=0}^{m} x_i^{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=0}^{m} x_i^n & \sum_{i=0}^{m} x_i^{n+1} & \cdots & \sum_{i=0}^{m} x_i^{2n} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m} y_i \\ \sum_{i=0}^{m} x_i y_i \\ \cdots \\ \sum_{i=0}^{m} x_i^n y_i \end{bmatrix} \tag{3}$$

In the text T_3, all the notional words in T_1 occur and their frequencies are higher than those in T_1. The words are distributed as $\{L, (H_1|H_3)\}$. In addition, the notional words in the text T_3 are 58 more than those in T_1, and the occurrence frequency more than once.

TABLE I. WEIGHT OF WORDS

L_{ij}		G_i	
Name	Formula	Name	Formula
Binary	$L_{ik}=\begin{cases}0, f_{ik}=0\\1, f_{ik}=1\end{cases}$	Normal	$G_k=\sqrt{\frac{1}{\sum_j l_{ij}^2}}$
H(term frequency)	L_{ik}=tfik	gf	$G_k=\frac{gf_k}{df_k}$
Log(H)	L_{ik}=log(tfik+1)	Log(G)	$G_k=1+\log_2\frac{N}{n}$
		entropy	$G_k=1+\sum_i \frac{p_{ik}\log_2 p_{ik}}{\log_2 N}$ $p_{ik}=\frac{H_{ik}}{gf_k}$

By comparing the data, we can find that most parts of text T_3 are overlapped with text T_1. Borrowing the concept of geometry, there are more coincidences at the same points, but the overlap area of two curves is not

large. The ratio of overlap area of texts T_1 and T_2 to the text T_1 is bigger than that of texts T_1 to T_3.

The method [2] for calculating the weight of words given by Dumains in local and full text is shown in table I.

From various local and full-field calculation, can draw the distribution ratio of entries appear in the text, and word discrimination in distinguishing the core word, namely high frequency word set:

$$\begin{cases} IG_k = H(FW) - H(FW \mid term_k) \\ H(FW) = -\sum_{FW} p(fw) \times \log_2 p(fw) \\ p(fw) = \dfrac{|wordset(fw)|}{\sum wordset(fw_1)} \\ H(fw \mid term_k) = -\sum_{fw} p(fw \mid term_k) \times \log_2 p(fw \mid term_k) \end{cases} \quad (4)$$

Combining the above calculation of local weight and full-field weight for notional words, find the formula for difference factor in word-number (η):

$$\eta = \frac{O_1}{O_2} \sum_{i=1}^{m} (\sum_{k=1}^{n} x_i^k - y_i)^2 + \sum_{k=1}^{m+n} ((u_1 + u_2)^k + (v_1 + v_2)^k) \quad (5)$$

where, O_1 and O_2 are total word counts of contrast text T_1 and T_2 respectively, m and n are the counts of notional words in contrast texts T_1 and T_2 after Chinese segmentation processing respectively, x_i and y_i are the occurrences number of notional words in the text respectively, u_1 and u_2 are the counts of function words in texts T_1 and T_2 respectively, v_1 and v_2 are the counts of special identifiers in texts T_1 and T_2 respectively, in which $O \geq m+u+v$.

In the sample texts T_1, T_2, and T_3, for T_1, O_1 is 130, m is 25, u is 40, v is 16; for T_2, O_1 is 156, m is 33, u is 57, v is 32; for T_3,O_1 is 470, m is 85, u is 110, v is 88. Take them into the formula, find $\eta(T_1,T_2) = 0.683$ and $\eta(T_1,T_3) = 0.477$. So $\eta(T_1,T_2) > \eta(T_1,T_3)$ is tally with the actual situation.

C. The Segmentation Method Based on the Special Identifier

Calculation method of Chinese text similarity based on difference of word number in this paper, use the knowledge about ontology, combined with the Chinese semantics Chinese dictionary Hownet, calculate primitive word similarity, then coupled with the word difference factor into semantic similarity calculation method of the whole text. Due to use of the knowledge about ontology, the following will introduced Ontology firstl.

(1) Introduction of Ontology

The word "ontology" comes from the philosophy as a part of metaphysic. Metaphysics systematic research on the principle under specific subject, mainly the essence of existence and experience, which is a newly appeared content in information technology engineering now. An ontology defines a common vocabulary and concept used for description and representation of knowledge of a field, ontology is also an engineering product, "specific vocabulary to describe the real world, and exploitation hypothesis set about the words and their meaning" [7].

Ontology includes Conceptualization, Explicit, Formal and Share. It's target is acquiring, describing and representing knowledge of relevant field, providing a common understanding of the knowledge of the field, determining the words of common understanding in the field, and Giving a clear definition of the relationship between these concepts and vocabulary of the form pattern from the different levels.

(2) Semantics calculation of vocabulary

In Ontology, semantic dictionary will use a tree structure to combined the used words together, as shown in the figure 3. In the tree structure, there is only one shortest traversal path between any two points, which represent the shortest distance between the two words. According to the semantic interpretation of the dictionary, it can show the difference between concepts or words that the two nodes represent to a certain extent, namely the semantic distance of two concepts [8].

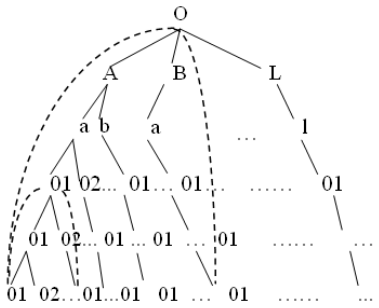


Figure 3. Hownet layer section structure view

In the arithmetic expression, set two words w_1,w_2, if w_1 has n senses s_{11},s_{12},...,s_{1n}, w_2 has m senses s_{21},s_{22},...,s_{2m}, set the similarity between w_1 and w_2 as similarity maximum of all the senses, namely.

$$simW(w_1, w_3) = \max_{i=1,...,n,\ j=1,...,m} simWS(s_{1i}, s_{2j}) \quad (6)$$

where, $simW(w_1, w_2)$ represents similarity between the two words, $simWS(s_{1i}, s_{2j})$ represents similarity between the two senses, sense is represented by sememe, according to reference [9] can get the similarity between the two sememe as:

$$simWP(p_1, p_2) = \alpha(d + \alpha) \quad (7)$$

where, p_1 and p_2 represent the two sememe, d, a positive integer, is the path length between p_1 and p_2 in sememe level system, here as the sememe distance between the two sememe. α is an adjustable parameter, which means the path length when the similarity is 0.5, here set α =1.6 on the depth of sememe tree. The similarity calculation formula between two notional is:

$$simWS(p_1, p_2) = \beta_1 simWP(p_1, p_2) + \sum_{i=2}^{4} \beta_1 simWP(p_1, p_2) \times \beta_i simWP_i(p_1, p_2) \quad (8)$$

D. Algorithm Flow Based on Difference in Word-Number

Basing on the above introduction, here put forward the semantic method of calculating text similarity based on difference in word-number. The task consists of the following steps:

- Determine the Chinese text to be compared.
- Segment Chinese texts into words.
- Collect statistics for word segmentation results according to the notional words, function words, and special markers.
- Rank the notional words in descending order according to the frequency they appear in the text.
- Calculate the difference factor in word-number.
- Map the Chinese segmentation results to Hownet.
- Calculate semantic similarity between text words according to formula.
- Extend lexical semantic similarity to the full text, combining the difference factor in word-number.

Firstly, judge whether the text is Chinese. Since various letters present in symbolic form in Chinese, when the main part of the text is not Chinese character, the following Chinese participle is unable to process the text. Thus affect the similarity calculation. After confirming the text is Chinese, do word segmentation process to the text. This paper makes use of Chinese word segmentation system (ICTCLAS).

Collect statistics for word segmentation results according to the notional words, function words, and special markers. The notional words include nouns, verbs, adjectives, numerals, classifiers and pronouns, while the function words include adverbs, prepositions, conjunctions, auxiliary, onomatopoeia and interjections. Rank the words in descending order according to the frequency they appear in the text, and combine the length of the word and the frequency it appears in the text into a two-dimensional array. Then calculate the difference factor in word-number according to formula.

Words in Hownet are composed of the notional words and function words. The notional words are described by independent description, relationship sememe description and symbolic sememe description. Independent description, namely using "the basic sememe" or "(the word)" to describe. Relationship sememe description, namely using " relationship sememe = basic sememe " or "relationship sememe=(the word)" or "(relationship sememe=the word)" to describe. Symbolic sememe description, namely using simple "{syntactic sememe}" and "{relationship sememe}" to describe "relationship symbols basic sememe" or "relationship symbols (concrete word)" function words.

In independent description, the first description is the basic sememe of words and the most basic, most important description of the notional words, and named as the first independent sememe description, since the independent sememe shows the most commonly used semantic of the words. Because notional words are specific parts in expressing the meaning in language, semantic similarity calculation of substantive part also need to be focus on, which would improve the efficiency of semantic calculation of notional words in Hownet, as the overall count reduces.

The notional words similarity calculation in Hownet consists of four parts: the similarity of the sense of the first independent sememe can be calculated by formula (9), denoted by $simWP_1(p_1,p_2)$, the similarity of other independent sememe are denoted by $simWP_2(p_1,p_2)$, since they may have multiple component, $simWP_2(p_1,p_2)$ is calculated by the method of the maximum combined weighted value:

$$simWP_2(p_1,p_2)=\sum_{k=1,\ldots,p}\lambda_i \max_{i=1,\ldots,m,\ j=1,\ldots,n} simWP(p_i,p_j) \quad (9)$$

The similarity of relationship sememe in relationship sememe expression is denoted by simWP3(p1,p2), take the maximum value of sememe combination as the similarity of relationship sememe:

$$simWP_3(p_1,p_2)=\max_{i=1,\ldots,m,\ j=1,\ldots,n} simWP(p_i,p_j) \quad (10)$$

The sense similarity in symbolic sememe description is denoted by $\mathrm{sim}WP_4(p_1,p_2)$, also take the maximum value of symbolic sememe combination as the similarity of symbolic sememe:

$$simWP_4(p_1,p_2)=\max_{i=1,\ldots,m,\ j=1,\ldots,n} simWP(p_i,p_j) \quad (11)$$

Synthesize all the above sense similarity, get the whole semantic similarity of two words as:

$$simWS(s_1,s_2)=\sum_{i=1}^{4}\beta_i simWP_i(p_1,p_2) \quad (12)$$

Among it, $\beta_i(1\le i\le 4)$ is controllable parameter, and $\beta_1+\beta_2+\beta_3+\beta_4=1$, $\beta_1\ge\beta_2\ge\beta_3\ge\beta_4$. From $simWP_1(p_1,p_2)$ to $simWP_4(p_1,p_2)$, the weight of which in overall similarity is decreasing. Considering the importance of independent sememe, the weight of the first independent sememe is usually set to more than 0.5. There is certain differences between the main semantic content and secondary content of words, thus the deviation caused by secondary meaning of words when the main semantic are different should also be taken into account in the sense calculation of two words.

Thus, after semantic weight contrast optimization, the calculation formula is improved as:

$$simWS(p_1,p_2)=\beta_1 simWP(p_1,p_2)+\sum_{i=1}^{4}\beta_1 simWP(p_1,p_2)\times\beta_i simWP_i(p_1,p_2) \quad (13)$$

To the result of text segmentation, there is a two-dimensional array about the length and frequency of notional words:

$$\begin{bmatrix} w_{11} & w_{12} & \ldots & w_{1n} \\ w_{21} & w_{21} & \ldots & w_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ w_{m1} & w_{n1} & \ldots & w_{mn} \end{bmatrix}$$

Among it, m is the length, n is the frequency, thus w_{11} represents the words that length are one, and only appear once in text. w_{mn} represents the words that length are m, and appear n times in text. Let N be a similarity matrix of text u and v, there is:

$$N = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{21} & \dots & u_{2m} \\ \dots & \dots & \dots & \dots \\ u_{m1} & u_{n1} & \dots & u_{mn} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{21} & \dots & v_{2m} \\ \dots & \dots & \dots & \dots \\ v_{m1} & v_{n1} & \dots & v_{mn} \end{bmatrix} \quad (14)$$

Among it,

$$u_{1i}v_{i1} = simW(u_{1i}v_{i1}) \quad (15)$$

In computation, firstly, traversal similarity matrix, combine the words of the greatest similarity, remove them from the rows and columns that they belong in the matrix. Continue to search similarity words combination and delete the matrix rows and columns, until the matrix is empty, then the maximum combination sequence of the text is:

$$\max L = \{simW\max_1, simW\max_2, ..., simW\max_k\} \quad (16)$$

In text T_1, T_2, word semantic similarity and difference factor in word-number are:

$$\begin{cases} sim(t_1,t_2) = \sum_{i=1}^{m+n} \beta_1 simS_i(t_1,t_2) \\ \eta = \frac{O_1}{O_2}\sum_{i=1}^{m}(\sum_{k=1}^{n} x_i^k - y_i)^2 + \sum_{i=1}^{m+n}((u_1+u_2)^k + (v_1+v_2)^k) \end{cases} \quad (17)$$

The formula for calculating text similarity based on difference in word-number is:

$$sim(t_1,t_2) = \frac{1}{m+n}\sum_{i=1}^{m+n} \eta_{(t_1,t_2)} simS\max_{m+n} \quad (18)$$

E. Analysis of Experimental Results

(1) The evaluation parameters of experimental results

Standard of evaluating Chinese text similarity algorithm consists of accuracy, high efficiency and applicability. Specific details are as follows:

Accuracy: Accuracy is the most important and most difficult index in text similarity calculation, since that it relates to the computer machine recognition to language, but the machine which can't read like human, can just correspond words to stored semantic according to the designed process. Therefore, a perfect similarity algorithm, must reflect human-readable text semantics of a text, and it's calculation results must be in line with standards of human cognition. As the basic work of information processing, influence of the similarity results in information retrieval, machine dialogue and other application areas is crucial.

High efficiency: Because of the basic position of text similarity calculation in information processing, similarity will be widely carried out in multi-level and multi-orientation, however, all the computer systems, their calculation speed are relatively limited, at the same time of computer multitasking, similarity calculation need to minimize occupation of the computation speed and system resources of the computer, so while ensuring the accuracy of the calculation result, reducing similarity calculation's occupation of computation resources, is the effective requirement of similarity calculation.

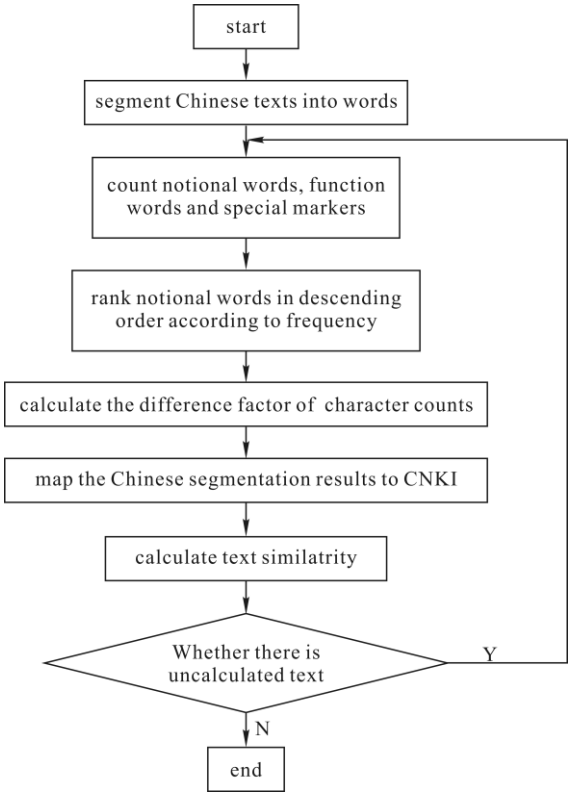


Figure 4. The calculation process of similarity based on the difference in word-number

Applicability: Nowadays, Computers are widely used. Because of the difference between economic level and the level of science and technology, the computer systems used in regions and industries also have difference. In addition to the difference on the computer hardwares, computer softwares are of more multifarious difference, so a good text similarity calculation algorithm must have good universality to the software and hardware of the computer, must have the same operation in different operating system. Similarity calculation results must be good support for further application, and have unit recognition and calculation level towards different types of text materials of different disciplines, won't generate errors caused by the expression of language or rare nouns of characteristic category, having a good adaptability to the computer environment, requiring similarity calculation of effective machine learning mechanism, in order to achieve good portability and stability of the results [10] .

(2) The result analysis and comparison

Assess the Chinese word segmentation system according to the above segmentation system evaluation standard. The experiment is completed on the computer with the CPU of Intel Conroe E6300 1.86GHz, 2GB memory and the operating system of Windows XP.

Take experimental data from self-constructed small text set as the test object, Including a total of 300 articles in 6 fields of economic, legal, news, chemical, computer, physics and astronomy, all of which were collected from field journals published materials, and paragraphs of between 50-500 words generally. Experimental materials drawn from arbitrary two pieces of material from them, calculate similarity by method 1, 2, 3 respectively, method 1 is the word semantic similarity calculation method from reference [11], method 2 is the word semantic similarity calculation method from reference [12], and method 3 is an improved calculation method of text similarity based on 3 layer structure. Adopt parameter values from reference [9] as the parameters in the formula. There are

$$\begin{cases}\alpha = 1.6, \gamma = 0.2, \delta = 0.2 \\ \beta_1 = 0.5, \beta_2 = 0.2, \beta_3 = 0.17, \beta_4 = 0.13\end{cases}$$

$$\Pr ecision = \frac{The\ number\ of\ correct\ results}{The\ number\ of\ all\ the\ return\ results}$$

$$\mathrm{Re}\, call = \frac{The\ identified\ correct\ results}{The\ total\ correct\ results\ in\ the\ database}$$

The experimental results is shown in Table 2.

TABLE II. THE CALCULATION RESULTS OF PARAGRAPHS SIMILARITY

Document category	Method 1		Method 2		Method 3	
	R/%	P/%	R/%	P/%	R/%	P/%
Economic	27	31	33	61	48	70
Legal	22	28	35	60	47	66
News	32	42	47	55	52	61
Chemical	26	31	37	41	46	53
Computer	29	35	38	44	51	57
Physics	26	40	39	48	48	52
Astronomy	19	39	24	36	39	41

The experimental results show: Compared with traditional Vector space model based on statistics [13] and The narrow sense similarity calculation based on lexical semantics, text similarity calculation method based on word difference factor has performed better that meets expectation in the model design process. This result shows that Chinese language characteristics embodied in text similarity calculation. Considering the essence of Chinese as semotactic language, the starting point for semantic calculation must be a meaningful words string combination after Chinese segmentation. In segmentation results, since a word can have different meanings, the word-number it contains and the number of times it appears in the text, show the importance of the words in the text to a certain extent. Basing on the statistics of the Chinese word segmentation results, calculate difference factor according to content words, function words and the number of special signs. After deterring the words and weights of semantic phrases, calculate the lexical similarity by Hownet, calculating similarity of all phrases in the two novels, and feature weighting. In accordance with the quantity and frequency of substantive results, technology on lexical similarity can improve the efficiency of calculation. Then combined with the word difference factor, expand the text similarity in terms of the similarity of the whole text [14].

The whole calculation process clearly represents the advantage of two different methods based on statistic and semantic. Method based on statistic is easy to Machine recognition for computer, while word frequency statistics can intuitively reflect the weight of phrases in the text and greatly improve the computational efficiency. Method based on semantic through the Hownet, can restore and combine the semantic category in the text to the greatest extent, calculate semantic similarity of text dynamically according to word frequency and length of unit word. By combining the advantages of two method, efficiency of similarity calculation is improved, in the meantime, accuracy of machine recognition in semantic is also ensured.

III. CONCLUSION

A. Summary

The algorithm for text similarity of word length differences can be applied to extensive aspects of Chinese information processing such as search engines, machine translation, machine proof-reading, speech synthesis, automatic classification, machine marking, artificial intelligence, speech recognition, and human-computer interaction. With extensive promotion of Internet and its high efficiency, the information spreading on Internet becomes the mainstream, which turns into the trend of development. Text similarity can be better used to calculate the similarity of web pages for the fields of advertising, and then target the ads to the concerned fields on all web pages with the same contents. At the same time, the calculation of text similarity is also essential in the distinction for text contents similarity of web pages. For different search targets, the search results can be sorted by text length in some algorithms. Text similarity in combination with sort by text length in mass similar search results has powerful value in use. This paper puts forward a Chinese text similarity algorithm based on word difference. After the research of many domestic and foreign literature, and further analysis and research on the current situation of the similarity calculation, put forward a new method of enhancing the performance of similarity-the combination of traditional methods based on statistic and semantic, the combination of the efficiency of method based on statistic and the accuracy of method based on semantic.

Among the methods based on statistics, vector space model is comparatively accurate, which through the statistics of words appear number in the text, treat words as characteristic mapping tospecific space, turn the text into space vector, and get the similarity of two text by calculating cosine for space vectorial angle. In addition, the generalized vector space model, latent semantic indexing model, method based on Attribute Theory, all belong to the calculation method based on statistics [15].

Among the methods based on semantic, there are the representative calculations based on Markov model and calculations based on lexical semantics. The Markov model method tries to combine the word frequency and

word order, be applied into the calculation of text similarity, using the longest common subsequence that reflectets word order, on the basis of the original word frequency, merge them into the compound words weight, perform data dimension reduction by using the state transfer matrix. Calculation based on lexical semantics, combining lexical semantic and text vector by Hownet, calculate text similarity by using the maximum matching algorithm [16].

Combining the advantages of two methods based on statistic and semantic also means having to face the problem of overcoming the shortcomings of the two methods. This paper attempts to take the difference in words number as the breakthrough point,Basing on the diversity of the number of Chinese words, combing with word frequency, number and the semantic of words, extend word similarity computation based on the Hownet to the text similarity calculation.

B. Further Work

Because of the limited research time and scientific research principles, this method also has some problems, which need further improvement.

- Statistics for word string combination, notional words, function words and the special identifier after Chinese text segmentation need further optimization, especially the notional word frequency-an important parameters for the text keyword coincidence, which can effectively improve computational efficiency.
- Lexical similarity calculation based on Hownet can be further strengthened by trying other semantic database or concurrent using of various semantic database, To provide more accurate, detailed meanings for lexical semantic switch.

All of the above is a pity this subject research, solution to the above problems and new attempt, may bring new calculation to similarity.

REFERENCES

[1] H. X. Li, "Study on the application of text classification retrieval based on complex network and ICA algorithm," *Journal of Multimedia*, vol. 8, no. 4, pp. 372-378, 2013.

[2] Y. Niu, and Y. C. Chen, "Research on the text length's effect of the text similarity measurement," ISIA 2010, *Springer*, pp. 112-117, 2010.

[3] Y. C. Chen, "Study of Chinese Text Similarity Based on Number Difference," Master Thesis, *Hubei University of Technology*, 2011.

[4] R. Sabirat, "Genetic approach for a flexible cell phone keypad with reduced keystrokes and key jamming for better human technology interaction," *Journal of Multimedia*, vol. 7, no. 5, pp. 341-352, 2012.

[5] S. Feng, Q. Wu, J. Wang, Y. Xu and G. Ding, "An Introduction of Cognition Information:From Form Aspect to Semantic Aspect," *Proceedings of 2013 22nd Wireless and Optical Communication Conference*, IEEE Chongqing University of Posts and Telecommunications, pp. 481-485, 2013.

[6] Z. Liu, X. Chen, "A Graph-Based Text Similarity Algorithm," *2012 National Conference on Information Technology and Computer Science*, Service Academy Conference Center(SACC-China), pp. 614-617, 2012.

[7] N. Guarino. "Formal ontology and Information Systems," *IOS Press*, pp. 3-15, 1998.

[8] S. Bin, L. Lin, "A Method of Computing the Semantic Similarity of Sentences Based on HowNet," *Computer & Digital Engineering*, vol. 2, no. 42, pp. 187-189, 2014.

[9] B. Jin, Y. J. Shi, H. F. Teng, "Text similarity algorithm based on semantic understanding," *Journal of Dalian*, vol. 45, no. 2, pp. 291-297, 2005.

[10] S. Dumais, "Improving the retrieval of information from external sources," *Behavior Research Methods, Instruments & Computers*, vol. 2, no. 23, pp. 229-236, 2001.

[11] G. Yu, Y. J. Pei, Z. Y. Zhu and H. Y. Chen, "Research on Text Similarity Based on The Semantic Calculation," *Computer Engineering and Design*, vol. 27, no. 2, pp. 241-244, 2006.

[12] Q. Liu and S. J. Li. "Word Meaning Similarity Computation Based On Hownet," *The Third Chinese Lexical Semantics Conference Collection*, pp. 59-76, 2002,

[13] K. Zhang, J. Luo and X. Chen, "Text Similarity Computing Based on Sememe Vector Space," *2013 IEEE 4th International Conference on Software Engineering and Service Science*, *IEEE IEEE Beijing Section*, pp. 208-211, 2013.

[14] Bin Li, Yuan Guoyong, "Improvement of TF-IDF Algorithm Based on Hadoop Framework, " *Computer Application and System Modeling International Conference Proceedings, Computer Science and Electronic Technology International Society*, pp. 0391-0393, 2012.

[15] B. Ning and J. Zhang, "Research on web information retrieval based on vector space model," *Journal of Networks*, vol. 8, no. 3, pp. 688-695, 2013.

[16] Z. Zhang and X. Yang, "Text Similarity Calculation Based on Language Network and Semantic Information," *Computer Engineering and Applications*, vol. 05, no. 804, pp. 33-38. 2014.

Sports Video Segmentation using Spectral Clustering

Xiaohong Zhao, Yanhua Qu, and Hong Zhang
Hebei Normal University of Science & Technology, Qinhuangdao, Hebei Province, 066004

***Abstract*—With the rapid development of the computer and multimedia technology, the video processing technique is applied to the field of sports in order to analyze the sport video. For sports video analysis, how to segment the sports video image has become an important research topic. Nowadays, the algorithms for video image segmentation mainly include neural network, K-means and so on. However, the accuracy and speed of these algorithms for moving objects segmentation are not satisfied, and easily influenced by the irregular movement of the object and illumination, etc. In view of this, this paper proposes an algorithm for object segmentation in sports video image sequence, based on the spectral clustering. This algorithm simultaneously considers the pixel level visual feature and the edge information of the neighboring pixels to make the calculation of similarity is more intuitive and not affected by factors such as image texture. When clustering the image feature, the proposed method: (1) preprocesses video image sequence and extracts the image feature. (2)Using weight function to build and calculate the similar matrix between pixels. (2) Extract feature vector. (3) Perform clustering using spectral clustering algorithm to segment the sports video image. The experimental results indicate that the method proposed in this paper has the advantages, such as lower complexity, high computational effectiveness, low computational amount, and so on. It can get better extraction effects on video image.**

***Index Terms*—Video Processing; For Sports Video Analysis; Sports Video**

I. INTRODUCTION

With the development of computer technology and multimedia technique, video analysis, as important research topic in the field of computer vision, has been widely used in image retrieval, object detection, image processing, and other fields. On the other hand, sports are usually transferred through video, therefore video processing technique then was applied to the sports. The image segmentation of athletes in the sports video has become the research focus in the digital image processing. At present, a lot of segmentation algorithms for motion objects are difficult to obtain good effects in sports video. The performance and speed of segmentation is not satisfied [1]. Therefore, the research on image segmentation method according to or approaching to the characteristics of sports objects has become the key issue.

The purpose of image segmentation is to divide the input image into some space adjacent, spectral similar homogenous regions and separate the target and background. As an important topic of image processing, the result of image has great influence to understand and analyze the subsequent image, pattern recognition etc. [2]. Among numerous image segmentation algorithms, threshold based segmentation and clustering based segmentation are two most popular algorithms in the field of image segmentation. Threshold based segmentation is to make use of intelligence algorithms as the optimizer and calculate optional value of the objective function under a certain criterion in an iterative way, obtaining the optimal threshold of split image [3]. Clustering based segmentation is the combination of optimization algorithm and image segmentation technique based on clustering, avoiding the local optimum problem as far as possible, and obtaining the optimal clustering as soon as possible. Then is achieves the efficient image segmentation [4]. Specifically, it extracts significant features on demand of the image or regional information. These features include gray value of pixel, object contour, reflection characteristics, vein, color, and such original features, and also include spatial spectrum histogram feature. Finally divide the image into the disjoint and consistent different areas, to make the features of these areas exist differences.

Although a variety of image segmentation methods have been developed, there is still no all-purpose and universal method. The image segmentation methods mainly include watershed algorithm, mean shift filtering algorithm, fuzzy theory, graph theory and active contour model, and so on [5]. The watershed method can achieve good results under certain conditions, but it is sensitive to noise and texture. Further, it does not consider such information as the difference of the area when segmenting the image. For mean sift filtering, it needs to build regional connection diagram, and then segment the image by the method of minimum spanning tree. Because of its simple principle, without pretreatment, less parameters, and many other advantages, it can achieve image adaptive filter and filter out the image noise, in the meantime adaptively reducing the smooth of significant edge information in local structure, It is a kind of edge-preserving filtering method. But this method just considers the gray variance of adjacent region, without fully considering other information.

The fuzzy theory based image segmentation method can express the uncertain degree of fuzziness through some certain collection. The consolidation of this method can express the fuzziness and randomness of human visual better and get a certain application in image

doi:10.4304/jmm.9.7.873-878

processing. However, it is usually inaccurate, changeful knowledge or information. . Graph theory based image segmentation method is to map the image into a weighted graph. Finally, according to the objective function, looking for which has the greater similarity among the image clustering, and the optimal partition with smaller similarity to segment the image. Active contours model based image segmentation method is a kind of dynamic two-dimensional closed curve model [6]. The principle of that is to use the combined action of external forces and content to extract the edge of the target, and achieves the convergence of the edge of the target through active contours, so as to realize the segmentation of the target image [7]. Nevertheless, this method has certain limitations, that is, need to set the closer desired solutions during initialization. For convergence, it is easy to cause the convergence of local solutions. For deep sunk local area, it is difficult to realize the local feature extraction due to the high computational complexity. Eventually it is hard to realize the real segmentation of target image.

Spectral clustering is an algorithm which can divide any-shape based on spectral graph theory. In recent years, it has been successfully applied in the fields of bioinformatics, information retrieval, image segmentation, and so on. Because the spectral clustering is a kind of clustering method based on the similarity, and graph theory, it needs to calculate the similarity between each pair of pixels in the process of image segmentation, which have larger space complexity and computational complexity. The use of spectral clustering for image segmentation can use similar matrix feature vector which reflects the similarity between data detect the internal structure of the data, and solve them with the standard linear algebra method. Compared with the traditional clustering methods such as K-means, it can cluster any sample space in any shape. Its essence is the transformation from the clustering problem into the optimal partitioning problem. Spectral clustering algorithm is to build an undirected graph and then make a multi-channel division. The principle of optimal partition is the largest similarity of in-house sub image and the minimal similarity between sub images. An effective solving method is to convert the original problem to the problem of eigenvalue and feature vector which can solving the Laplace matrix. It adopts the continuous relaxation form. Using these feature vectors construct a simplified data space, which make the data distribution structure more obvious in this space. The mean shift of spectral clustering (MSSC) algorithm [8] has the automaticity and accuracy, but lack rapidity, and has higher complexity. It combines mean shift and spectral clustering algorithm for image segmentation, and focuses on the analysis of the influence of the segmentation results of different kernel functions, without considering improvements of spectral clustering algorithm.

The image space characteristics in the conventional spectral clustering algorithms are not considered, but construct similarity matrix by using the grey value or the wavelet coefficients. Therefore, it is susceptible by oversize similarity matrix and scale parameters in the similarity measure. When image resolution is high, the spectral clustering method will lead to overlarge adjacency matrix which is beyond the scope of computer memory caused by overmuch vertex. To solve this problem, there are mainly three kinds of solution applied to spectral clustering algorithm [9]: (1) make the adjacency matrix turn into the sparse matrix in a certain way, and then calculate it. This method mostly uses the KNN algorithm [10]. (2)Pre-segment the image. That is, dividing the image into lots of small homogenous areas, taking them as new vertexes and using spectral clustering algorithm to get the final results of segmentation. This method usually uses watershed segmentation algorithm or K-means algorithm. (3) Use spectral clustering algorithm based on Nystrom estimates.

In view of the disadvantages of slow speed of existing motion object segmentation and the accuracy of the segmentation is susceptible by irregular motion of the object and the illumination. Given this, this paper proposes a new video sport segmentation method based on spectral clustering to improve the quality of image segmentation and reduce computational complexity. This algorithm also considers its own characteristics of the pixel and the visual feature and edge information of the neighboring pixels at the same time to make the calculation of similarity is intuitive and not affected by factors such as image texture. The proposed method is composed of three steps: (1) preprocess the extraction of video image sequence, that is using weight function build and calculate the similar matrix between pixels. (2) Extract the image features over time sequences. (3) Clustering the feature vector using spectral clustering for image segmentation, and then realize the processing of image difference, to get moving objects.

The experimental results indicate that the method proposed in this paper has the following advantages. (1) Low computational complexity which enables it be applied to large scale problem. (2) High time-validity with which can be applied to real applications. (3) It achieves state-of-the-art performance on the standard evaluation dataset. Further, it can be potentially applied to other tasks of image segmentation and sports analysis.

II. SPECTRAL CLUSTERING FOR SPORTS VIDEO ANALYSIS

Motivated by the limitations of existing approaches and the above discussions, in this section, we will present a substituted approach which is based on spectral clustering. Firstly, we will give the description of the spectral clustering algorithm. Then the overall approach is shown in Figure 1 [11].

The spectral clustering algorithm [7] first constructs a similarity matrix $S \in R^{n \times n}$ for n given data points $x_1, ..., x_n$ where $S_{ij} \geq 0$ measures the relationship between x_1 and x_j. Then, group these data points into k clusters using similarity information. Because several variants of spectral clustering are proposed, we consider the commonly used normalized spectral clustering [8].

See [9] for the detailed survey of spectral clustering. An example similarity function is the Gaussian:

$$S_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (1)$$

where σ is a scaling parameter for controlling how rapidly the similarity S_{ij} reduces along with the distance between x_i and x_j. Here we consider the following normalized Palladian matrix [7]:

$$L = I - D^{-1/2} S D^{-1/2} \quad (2)$$

where D is a diagonal matrix with $D_{ii} = \sum_{j=1}^{n} S_{ij}$. Note that $D^{-1/2}$ is the inverse square root of D.

It is easily to give the explanation for any S with $S_{ij} \geq 0$. The Laplacian matrix is symmetric positive semi-definite. In the ideal case, the data in one cluster have no relationship with those in others, nonzero elements of S and hence L can only appear in a block diagonal form:

$$L = \begin{bmatrix} L_1 & & \\ & \ddots & \\ & & L_k \end{bmatrix}$$

As we all known that L has k zero-eigenvalues, which are also the k smallest ones [9]. Their matching eigenvectors, written as a $R^{n\times k}$ matrix, are

$$V = [v_1, v_2, ..., v_k] = D^{1/2} E \quad (3)$$

where $v_i \in R^n, i = 1,...k,$ and $E = \mathrm{diag}(e_1, \cdots, e_k)$, where e_i in different length are vectors of all ones. Because $D^{1/2}E$ and E have the same structure, one can easily cluster the n ones into k groups through the simple clustering algorithms such as K-means. Therefore, one has to find the first k eigenvectors of L, that is, the eigenvectors corresponding to the k smallest eigenvalues. Whereas, in fact we obtain the eigenvectors in the form of $V = D^{1/2}Q$, where Q is an orthogonal matrix. [8] give a proposition for normalizing V,

$$U_{ij} = \frac{V_{ij}}{\sqrt{\sum_{r=1}^{k} V_{ir}^2}}, \; i = 1,...,n, \; j = 1,...,k. \quad (4)$$

Each row of U has a unit length. Because of the orthogonally of Q, Equation (4) is equal to

$$U = EQ = \begin{bmatrix} Q_{1,1:k} \\ \vdots \\ Q_{1,1:k} \\ Q_{2,1:k} \\ \vdots \end{bmatrix} \quad (5)$$

where $Q_{i,1:k}$ displays the i^{th} row of Q. Then the rows of U correspond to k orthogonal aims on the unit sphere. By means of the K-means [8] or other simple techniques [10, 11], we can easily and effectively cluster the n rows of U.

The spectral clustering algorithms also can be derived from the point of view of graph cut, apart from analyzing properties of the Laplacian matrix. That is, in accordance with the relationship between points, we can partition the matrix. Normalized Cut, Min-Max Cut, Ratio Cut [12] are representative graph-cut methods.

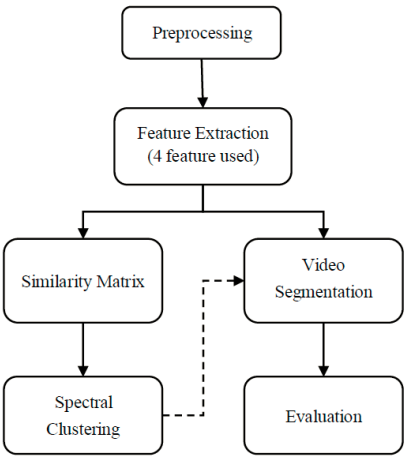


Figure 1. The flowchart of the proposed approach

III. EXPERIMENTAL RESULTS

This section will empirically assess our proposed spectral clustering based on spectral clustering for sport video analysis. The experiment steps are as follows. First, collect data according to experiment design; second, model the data and extract the feature; third, train the model and evaluate its performance. The experimental procedures are shown in Figure 1. This section will sequentially summarize the dataset, assessment standard and experimental results [12].

A. Experimental Dataset

The Vicon Physical Action data used in the experiments were collected from Theo the odorid is of University of Essex. It is used to separate for each physical activity. Seven male and three female subjects of age 25 to 30, who have experienced aggression in scenarios such as physical fighting, took part in the experiment. With area 4x5.5m, the ten subjects expressed normal and aggressive physical activities at random locations. The duration of each action was approximately 10 seconds per subject, which corresponds to a time series of 3000 samples, with sampling frequency of 200Hz. Within this performance time, approximately 15 action trajectories were extracted counting in average 15 normal, and 15 aggressive actions. The dataset includes 3000 samples with 27 attributes, and these attributes describes as Head, Left arm, Right arm, Left leg, Right leg, Marker with a pair of markers is attached at each body segment for 3D data acquisition. Coords with the 3

coordinates define the 3D position of each marker in space. The categories are shown in Table I.

TABLE I. CATEGORY DISTRIBUTION OF SAMPLES

Categories	Number of samples
Running	800
Jump	1200
Long jump	1000
Total number of samples	3000

B. Evaluation Criterion

In order to comprehensively validate the approach proposed in this paper and the related algorithm, in the experiment of sport video analysis. The classification accuracy and classification recall are take advantage of as the verification standard [13]:

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{tn} + \text{fp} + \text{fn}}$$

$$\text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}}$$

where tpindicatestrue positive which expresses the positive data with true label; tnrepresentstrue negative that expresses the negative example with true label; fprepresentsfalse positive; and fnimpliesfalse negative. Although other standard like recall can also be take advantage of to verify the approach for sport video analysis, accuracy is the most widely used evaluation standard at present. Considering two class identification problems of sport video analysis, the two categories of data are respectively marked as positive instance and negative data. The above equation is defined as the ratio of all true examples on all the test data. For multiple class identification problems of sport video analysis, they can be directly converted to a group of two class identification problems of sport video analysis.

C. Main Results

In the first experiment, we evaluate our proposed spectral clustering algorithm for sport video analysis, over the Vicon physical action dataset. We use two comprehensive standards, accuracy and recall, for experimental assessment. Identification accuracy and recall are two typical and popular measures for the correctness of the identification model. The experimental procedure is report in the experiment section. The preprocessing procedure and feature extraction step are important to experiment since they carry discriminant information. The proposed method spectral clustering is train using the above approach, and some parameters of spectral clustering are obtain by two strategy: hand-specification strategy and cross-validation strategy. We do the test multiple rounds, where in each round we randomly split the dataset to training set as well as test set [14].

The test experiment of the proposed spectral clustering method is run for 5 rounds over the randomly divided training sets and test sets. The recognition performance for different configuration is report in Table II. From the experiment results, we could find that, for different experimental setting, for distinct evaluation standard, accuracy and recall, the proposed approach is higher than GMM, where our approach is within the range of 77.08%~84.62% and 74.99%~81.62% respectively. We also find that the average accuracy is 80.58. These results are consistent with the previous paper, which mean that accuracy is a reliable measure for sport video analysis using spectral clustering [15]. The reasons for these results are mainly threefold as follows. (1) The spectral clustering method can be applied to the conditions that data is large scale, high dimension, with a large number of heterogeneous information. (2) The cross-validation method is according to the distribution information of the input data to select the model parameters of the spectral clustering, which makes the spectral clustering having better adaptability. (3) The processing step for data is able to remove noise and keep useful information effectively, and the element steps of our algorithm could cooperate.

TABLE II. THE IDENTIFICATION RESULTS OF SPORT VIDEO ANALYSIS USING SPECTRAL CLUSTERING

Experimental round	Evaluation standard	Approach	
		GMM	Spectral clustering (ours)
1	Accuracy	75.30	80.75
	Recall	72.76	74.99
2	Accuracy	75.46	77.08
	Recall	73.73	75.52
3	Accuracy	74.71	81.26
	Recall	74.76	77.19
4	Accuracy	73.20	80.10
	Recall	75.52	77.79
5	Accuracy	76.13	80.91
	Recall	71.95	81.62

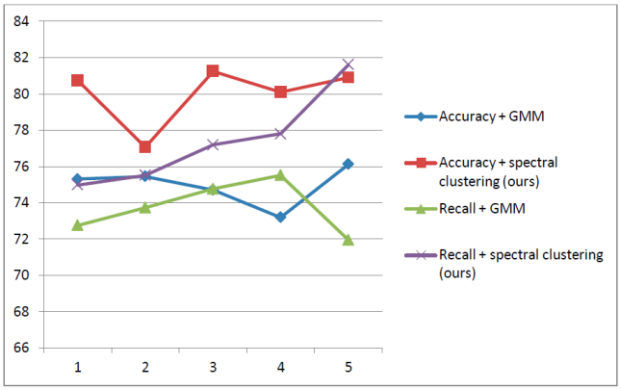


Figure 2. The identification results of sport video analysis

In the second experiment, we target to verify the advantage of our proposed spectral clustering for sport video analysis, and the superiority of the feature extraction algorithm. The experiments are performed over Vicon physical action dataset. The data were collected from human real action, by Theo Theodoridis of University of Essex. The dataset includes 3000 samples with 27 attributes, which used to separate for each physical activity. The used evaluation standards are accuracy and recall where Identification accuracy and recall are two typical and popular measures for the

correctness of the identification model. The experimental step is summarized in the above section, where the spectral clustering is train using the standard approach, with the parameters set by authors.

We extensively compare our proposed spectral clustering approach for sport video analysis with three algorithms, K-means, GMM and IGMM. The results are summarized in Table II. These experimental results indicate that: (1) the proposed spectral clustering method outperforms all three compared algorithm significantly, for different experimental configurations, different number of training sample, and distinct evaluation criterions. (2) The proposed algorithm show robustness against the round of experiments, as well as the evaluation criterion, which no wonder suggests that the proposed approach could be used to a lot of tasks. The possible reasons are three folds. (1) The spectral clustering method can be applied to data with large scale, high dimension, and heterogeneous information. (2) The parameter selection approach is according to the distribution information of the input data to choose the model parameters of the spectral clustering, which makes the spectral clustering having better adaptability. (3) The framework of the proposed approach contains a group of comprehensive procedures which sequentially maximizes the identification ability.

TABLE III. PEFFORMANCE COMPARISION FOR SPORT VIDEO ANALYSIS

Experiment round	Approach	Evaluation Criterion	
		Accuracy (%)	Recall (%)
1	K-means	74.64	71.26
	GMM	75.30	72.76
	IGMM	79.11	73.78
	Spectral clustering	80.75	74.99
2	K-means	72.84	68.75
	GMM	75.46	73.73
	IGMM	78.01	70.94
	Spectral clustering	77.08	75.52
3	K-means	72.64	68.90
	GMM	74.71	74.76
	IGMM	75.83	73.05
	Spectral clustering	80.58	77.35

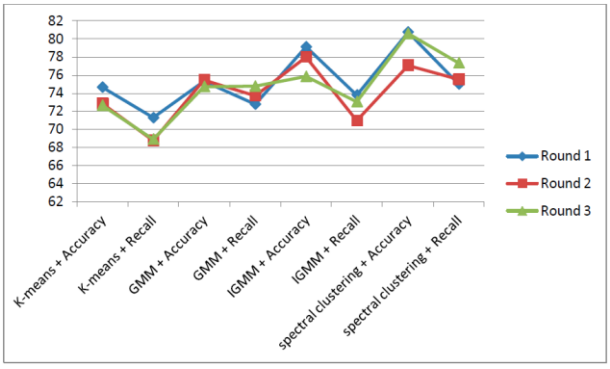


Figure 3. Pefformance comparision for sport video analysis

In the third experiment, we extensively compare our proposed spectral clustering algorithm for sport video analysis with three approaches, K-means, GMM and IGMM. The identification results are present in Table III. These experimental results show that: (1) the proposed spectral clustering outperforms all three compared algorithm significantly, under varying experimental settings, different number of training example, and varying assessment standards. (2) The proposed algorithm exhibits robustness against the round of experiments, and the evaluation standard, which obviously suggest that the proposed approach could be utilized to a number of tasks. The reasons are three folds. (1) Comparing with the traditional machine learning methods the spectral clustering can be well applied to the situation of large scale data, high dimension, and a large number of heterogeneous information. (2) The parameter selection approach is based on the distribution information of the input data to select the model parameters of the spectral clustering, which enables the spectral clustering having better adaptability. (3) The framework of the proposed approach is composed of a set of comprehensive procedures which maximizes the recognition ability sequentially.

TABLE IV. THE IDENTIFICATION RESULTS OF SPORT VIDEO ANALYSIS USING SPECTRAL CLUSTERING

Training samples	Evaluation Standard	Approach	
		Spectral clustering	GMM
30%	Accuracy	72.33	67.37
	Recall	69.50	62.33
40%	Accuracy	75.28	70.74
	Recall	74.13	65.30
50%	Accuracy	81.69	73.26
	Recall	80.01	70.74
60%	Accuracy	83.52	77.55
	Recall	84.18	76.59
70%	Accuracy	90.09	82.30
	Recall	86.05	79.21

We conduct experiments over Vicon physical action. The data were collected from human real action, by Theo Theodoridis of University of Essex. The dataset includes 3000 samples with 27 attributes, which used to separate for each physical activity. This experiment will verify the advantage of spectral clustering in sport video analysis, as well as optimization. It adopts the approach present in above part to train spectral clustering. The evaluation criterions are respectively accuracy and recall. The test was conduct for 10 rounds on this method, and the overall results of distinct experimental configuration are summarized in figure 3. As report in figure 3, the value of accuracy is around 81.69%, which consistently outperforms other compared approaches. Further, in different experimental rounds, the accuracy of the proposed approach also outperforms other compared algorithm. These results are consistent with the previous paper, which suggest that accuracy is a reliable measure for sport video analysis and spectral clustering. The reasons are from the following three aspects. Firstly, the spectral clustering method can be applied to the large scale and uneven distributed data. Secondly, the parameter selection approach is on the basis of data distribution of the input data to determine the model parameters of the spectral clustering, which improves the adaptability of spectral clustering. Thirdly, the

experimental procedures of the proposed algorithm could provide informative features and could maximize the discrimination ability.

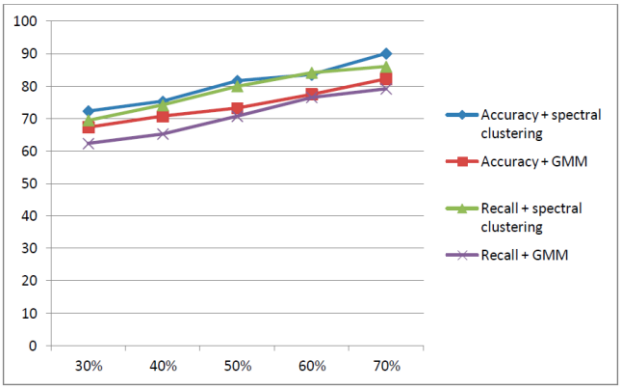


Figure 4. The identification results of sport video analysis

IV. CONCLUSION

This paper adopts the method of spectral clustering to implement image segmentation technique in sports video image. The process mainly includes the following steps: (1) Preprocess the sports video image and get image feature vector matrix. (2) Extract the color feature, texture feature and shape feature etc. of the image. In the process of each extraction of image features, using different feature extraction technique can realize feature extraction under certain conditions, such as the extraction of texture feature can be treated with wavelet, and then take the extracted feature information as the image feature and input during the processing of image classification. (3) Calculate the image similarity with the extracted feature matrix. Finally the experimental results show that: (1) The algorithm is compared with several typical algorithms. It can be seen that spectral clustering considering the pixel itself feature, visual feature and edge information, which can greatly improve the accuracy. (2) Compared with the existing methods, the proposed algorithm has low complexity and strong real-time capability. (3) Finally get the image features extraction through K-means clustering algorithm for image segmentation. (4) With regard to the higher time and space complexity of image segmentation of motion image sequence, the proposed algorithm can better solve the problem of processing large image data and reduce the time and space complexity of clustering in the image segmentation. The proposed algorithm can better extract the sports video image sequence, but still have a lot of difficulties in view of the feature extraction of image segmentation. So how to improve the extraction of more image feature information is still the focus of the further research.

REFERENCE

[1] AdrienBartoli, NavneetDalal, Biswajit Bose, Radu Horaud from Video Sequences to Motion Panoramas, 2002.
[2] Y. Jiang, The Research of Sports Video Object Tracking Based on Mean-Shift and Color Histogram Tracking Algorithm, *Journal of Soochow University Engineering Science Edition*, Vol. 2, No. 32, pp. 33-36. 2012
[3] B. Song, G. Song, Study and Implementation of a Video-Based System for Athletic Movement Analysis, *Electronic Science and Technology*, Vol. 11, No. 22, pp. 76-80, 2009.
[4] M. He, A Study of the Data-Collection Technology of the Video-Image in Motion in Track and Field Events, *Bulletin of Sport Science &Technology*, Vol. 2, No. 21, pp. 130-132, 2013.
[5] X. Cai, G. Dai, L. Yang, Survey on Spectral Clustering Algorithms, *Computer Science*, Vol. 7, No. 35, pp. 14-18. 2008.
[6] J. Li, J. Zhou, J. Guan, A survey of clustering algorithms based on spectra of graphs, *CAAI Transactions on Intelligent Systems*, Vol. 5, No. 6, pp. 405-412. 2011.
[7] X. Ma, L. Jiao, Image Segmentation Based On Watershed and Spectral Clustering, *Journal of Infrared and Millimeter Waves*, Vol. 6, No. 27, pp. 452-456. 2008.
[8] A. Y. Ng, M. I. Jordan, and Y. Weiss, On spectral clustering: Analysis and an algorithm. *In Proceedings of NIPS*, pp. 849-856, 2001.
[9] U. Luxburg, A tutorial on spectral clustering. *Statistics and Computing*, Vol. 17, No. 4, pp. 395-416, 2007.
[10] S. X. Yu and J. Shi, Multiclass spectral clustering, *In Proceedings of ICCV*, pp. 313, 2003.
[11] H. Zha, C. H. Q. Ding, M. Gu, X. He, and H. Simon. Spectral relaxation for k-means clustering. *In Proceedings of NIPS*, pp. 1057–1064, 2001.
[12] L. Hagen and A. Kahng, New spectral methods for ratio cut partitioning and clustering, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 11, No. 9, pp. 1074-1085, 1992.
[13] Jian Wu, Zhi-ming Cui, Jian-ming Chen, Guang-ming Zhang, A Survey on Video-based Vehicle Behavior Analysis Algorithms, *Journal of Multimedia*, Vol. 7, No. 3, pp. 223-230, 2012
[14] Wenhui Li, Yifeng Lin, Bo Fu, Mingyu Sun, Wenting Wu, Cascade Classifier Using Combination of Histograms of Oriented Gradients for Rapid Pedestrian Detection, *Journal of Software*, Vol. 8, No. 1, pp. 71-77, 2013
[15] Aiying Lu, Wenrui Ding, Hongguang Li, Multi-information Based Safe Area Step Selection Algorithm for UAV's Emergency Forced Landing, *Journal of Software*, Vol. 8, No. 4, pp. 995-1002, 2013

Application of Tunnel Seismic Image Approach to the Advanced Geological Prediction for Tunnel

Guo Jinmin and Luo Congshuang
School of Transportation Engineering, Henan University of Urban Construction, Pingdingshan, Henan, China

***Abstract*—With the large-scale construction and rapid development of underground engineering, a large number of underground engineering construction, such as tunnel and subway emerged. The geological structure of the tunnel is complex and the roc is fragmented. Tunnel seismic image is a new geophysical technique. This method has advantages such as high resolution, high reliability and obvious image characteristics. Therefore, to insure the safety of the construction and to eliminate geological disasters, the advanced prediction technology of seismic image in tunnel detection is applied. In the paper, the seismic image was used to detect the Huangzhuang tunnel geological condition. The reflected waves, the refraction wave and the surface waves can reflect the same geological conditions, and the result is in accordance with the drilling. Tunnel seismic image can effectively and safely guide the excavation of the tunnel section working surface in combination with reconstructed images and excavation technology.**

***Index Terms*—Tunnel; Seismic Image; Geological Conditions Reflected Waves; Surface Waves**

I. INTRODUCTION

In recent years, with development of society, economy has increased greatly, the highway, high-speed railway, large water conservancy and other various infrastructure construction of rapid development, specially the tunnels, subways and so on, the large-scale construction of underground projects. Because of the pecularities of tunnel complicated hydrogeological conditions and the characters of underground engineering projects, a series of environmental geotechnical problems, such as fault, ground subsidence, karsts, crushing zone, collapse, high geostress area, water bursting, instability of cave, water irruption and all kinds of bad geological body become more conspicuous and frequent [1-6]. The geologic phenomenon is often encountered during the tunnel construction. And these incidents will not only influence the construction of underground engineering projects progress but also bring severe geological disasters, and even cause serious casualties and economic loss [7-9]. To ensure construction safety and avoid such accidents, the geological conditions and rock mass in front of working face must be done before the excavation.

We should find out the geological diseases that may meet in tunnel excavation process and forecast possible accident and danger under the complicated geological conditions. In order to take timely measures to ensure construction safety, higher application requirements and countermeasures must be prepared beforehand. So, it needs faster development of tunnel prediction technology [10-17]. The advance of geological forecast is of great significance in tunnel construction.

Geological prediction in tunnel construction is a leading international topic and problem all over the world [18-24]. The tunnel geological prediction becomes a difficult problem with high demand and difficult technology for complicated geological conditions and frequent engineering accident of tunnel. Advanced geological forecast methods includes geological analysis, advanced drilling method, the seismic reflection wave method, geological radar method, transient electromagnetic method, land sonar method, seismic wave tomographic imaging method and so on [25-28]. At the current stage, tunnel prediction is mainly based on seismic reflections, assisted by geological radar. In order to effectively predict geological problems in the front of tunnel face using the seismic reflection signals, special prediction technologies such as Negative Apparent Velocity, Horizontal Seismic Profile (HSP), Tunnel Seismic Prediction (TSP), True Reflection Tomography (TRT), and Tunnel Seismic Tomography (TST) have been developed using different observation patterns and data processing techniques [29-31].

This paper discusses the application of seismic methods under the tunnel construction. We use multichannel system enhancement seismographs, which has 48 channels that can meet the test. The seismic image can collect the continuous profile data, and the microcomputer can process these data rapidly.

II. PROJECT SUMMARY AND GEOLOGICAL CONDITIONS

A. Project Summary

Huangzhuang tunnel is a highway tunnel, the starting pile number is k145+030, the tunnel at the end of the pile number is k145+958, its total length is 0.928 km. The road standard is secondary and the design speed is 40 km/h. The clear height is 5m, and the width is 9m. Tunnel portal design is for terminal wall. The tunnel is shallow, buried from K145+256 to k145+418 and the buried depth from 18.5 to 31.3 m.

B. Geological Conditions

Tunnel hole body is mainly through the late yanshanian granite and the import section is mainly

doi:10.4304/jmm.9.7.879-886

cretaceous andesite. The exposed rock from the new to old is Q_4^{el+al}, Q_4^{el+pl} and Ptn.

C. The Basic Principle of Tunnel Seismic Image

Seismic waves, propagating in the tunnel wall rock, spread widely and are reflected and refracted when encountering interfaces between rocks with different acoustic impedances. Reflected waves returning to the receivers are recorded. Surface waves are very strong and spread along both the axial and radial directions. The principle of tunnel seismic image is in Fig. 1. From the below figure, you can see the surface wave (direct wave), reflected wave and refraction wave in the seismic image.

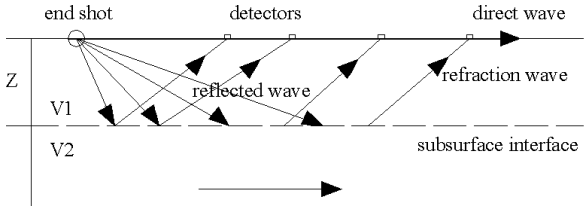


Figure 1. Principle of tunnel seismic image

III. RESULTS AND ANALYSIS

A. The Reflected Waves

The complexity of the reflection seismic waves inside the tunnel is due to surface waves and converted waves near the tunnel surface. The signal waveform of reflection waves observed in the tunnel and its time series can be expressed as [32]:

$$A(t,\ m_i,\ m_j)=\sum^{n} R(m_k)\delta(t-t_k) \tag{1}$$

where $A(t,\ m_i,\ m_j)$ is the time series of the reflected signal recorded at m_i by a shot point at m_j, n is the number of reflection events, $R(m_k)$ is the reflection strength of the point m_k in the front of tunnel face, and is in the range of $-1\le R(m_k)\le 1$, and δ is the travel time of the reflected wave consisting [33]:

$$t_k=\int_{mi}^{mk}\frac{ds}{v}+\int_{mk}^{mj}\frac{ds}{v} \tag{2}$$

The reflection strength $R(m_k)$ is a synthetic variable and is also the arithmetic product of the refraction coefficient of the forward wave (E_{1i}), the refraction coefficient of the backward wave (E_{ij}), and the reflection coefficient of the radial tip (R_K).

$$R(m_k)=E_{11}E_{12}\bullet\bullet\bullet E_{1K-1}R_K E_{2K-1}E_{2K-2}\bullet\bullet\bullet E_{21} \tag{3}$$

Both of the refraction coefficients of the forward and backward waves are in the range 0-1. Therefore, the reflection strength $R(m_k)$ is determined by R_K. Under the condition of small differences of wave impedance among the media and fewer transmission layers, the strength and polarity of the reflected wave $R(m_k)$ is mainly determined by R_K, so approximately [34]:

$$R(m_k)=R_K \tag{4}$$

It represents the differences of the geologic framework in the front of the tunnel face and the main characteristics of geological structure. Seismic migration imaging obtains the positions of the reflection planes in the front of the tunnel face, the strength $R(m_k)$, and velocity distribution V_m. A profile of $R(m_k)$ can be used to express the characteristics of geologic texture and structure and to compute the distribution of every reflection coefficient R_K if needed.

The distribution of reflection strength $R(m_k)$ in the front of the tunnel face can be inverted on the basis of synthetic data [35]:

$$R(m_k)=\frac{1}{LP}\sum_{l=1}^{L}\sum_{j=1}^{p}A(t_i,\ m_j,\ m_1) \tag{5}$$

where L and P represent the number of shot points and receiver points respectively, and t_i is from the travel time equation (2). The distribution of V can give equation (5) extreme values and can be determined by velocity scanning. The seismic reflection of the image map is shown in the Fig. 2. The seismic reflection of the CDP section is in Fig. 3. The seismic reflection of the CDP section energy spectrum is in Fig. 4.

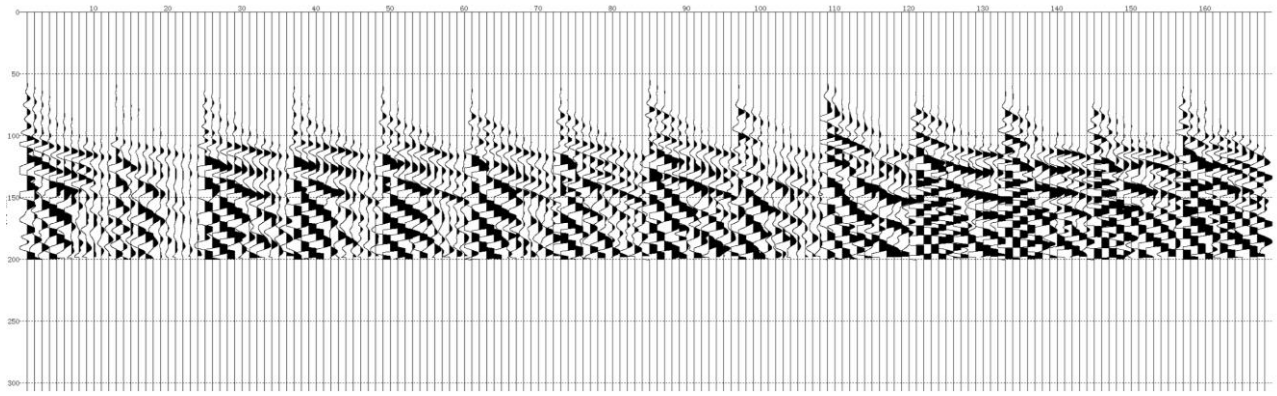


Figure 2. The seismic reflection of the image map

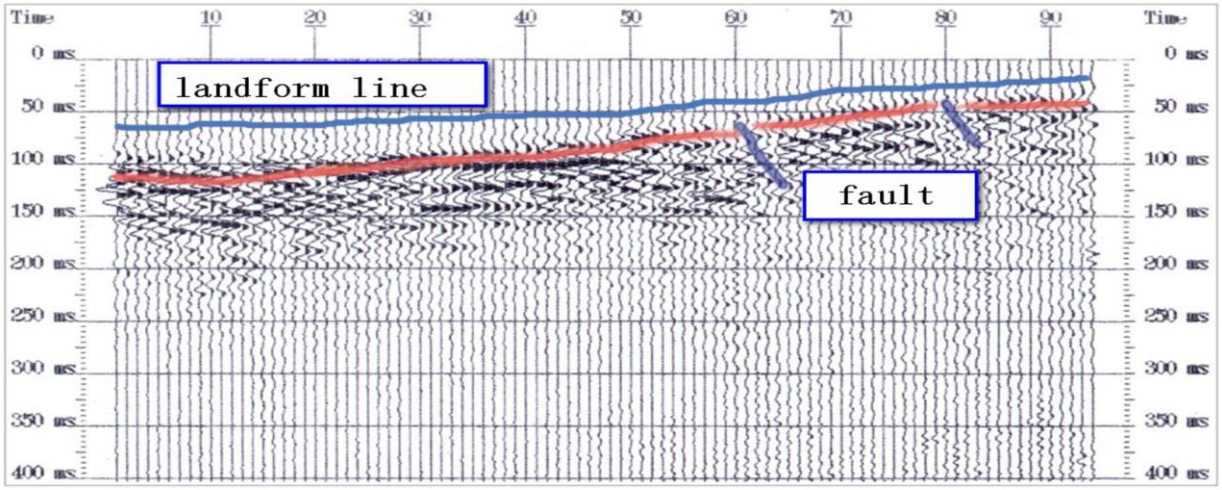


Figure 3. The seismic reflection of the CDP section

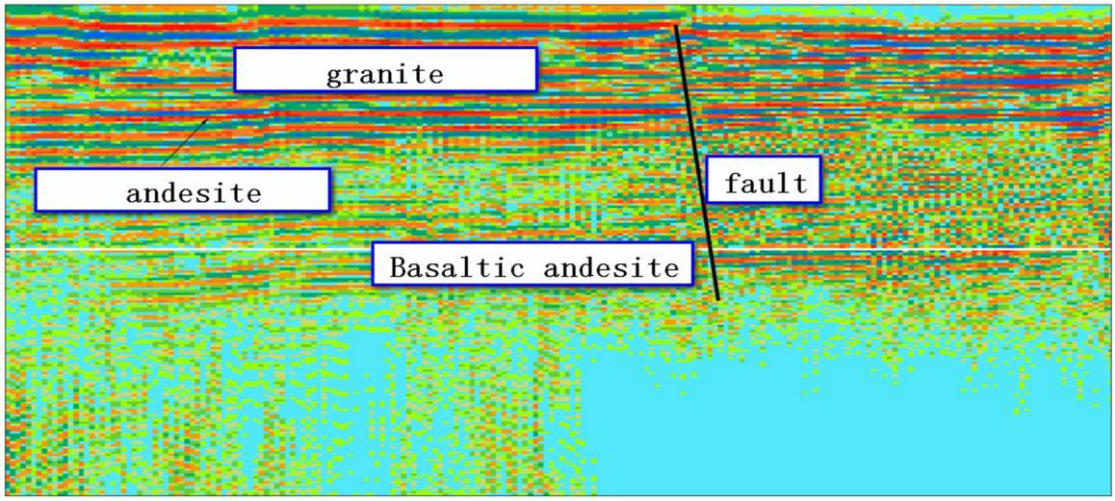


Figure 4. The seismic reflection of the CDP section energy spectrum

From the Fig. 2, Fig. 3 and Fig. 4 we can see that the geological conditions can be predicted before the tunnel excavation. The landform line and the fault are clearly visible. The rock interface can be identify.

B. The Refraction Wave

The single channel seismic image is in the Fig.5.The single channel seismic energy spectrum is in Fig. 6. From Fig. 5 and Fig. 6, we can see that the shallow geological conditions are relatively good. Local presence of cavity and discontinuous medium.

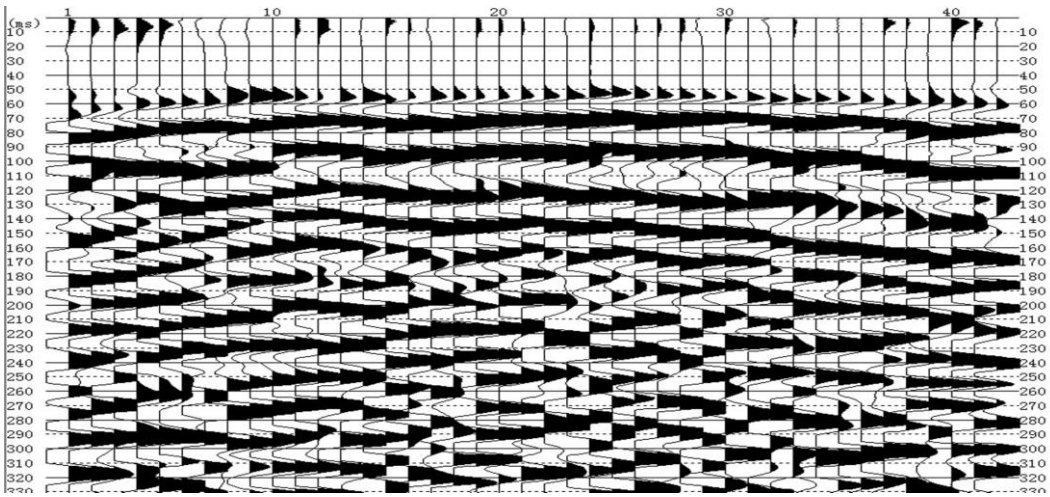


Figure 5. Single channel seismic image

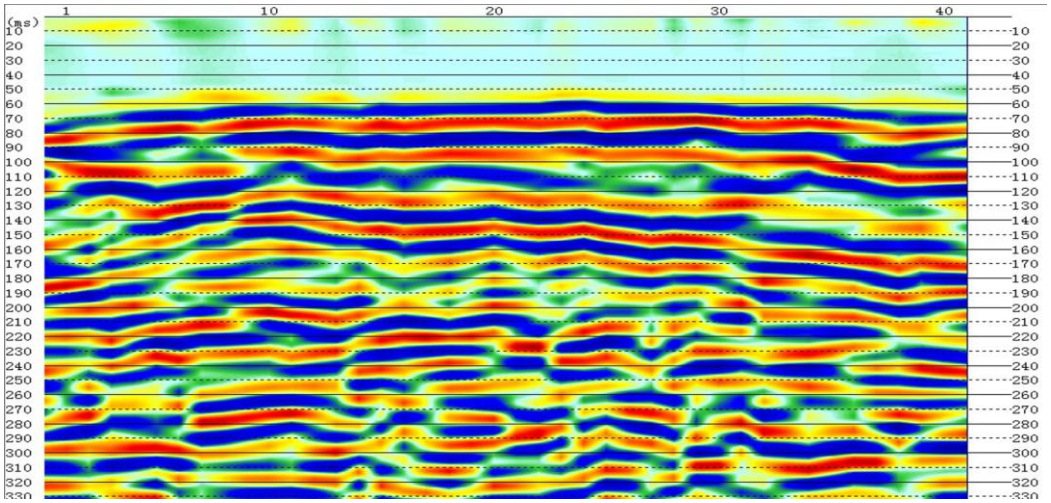


Figure 6. Single channel seismic energy spectrum

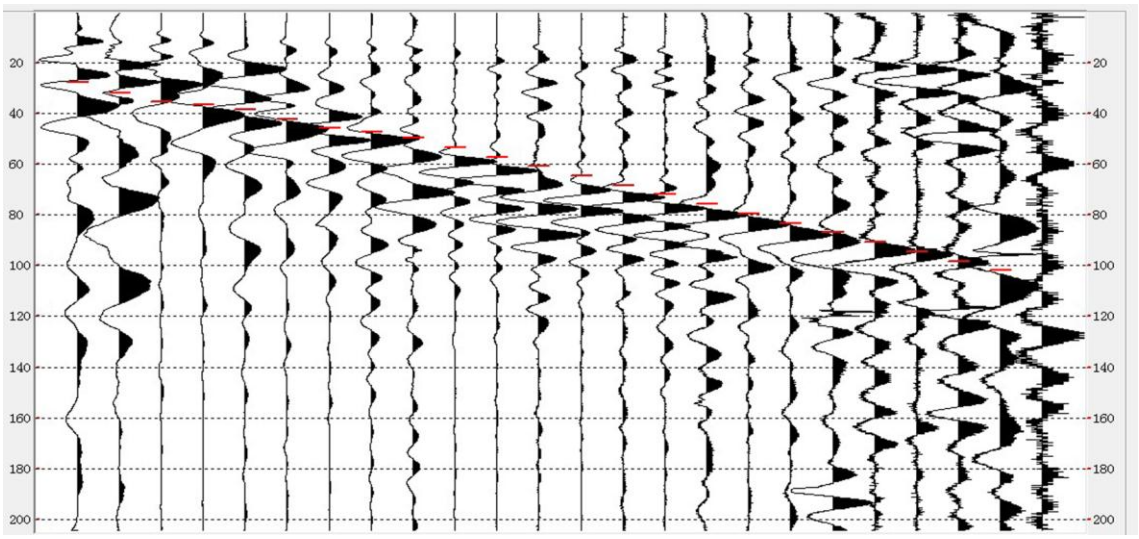


Figure 7. P-wave seismic image map

C. Surface waves (P-waves and S-waves)

Surface waves are very strong and spread along both the axial and radial directions. As the affected depth inside the wall rock is about one wavelength, geophones will record strong surface waves if shallowly buried. P-and S-waves spreading inside the wall rock will generate converted waves on the surface around the tunnel, P-waves are converted to S-waves and S-waves are converted to P-waves. These wave groups are recorded all at the same time [36].

Figures 1 and 2 show wave front profiles of axial and radial sections, from which the travel time characteristics of P-waves, S-waves, and different kinds of converted waves can be clearly seen.

The dependence of P- and S-wave velocities on elastic properties and density are reviewed along with the significance of obtaining both P- and S-wave velocities for engineering purposes.

However, in this case, the particle motion is transverse to the direction of travel of the wave. S-waves are occasionally called transverse waves or shear waves.

Seismic velocity is a function of the density of the earth material and the elastic properties of that material. The expression is as follows:

$$V = (K / \rho)^{1/2} \tag{1}$$

where V is P- or S-wave velocity, K is the effective elastic parameter, andρis the density of the medium, relates velocity, density, and the elastic parameters. K is a function of Lamé's constants, λ and μ, which are related to how a material responds to normal and shearing stresses [37]. The effective elastic parameter for the P-wave velocity is:

$$K = (\lambda + 2\mu)^{1/2} \tag{2}$$

and for S-wave velocities, it is

$$K = \mu^{1/2} \tag{3}$$

Several elastic constants are also expressed in terms of Lam ẽs constants. One of the most important parameters in seismic work is Poisson's ratio σ. If a cubic element of isotropic material is stretched along one axis of the

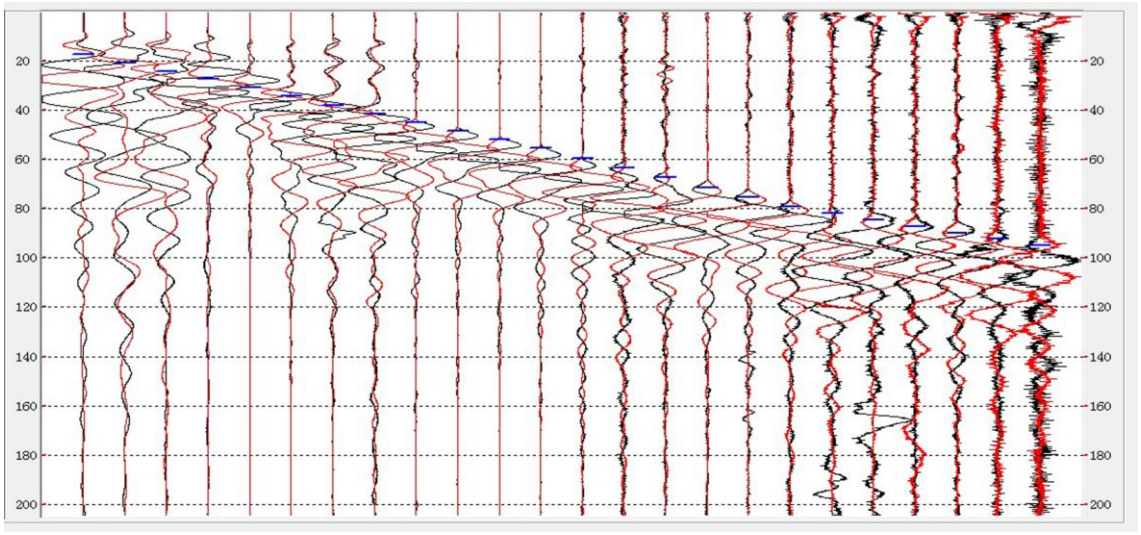


Figure 8. S-wave seismic image map

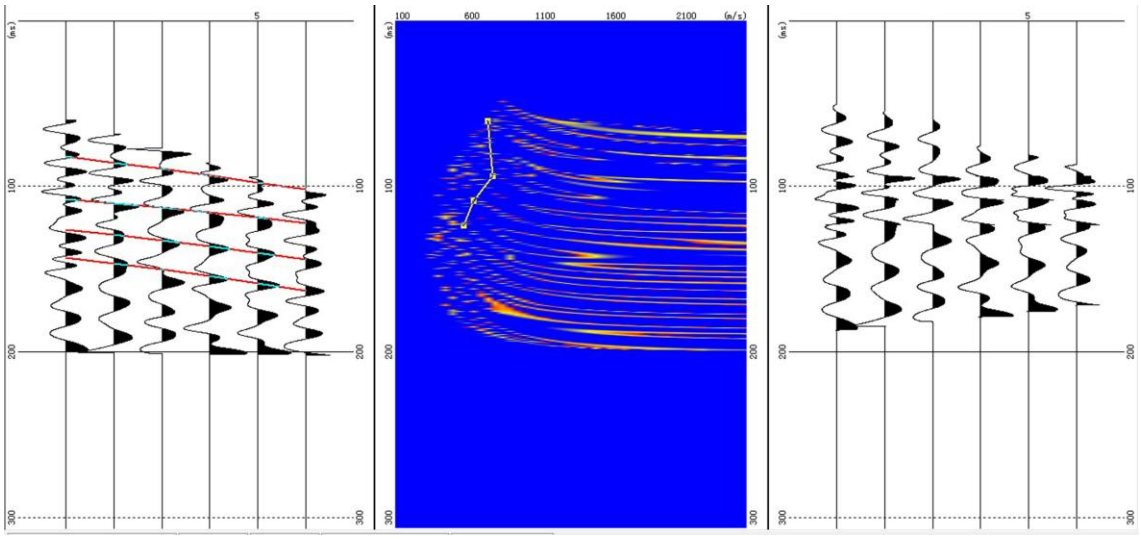


Figure 9. Velocity analysis seismic image

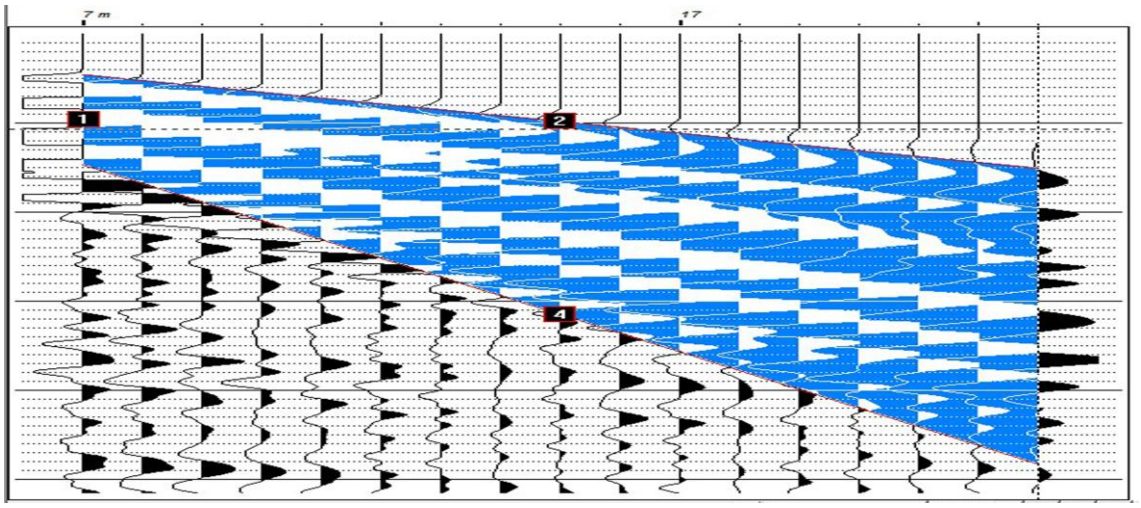


Figure 10. Surface wave seismic image

cube, the dimensions of the cube along the other two axes will decrease. This diagrammatically in two dimensions and defines the strains ε_x and ε_y. Poisson's ratio is

$$\sigma = \varepsilon_y / \varepsilon_x \quad (4)$$

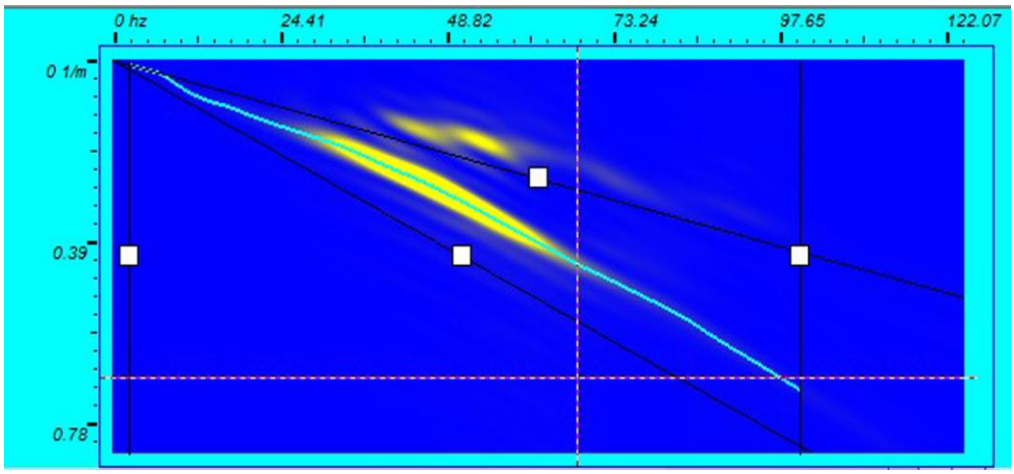


Figure 11. Surface wave seismic image energy spectrum

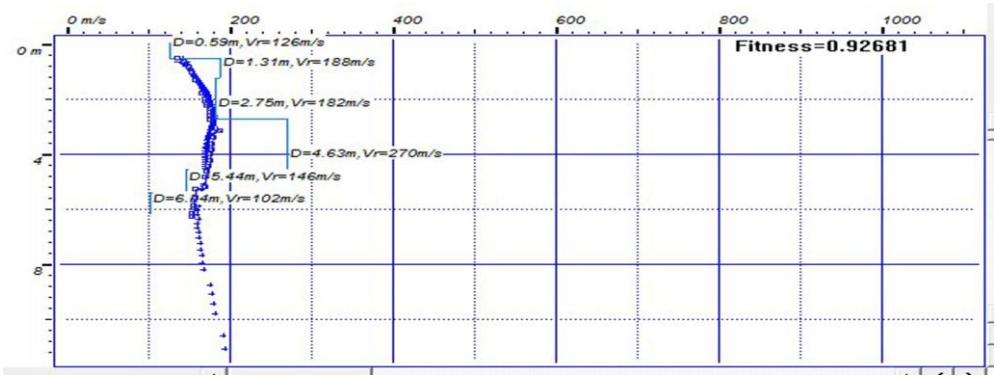


Figure 12. Velocity of Surface wave

TABLE I. SURFACE WAVE VELOCITY ANALYSIS

Serial number	Depth(m)	Level number	P(m)	S(m)	P(m/s)	S(m/s)	Density (g/cm3)	Poisson ratio	Shear modulus (Mpa)	Young modulus (Mpa)	Bulk modulus (Mpa)
1	77.0	1	27.9	17.6	2760.8	4376.5	2.650	0.0	50756.9	0.0	-3055.0
2	78.0	2	32.2	20.9	232.6	303.1	2.650	0.0	243.5	0.0	-100.0
3	79.0	3	35.6	24.3	294.2	294.2	2.650	0.0	229.4	0.0	-0.4
4	80.0	4	36.8	27.6	833.4	303.1	2.650	0.4	243.5	693.3	1597.0
5	81.0	5	38.7	31.0	526.4	294.2	2.650	0.3	229.4	583.9	505.0
6	82.0	6	42.5	34.5	263.2	285.8	2.650	0.0	216.4	0.0	-32.1
7	83.0	7	46.0	38.0	285.8	285.8	2.650	0.0	216.4	0.0	-0.4
8	84.0	8	47.5	41.5	666.8	285.8	2.650	0.4	216.4	600.5	961.1
9	85.0	9	49.8	45.0	434.8	285.8	2.650	0.1	216.4	484.7	284.2
10	86.0	10	53.5	48.5	270.3	285.8	2.650	0.0	216.4	0.0	-22.4
11	87.0	11	57.2	52.0	270.3	285.8	2.650	0.0	216.4	0.0	-22.0
12	88.0	12	60.9	55.6	270.3	277.8	2.650	0.0	204.5	0.0	-10.2
13	89.0	13	64.7	59.5	263.2	256.5	2.650	-8.9	174.3	-2746.9	9.3
14	90.0	14	68.4	63.5	270.3	250.0	2.650	-2.5	165.7	-485.0	28.1
15	91.0	15	72.1	67.4	270.3	256.4	2.650	-4.0	174.3	-1047.2	19.2
16	92.0	16	75.9	71.4	263.2	250.0	2.650	-4.1	165.7	-1036.8	17.2
17	93.0	17	79.6	75.3	270.3	256.4	2.650	-4.0	174.3	-1047.1	19.2
18	94.0	18	83.3	79.3	270.3	250.0	2.650	-2.5	165.7	-485.0	27.0
19	95.0	19	87.0	81.9	270.3	384.6	2.650	0.0	392.1	0.0	-198.2
20	96.0	20	90.8	84.5	263.2	384.6	2.650	0.0	392.1	0.0	-208.1
21	97.0	21	94.5	87.2	270.3	370.4	2.650	0.0	363.6	0.0	-170.0
22	98.0	22	98.2	89.8	270.3	384.6	2.650	0.0	392.1	0.0	-198.2
23	99.0	23	102.0	92.4	263.2	384.6	2.650	0.0	392.1	0.0	-208.1
24	100	24	0.0	95.1	0.0	370.4	2.650	0.0	363.6	0.0	0.0
Last	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen	GenPen

Because the strains are dependent on the elastic constants, Poisson's ratio can be expressed in terms of Lamés constants,

$$\sigma = \lambda / 2(\lambda + \mu) \tag{5}$$

Poisson's ratio can be expressed in terms of P- and S-wave velocities, VP and VS, respectively, by

$$\left(V_S / V_P\right) = (1/2 - \sigma)/(1 - \sigma) \quad (6)$$

The P-wave seismic image map is in Fig. 7. The S-wave seismic image map is in Fig. 8. The Velocity analysis seismic image is in Fig. 9.

From Fig. 8, Fig. 9 and Fig. 10, we can see that the p-waves and S-waves reflect certain phenomenon. Through our processing and analysis of the map, P- and ave studies, can be made on the surface or in a borehole.

Through the analysis of surface wave in seismic image and processing, we can find the surface wave seismic image in Fig. 10, Surface wave seismic image energy spectrum in Fig. 11, Velocity of Surface wave is in Fig.12. The Surface wave velocity analysis is in Table I.

From Fig. 10, Fig. 11, Fig. 12 and Table I. we can see that the rock interface around the tunnel is more. Through the image we can calculate the thickness of rock layers and investigate the situation of the geological anomaly area.

IV. CONCLUSIONS

The seismic image was used for Huangzhuang Tunnel and it worked well. The forecast results proved that by using the image method, the geological conditions of the area in front of the working face of the tunnel could be well known timely, and also the forecast result could be sent timely. Meanwhile the applicable procedures of seismic image method were so simple that it was possible to carry on the forecast work without the influence on the tunnel construction. With the seismic image, the basis could be provided to the tunnel construction and parameter adjustments to guarantee the construction safety. The reflected waves, the refraction waves and the surface waves can reflect the same geological conditions, and the result is the same with the drilling. Tunnel seismic image can effectively and safely guide the excavation of the tunnel section working surface in combination with reconstructed images and excavation technology.

REFERENCES

[1] Lankston R W. High-resolution refraction seismic data acquisition and interpretation. *Geotechnical and environmental geophysics,* 1990, 1 pp. 45-73.

[2] Zhao Y, Jiang H, Zhao X. Tunnel seismic tomography method for geological prediction and its application. *Applied Geophysics,* 2006, 3(2) pp. 69-74.

[3] Zhao-huang Z. Prediction ahead of the tunnel face by the seismic reflection methods. *Chinese Journal of Geophysics,* 1994, 2.

[4] Cosma C, Enescu N. Characterization of fractured rock in the vicinity of tunnels by the swept impact seismic technique. *International Journal of Rock Mechanics and Mining Sciences,* 2001, 38(6) pp. 815-821.

[5] Hanson D R, Haramy K Y, Neil D M. Seismic tomography applied to site characterization. *Geotechnical Special Publication,* 2000 pp. 65-79.

[6] Suzuki K, Nakata E, Minami M, et al. Estimation of the zone of excavation disturbance around tunnels, using resistivity and acoustic tomography. *Exploration Geophysics,* 2004, 35(1) pp. 62-69.

[7] Dong S H, Wang Q. Application of Tomography in Radio Wave Tunnels Perspective. *Journal of China University of Mining & Technology,* 2003, 5 pp. 024.

[8] Riddle G I, Hickey C J, Schmitt D R. Subsurface tunnel detection using electrical resistivity tomography and seismic refraction tomography: a case study//*23rd EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems.* 2010.

[9] Gong X B, Han L G, Niu J J, et al. Combined migration velocity model-building and its application in tunnel seismic prediction. *Applied Geophysics,* 2010, 7(3) pp. 265-271.

[10] Lüth S, Bohlen T, Giese R, et al. Seismic Tomography and Monitoring in Underground Structures: Developments in the Freiberg Reiche Zeche Underground Lab (Freiberg, Germany) and Their Application in Underground Construction (SOUND)//*Tomography of the Earth's Crust: From Geophysical Sounding to Real-Time Monitoring. Springer International Publishing,* 2014 pp. 115-133.

[11] Babacan A E, Gelisli K, Ersoy H. Seismic tomography and surface wave analysis based methodologies on evaluation of geotechnical properties of volcanic rocks: A case study. *Journal of Earth Science,* 2014, 25(2) pp. 348-356.

[12] Kotyrba B, Schmidt V. Combination of seismic and resistivity tomography for the detection of abandoned mine workings in Münster/Westfalen, Germany: Improved data interpretation by cluster analysis. *Near Surface Geophysics,* 2014, 12(3) pp. 415-425.

[13] Sherman C S, Glaser S D, Rector J. Simulation of Seismic Tunnel Detection Experiments in Heterogeneous Geological Media//*AGU Fall Meeting Abstracts.* 2013, 1 pp. 1691.

[14] Hosseini N, Oraee K, Shahriar K, et al. Studying the stress redistribution around the longwall mining panel using passive seismic velocity tomography and geostatistical estimation. *Arabian Journal of Geosciences,* 2013, 6(5) pp. 1407-1416.

[15] Octova A. Implementation of Crosshole Seismic Travel Time Tomography for Predicting Near-surface Geological Structure During the Develoment of SMART Tunnel in Kuala Lumpur–Malaysia. Bina Tambang, 2013, 1(2).

[16] Dolgin B P, Davis C P. Method for detecting underground tunnels: U.S. Patent 8,510,048. 2013-8-13.

[17] Zhdanov M S, Cox L H. Multinary Inversion for Tunnel Detection. *Geoscience and Remote Sensing Letters, IEEE,* 2013, 10(5) pp. 1100-1103.

[18] DUAN C, YAN C, XU B, et al. The application of Cross-hole Seismic CT Method in the Karst Cave Exploration of Metro Engineering Construction. *Geological Review,* 2013, 6 pp. 027.

[19] Lin L, Chuan H, Jing Z, et al. Seismic Response Analysis for Shallow Tunnel in Different Earthquake Intensity. *Advanced Materials Research,* 2013, 680 pp. 161-165.

[20] HU J, LIU Y. Seismic dynamic response analysis of three-dimensional crossing tunnel on soft foundation. *Sichuan Building Science,* 2013, 2 pp. 055.

[21] DUAN T, ZHAO H, HU Y, et al. Advance Detection Technology with Seismic Reflection Wave and Its Application in Coal Mine Roadway Excavation. *Mining Safety & Environmental Protection,* 2013, 2 pp. 023.

[22] HAN R, ZHANG Y, CHEN F. Rock-burst Danger Prediction of Mining Face Based on Seismic CT Transmission Technology. *Coal Mining Technology,* 2013, 3 pp. 035.

[23] DOU L M, CAI W, GONG S Y, et al. Dynamic risk assessment of rock burst based on the technology of

seismic computed tomography detection. *Journal of China Coal Society*, 2014, 39(2) pp. 238-244.
[24] Huazhong Y, Xiongwu H, Pingsong Z. Numerical Simulation of Advanced Detection by Direct Current Electrical Method in Tunnel. *Chinese Journal of Engineering Geophysics*, 2013, 2 pp. 016.
[25] GUAN A, ZHANG H, CHEN J. Analysis on the Application Effectiveness of TGP System to the Advanced Geological Prediction for Dapingshan Tunnel. *Safety and Environmental Engineering*, 2013, 4 pp. 033.
[26] Ma H, Chen S, Tan X, et al. Advanced geological detection for tunneling in karst area. *Bridges*, 2014, 10: 9780784412121.340.
[27] Meng Y, Ji F, Jia L, et al. A new approach for forecasting the appearance of sinkholes near the Jinshazhou tunnel. *Environmental Earth Sciences*, 2014, 71(8) pp. 3339-3347.
[28] Du S, Zhang J, Deng Z, et al. A New Approach of Geological Disasters Forecasting using Meteorological Factors based on Genetic Algorithm Optimized BP Neural Network. *Elektronika ir Elektrotechnika*, 2014, 20(4) pp. 57-62.
[29] Pingsong Zhang, Shengdong Liu, Rongxi Wu, Observation of overburden failure of coal seam by CT of seismic wave. *Chinese Journal of Rock Mechanics and Engineering*, 23 (15), 2510–2513 (in Chinese) (2004)
[30] Cheng X, Xu W, Yue C, et al. Seismic response of fluid–structure interaction of undersea tunnel during bidirectional earthquake. *Ocean Engineering*, 2014, 75 pp. 64-70.
[31] Kouretzis G P, Andrianopoulos K I, Sloan S W, et al. Analysis of circular tunnels due to seismic P-wave propagation, with emphasis on unreinforced concrete liners. *Computers and Geotechnics*, 2014, 55 pp. 187-194.
[32] Benvenuti A, Moscariello A. Seismic Geometry and Facies Analysis of a Quaternary Tunnel Glacial Valley Infill in the Dutch North Sea: Preliminary Results//*STRATI 2013*. *Springer International Publishing*, 2014 pp. 781-785.
[33] Chen Z Y, Shen H. Dynamic centrifuge tests on isolation mechanism of tunnels subjected to seismic shaking. *Tunnelling and Underground Space Technology*, 2014, 42 pp. 67-77.
[34] Tsinidis G, Heron C, Pitilakis K, et al. Physical modeling for the evaluation of the seismic behavior of square tunnels//Seismic Evaluation and Rehabilitation of Structures. *Springer International Publishing*, 2014 pp. 389-406.
[35] Gomes R C. Numerical simulation of the seismic response of tunnels in sand with an elastoplastic model. *Acta Geotechnica*, 2014 pp. 1-17.
[36] Lanzano G, Bilotta E, Russo G, et al. Experimental and numerical study on circular tunnels under seismic loading. *European Journal of Environmental and Civil Engineering*, 2014 (ahead-of-print) pp. 1-25.
[37] Unutmaz B. 3D liquefaction assessment of soils surrounding circular tunnels. *Tunnelling and Underground Space Technology*, 2014, 40 pp. 85-94.

A Research on Personalized Retrieval of Mobile Learning Resources

Lin Hu, Honghua Xu*, and Xinfang Wei
Changchun University of Finance and Economics, Changchun, Jilin Province, China
*Corresponding author, Email: huhu315@sina.com, honghuax@126.com

***Abstract*—With the widespread use of mobile devices and fast development of wireless network, m-learning has become another hot topic except e-learning. Considering the less utilization of traditional learning resources and the poor-directed users, this paper puts forward an improved method of m-learning. Due to oceans of m-learning resources, this paper suggests an integration platform of information resources based on granular in order to improve learning efficiency. Meanwhile, personalized concept is introduced into the system. This paper studies personalized problems of hierarchical retrieval model and the methods of how to dynamically obtain users' interest. According to the hierarchical features of website structure, the system can actively obtain users' interest as well as output retrieval results based on different users' backgrounds under the help of ant colony algorithm. This method is easy to realize and can effectively trace users' short-term and long-term interest changes. It is suitable for the complicated and changeable network environment.**

***Index Terms*—Mobile Learning; Personalized Retrieval; Hierarchical Model**

I. INTRODUCTION

Because the fast development of internet technology, there are more and more forms of resources expressions. From traditional plain text to multimedia, the internet environments where resources exist are different. Because the internet information resources boom; the sorts of resources greatly are enriched; retrieval systems of database and information resources are more and more; retrieval ways and means are various, data redundancy is resulted. The degree of interrelatedness is low. Abundant information isolated islands appear, meanwhile, the retrieval burden of users is gradually aggravated, whereas, there exist many problems in current information system such as module dispersion, resources isolation, inefficient integration. How to gather, order and correlate these information resources which are isomerous and sourced differently. How to integrate retrieval systems to make clients know where to find the needed information; how to search for these information and how to filter retrieval results. The above mentioned are the ultimate purposes of information resources integration.

Through information resources integration, setting up an information integration visiting interface and a logically-centered but physically-distributed information resources integration platform will more effectively realize information resources share and key data reuse.

Information resources integration refers to integrating information resources which are originally discreet, polynary, isomerous and distributed in a certain range in a logical or physical way in order to make it easy to manage, utilize and serve [1, 2]. Information resources integration can also be understood as gathering discreet resources, changing unordered resources into ordered ones to make users easy to search information and to make it more convenient to serve users. This is a general definition of information resources integration. It includes the processes of information picking up, organizing, processing and serving. Learning resources integration is mainly about m-learning. Integrating relative and discreet learning resources is convenient for retrieving, managing, using and realizing personalized study.

Learning information resources integration generally includes the following there layers [3]: The first is data layer. Information resources all exit in various high-capacity servers. The data's forms of expression are various due to different database types. The second is operation layer, that is, relative learning information resources are utilized through software or platform. The third is application layer, that is, it is an all-round integration of learning data content, software systems and basic setup.

After integrating learning information resources, the integrated platform provides a single-stop-style visiting method for all the resources. Moreover, integrated accessible site categorizes all the learning information resources, from the originally addressable service to active service to make visiting more convenient and satisfy users' demands.

The current era is the one with information explosion. Facing oceans of information, most people are at a loss, not clearing about their actual demands. During the process of learning and retrieval, users should be more well-directed to visit the websites of learning information resources [4, 5]. Through adopting the integrated platform of learning information resources, learning information resources can be reasonably categorized; the needed resources can be classified and categorized more briefly and the users' response methods are more personalized.

Personalized retrieval may quickly solve the problem of "one for all users". It can offer corresponding advices

doi:10.4304/jmm.9.7.887-894

according to the differences in users. By means of hierarchical information of the websites, users' visiting records can be classified in order to find out the changes of users' interest. Thus, personalized retrieval method can be realized based on users' interest.

On the basis of m-learning resources integration, users may quickly find out the information they need with the help of personalized retrieval technique. Therefore, the visiting efficiency can be improved and the time users need to visit information websites can be reduced.

II. Granular Computing and Quotient Space

Granular computing [6, 7] with other subjects' philosophic thinking and methodology in return abstracts them into methods and strategies irrelated with concrete fields. The process of granular computing is the process of solving to complicated problems. The comprehensive understanding to complicated problems usually has multiple points of view [8, 9]. Each angle is multi-stratal. The result of granular computing presents a granular structure with multiple points of view and multiple levels. This structure is a systematic and approximate description and solving to complex problems.

The granulation of problem space refers to decompose concrete problem into several small ones and the understanding to the whole problem into analyzing in from different aspects [10-12]. The granulation of problem space may use the principles of approximate decomposition and loose coupling and it also may adopt the top-down method or bottom-up method. Granular computing can solve problems through visiting granular structure including upward and downward interactions in hierarchical structure and the moving way in the same hierarchy. There are two ways: One is the transformation and reasoning among granularities in the same granular layer. Two is the transformation and reasoning among particles in different granular layers. The connection among different granular layers is shown by mapping. The same problem on different granular layers is shown by different details and granularity.

Quotient space [13, 14] is a common method of granular computing. In the theory of quotient space, there are triples (X, F, T) to represent the objects to be studied. Among the three elements, x stands for universe [9]. T stands for topology structure on X and F stands for property function. Suppose an equivalence relation R on the triples (or an equivalent division). The question of projection to be discussed by the theory of quotient space is to consider quotient structure and quotient property function on quotient set X defined by equivalent relation R, that is, the construction problem of [X], [F], [T]. In fact, it is the transformation problem from different layers of granulation.

The acknowledged feature of human beings is that we can observe and analyze the same question from extremely different granulations. The so-called granulation is to categorize the elements with similar features to a new element, then to study it. If a subset is considered as an element, the elements in the subset are considered as equivalent ones, which is in accordance with the mathematic concept of equivalence relation which is also called equivalence relation or indistinguishable relation. According to the relation mentioned, all the equivalent elements can be called as new elements. All these new elements compose a set called quotient set. So granulation analysis is changed into the object analysis of certain quotient set.

In the study of different granulations, we find that the granulation of each question has one-to-one correspondence with its indistinguishable relation. Faced with a complicated question, people can't consider all the details of the question; therefore, they need to analyze the question from multi-level granulation. In order to find out multi-level granulation solving, thinning the question gradually may greatly improve the computing efficiency.

To begin with, it is necessary to find out the most suitable division rules which include property division, structure division, constraint division and projection division [15, 16]. There are many kinds of methods to divide granulation. It is quite important to choose proper division method according to different granulation. In the solving process of actual granulation division, there needs to adopt quotient space reasoning model to reason and analyze on high-level granulation to get a certain property. Then we can thin the question gradually in order to solve the problem.

III. Learning Information Recources Retrieval on Multi-layer Hierarchical Model

A. The Retrieval Model of Learning Resources

With the fast development of m-learning, the amount of m-learning resources is quite large. The types of learning resources are richer and richer. Efficient retrieval for learning information resources is the technique which can make learners realize study at any time, at any place. Brief m-learning process is shown in Fig. 1.

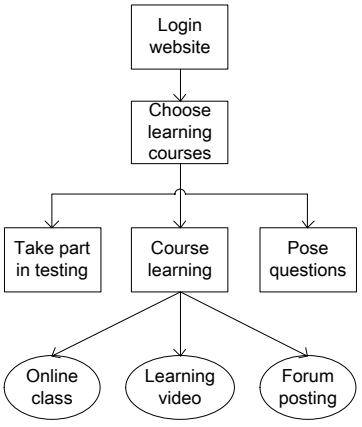


Figure 1. M-learning Process

However, there exist a lot of isomerous learning resources on net, which brings difficulty in retrieval and integrating learning information resources. Learning information resources which are from different places (online class, learning video, forum etc.) and shown by different ways (testing, class learning, posing questions etc.) may decrease m-learning efficiency. Therefore, this

thesis puts forward to using quotient space to realize m-learning resources retrieval technique to integrate numbers of resources to improve learning efficiency.

Learning resources retrieval is an important part in the process of resources integration. The purpose of integrating isomerous learning resources which are shown in various formats stored in different places and from different databases is to make users more convenient and more reasonable to make use of resources. How to provide users with high-efficiency resources retrieval service to make them more convenient and faster to find out the needed resources is a problem we must face in the process of resources. If a united retrieval platform is needed for information resources from different servers and different databases, an agent program is also needed to transform the inputs of the users into inquire statements of various databases and then search base by base. The agent program transforms the searching results into a united format (such as XML etc.) and then sends them to browser terminal in order to show the users.

To better find out the learning resources and information the users need and simultaneously to avoid the unnecessary repeated labor and resources waste, resources retrieval platform may provide the function of synonym extension, that is, similar meanings or similar language expressions are considered as synonyms. For example, while retrieving word "language", various expressions containing "language" can be found. We may take the current hot learning keys as initial granularity to perform equivalent divisions on different layers in order to form equivalence classes. Because there is no problem for human beings' thoughts to switch among different granularity layers, we can construct quotient space based on universe to go back and forth freely on various granularity in the actual application.

Fig. 2 takes key "language" as an example, illustrating how to divide equivalently and how to fulfill resources retrieval on different granularity layers according to the choices of users.

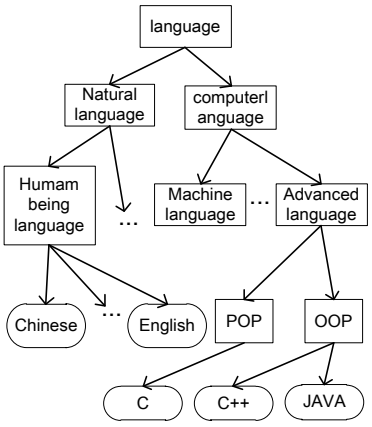


Figure 2. The retrieval process of learning information resources

When users search key "language", they may not be quite clear about their intentions. So related searching help is needed. On different layers (granularities), related searching obtains granularities through drawing granularities from universes. When users find out proper granularity, the key the user had chosen will be considered as retrieval condition on the granularity. The system is to construct different SQL statement to search for information and then return results to users.

The research/study applies common related methods of quotient space in granular computing to the retrieval process of learning information resources integration. It directs users to look for definite purposes. It helps find out the needed resources and improve the efficiency of learning resources retrieval to make m-learning process more efficient.

B. The Algorithm of Learning Resources Retrieval

There exist large amount of m-learning resources on net. How to design efficient retrieval algorithm of learning information resources is the key to realize study at any time, at any place. Therefore, this part puts forward a learning information resources retrieval algorithm based on quotient space after studying relative methods of granular computing and quotient space and also puts forward an integration technique based on granular computing. In the retrieval algorithm, because learning resources are obtained from web, a united URL class of returning results should be designed [17-19]. These returning results contain the struts of the retrieved linking address URL, retrieval keys, corresponding URL's introduction and the retrieved number.

Firstly, according to retrieval keys to set up various layers' granularities to be considered as initial granularities, the system performs equivalent division of various layers to form equivalence classes. Because human beings' thoughts can freely switch among various granularity layers, we may construct quotient space according to granularities drawn from universes. But, sometimes users are not quite clear about what they intend to retrieval and what they express. Therefore, the related results in many databases through many SQL statements should all be displayed for users to choose from. The related searching should be carried out on various layers.

Secondly, while providing intelligent displays for clients, if there is a certain display that is suitable for a user's demand, the key that the user has chosen will be an retrieval condition on that granularity. The results will be returned to users through constructing various SQL statements. If there are too many display records returned, users may choose one as child key. Then the procedures above can be repeated. Through the repeated retrieval, the granularity may become smaller and smaller with repeated retrieval.

At last, because of considering learning information resources integration, in the process of designing algorithm, the problem is that how to organize different types or various constructions of data should be considered, that is, the problem is how to convert among different formats such as doc, pdf, html, txt etc. in order to use a united and merged way to show users.

In Fig. 3, when the user inputs keys on the interface web, there will be corresponding intelligent clues provided for users to choose from. Here, there are two

situations. One is that the needed information will be found in the first retrieval. Then it will be transformed and merges to show the user. The second is that the needed information won't be found in the first retrieval. Then the sub-retrieval will occur until the satisfying granularity can be retrieved.

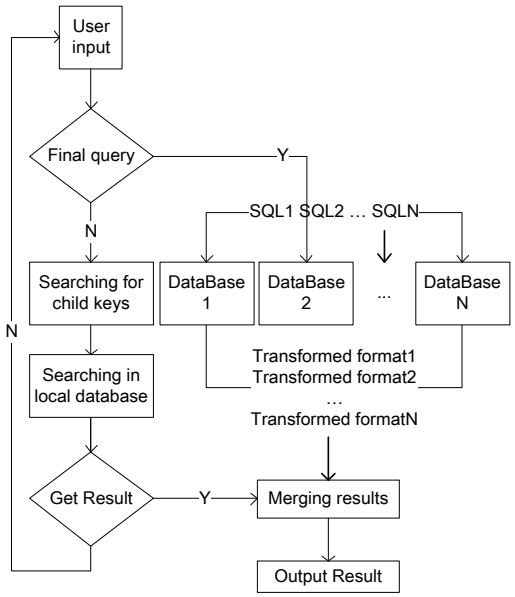


Figure 3. The algorithm of learning information resources retrieval

IV. Personalized Retrieval Based on Users' Interest

When people search for information materials, they always hope to find out abundant and accurate information, that is, they hope that the result information can satisfy the following two demands: One is that the result information can include all the relative content concerning the retrieval theme. Two is that the result information should exclude the irrelevant content concerning the retrieval theme.

In order to more accurately express the retrieval theme of the user to satisfy the demands of different users, we put forward a personalized retrieval model, that is, when information is retrieved, the retrieval not only retrievals according to retrieval statements but also considers the user's background and interest.

A. User's Interest Model

The key technique of personalized retrieval is to obtain user's interest. The obtaining methods include two sorts: explicit and implicit. Displaying obtainment needs users to actively provide interest modes. Generally there are two ways: One is to express their preferences and retrieval intentions through the method of selecting field [20, 21]. Two is to evaluate the retrieved results and feedback (including positive feedback, negative feedback) through certain method. Displaying obtainment can be easily realized and can get the actual interest of the user but it adds some extra operations to the user, so it is not easy to use. Implicit obtainment is transparent to users. Through tracking the users' scanning actions, potential users' interest mode can be mined out by means of web servers, users' terminals and other personal devices [22-24]. These information containing many content of user's scanning including the data such as the webpages users scan, the time spent on the webpage, collecting or not, file length, visiting frequency, URL address, turning the page etc. Through analyzing these information, feature data of the users can be mined out to be taken as users' interest. Mining data mainly include web content mining, web structure mining and web usage mining.

According to different time span, user's interest mode can be classified as long-term mode and short-term mode. Users' long-term interest and background are the main elements of their personalized information and it is called long-term interest mode. During a short period of time or in a process of retrieval, there is a definite retrieval aim which is not changeable and it is called short-term interest mode. Because users scan less and record less in short period of time, it is quite hard to find out users' short-term interest model.

Personalized retrieval can be conducted according to users' different interest. Through modifying or filtering retrieval statements or retrieval results by means of users' interest information, retrieval complexity can be effectively decreased and retrieval accuracy can be increased. As for the layer retrieval based on the theory of quotient space, retrieval complexity can be decreased and retrieval accuracy can be increased by means of a gradual thinning tree.

Users' interest mode needs to be vectorizably handled in personalized layer retrieval. When the users' interest mode is vectorizable, weight is users' interests. Meanwhile, to make retrieval convenient, when user has several interest preferences, several vectors can be added and the result shows the user's interest.

The core thinking of personalized layer retrieval based on user's interest is that in layer retrieval, layered user's interest is treated as equivalent class of corresponding layer. According to retrieval mode and user's interest, weight value of each equivalent class can be calculated. The weight value of equivalent class is above the given threshold value, it is the ultimate retrieval field.

B. The Way of Obtaining User's Interest Model

In order to quickly and accurately retrieval the needed information from oceans of data, people have made large amount of rewarding researches in the aspect of retrieval accuracy of retrieval system. In fact, personalized retrieval mode based on user's interest mode is an effective method of improving precision ratio, whereas, it is the basis of personalized retrieval for the mining of user's interest model.

Currently, internet websites all have certain layer structure. As is shown in Fig. 5, the highest layer is home (as opposed to home page). The child nodes are called inner nodes (as opposed to the first-level module/category). Inner node can also contain smaller inner node (as opposed to smaller module). The lowest layer is leaf node (as opposed to content page of specific information). We define root node as the layer 0, whose child node is layer 1. So, content page is the layer n.

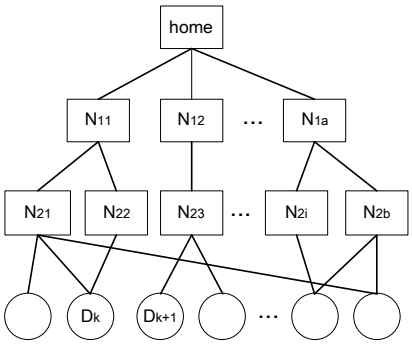


Figure 4. The Structure Model of Hierarchical Web Site

As for the website of hierarchical structure, we consider each layer as a collection of several class nodes. For example, as is shown in Fig. 4, the nodes on the first layer are made up of n nodes (N_{11}, N_{12} …, N_{1a}), the nodes on the second layer are made up of b nodes (N_{21}, N_{22} …, N_{2b}). When users visit leaf nodes (class nodes), each class node on the path is considered to be visited. Considering the nodes on higher layer, the visiting numbers are more because of their positions. Therefore, when users visit certain node, the feedback value of each node on the path shows exponential decrease layer by layer from bottom to top.

Essentially speaking, users' interest mode reflects their scanning interest and habits. For the website of hierarchical structure, users' interest mode also bears layer features. And certain interest mode can only be expressed by the visiting sequence of the layer's class node. Therefore users' interest mode is different from the aspect of each layer. As is shown in the website in Fig. 1, for example, a visiting sequence of interest mode in leaf node layer is D_kD_{k+1}. The corresponding visiting sequence of layer 1 is $N_{11}N_{12}$, the visiting sequence of layer 2 is $N_{22}N_{23}$.

When users scan webpages to obtain information, they may first log on the home. Then they will scan N_{11}, N_{12} and the content page D_k. They also may open N22 through other ways such as favorites. Then they scan content page D_k. No matter what is the visiting object, class node or content page, log records have the same format on the server.

V. Algorithm of Obtaining User's Interest

According to the website's fixed hierarchical feature, we put forward a new obtaining method of multilayer users' interest mode by means of ant colony algorithm [25-27]. This method can dynamically capture the latest users' interest mode of various layers to satisfy the demand of layered user's interest in hierarchical retrieval model.

The thought of ant colony algorithm roots from swarm intelligence when ants seek for food. Without the help of vision, ants can find the shortest path from food sources to ant nest through collaboration. In the process of foraging, ants may release as well as sense pheromone which is in proportion to the quality of food sources. When seeking foods, they tend to move toward the direction of higher concentration pheromone. Therefore, ants' collective behavior shows us the phenomenon of positive feedback. The more ants some path is chosen, the higher probability the followers will choose it. So the food sources of higher quality and short distance will attract more and more ants. The pheromone density will increase more quickly. In order to apply ant colony algorithm to hierarchical pattern recognition of users' interest, several parameters such as ant colony algorithm's support, pheromone density and preference function etc. should first be defined:

Support degree ηl stands for the probability of users' visiting on path 1, which is represented by visiting frequency of this path. Suppose Nij and Nim are two nodes in layer i. The path l(i, j, m) represents some path from node Nij to node Nim. The users' visiting number on the path l(i, j, m) is recorded as Cl(i,j,m). Then for the hierarchical interest pattern, the support degreeηl (i, j, m) is defined as

$$\eta_l(i,j,m) = \frac{C_l(i,j,m)}{\sum_{k=1}^{n} C_l(i,j,k)} \quad (1)$$

Pheromone density τl (t) stands for users' visiting interest in some path 1, which decreases as time passes by and increases as users visit to path 1.

$$\tau_l(t+1) = (1-p)\tau_l(t) + \delta\Delta\tau_l \quad (2)$$

In Eq.2, p is the mobilization factor of pheromone. δis the constant which is related with pheromone density. Δ τl is pheromone value added.

$$\Delta\tau_l = \frac{\tau_l(t)\times t + F_b}{t-1} - \tau_l(t) \quad (3)$$

where Fb is the feedback value of the node which users visit at moment t+1.

Preference function Pl(t) represents the users' preference to visiting path 1, which includes pheromone and support level. Supposeτk (t) and ηk stand for corresponding pheromone and support level opposite to path K. Then preference function Pl(t) is defined as:

$$p_l(t) = \frac{[\tau_l(t)]^{\alpha}.[\eta_l]^{\beta}}{\sum_{k=1}^{n}[\tau_k(t)]^{\alpha}.[\eta_k]^{\beta}} \quad (4)$$

In Eq. 4, n is the total number of paths in this layer. If the threshold path node can be checked layer by layer according to Eq. 4. Then hierarchical interest model of group users can dynamically be obtained.

Considering that the interests of web users are gradually changing, the mining of users' interest not only needs to find out their long-term interest but dynamically adjusts according to users' interest. So the mining process is dynamic. Meanwhile, website bears hierarchical feature. Users' interest bears hierarchical feature as well. Mining users' interest from various layers can provide higher-level and more information of users' interest and also can provide proper users' information retrieval based on hierarchical retrieval method. Through comparing people's information retrieval to evolving ant colony's foraging action, dynamic obtainment of multi-layer users' interest model can be better realized.

The basic thinking of the algorithm is as follows: comparing people's information retrieval to evolving ant

colony's foraging action, comparing once scanning process of various nodes on website to ant's once foraging action period. All the actions of users' scanning are recorded in log files, therefore we can get the value of preference function $p_l(t)$ of all the path nodes layer by layer by means of ant colony algorithm. Then we can dynamically identify users' interest mode quickly and accurately.

VI. Simulation and Analysis of the Experiment

Users' interest can be divided into long-term and short-term. Long-term interest of users refers to the permanent and stable interest demonstrated in the process of long-term online process. Short-term interest of users refers to the interest in a relatively short period of time. Short-term interest changes quickly. Long-term interest is relatively stable. It can be separated based on web browsing activity of users.

In order to inspect the applicability of obtaining users' interest, we conduct a simulation experiment on the accuracy of long-term interest and the sensibility of short-term interest change.

The process description of the algorithm is as follows:

Input: the threshold of the preference function is Tp after preprocessing m log records.

Output: internet pattern is H at all levels on the website.

Step 1: Initialize the parameters of ant colony algorithm, t=0, $\tau_l(t)$ =0, η_l =0, C_l =0. Initialize the structure tree of the website Tree=NULL.

Step 2: Build up the website's hierarchical structure tree "Tree" according to log records.

Step 3: Find out the path 1 which is corresponding to N_i log record in Tree. Work out all the feedback values F_b of all the nodes on path 1. Then update $\tau_l(t)$ and C_l. Repeat the operation above until processing all the records.

Step 4: Calculate all path nodes η_l according to Eq.1.

Step 5: Initialize α and β.Work out all the $p_l(t)$ on N_i layer again. If $p_l(t)$ is larger than the threshold value T_p of preference function, it is the interest pattern on each layer.

The experiment data adopt three groups. The training sample of the first group adopts visiting log of one month. The training sample of the second group adopts visiting log of one week. The training sample of the third group adopts visiting log of one day.

We conduct a simulation experiment on the data above. We only consider the first four layers of the website. The accuracy comparison of each layer's users' interest obtainment is shown in the following Fig. 6.

From Fig. 5, the obtainment of users' long-term interest is more stable. Meanwhile, as the inspection layer is higher, the testing rate of accuracy is also increased. This situation is not at random but depends on the nature of users' interest. The longer the time is, the higher the layer is. The hit rate of prediction is higher.

The other advantage of this method is that it may dynamically find out the interest changes of users. The following is a simulation experiment which shows the interest changes of users. The data are from the log files of certain company that record the users' visiting traces of four products A, B, C and D. Over time, users continuously visit product A and demonstrate keen interest on it. But recently users don't visit product A, which shows that users' interest has slightly changed. Part of log records is as follows:

../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_A/pagex/..
../company/products/product_B/pagex/..
../company/products/product_B/pagex/..
../company/products/product_A/pagex/..
../company/products/product_C/pagex/..
../company/products/product_B/pagex/..
../company/products/product_D/pagex/..
../company/products/product_B/pagex/..
../company/products/product_B/pagex/..
../company/products/product_B/pagex/..
../company/products/product_A/pagex/..
../company/products/product_C/pagex/..
../company/products/product_D/pagex/..
../company/products/product_D/pagex/..

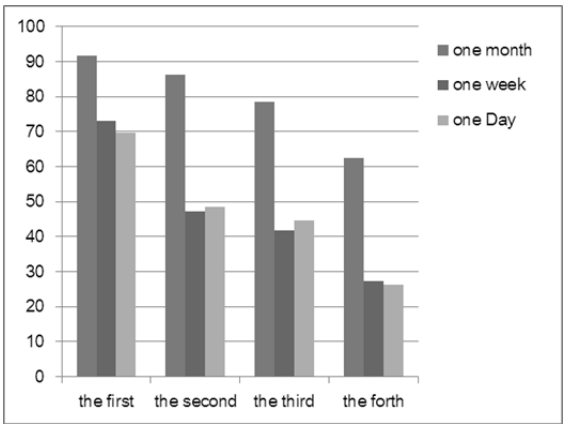


Figure 5. The accuracy comparison of interest obtainment

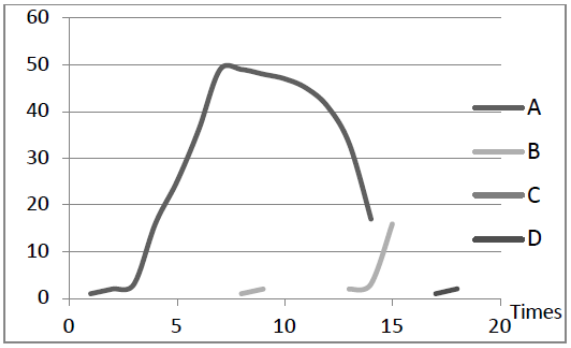


Figure 6. The diagram of product preference function

From Fig. 6, we may easily find that at the initial stage, the preference function value of product A node presents exponential growth because users continuously visit it. The preference function values of other manufacturers' nodes are zero. In the latter half of the experiment, as the visiting number of product A decreases, the preference function values of other nodes increase because they are frequently visited by users especially product B. As a

result, this method may dynamically capture users' interest pattern when their interest has changed.

VII. CONCLUSION

Because of a big amount of m-learning resources, the thesis puts forward the technique of learning resources integration based on granular computing. After studying relative methods of granular computing and quotient space, this thesis also puts forward an retrieval process of learning information based on quotient space. By means of retrievaling keys to set up granularities of various layers, which are considered as initial granularities, equivalent divisions of various layers are carried out. Learning information resources integration is realized through organizing various types or different constructions of data in a certain way. Then introducing personalized concept into the system, the paper puts forward a personalized layer retrieval method based on users' interest. Users' information retrieval is opposed to ant colony's foraging action. One-time scanning process of each node of the website is opposed to once ant colony's foraging action period. According to users' scanning log information, users' interest mode can be dynamically identified. This method is easy to realize and can capture the short-term and long-term changes of users' interest quickly and accurately.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of the project of Educational Science the twelfth Five-year Plan of Jilin Province "The Theoretical Research and Practice of Educational Model based on Cloud Computing "

REFERENCES

[1] Lundin, J., & Magnusson, M. Collaborative learning in mobile work. *Journal of Computer Assisted Learning*, 19(3), 273-283, 2008.

[2] Koeniger-Donolue, R. Handheld computers in nursing education: A PDA pilot project. *Journal of Nursing Education*, 47(2), 74-77, 2008.

[3] S. B. Zhong, D. H. Zou, Web Page Classification using an ensemble of support vector machine classifiers, *Journal of Networks*, Vol. 6, No. 11 (2011), 1625-1630, Nov 2011.

[4] Y. Xu, D. B. Zhang, Y. N. Wang, Active Diverse Learning Neural Network Ensemble Approach for Power Transformer Fault Diagnosis, *Journal of Networks*, Vol. 5, No 10 (2010), 1151-1159, Oct 2010.

[5] M. L. Chen, Y. Ma, Mobile Learning System based on Cloud Computing, *Journal of Networks*, Vol. 8, No. 11 (2013), 2572-2577, Nov 2013

[6] Gang Xie, Tianrui Cao, Texture Features Extraction of Chest HRCT Images Based on Granular Computing, *Journal of Multimedia*, Vol. 5, No. 6 (2010), 639-647, Dec 2010

[7] Fu-gui He, Yan-ping Zhang, Multi-granularity-based Routing Algorithm for Dynamic Networks, *Journal of Networks*, Vol. 9, No. 5 (2014), 1333-1338, May 2014

[8] Bargiela, A. and PedryczW. Granular Computing: an Introduction, *Kluwer Academic Publishers*, *Boston*, 2002.

[9] Hobbs, J. R. Granularity, Proceedings of the Ninth International Joint Conference on Artificial Intelligence, 432-435, 2011.

[10] Inuiguchi, M., Hirano, S. and Tsumoto, S. (Eds.) Rough Set Theory and Granular Computing, Springer, Berlin, 2003.

[11] C. H. Huang, S. L. Chin, L. H. Hsin, Jason C. Hung, A Web-based E-learning Platform for Physical Education, *Journal of Networks*, Vol. 6, No. 5 (2011), 721-727, May 2011.

[12] Pedrycz, W. (Ed.) Granular Computing: An Emerging Paradigm, *Physica-Verlag, Heidelberg*, 2001.

[13] Yuanbin Mo, Xinquan Zhao, Shujian Xiang, Application of Particle Swarm Optimization in Figuring out Non-differentiable Point of Function, *Journal of Computers*, Vol 7, No 10 (2012), 2542-2548, Oct 2012

[14] Nasiroh Omar, Colin Higgins, Colin Harrison, Diego Campo Millan, Internet Research Support System: An Application for Immediate Feedback in Authentic Internet Research, *Journal of Computers*, Vol. 2, No. 7 (2007), 68-74, Sep 2007

[15] Skowron, A. and Stepaniuk, J. Information granules: towards foundations of granular computing, *International Journal of Intelligent Systems*, 16, 57-85, 2001.

[16] Aderinoye, R. A., Ojokhela, K. O. & Olojede, A. A, "Integrationg mobile learning into nomadic education programmes in Nigeria: Issues and perspectives", *International Review of Research in Open and Distance Learning*, vol. 8, no. 2, pp. 1–17, 2007.

[17] Cook, J., Bradley, C., Lance, J., Smith, C., &Haynes, R,, "Generating Learning Contexts with Mobile Devices", *In N. Pachler (Ed), Mobile Learning: towards a research agenda. Occassional papers in work-based learning. WLE Center for Excellence, London:* 2007.

[18] Laurillard, D. Rethinking University Teaching: a Conversational Framework for the Effective Use of Learning Technologies, *London: RoutledgeFalmer*, 2002.

[19] M. Dorigo, A. Colorni and V. Maniezzo, "The Ant System: optimization by a colony of cooperating agents," *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, vol. 26, no. 1, pp. 29-41, 2010.

[20] M. Dorigo and G. Di Caro, "The ant colony optimization meta-heuristic," *In: New Ideas in Optimization, D. Corne, M. Dorigo and F. Glover Eds. London, UK: McGraw Hill*, pp. 11-32, 1999.

[21] H. X. Li, Text Classification Retrieval Based on Complex Network and ICA Algorithm, *Journal of Multimedia*, Vol 8, No 4 (2013), 372-378, Aug 2013.

[22] Alberto Pinto, Multi-model Music Content Description and Retrieval Using IEEE 1599 XML Standard, *Journal of Multimedia*, Vol. 4, No. 1 (2009), 30-39, Feb 2009.

[23] X. Z. Liu, Y. W. Peng, Y. H. He, Z. K. Tang, Uniform Storage Model-based Update Scheme of On-line Information Retrieval System, *Journal of Networks*, Vol. 8, No. 9 (2013), 2179-2185, Sep 2013.

[24] V. Dhar, D. Chou and F. Provost, "Discovering interesting patterns for investment decision making with GLOWER - a genetic learner overlaid with entropy reduction, " *Data Mining and Knowledge Discovery*, vol. 4, no. 4, pp. 251-280, 2000.

[25] M. Dorigo, E. Bonabeau, and G. Theraulaz. Ant algorithms and stigmergy. *Future Generation Computer Systems*, 16(8) pp. 851–871, 2003.

[26] C. C. Cheng, P. L. Huang, K. H. Huang, Cooperative Learning in Lego Robotics Projects: Exploring the Impacts of Group Formation on Interaction and Achievement,

Journal of Networks, Vol. 8, No. 7 (2013), 1529-1535, Jul 2013

[27] Yong-sheng Wang, A New Image Threshold Segmentation based on Fuzzy Entropy and Improved Intelligent Optimization Algorithm, *Journal of Multimedia*, Vol. 9, No. 4 (2014), 499-505, Apr 2014

Node Self-localization Algorithm Based on Received Signal Strength in Sensor Networks

Shi ZhiYong [1], Shao Dan [2], and Wu Boqiao [3]
1. Mathematics & Science College, Shanghai Normal University, Shanghai, China
2. Department of Public Security Technology, Railway Police College, Zhengzhou, China
3. Computer Department, Hunan Vocational Institute of Information Technology, Changsha, China

***Abstract*—The traditional wireless sensor network localization algorithm can be divided into the positioning algorithm based on distance and has nothing to do with the distance algorithm, due to the positioning algorithm based on distance is not suitable for the application of low power consumption and low cost areas and distance has nothing to do with poor precision problems of algorithm, this paper proposes a node self-localization algorithm based on received signal strength. Based on node location database in the grid topology analysis of the known conditions, and from the Angle of practical evaluates the localization algorithm scalability and fault tolerance. Finally with the traditional wireless network positioning algorithm do the experiment comparisons. The experimental results show that the proposed algorithm is less than the traditional algorithm in the interference in the environment of high precision, low cost, less energy consumption.**

***Index Terms*—Wireless Sensor Networks; Load Balancing; the Network Topology; Regional Extension**

I. INTRODUCTION

Sensor network compared with the traditional wired networks and wireless networks, have something in common, so many existed methods were transplanted to the sensor network. But compared with wired networks, there are significant differences in perception of the nodes, self-organization communication and deployment methods. Compared with the self-organized wireless network, sensor network has less infrastructure support, more stringent resource constraints and more large-scale deployment and so on.

Wireless sensor network (WSN) technology can be widely used in the military, civil, industrial, commercial, and other fields. Because of the great potential of their own, since it appears from a subject of great concern to the military, industry and academia [1]. Generally believed that the United States was the first country involved in this research field [2]. The wireless sensor network combines the sensor technology, the MEMS technology, the wireless communication technology, the embedded computer technology and the distributed information processing technology, and it accomplishes the data acquisition, data processing and the communication by the interaction between the sensor and the outside world, and it is widely applied in environment, transportation, military, aviation and medical health etc [3-5]. As the key technology of the wireless sensor network, location is significant for the study of network protocol based on the geographical location and the basic application of the wireless sensor network. American magazine Technology Review ranked wireless sensor network as the first ten big emerging technologies which will have the most profound influence on human life in the future. Business Week predicts wireless sensor networks are on the list in the four new technologies of future [6]. In 159 technical issues of "Foresight research over the next 20 years" in China, there are eight items that directly describes sensor network. In early 2006, issued by the national platform for medium-and long-term science and technology development planning for information technology to determine the three mentioned in the forward direction IntelliSense technology and ad-hoc network technologies, both directions are directly relevant to the study of wireless sensor networks [7-10]. Therefore, it can be seen wireless sensor network technology occupies an important position in current technology development and its widespread use has not only led to new information revolution, but also will bring enormous changes to the human society. Compared with traditional wireless networks, wireless sensor network has its unique advantages. Wireless sensor has the characteristics of no fixed equipment support and network deployment is simple, fast, without direct human manual operation, which can be placed directly by plane aerial seeding, etc, applying to the specific, severe environment that the human is unable to reach. Sensor node is small in size and the price is low, which may use in large quantities of relatively broad environment, the average number of sensor nodes in wireless sensor networks, distribution density is big, make to collect more comprehensive data, relative precision is more high. Computing power of wireless sensor nodes and wireless communication capabilities make sensor networks can be re deployed network, make a timely response to environment change or sensor networks and network control instruction, even if the environment is destroyed, still can realize the automatic recovery network to transfer information, etc.. Monitoring these features of wireless sensor networks wireless sensor network is suitable for complex environment, this application also has been in wireless sensor networks is a hot research.

doi:10.4304/jmm.9.7.895-901

For the research of wireless sensor network is still at the early stage in China, more exploration of the concepts, systems and technology, research subjects is relatively concentrated in institutions and research institutes, less involved in the enterprise. However, it should be noted, the state has invested efforts in the relevant field of research increased year by year. Since 2003,a series by the National Natural Science Foundation of the key technologies for sensor networks to support research projects have been launched, made a series of achievements and formed a research boom related technologies. In recent two years, a number of important research results related fields in the domestic core journals have been published [11]. Wenchao Meng and Li Yu Meng proposed the extended kalman localization algorithm based on optimal beacon group [12]. By analyzing the distance between target signal source and the beacon node, collinearity between beacon nodes, this paper proposed a kind of optimal beacon group selection mechanism to reduce the complexity of the algorithm and communication overhead. Dan Zuo and Jianjun Mo proposed a mobile beacon auxiliary node based on incremental positioning[13-18]. Zhongdong Hu and Fangfang Jia have researched on ISO location algorithm for Wireless sensor network based on angle judgment [19-24]. The position of the unknown node is judged by the angular sum to improve the positioning accuracy of the algorithm node ISO relatively sparse. Point positioning algorithm uses mobile beacon transponder to locate information in different locations, and expand the coverage of centroid localization algorithm.

The Slef-positioning of sensor nodes is the foundation and prerequisite of all sensor network applications. Wireless sensor network positioning technology, according to the different objects are positioned into two: one is sensor nodes self-positioning, the second is the target signal source localization. Database network of wireless sensor network technology shown in figure 1. Wireless sensor network applications in all aspects of research have made great progress, especially in major natural disasters, emergencies, and the environment, education, health and other fields have a wide range of specific applications.

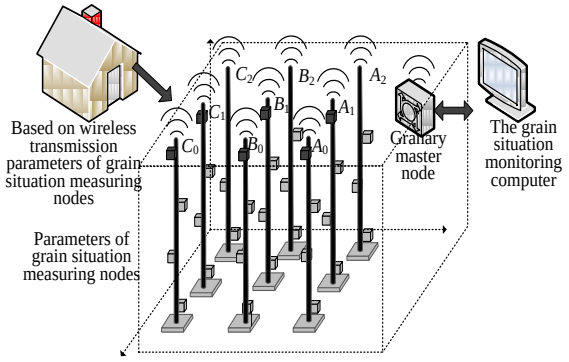


Figure 1. The database of sensor network

This paper mainly made expand and innovative work in the following areas to:

(1) Proposed a wireless signal location tracking algorithm based on wireless sensor network. The algorithm can locate the target node's position and loss statistics based on the existed data network, and the gives specific steps of the algorithm to further proof the accuracy of the algorithm and extensible compatibility.

(2) In order to further verify the accuracy of the algorithm in sensor networks, the algorithm verified under various sensor network model. The experimental results show that: this paper was used to optimize the network node target tracking algorithm, which compared with traditional algorithm can be more effective, fast rates to find the target position, and to be more accurate, to a rate faster, have convergence and positioning.

II. WIRELESS SENSOR NODE LOCALIZATION TRACKING ALGORITHM PRINCIPLE

Wireless sensor network node tracking algorithm will be taken into account the introduction of a relay node communication protocol using AF. Through the temporary relay nodes to upgrade for beacon node method completes the target signal source location, the algorithm principle is shown as figure 2:

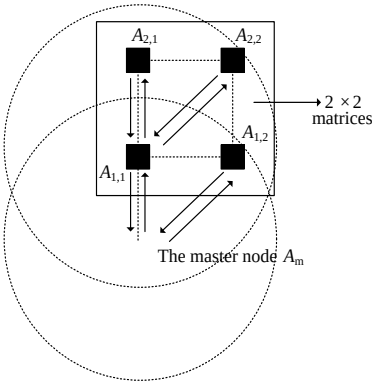


Figure 2. The sketch map of the node localization principle

In figure 2, the signal of target signal source MS (x, y) was accepted successfully by available beacon nodes $VSet(c,g,r)$ and $VSet_2(c,g,r)$, and set the reception time as x_2 and $c \prec g \prec r)$. The 4 squares $A1$, 1, $A1$, 2, $A2$, 1, $A2$, 2 in figure 2 are the wireless network nodes; they are set on the square mesh randomly; every node stores the location database; but they do not know their exact location information. A master node Am is added to the whole wireless network system to realize the self-localization; the master node is set as shown in figure 2 to simplify the system model; the distance between it and the node $A1$,1 is the lattice distance of this square mesh; the master node knows its location information.

In wireless sensor networks, there are many kinds of algorithm is our common. Different algorithms performance results is different, as shown in table I.

From the table I we can see that the common TDOA algorithm is highest, centroid algorithm and the CPA is the worst in position accuracy. At the same time, the TDOA algorithm also has its disadvantage that large communication overhead, cannot good complete the

TABLE I. THE COMPARISON OF POSITIONING PERFORMANCE RESULTS FOR SEVERAL POSITIONING TRACK ALGORITHM

Algorithm / Evaluation	TDOA	RSSI	DOA	Centroid algorithm	CPA
Positioning accuracy	Higher	Good	Higher	general	general
Computation	larger	Medium	larger	Smaller	Smaller
beacon node density	Little influence	Great influence	Little influence	Great influence	Great influence
average connectivity	Little influence	Great influence	Little influence	Great influence	Great influence
Communication overhead	larger	Smaller	Smaller	Smaller	Smaller

tracking target, the lack of adequate available beacon nodes. However this algorithm can be adopted to complete the target positioning for wireless sensor network notes. And it can effectively alleviate the network node degradation, interrupt, network itself losing the ability to locate the problem. Consequently it can expand the area covered, save operation, as well as the practical application of the hardware cost and communication costs, also it has extensive applicability.

In wireless sensor networks, the signal reception process can be come true through the target tracking algorithm, the whole process of implementation is shown in figure 3.

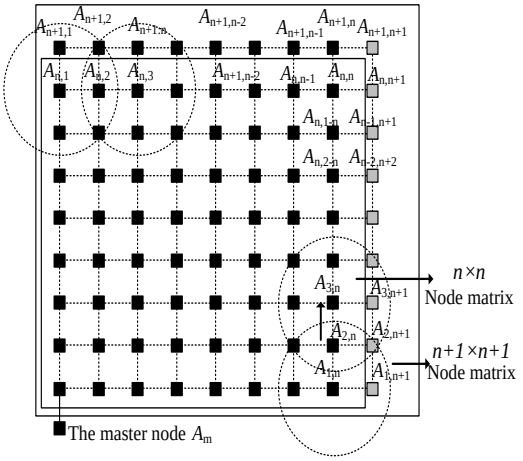


Figure 3. The Implementation Process of Wireless Sensor Network Location Target

From the figure 3 we can see that, when $n \times n$ is even number, the positioned node number of unknown node in a horizontal orientation and in a vertical row are in the different change. When the node An + 1, n, and the A n ,n+1 distinguish, node An will change with moves. When a node An +1, n-2, An +1, n-1 the node location tracking, we then found in the neighboring nodes known An, n-1 did not send him to locate relevant information, and the other is sent directly to the location information. After receiving the information, the process immediately reply message, the node information stored in its data by comparison, access to their own position location information. Finally then carry out location the unknown node An +1, n, An +1, n +1, all the nodes in the network location information showing knowledge will also be the end of the work. According to the above-mentioned principles positioning sensor nodes in a wireless network algorithm, showing that the localization algorithm is feasible, simple and easy to understand the whole process of positioning, no more money to spend to buy a lot of external auxiliary positioning equipment. It indicates the positioning approach than traditional methods more simple and easy to implement, and the cost of low loss.

III. ALGORITHM LOCATION PARTICLE WAVE MODEL

Particle filter algorithm is applied to wireless sensor target tracking, the tracking motion model is through the target motion state equation and measurement equation to draw.

(a) Target location state equation

In this paper the main consideration is the network nodes target tracking problem which do uniform motion in the two-dimensional plane, therefore the target network movement state equations is linear equation, as follow formula (1):

$$T_a=(c_1,g_1,r_1).(c_2,g_2,r_2) \tag{1}$$

where, $T_a=(c_1,g_1,r_1).(c_2,g_2,r_2) \in y_i^u$ is state vector, (c_1,g_1,r_1) is the target position coordinates at t time, (c_2,g_2,r_2) is the motion speed along x direction and c direction at g time. Matrix (c_2,g_2,r_2) is the state transfer matrix of target.

Considering the target positioning in network node do uniform motion, namely target coordinates of x axis direction and speed at t time respectively $T_a=(c_1,g_1,r_1).(c_2,g_2,r_2) \in y_i^u$ and $T_a=(c_1,g_1,r_1). \in y_i^u$. Because of the influence of noise, the system state transfer matrix is

$$T_a=\begin{Bmatrix} 1 & o & p \\ 0 & 2 & 8 \\ 2 & 1 & o \end{Bmatrix}.h_2 \begin{Bmatrix} \frac{t_1^2}{0} & 3 & 1 \\ c_2 & 0 & 1 \\ g_2 & 2 & 1 \end{Bmatrix}$$ is the system sampling period.

(b)Target measurement equation

The system target measurement equation is nonlinear equation, as follow formula (2):

$$\begin{aligned} & b \in Z_K \\ & h \in \{(g_1, g_2, \cdots, g_l, *^{q-l}) \mid g_i \in Z_2 . 1 \le i \le l \prec q\} \\ & m \in Z_q \end{aligned} \tag{2}$$

where, h is the measurement map; Z_K is measurement noise, as a general rule, which is gaussian distribution.

Specific target tracking application of wireless network, according to different sensor nodes to collect the amount

of map information may be measured relative distance measurements, received signal strength measurements, the relative angle measurements, measuring and mapping the relative relationship between velocity measurements, etc..

This paper is adopted by the received signal strength measurements, the formula (3) as follow:

$$h \in \{(g_1, g_2, \cdots, g_l, *^{q-l}) \mid g_i \in Z_2.1 \qquad (3)$$

where g_i is the target position at time. g_i is the i_{th} sensor position. $g_i \in Z_2.1$ is decided by wireless information transmission environment.

IV. Node Tracking Method of Target Signal Sensor Network

In the positioning method tracking in the sensor node network target signal is mainly filtrated the useful information by the node consumed and node information in the network. A rejection of unwanted node information is to the node information selected with advantage. According to the online user information query, sensor networks will be every once in a while for target tracking location information for a report. The central network node may be replaced, each central node collected user information and the received node information. By particle swarm algorithm to predict the trajectory of possible targets appear, so select the next central node user network to get the most useful information valuable. From figure 4 can be seen as a mobile user target network nodes, network nodes dynamically transmitted in sub-tree node increases or deletes some useless information networks. So that it not only can effectively control overhead communication do not waste, but also can reduce energy consumption of network nodes user information to improve the positioning of the sensor network node target tracking accuracy.

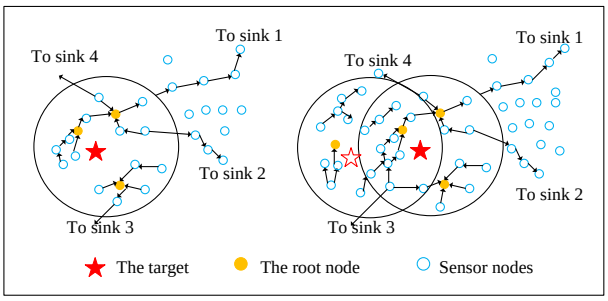


Figure 4. Sensor node network user based on target tracking algorithm diagram

In figure 4, sensors gather the target node, the root node, sensor node three network models. When network users input R to inquire target nodes information, it can inquire the central node through sub-tree node. Then the initial position of target tracking Z_K can be estimated by close around node. And it also can inquire the next user's central node b, that is $h \in \{(g_1, g_2, \cdots, g_l, *^{q-l}) \mid g_i \in Z_2.1$. The target information transmitted to node b, which Z_K is the i_{th} sensor node detection information. The node information Z_K and its measurement information Z_K, By particle swarm algorithm to achieve the purpose of the update sensor network node location $\mid g_i \in Z_2.1$. At the same time, use the same method can find the next node Next=N. By analogy, until the user target node to leave the scope of sensor network nodes can be detected. In the figure the central node and the skin 4 and skin 3 respectively through the shortest path to the target position at the moment the information is transferred back to the estimated query node Q. It indicated that the algorithm uses the prediction mechanism to reducing the amount of traffic between nodes, and the communication consumption, improve the accuracy of targeting nodes.

V. Experimental Results

In order to examine the algorithm for sensor network node localization target, with higher accuracy, the paper employ several different algorithms error performance on the target node of the network, as well as accurate positioning of nodes in the network data were simulated.

A. Simulation Experiment

The main evaluation indicators of the wireless sensor network localization algorithm include localization cost, energy consumption, localization accuracy, fault tolerance and adaptability etc. Aiming at the network model with known location database and removing the dynamic environment's impact, the algorithm this text proposes can localize all nodes accurately.

The particle swarm localization algorithm in this topic, the positioning process is to start from the main control node, the node is initialized to unknown areas, and then gradually outwards tracking sensor nodes in the network. Experiments were used to measure TDOA location algorithm and regular trilateral positioning calculation algorithm for multiple positioning. Compared with the mean of positioning error. Figures 5 and 6, 7 are used to reflect the change of the beacon with the sensor network and the number of nodes in the noise variance. The initial states of all the nodes to be localized are in the receiving state to reduce the localizing time. The localized nodes which are not close to the unknown nodes turn into the dormant state with lower energy consumption so as to reduce the energy consumption of the whole network. The realization of the algorithm is that the localized node sends Hello to its neighbor unknown nodes, then it judges the distance according to the returned TDOA value to realize the localization of the unknown node. The localizing time can be calculated according to the number of sending data frames; time in getting channel by the back off mechanism is ignored. Figure 5 shows the localizing time of different grids by using the simulation software MATLAB; the time unit is the time of sending one data frame.

From figure 6, the localizing time and the number of unknown nodes are linear increased; from the fitting curve of data, the average time in localizing a node is 13 time of data frame.

Figure 7 uses the simulation software to get the sending data frame of setting different nodal grids; it

shows that the data frame the localization needs and the number of unknown nodes are also linear increased. From the fitting curve of data, localizing one node needs to send 24 data frames averagely. The energy consumed in the localization depends on the localizing time of the network system and the number of data frames needed. The initial states of nodes to be localization are the receiving state; the nodes close to the master node can be localized first and then turn into the dormant state, thus the energy consumption in localization is less. Then the nodes far from the master node will be localized; their waiting time is long, thus they consume more. The above analysis summarizes the differences in property between this text's algorithm and other algorithms; the result is shown as table 1.

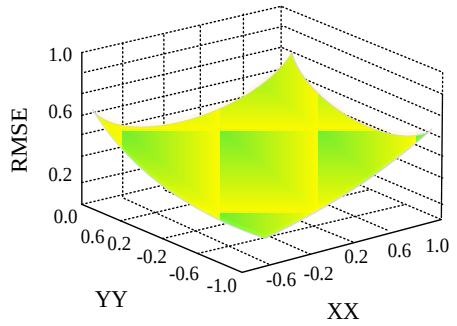


Figure 5. TDOA three boundary value measurement algorithm

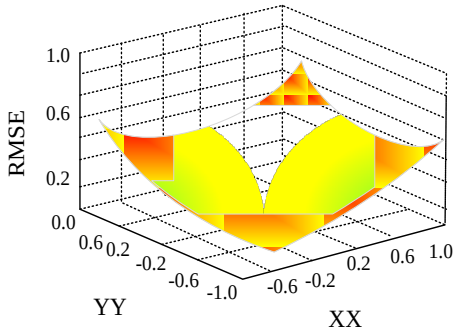


Figure 6. Error TDOA perpendicular bisector localization algorithm

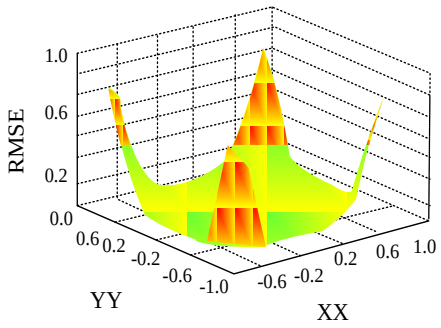


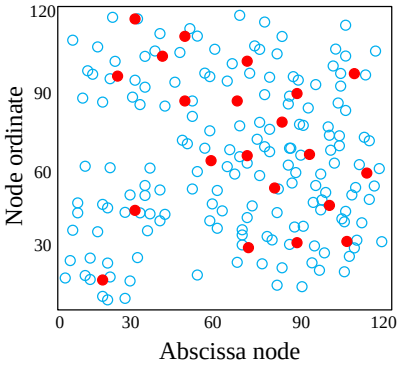
Figure 7. Node localization error of particle swarm optimization value

From figure 5, 6, 7, the algorithms based on the distance have higher accuracy and need extra hardware; the cost is higher. The range free algorithms have lower accuracy and do not need the extra hardware; the cost is lower, but the nodes need more information interaction, thus they consume more. TDOA and this text's algorithm use the RIOS ranging function which is included in the wireless communication platform, so they do not need extra hardware; but the wireless communication platform TDOA uses is based on the wireless local area network; the cost of the hardware and the energy consumption are high. The localization algorithm based on the wireless transceiver chip with low energy consumption has low cost and consumes less. While the error value of particle swarm algorithm in wireless sensor network is the minimum. For the target network node localization, the cost is lowest and energy loss is minimum, so that it further proves its superior performance.

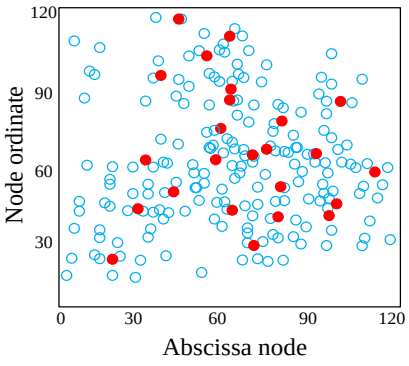
B. Error Performance of Target Localization Particle Swarm Algorithm in Wireless Sensor Network Nodes

When the wireless sensor network node localization exist errors, this experiment were used in polynomial curve fitting algorithm and particle filter algorithm for target tracking. Finally, using the mean square error RMSE represent tracking performance.

Specific simulation parameters used in this experiment are: network node coverage area: 100m × 100m. Setting the number of nodes is 38. the Target signal strength is $yt_d = z_q + 1$, when the close-set distance is $t_d = 1/m$. Path loss factor is $h = 5$. Wireless sensor nodes in the network to the destination network to track speed 1m / s to do uniform motion. Sampling particle number: N = 108 particles. Polynomial curve fitting order: K = 10. With the experimental parameters changes in the network, different wireless toward the target node sensor network error and sensor network node curve tracking algorithm presented will be different, as shown in Figure 8:



(a) Polynomial curve fitting method of node positioning error trend



(b) Tracking node localization to Li Zibo algorithm

Figure 8. To the sensor network node error and the relationship between two algorithms

Through the experiment, from figure 8 (a, b), we can easily see, the use of network nodes target tracking algorithm based on particle wave carried out with an accuracy better than using polynomial curve fitting algorithm to a network node localization information high. The latter algorithm tracks the target network node positioning, whose sensor network node positioning error is relatively large, as shown in figure 8 (a, b) the blue curve. When the node localization exists error, the particle swarm target tracking is 14.321 , and polynomial curve fitting node value target tracking algorithm has reached 27.632, the latter is the target location error difference of 13.331%. On the contrary, the use of particle wave algorithm to track the target node localization, the error rate is reduced 1.3331%, showing the improved positioning accuracy. Figure 8 (a, b) error value and the relationship, it is clear that when the target node positioning exist error in the network, no matter what kind of algorithm, the track value will be affected. Only can we do just minimize the target tracking value to a minimum. Thus improving target sensor network node is still one of the sensor network nodes to improve the accuracy and effective way, by this method, effectively improve the border region node localization big error. But also it provides for the development of sensor network node provides a challenge.

C. The Network Localization Experiment

The design thought of the wireless network sensor node is to reduce the energy consumption as much as possible on the condition that it has satisfied the property. The microcontroller the wireless sensor node selects is C8051F121 from the silicon company; the r_f chips are CC2420 and CC1020 from the TI company. In the test, CC1020 can transmit beyond 500m in the open, and its ability of penetrating barriers is strong. The nodes are designed in hierarchical thought to give nodes better robust and expansibility; the node system includes 4 layers; every layer has a shared interface; the position of every layer can interchange; the shared interface uses the air interface PC104, and the R_f module layer uses different keysets, thus CC1020 and CC2420 can exchange. The operating system of the software platform uses $\upsilon\cos - II$.

To test the validity of the localization algorithm this text proposes, the self-localization experiment of the 3×3 grid made up by 9 unknown nodes is designed. In the lab, the 9 nodes are set randomly to the 3×3 grid, and the distance is 0.9m; the unknown nodes are localized by the master node which is linked to the computer; the experiment scene is shown as figure 9:

The locating platform of wireless sensor network is developed to observe the localization result; the left part of the platform shows the coordinate of node localization; the gray square represents the nodes to be localized; after localization, the node coordinate will show in the square to test the correctness of the localization. After starting the network system, the unknown nodes begin to localize; after being localized, the node will show its coordinate in the network platform through the master node. The practical experiment shows the unknown node can realize its localization accurately, and the localizing time is short; the preliminary results from a small network show that the success rate of localization is 100%, which can match the location database in the real database.

Figure 9. The experiment scene of the sensor localization with known location database

The experiment shows the unknown node can localize itself correctly, and it has short localizing time, low energy consumption. To test the anti-interference ability of the localization, people walking in the nodes will lead to wrong localization during the localization, which shows the localization is easy to be affected by the dynamic environment.

Because of the limitation of conditions, this text just conducts a 3×3 nodal matrix experiment. This experiment proves the viability of this text's localization algorithm. Of course, there are many practical issues to be solved in the application.

VI. CONCLUSION

From the practical application of the sensor network, this text proposes a node self-localization algorithm based on the received signal strength; it calculates the nodal location according to the network model with known location database, and it analyzes the expansibility, the fault tolerance and the best transmit power of nodes. The simulation experiment and the real nodes experiment show this text's algorithm has the features of high accuracy of localization, low cost, short localizing time and low energy consumption, which provides a new type practical method and solution for the node location of sensor network.

REFERENCES

[1] Xin Huang, Xiao Ma, Bangdao Chen, Andrew Markham, Qinghua Wang, Andrew William Roscoe. Human Interactive Secure ID Management in Body Sensor Networks. *Journal of Networks*, Vol. 7, No. 9 (2012), 1400-1406

[2] Chen, Po-Wei; Ou, Kuang-Shun; Chen, Kuo-Shen. 2010. IR indoor localization and wireless transmission for motion control in smart building applications based on Wiimote technology. *SICE Annual Conference* 2010.

[3] Kuo-Feng Huang, Shih-Jung Wu, Real-time-service-based Distributed Scheduling Scheme for IEEE 802. 16j

Networks. *Journal of Networks,* Vol. 8, No. 3 (2013), 513-517

[4] Y. Geng, J. Chen, K. Pahlavan, Motion detection using RF signals for the first responder in emergency operations: A PHASER project, *2013 IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, Britain Sep*. 2013

[5] Y. Geng, J. He, K. Pahlavan, Modeling the Effect of Human Body on TOA Based Indoor Human Tracking, *International Journal of Wireless Information Networks* 20(4), 306-317

[6] Mao X. Nguyen, Quang M. Le, Vu Pham, Trung Tran, Bac H. Le, "Multi-scale Sparse Representation for Robust Face Recognition", *Proceeding(s) of Conference on Knowledge and Systems Engineering (KSE),* pp. 195-199, 2011.

[7] PECES Consortium, PECES Communication Mechanisms and Registry Interface Specification, *Deliverable D. 3. 2, PAS, http: //www. ict-peces. eu*, last accessed June 2012

[8] Yikui Zhai, Junying Gan, Jingwen Li, "Study of occluded robust face recognition approach based on homotopy algorithm and color information fusion", *Signal Processing*, vol. 21, no. 11, pp. 1762-1768, 2011.

[9] Snoonian, D, Smart buildings. 2003. Spectrum, IEEE Volume: 40, Issue: 8 Digital Object Identifier: 10. 1109/MSPEC. Page(s) pp. 18 – 23

[10] Kasman Suhairi, Ford Lumban Gaol, The Measurement of Optimization Performance of Managed Service Division with ITIL Framework using Statistical Process Control. *Journal of Networks,* Vol. 8, No. 3 (2013), 518-529

[11] Haiping Huang, Hao Chen, Ruchuan Wang, Qian Mao, Renyuan Cheng. (t, n) Secret Sharing Scheme Based on Cylinder Model in Wireless Sensor Networks. *Journal of Networks*, Vol. 7, No. 7 (2012), 1009-1016

[12] C. Becker, G. Schiele, H. Gubbels, K. Rothermel: BASE - A Micro-broker based Middleware For Pervasive Computing, *In Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications*, pp. 443-451, Fort Worth, USA, March 2003

[13] Zhao Liangduan, Zhiyong Yuan, Xiangyun Liao, Weixin Si, Jianhui Zhao. 3D Tracking and Positioning of Surgical Instruments in Virtual Surgery Simulation. *Journal of Multimedia*, Vol. 6, No. 6 (2011), 502-509

[14] D. Garlan, D. Siewiorek, A. Smailagic, P. Steenkiste, "Project Aura: Towards Distraction-Free Pervasive Computing", *IEEE Pervasive Computing,* vol. 1, no. 2, pp. 22-31, April-June 2002. http: //dx. doi. org/10. 1109/MPRV. 2002. 1012334

[15] Guang Yan, Zhu Yue-Fei, Gu Chun-Xiang, Fei Jin-long, He Xin-Zheng, A Framework for Automated Security Proof and its Application to OAEP. *Journal of Networks*, Vol. 8, No. 3 (2013), 552-558

[16] Pradeep Macharla, Rakesh Kumar, Anil Kumar Sarje, "A QoS routing protocol for delay-sensitive applications in mobile ad hoc networks," *COMSWARE* pp. 720-727, 2008

[17] Hasler M, Schimming T. Chaos communication over a noisy channe, Int J Bifur Chaos, 2000, 4(10) pp. 719-736.

[18] Yassen M. T. Chaos synchronization between two different chaotic systems using active control. Chaos, *Solitons & Fractals:* 2005, 23(4) pp. 131-140.

[19] Pehlivan I, Uyaroglu Y. Simplified chaotic diffusionless Lorentz attractor and its application to secure communication systems. *Communications,* 2007, 1(5) pp. 1015-1022.

[20] Zhang R X, Yang S P, Adaptive synchronization of fractional-order chaotic systems via a single driving variable. *Nonlinear Dynamic,* 2011, 66(4) pp. 831-837.

[21] Sarasola C, Torrealdea F J, Anjou A D. Feedback synchronization of chaotic systems. *Int J Bifurcation and chaos,* 2003, 13(l) pp. 177-191.

[22] B. Johanson, A. Fox, and T. Winograd, "The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms", *IEEE Pervasive Computing,* pp. 67-74, April-June, 2002 http: //dx. doi. org/10. 1109/MPRV. 2002. 1012339

[23] Aricit, D Ikbas S, Altunbasak Y. A histogram modification framework and its application for image contrast enhancement. *IEEE Trans on Image Processing,* 2009, 18(9) pp. 1921-1935

[24] orgu O, Solak E. On the synchronization of chaotic systems by using state observers. *Int J Bifurcation and choas,* 2007, 7(6) pp. 1307-22.

[25] Liu F, Shna X, Qiu Z. A linear feedback synchronization theorem for a class of chaotic system. Chaos, *Solitons and Fractals,* 2012, 13(4) pp. 723-30.

[26] Liao T L. Adaptive synchronization of tow Lorenz system. *Chaos solitons and fractals,* 2012, 9(9) pp. 1555-1561.

Stadium Evacuation based on Multi-agent System

Cao Ai-chun [1], Yang Xiao-ting [2], and Hou Xu-dong [3]
1. Nanchang University, Nanchang, China 330029
2. Jiangxi Normal University, Nanchang, China 330027
3. Shinsoft Information Co., Ltd. Shanghai, China 200233

***Abstract*—In order to design a reasonable pedestrian evacuation and exclude the security risks in large stadiums, this thesis proposes the research of security model of evacuation in stadiums combined with multi-agent and cellular automata. The research is based on cellular automata model and the process makes extended analysis to the cell's behavior of autologous, and then it makes simulation experiments of the process of simulating the evacuation in large stadiums. Simulation results finally show that cell Agent combines the advantages of multi-agent and cellular automata, which fully considers the individual internal factors. Compared to traditional cell cellular automata, it is more close evacuation situation of the realistic major sports stadium, and it shortens the evacuation time and improves the safety.**

***Index Terms*—Node Stadium; Multi-agent System; Simulation**

I. INTRODUCTION

Crowded public places such as theaters, stadiums, railway stations, docks, business center and supermarkets have characteristics of high t crowds gathering and mobility [1-3]. In case of when unexpected disasters like fire, explosion and riot occur, serious life and property damage will be caused.

Since the 90s, the construction of sports stadium has achieved brilliant achievements; there is a significant improvement from the shape of stadium, or the quality and quantity. In recent years, the large-scale international events like Beijing Olympic Games, Asian Games in Guangzhou, Shenzhen Universiade sports are held in China, so many massive stadiums were built in many cities, which is with large scale unmatched with the usual scale. But the modern stadiums presents some new features, mainly reflected in two aspects: Firstly, increasing the size of the venue, the modernization level has been improved significantly; Secondly, the evacuation characteristics is the basis of designing the Square, shopping mall, entertainment and so on, but also it is an important basis for establishing pedestrian evacuation facilities. Pedestrian evacuation infrastructure is one of the main pedestrian traffic facilities. Generally the crowded place is pedestrians' necessary places. Fire accident of arena fire is the accident, which is super dangerous, and it is not easy, to find out.

Dangerous, sudden, transient and other characteristics, and also the sports venues is with special internal functions, numerous equipment lines, personnel highly concentrated and other features, so it is resulting in a fire more incentive. If a fire occurs, the consequences could be disastrous. Unlike the general architectural space, the stadiums interior space is different. When the fire occurs, the crowd is easy to be panic because the directions are not clear in unexpected situation, so it is easy to lead to outlet obstruction, crowd stampede, and more affect the safety of personnel evacuation. The final man-made danger of building fires or large numbers of people is the serious consequences jostlement caused by gathering disasters. As shown in Figure 1. While through the study the characteristics of reasonable crowd gathering and crowd evacuation to the large crowded distribution center, such as subway hub, rail transportation stations, stadium course, large amusement parks, etc., the theoretical basis for design of crowd gathering and crowd large distribution. At the same time, it provides evacuation strategies and advice to the relevant departments. In case of when under an emergency case, the crowd can be reasonable evacuated to avoid casualties and property losses, and it is currently an important issue need to be resolved.

Figure 1. Stadium crowd stampede when fire break

Thus whatever the stadiums fire or crowd riots, it is directly or indirectly caused by the crowds stampeded, but the main reason of crowded stampede occurred mainly due to the excessive population density of congestion and pedestrian accidents caused by a fall or squat stampede accident. Effectively making quick and safe evacuation of personnel to the secure area will be able in some way reduce the danger of accidents; therefore the design of stadiums evacuation is very important.

With the development of modern design and technology, the size and capacity of sports stadium continues to develop [4]. The Appearance of great density of crowd in stadium raises the possibility of potential

doi:10.4304/jmm.9.7.902-909

accident casualties. FIFA records that between 1913 and 2000 there are 23 casualties worldwide and the total death is at least 1,380 [5-8]. The accident statistics such as Kristin in Canada and Fruin in New Zealand show that there are 3,000 casualties in the last 10 years in the world sports arena, which is larger that the casualties probability of the frequently occurred accident (such as car crash). Therefore, study on stadiums evacuation has important and practical significance, and it is always the hot point in the research field of public safety [9-13].

Stadiums evacuation is an organic whole consisted with many interrelated and interacted elements, which is a complex system with the characteristics of complex, multi-level nature and feedback, and it is with strong spatial and chronological features. For this problem, many domestic and foreign scholars have done a lot in-depth researches and proposed the stadium evacuation models like cellular automata, multi-agent model (MAS) [14-18]. The museum crowd has strong autonomous and intelligent features, which has become an important direction of current research.

For the researches in this field there are three types of parties, namely the rule-based approach, social force of approach and cellular automata approach. Rule-based approach at a lower population density or moderate cases can simulate the realistic movement of people, but can not handle the contact between people, so it can not simulate the squeezed results [19-23]. Helbing D et al on the basis of social force method reappear the extrusion pressure when the crowd squeezing through analyzing the repulsion between man and man, the repulsion between man and static obstacles between man and target and the gravity and friction between man and target and establish the simulation model of crowd evacuation under a state of panic [24-25]. The disadvantage of this model is that when the simulation is with high-density crowd, crowd shaking phenomenon will be appeared. Cellular automata method restricts the movement of people within discrete boxes, and the state of the cellular at some certain moment is according to the state of the last moment and the entire neighbor cellular to locally update, so the performance of the whole cellular space is represented as the changes in discrete time dimension changes. In the study of crowd evacuation simulation on fire scenario, the cellular automaton model is combined with the individual behavior, and this the commonly used methods. Therefore, stadiums evacuation study has important practical significance, and the public safety has been a hot research field. Stadiums crow evacuation is an organic unit consisting of many interrelated and interaction elements. It is a complex system with natures of complexity, multi-level, feedback and other characteristics, which has strong spatial and chronological features. To solve this problem, a number of scholars at home and abroad have in-depth and numerous researches and they found that the crowd in stadium has a strong autonomy and intelligent features, so it has become an important direction of current research.

Mr Young et al introduce the concept of risk level of location based on cellular automata method to determine the movement of pedestrians and simulate the escaping process of pedestrians when the fire broke. Mr Young et al adopt the method using cellular automata and combined with five different kinds of acting force of individuals under emergency situations. Under the simulating emergency evacuation, Yue Hao er al use two dynamic parameters and combine cellular automata to reflect the pedestrian evacuation. Zhou Jinwang et al consider the effect of side by side pedestrian in pairs, around in pairs and mixed pairs to the process of pedestrian evacuation.

Cellular automata are a cross-cutting science. It is the result of combination of mathematicians, physicists, biologists and computer scientists. Therefore, different scholars have different explanation on cellular automata: mathematicians think that cellular automaton is a mathematical model with a discrete space-time; biologists think it is an abstract phenomenon of life; Physicists will treat it as a discrete dynamical system. In cellular automata, the space is divided into many units with regular grids. Each cell in the ruled grid is called cellular and all cellular follow the same rules to be updated.

This thesis combines the advantages of agent and cellular automata model; more realistically simulates the process of stadium evacuation behavior and proposes the crowd evacuation model in large stadium combined with multi-agent and cellular automata. Simulation results show that fully considers the individual internal factors, and it is closer to the crowd evacuation situation in realistic large stadium.

This paper mainly makes the extensive and innovative work in the following areas:

(1) The current studies on wireless sensor network resource optimization tend to focus on the MAC or strategies on transport layer to achieve a single optimization goal. In order to improve node energy, storage and computing power, this paper proposes the cross-layer resource optimization algorithm. The algorithm introduces the power control model associated with the probability of link connection and establishes the network collaborative optimization mechanism based on congestion cost and energy cost through combining the power control on physical layer, access control on Mac layer, flow control model on transport layer and using the network utility and life as optimized goal. The vertical decomposition theory is used to realize the decomposition with step-by-step of the optimized goal, and then the updated formula based on the power consumption of nodes and probabilities and rate of link access is deduced to achieve the distributed optimization solutions.

(2) In order to further verify the correctness and effectiveness of the proposed the optimization algorithm with cross-layer resources, the validation experiments cross-layer optimization algorithm in wireless sensor network based on network utility and lifetime are conducted. Wireless sensor network is consisted with six nodes, and after a brief fluctuation each node is convergent, and the probability of node 5 accessing to

aggregation node 6 is 1. After a brief fluctuation the congestion cost is close to zero, which indicates that under the experimental environment the algorithm can alleviate and eliminate the network congestion status. Energy cost is also convergent varies with the iteration time, in which due to the node 5 bears more data transfer tasks, the energy consumption is higher and its energy cost parameters converge is a nonzero value. The simulation results show that: cross-layer resource optimization algorithm is effective and has fast convergence.

II. CELLULAR AUTOMATA AND MULTI-AGENT

A. Cellular Automata

Cellular Automata (referred to as CA) is the mathematical model based on the evolution of cell, which is both discrete on time and space. From a historical perspective, cellular automata are first used by mathematician and physicist John Von Neumann and Stanislaw Ulam in 1940. As a mathematical model, cellular automata have a strict scientific definition: cellular automaton is defined in the cellular space consisting of discrete and finite-state cellular composition, and according to some local rules, it is the dynamical system evaluated in discrete time dimensional. Cellular automata system can be expressed by a four-tuple:

$$A = \{d, S, N, f\} \tag{1}$$

where in, d is the number of dimension of cellular automata; S represents the discrete and finite set of states of the cellular; N represents the combination of cellular space in the neighborhood; f maps Sn to a partial conversion function on S.

Theoretically, cellular automata can be expressed as any dimension. The current researches are concentrated on one dimension or two-dimension, while a lot of phenomenon are represented by concrete or abstract distributed two-dimensional, so the two-dimensional cellular automata is most widely used. Two-dimensional cellular division dimensional is generally divided as equilateral triangle, regular hexagon or regular square grid (shown in Figure 2).

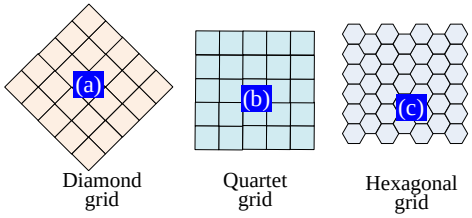


Figure 2. Meshing form of two-dimensional cellular automata

Cellular automata can be used to study many general phenomena, such as Information transfer, evacuation analysis and urban spatial change. When these studies use CA, there is an obvious common – the model used for large graphical display is based on two-dimensional cellular automaton model. This is because that in the research of searching the change rules of the target, two-dimensional CA can be sufficient to facilitate the expression of computer graphics on the plane.

Stadium crowd evacuation is different from the crowd evacuation in other large-scale venues (such as a large plaza, public entertainment, subway stations, etc.) – the audience stands in stadiums is designed as multi-steps, so when the site map of stadium is mesh divided of cellular automaton, the two-dimensional plane situation can not be simply considered. Because when the emergencies and the audience need to be evacuated, the majority of audience will be eager to find exit due to the panic mentality, the audience stands are the stepped design, so people often will firstly escape to the down side. Therefore, the elevation factor of the sports stadiums terrain plays an important role in the cellular properties - when the situation of audience choosing escape to the bottom is virtual, for the cellular around the currently occupied cellular automata as long as the current elevation value is high it will not be chosen.

B. Multi-agent Crowd Evacuation Model

In the framework of simulating the crowd movement the crowd evacuation situation is considered, and the framework of crowd evacuation model based on multi-agent can be simplified as in Figure 3; the model components include intelligent body, simulation model, map environment and the observer.

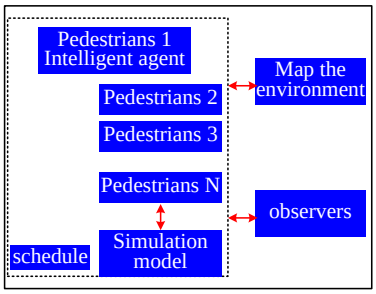


Figure 3. Crowd evacuation model based on multi-agent

(1) Space Environment Z=Rp-gon, Rp-line,; B is the test area

The bounded closed set of the inside entire area; planar road collection Rp-gon = Rp1, Rp2, ..., Rpn, which is possible value range of the location of the audience and also is the constraint of the agent audience movement. Linear path collection Rp-line = Rl1, Rl2, ..., Rln, which is the reference direction of the action line of one point when audience is leaving; other environmental collection B = {Build, Block}, which is combined with the collection of buildings and collection of obstacles, and both are polygons.

(2) Audiences agent

$$P-Agent = \begin{Bmatrix} Id, P_o, P_T, P_N, V, P, A, R, D_{Va}, \\ D_a, Rule, Query, SNext, Esti, Act \end{Bmatrix} \tag{2}$$

where in, Id is the unique identifier for the audience; PO, PT, PN are the starting position, target position and current position of the agent; V is the mobile speed of the agent; P is the perceived state set; A is the set consisted

by all the possible behavioral programs of the agent; R is the set of all the results corresponding to the behavior program; DVa is the rule set corresponding to the sensing status and programs set; Dar is the correspondence between the scheme set and results state set; Rule is the specific decision law; Query defines the process of Agent sensing the surrounding environment; SNext is the latest sensing status, which combines with the corresponding rule of perception state and the program set to determine the progress of the current results set; Esti is the given program set, which according the corresponding rule of perception state and the program set to determine the progress of the possible results state; Act is different decision rule adopted by the set of results state corresponding to one program using state set, which selects a particular program from the set of proposed program.

Population $A^t = \{P_{a1}^t, P_{a2}^t, ...P_{an}^t\}$ is the entire set at time t of the simulated audience. In accordance with the different value PN of the current position the elements of A^t can be divided into the set which is reached the destination and not reached the destination A_{in}^t and A_{out}^t ; If PNi-POi>ε, Pai $\in A_{in}^t$; If PNi-POi<ε, then Pai $\in A_{out}^t$. Clearly, $A_{out}^t \bigcup A_{in}^t = A^t$ and $A_{out}^t \cap A_{in}^t = \cong$. When the simulation starts, $A_{in}^t = A^t$; when the simulation is end, $A_{out}^t = A^t$.

The time when the simulation is complete T={t $A^t = A_{out}^t$ }, thus the evacuation time when evacuating the crowd at emergency can be obtained as T = inf t, that is the minimum of t when t $A^t = A_{out}^t$.

The set of the entire audience agent at t moment all audience location is as follows:

$$\vec{\boldsymbol{P}}_{t+1}(x,y) = \vec{\boldsymbol{P}}_t(x,y) + \vec{\boldsymbol{v}}_{prefer}.\Delta t \qquad (3)$$

Vector PA^T Describes the position information of all the simulation objects at time t; the rule of PA^T changing with the time describes the dynamic process of the crowd movement, which becomes the core state variables of the simulation model, and it is the s fundamental object observed by the observer in process of model simulation. The ultimate goal of simulation is to find out the change rule of the state variables.

III. CROWD EVACUATION MODEL

A. Model Assumption

Each individual is always constantly awareness the around situation, and then make their own judgment to make their behavior decisions. All of the crowd evacuation people are constantly on the individual behavior. All of these individual behaviors are interaction, overlapping, and form the whole crowd behavior with the scene interaction of the crowd. In this process, many factors affect the behavior of individuals and groups, which must be taken into account when modeling.

(1) When the crow is evacuating in stadium they are always toward the exit, and adjust the next behaviors and decisions based on the environmental risk and their own circumstances.

(2) When the crow is evacuating in stadium, there are some complex phenomena like together, herd, hinder, “haste makes waste” and so on.

(3) Evacuation speed of the crowd in stadium is influenced by their own psychological factors, physiological factors and population density, and different types of people have different speeds.

(4) Crowd in stadium can be divided into agent, and there are three types of agents: elderly, adults and children. The elderly and adults occupy 16 cellular and children occupy nine cellular, which are respectively represented with yellow, blue, purple, and each agent identifies the respective numbers.

B. Model Construction

Based on multi-agent technology, stadium crowd is regarded as autonomous agent, which has complex decision making ability, their knowledge and experience. They can adjust their behavior according to the outside information and internal factors, and the general structure of its models is shown in Figure 4.

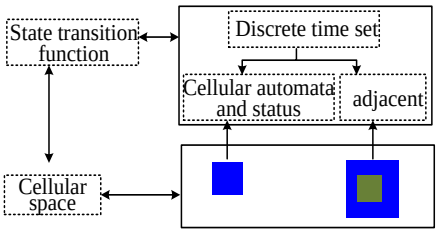


Figure 4. Agent structure

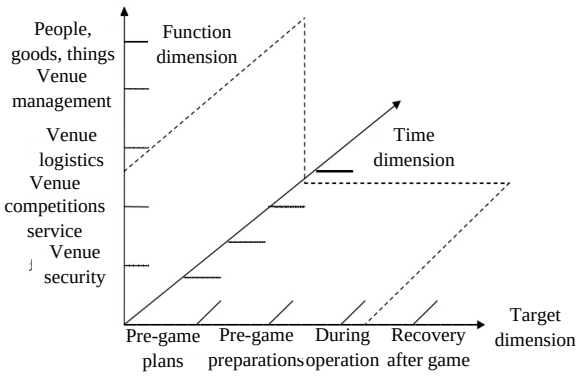


Figure 5. Stadium three-dimensional risk diagram of event management structure

C. Crowd Structure

Large stadium has perfect function, complex structure, and different kinds of internal facilities and equipment, so any a part in the design, purchase, construction, use and maintenance process can lead to the venue risk. Permanent structures, temporary works, building facilities, competition system, the special material and special material are the material basis of sporting event. Stadium is the venue of a sporting event and the facilities

equipment quality and maintenance guarantee the events go smoothly and athletes safety. At the same time effective equipped with emergency equipment can alleviate the damage of sudden emergencies. Therefore, we need systematic analyze the venue of risk source classification from the factors of material, to find the risk source damage to the venue, thus affect the events go smoothly and personnel safety. As shown in figure 5.

D. The rules of the Model Operation

In this paper, the model is on the basis of two-dimensional cellular automata, cellular automata is the discrete dynamic system that composed of finite state variables on uniform grid, its operation rule is:

(a) All of cellular automata state is happening simultaneously;

(b) The state of the ith cell of time t + 1 is determined by the moment / 3 y cell and adjacent distance less than y cell state.

Unlike regular cellular automata model, the model is defined in the two-dimensional plane, two-dimensional plane of building or other evacuation facility is divided into $L \times L$ uniform grid, each lattice point can only be empty or occupied by a person, where Z is the system size. Because the crowd density is big in the crowd evacuation , therefore, each cell model corresponding to $\frac{1}{3}m \times \frac{1}{3}m$ space. This space give full consideration to the model applicability in high density conditions, because of the high density case, individuals are tightly interleaved squeezed together, and pedestrians will deformation and compression because of extrusion, reached 8people $/m^2$ in the limit case. The maximum stand density was 9people $/m^2$ in our country, if not considering this problem, the density is too large, it may lead to failure simulator. Therefore, this paper cellular space divided into 9people $/m^2$ as the standard space partition. As shown in figure 6.

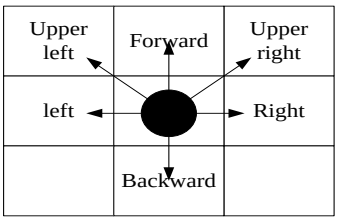


Figure 6. Stadium pedestrian movement pattern

As shown in figure 5, each cell of cellular automata is dependent on its field of all cellular automata, so it is very important to choose field. In this paper, the model use Moore type and field radius is 2, pedestrians can forward, backward, left, right, upper left, upper right, forward, left, right speed is Im/s, upper left, upper right speed is ()and backward speed is-1mA.

The evolvement of this model is:

(1) In a step T, whether single or travel in pairs personnel can only move one step in the model, can choose to move forward, backward, left, right, upper left, upper right. All personnel determine the next step lattice according to the state of the field and direction selection rules;

(2) When choosing the next target, according to target lattice formula to choose, pairs of pedestrians lattice selection according to the size of the sum of two ground field strength.

$$T_{ij} = \exp(-R_{ij})\exp(-F_{ij})\exp(D_{ij})E$$

$$E = \begin{cases} E_0 \exp\left\{-\min\left[\sqrt{(x_i - x_0^m)^2} + \sqrt{(y_i - y_0^m)^2}\right]\right\} & (x_i, y_i \text{ is not obstacle}) \\ 0 & (x_i, y_i \text{ is obstacle}) \end{cases}$$

(3) In the process of evolution, people can select a lattice is taken by previous step as their next target. However, only when two people choose the location is the next position, can swap. Otherwise, both keep still. People walking towards exchange ratio will be relatively high.

(4) After pedestrians move to the target lattice, the person according to the position again recalculates the next step target lattice, to determine the next movement direction;

(5) All personnel are updated simultaneously, all completion transfer to (1), and do the next update.

(6) After both sides personnel evacuation, stop the evolution process.

E. Behavior Execution

The speed obtained by the decision of other agents will update the location of each agent, the specific implementation process is shown in equation (8) as follows:

$$\text{P} \quad A^t = \left\{PN_1^t, PN_2^t, ...PN_n^t\right\} \tag{8}$$

where in, $\vec{\boldsymbol{P}}_t(x, y)$ is the current agent position; $\vec{\boldsymbol{P}}_{t+1}(x, y)$ is the position for the next time step; $\vec{\boldsymbol{v}}_{prefer}$ is the speed of agent after decision according to the environmental information and self status; Δt is the time step. Flow chart of the specific algorithm is shown in Figure 5.

Figure 7. Stadium

IV. SIMULATIONS

A. Simulation Environment

To test the usefulness of Cell Agent model, it is applied to evacuation plan of Changsha Stadium as shown in Figure 7. There are 245 nodes and 5 exports. Simulation experiments are implemented in the computer with PIV 3.3GHZ CPU, 3GB RAM, Windows XP. The algorithm programming language is as VC6.0 + +, and single CA model and MAS model are used for comparative experiment.

B. Results and Analysis

Under the circumstance of the different physical decline value power, the curve changes of the evacuation crowd and evacuation time are shown in Figure 7. As can be seen from Figure 8, the physical value decreased value power of has a greater impact on evacuation time.

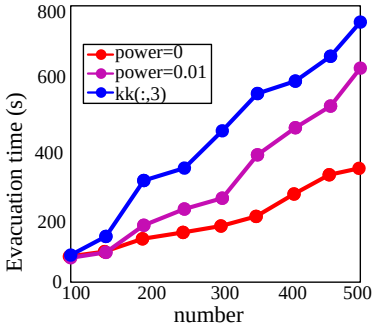


Figure 8. Flow chart of Cell Agent algorithm

Considering when the not panic value Cpanic, the curve changes of the evacuation crowd and evacuation time is shown in Figure 8. As can be seen from Figure 9, when considering the panic value Cpanic, the crowd evacuate faster at the beginning stage, but when the number of people is large, there are a large crowd, which causes the panic herd behavior, and then the phenomenon of the long evacuation time and the "haste makes waste" will be appeared.

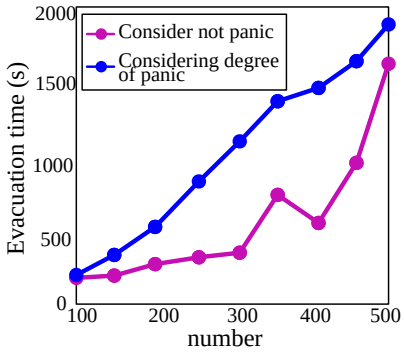


Figure 9. Impact of the level of panic on evacuation behavior

Under the circumstance of different Evacuation crowd, the stadium crowd evacuation time of single CA model, MAS model and Agent model are shown in Figure 10. Figure 10 shows that compared with the single model and MAS model, the total time of cell Agent model is relatively stable, and the crowd evacuation time is relatively small, which greatly save the total time of pedestrian evacuation. Comparative results show that Cell Agent model is good to overcome the shortage of personality differences caused by simple using the cellular automata, and at the same time avoid the computational complexity generated by simply use the multi-agent model and the macro problem of can not sufficiently presenting the crowd movement. The traditional one-dimensional cellular automata, when the time step T = 12, 88 people not evacuated, as shown in figure 9; The two-dimensional cellular automata agent, when the time step T = 34, only 10 people evacuation, as shown in figure 10. Obviously, the traditional one-dimensional cellular agent is much worse than two-dimensional performance in evacuation, new two-dimensional agent cell is superior to the traditional.

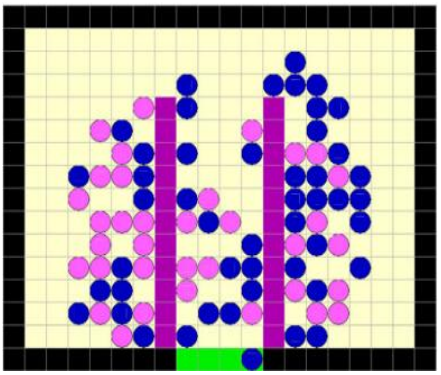


Figure 10. One-dimensional cellular automata evacuated figure

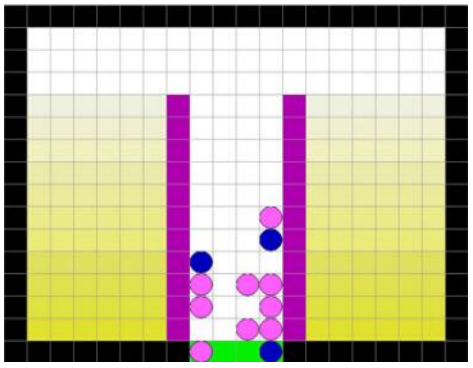


Figure 11. Two-dimensional cellular automata evacuated figure

The relationship between the number of evacuated crowd and time is shown in Figure 11; when considering the individual response time, the average response time is set as 60 s, the proposed model is adopted to re-simulate; the relationship between the number of evacuation with double outlet and time is shown as curve a in Figure 11. When considering the multi-agent movement, the relationship between the number evacuation and time is shown as curve a3 in Figure 7. Figure 11 shows that the curve a1 and curve a2 are approximate linear; a3 is nonlinear; the number of evacuation and time curve is nonlinear and they are comparatively real. Thus, by adding multi-agent and cellular automata in simulation, the phenomenon of crowd evacuation can be better simulated.

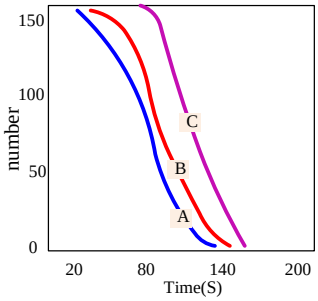
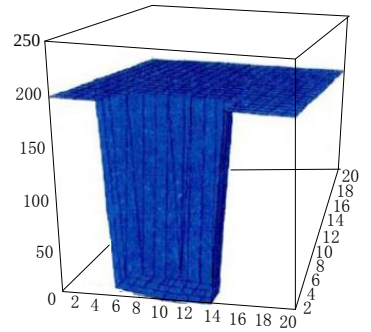
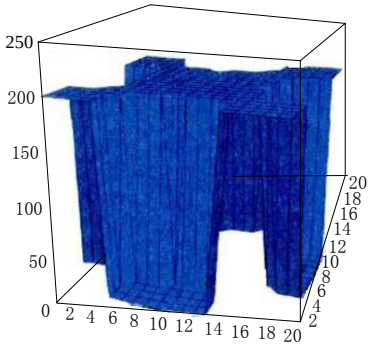


Figure 12. Relation curve between the number of exit s and evacuation time



(a)A single door evacuation



(b) Multiple door evacuation

Figure 13. Progress of dynamic simulation of crowd evacuation with emergency event

Assume that the emergency occurs, one road (or several rode) of crowd evacuating is closed for some reason, and the evacuation movement is simulated. At the beginning, the basic shape remains the same as when the roads are not closed, and the crowd inside the stadium began to take shape of the annular flow patterns, which are not significantly different displays, and it is the same with the actual situation consistent with audience in place; when in the progress of stimulating the crowd to spread and there are two outlet adjacent to the closed exit, the flow of people obviously increased, and the population around the closed exit is substantially gather at the nearest exit adjacent to these two exits (shown in Figure 12，13). In the subsequent diffusion process, in several second nearest exits the flow of people also have significant increase, which show that when the road is closed, the multi-agents in order to reach the original goal they have to take the bypass road, which will increase the flow of people in neighboring exits and roads. The duration of the entire simulation process is significantly longer, the complexity of the temporal and spatial variation is greatly increased, and also the complexity and difficulty of the treatment of emergency are significantly increased.

V. CONCLUSION

In order to eliminate the security risks in large stadiums, this paper proposes the crowd evacuation safety model based on multi-agent and cellular automata in large stadiums. The model is based on the characteristics and laws of crowd evacuation in large stadiums and combined with multi-agent and cellular automata. This model is closer to crowd evacuation situation in real large stadiums, which shortens the evacuation time and provides some reference for the scientific and rational analysis of crowd evacuation behavior in large stadiums.

REFERENCES

[1] D. Xu, Z. Y. Feng, Y. Z. Li, et al. Fair Channel allocation and power control for uplink and downlink cognitive radio networks. *IEEE., Workshop on mobile computing and emerging communication networks*, 2011 pp. 591-596

[2] Haoting Liu, Guohua Jiang, Li Wang, “Multiple Objects Tracking Based on Snake Model and Selective Attention Mechanism”, *International Journal of Information Technology*, vol. 12, no. 2, pp. 76-86, 2006.

[3] Chuanxu Wang, “Face Segmentation Based on Skin Color in Complicated Background and Its Sex Recognition”, *Journal of Software*, vol. 6, 2011, pp. 1209-1216.

[4] W. C. Kao, M. C. Hsu and Y. Y. Yang, “Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition” *Pattern Recognition*, vol. 43, 2010, pp, 1736-1747.

[5] Yinghua Xue, Hongpeng Liu, Intelligent Storage and Retrieval Systems Based on RFID and Vision in Automated Warehouse. *Journal of Networks*, Vol. 7, No. 2 (2012), pp. 365-369

[6] LI La yuan, LI Chun lin, "A multicast routing protocol with multiple QoS constraints," *Journal of Software*, vol. 15, No. 2, 2004, pp. 286-291.

[7] H. Tang, M. C. Chen, Y. S. Sun, et al. A spectral efficient and fair user-centric spectrum allocation approach for downlink transmissions. *IEEE., Globecom.*, 2011 pp. 1-6

[8] Muhammad J. Mirza, Nadeem Anjum. Association of Moving Objects across Visual Sensor Networks. *Journal of Multimedia*, Vol 7, No 1 (2012), 2-8

[9] Y. Geng, J. Chen, K. Pahlavan, Motion detection using RF signals for the first responder in emergency operations: *A PHASER project, 2013 IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, Britain Sep*. 2013

[10] S. Li, Y. Geng, J. He, K. Pahlavan, Analysis of Three-dimensional Maximum Likelihood Algorithm for Capsule Endoscopy Localization, *2012 5th International Conference on Biomedical Engineering and Informatics (BMEI), Chongqing, China Oct*. 2012 (page 721-725)

[11] Y. Geng, J. He, H. Deng and K. Pahlavan, Modeling the Effect of Human Body on TOA Ranging for Indoor Human Tracking with Wrist Mounted Sensor, *16th International Symposium on Wireless Personal Multimedia Communications (WPMC), Atlantic City, NJ*, Jun. 2013.

[12] Y. Geng, J. He, K. Pahlavan, Modeling the Effect of Human Body on TOA Based Indoor Human Tracking, *International Journal of Wireless Information Networks* 20(4), 306-317

[13] S. Kudekar, T. Richardson, R. Urbanke, "Threshold saturation via spatial coupling: why concolutional LDPC ensembles perform so well over BER," *IEEE*

[14] T. Gao, Z. H. Liu, J. Zhang, "Redundant Discrete Wavelet Transforms based Moving Object Recognition and Tracking", *Journal of Systems Engineering and Electronics*, vol. 20, no. 5, pp. 1115-1123, 2009.

[15] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for On-Line Nonlinear/ Nongaussian Bayesian Tracking", *IEEE Trans. Signal Process.*, vol. 50, no. 2 pp. 174-188, 2002.

[16] D. Cai, X. He, J. Han, and H. -J. Zhang, "Orthogonal Laplacianfaces for face recognition," *IEEE Transactions on Image Processing*, vol. 15, pp. 3608-3614, November 2006.

[17] L. Y. Ren, "Study on Scheduling Optimization in Construction Project of Lagerstroemia Hope City," *Xi'an University of architecture & techinology*. vol. 6, No. 2, 2011, pp. 12-17.

[18] Transactions on Information Theory, 2011, 57(2) pp. 803-834. http: //dx. doi. org/10. 1109/TIT. 2010. 2095072

[19] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, pp. 1373-1396, June 2003.

[20] R. Berangi, S. Saleem, M. Faulkner, et al. TDD cognitive radio femtocell network (CRFN) operation in FDD downlink spectrum. *IEEE, 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, 2011 pp. 482-486

[21] Pearson S. Taking account of privacy when designing cloud computing services. In CLOUD '09: Proceedings of the 2009 ICSE workshop on software engineering challenges of cloud computing, IEEE Computer Society, Washington, DC, USA, 2009. pp. 44–52. http: //dx. doi. org/10. 1109/CLOUD. 2009. 5071532

[22] Y. Yona, M. Feder, "Efficient parametric decoder of low-density lattice codes," *IEEE International Symposium on Information Theory: June 28-July 3, 2009, Seoul, Korea. New York, NY, USA: IEEE*, 2009, 8 pp. 744-748.

[23] Zhichao Lian, Meng Joo Er and Juekun Li, "A Novel Face Recognition Approach under Illumination Variations Based on Local Binary Pattern", *the 14th international conference on Computer analysis of images and patterns*, vol. 6855, 2011, pp. 89-96. S.

[24] A. Li, S. Shan, and W. Gao, "Coupled bias-variance tradeoff for cross-pose face recognition," *IEEE Transactions on Image Processing*, vol. 21, pp. 305-315, January 2012.

[25] K. I. Kin, K. Jung, and H. J. Kim, "Face recognition using kernel principal component analysis," *IEEE Signal Processing Letters*, vol. 9, pp. 40-42, February 2002.

[26] J. He, Y. Geng and K. Pahlavan, Modeling Indoor TOA Ranging Error for Body Mounted Sensors, *2012 IEEE 23nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sydney, Australia Sep.* 2012 (page 682-686)

Occluded Fingerprint Recognition Algorithm Based on Multi Association Features Match

Liu Wei-Chao and Guo Hong-tao
Academy of Information Technology, Luoyang Normal University LuoYang, China

***Abstract*—Recognition rate of mainstream fingerprint recognition algorithm is very low for occluded fingerprint image. In order to solve this problem, based on multi association matching features this thesis proposes the recognition algorithm (RA-MAMF). Firstly, image is pre-processing, the fingerprint image's Gabor filter is enhanced, and bi-narized and thinning pretreated are involved; secondly, the image is divided into multi homogeneous subsets, in which statistical association features and the bifurcation points of each subset are respectively extracted; finally on the basis of fingerprint, images are to be recognized, they are compared and match with the subset. Complete and occluded fingerprints data sets are used to make tested; the recognition algorithm based on multi association matching features has achieved excellent recognition accuracy; the RA-MAMF algorithm does not significantly increase the operating time, and this method effectively solves the low accuracy of traditional identifying the Occluded fingerprint image.**

***Index Terms*—Association Characteristics; Multi Features; Occluded Fingerprint Identification; Recognition Algorithm**

I. INTRODUCTION

As a tool for authentication, fingerprint has a long time. Between 7000 BC and 6000 BC, fingerprint has been used as the tools for identity authentication in ancient China and ancient Syria using [1] In 1923, Purkinje firstly classified fingerprints; in the late of 19th century, F.Glton carried out an extensive research on fingerprint; in 1960, FBI, Home offiee in the UK and Paris police invested heavily in research and development of Fingerprint Identification System together [2]; in 1975, the FBI's AFIS system and the Japanese NEC company's fingerprint identification systems are formed; in 1998,JAIN et al proposed the results integration of fingerprints and face recognition; in 2000, the methods if identifying the specific parameters of each user, which integrates the recognition results of fingerprint, face and gesture; in 2001, an overview of the multi-biometric is proposed [3]. In early 2006, Australia has successfully issued the world's first biometric passports. On November 2007, U.S. Department of Homeland Security announced that all non-US citizens should take digital camera pictures and should be scanned their ten fingers when entering the United States. Fingerprinting will be in rapid development period of popularity [4]. In 2009, the U.S. successfully update the fingerprint identification systems. In China, Tsinghua University began to study the fingerprint recognition in the 1980s [5]. Automatic Pattern Recognition Institute of Chinese Academy of Sciences National Laboratory has been committed to the research on "based biometric identity verification" since the 1990s, which has great achievements on fingerprint, iris, face recognition and so on. Perception National Laboratory of Peking University has undertaken the national scientific and technological projects, and has A long-term basic research on fingerprint identification, and has proposed a set of original theories and practical efficiency algorithm [7].

The original fingerprint recognition uses the artificial way, i.e. the fingerprint is got by hand-made; the fingerprint card is established and then the fingerprint is stored in the database. If there is a need for the print, it can be manually found in the fingerprint database and the fingerprint cards is compared by experts in artificial. After the 1960s [8], the optical scanners and other tools are used to transfer the fingerprint image into digital fingerprint, and then the advanced Automated Fingerprint Identification System (AFIS) based on digital image processing, pattern recognition, expert systems, artificial intelligence and other areas is gradually developed [9]. The computer is used to replace the manual, which can automatically process and match the fingerprint identification. The emergence and development of AFIS greatly reduces the required time of fingerprints saving, retrieving and identifying, which makes the fingerprint recognition technology as a safe and convenient authentication tool. It has been widely used in bank credit cards, computer user authentication systems, access control systems, ID card, network transmission security encryption and so on. Pattern Recognition (PR), also known as pattern classification, objects recognition and so on [10]. Through analyzing various information and data of object or phenomenon, the ubiquitous features set of the object properties are summarized and identified, and then the identify objects are identified (classified) according to the degree of similarity of characteristics. Factor collection which can express and summarize the identification object is called the pattern of object or phenomenon and the factors of the consisting pattern are known as the descriptor of the patterns [11], also is known as pattern characteristics. Thus, the pattern is a collection consisting of the representative characteristics of object (or sample). Pattern recognition includes two interconnected stages, namely the learning phase and the identification phase. The process forming and

doi:10.4304/jmm.9.7.910-917

summarizing the pattern is called the learning phase of model recognition. Learning phase makes feature selection for the sample set (also known as learning set or training set) to find the law of classification. After that the identification phase takes the classification and decision-making for the recognition object according to the formed mode, thus the process of object automatic recognition and classification is completed [12].

The existing mainstream fingerprint recognition algorithm and products are for complete fingerprint image, but in practical applications, due to uneven pressing efforts, over or under exposure and other factors the image brightness and strong lines contain noise or distortion; especially in Criminal Investigation fingerprint often is Occluded, damaged and Occluded, and therefore how to accurately identify the Occluded fingerprint has important practical significance [13]. The traditional feature extraction method and decision-making method can not effectively solve the fingerprint distortion, fault, Occluded and dislocation, and the filled interpolation method also can not correctively reflect the texture in torn region, especially for Occluded cases with a wide range. To solve the problem of difficult identification the Occluded fingerprint, some scholars from different angles makes improvements on fingerprint recognition algorithm. Literature [14, 15] based on direction filtering, median filtering, Gabor filtering, fuzzy operators and other methods improve the fingerprint image contrast, which enhances the image effect, but such methods are mainly to solve the problems of fingerprint image with noise, distortion and other issues, and also the calculation has higher complexity and it can not really solve the identification of Occluded fingerprint image. Literature [16] based on nonlinear interpolation and curve evolution interpolation enhance and complement the Occluded fingerprint image; literature [17] respectively based on fuzzy C-means clustering and watershed regions steer the image interpolation algorithm, which realizes the segmentation and enhancement of the Occluded image, but this type of interpolation algorithms are based on the original fingerprint image area to make supplement for Occluded areas, which is suitable for smaller and can not really reflect the real characteristics of the lines for the Occluded region, and it is with a strong contingency [17]. The minutiae feature points like bifurcation point and endpoint to some extent reflect the uniqueness of fingerprint images, which widely used in the complete fingerprint recognition algorithm [18], but it can not reflect the related information between each pixels, and thus it can not be directly applied to Occluded fingerprint identification. Statistical uses the regionalized variables as theory and semi-variogram as tool to study the spatial structure and correlation of objects [17, 19].

Semi-variogram function based on statistical extract the fingerprint image features, which can focus reflects the association information between the fingerprint pixels to make up for the shortage of minutiae points. In addition, based on global features matching algorithm, the blank portion of Occluded fingerprint image will extract the global characteristics, which causes a mismatch. If the fingerprints are from the same source, although Occluded fingerprints are not match with the global features of complete fingerprints, the local features be matched and the Occluded fingerprint image can be divided into a plurality of subsets. They determine the last match results based on the matching rate of subset features, which can prevent characteristic variation lead by fingerprint gaps. Based on the above two points, the Occluded fingerprint recognition algorithm based on association features and multi-subsets match is developed; the new method makes improvement to fingerprint recognition algorithm from the angles of features extraction and multi-regional subset match and applies this algorithm into the complete fingerprint database and Occluded fingerprint database identification, which achieves optimum results and demonstrates the effectiveness of the algorithm.

II. FINGERPRINT RECOGNITION MODEL OF RA-MAMF ALGORITHM

In the fingerprint identification process, due to various reasons the input fingerprint image is a gray image with lots of noise. The purpose of preprocessing is to remove the noise in the image to make the image and edge more clear, and then make the image to be a graph with clear point line, which offer convenience for extracting the correct fingerprint feature. The preprocessing of fingerprint image is important in the whole fingerprint recognition system, which will have a direct impact on the results of fingerprint identification. Preprocessing is generally divided into four steps: image segmentation, image filtering, binarization and thinning.

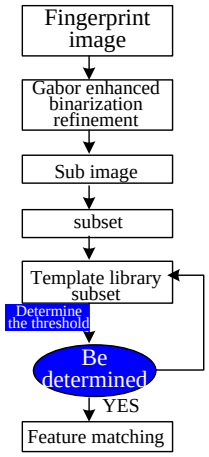


Figure 1. Flow chat of Occluded fingerprint identification

Ideas of fingerprint recognition based on RA-MAMF algorithm are need to keep some rules. Gabor filtering is used to de-noise and enhance the fingerprint image, and the binary processing and refining treatment are used. The whole image is evenly divided into N sub-images, and the non-directional characteristic like statistical correlation characteristics, the bifurcation point and end points of each sub-image are extracted to generate the feature subsets. Based on the fingerprint image need to be tested, feature subsets of the fingerprint image in database are matched, and the percentage of match subset is used

to determine the belonging. Flow chat of RA-MAMF algorithm is shown in Figure 1.

III. FINGERPRINT RECOGNITION ALGORITHM DESIGN OF RA-MAMF

A. Image Pretreatment

The most important step of Fingerprint preprocessing process is the filtering for fingerprint image, which is the core issue to be resolved in fingerprint image preprocessing. The purpose of filtering is enhancing the contrast of the ridge valleys structure and suppressing the noise at the same time, connecting the breaking ridge line and separated ridge line, highlighting some of the information according to the specific image needs, and weakening or removing some unwanted information.

Due to the light intensity, temperature, intensity and other factors, fingerprint extraction is with vary degrees of noise, which affect the fingerprint feature extraction and fingerprint recognition, so the fingerprint image is preprocessed, which can eliminate the noise to improve the quality of fingerprint image recognition. (1) Gaussian smoothing. Initial fingerprint image is convolved with the Gaussian function, which can effectively suppress the noise, so that the fingerprint image is smoother. Assuming the original image is $F(X_t|\lambda)$ and X_t, λ are the rows and columns of image, Gaussian smoothing formula is as follows:

$$F(X_t|\lambda)=\sum_{i=1}^{M}\omega_i G(X_t|\mu_i,\Sigma_i) \tag{1}$$

where in, ω is the variance of the Gaussian function. Calculating the direction of the field. The image is divided into k × w pieces; the pixel is used as the unit; and Sober operator calculate the gradient $L(X/\lambda)$ pixel:

$$G(X_t|\mu_i,\Sigma_i)=\frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}}\exp[-\frac{1}{2}(X_t-\mu_k)^T\sum_i^{-1}(X_t-\mu_k)] \tag{2}$$

$$L(X/\lambda)=\log\prod_{t=1}^{T}P(X_t/\lambda)=\sum_{t=1}^{T}\log P(X_t/\lambda) \tag{3}$$

where in, λ and X are the coordinates of the pixels. The direction field value of each sub-block center pixel (i, j) is calculated as follows:

$$\begin{aligned}L(X/\lambda) &=\sum_{t=1}^{T}[\sum_{i=1}^{M}g(i|X_t,\lambda)]\log P(X_t/\lambda)\\ &=\sum_{t=1}^{T}q_t(\Theta;X_t)\end{aligned} \tag{4}$$

$$g(i|X_t,\lambda)=2I(i|X_t,\lambda)-h(i|X_t,\lambda) \tag{5}$$

$$h(i|X_t,\lambda)=\frac{\omega_i G(X_t|\mu_i,\Sigma_i)}{P(X_t|\lambda)} \tag{6}$$

where in, ω_i is the direction field value of each sub-block center pixel (i, k) of the ridge line of least squares. Calculating the frequency field. Because the grey value of the fingerprint ridge lines in the direction perpendicular to the line shows the gradation generally sine law, through calculating the average number of pixels p (i, λ) between adjacent peaks to get the final ridge frequency h(i , k) = 1 / p (i, λ). Enhancing Gabor filtering. After obtained frequency field and orientation field of fingerprint image, two-dimensional Gabor filtering operator can be constructed to get to the enhanced de-noising fingerprint image. Two-dimensional filtering operator can be expressed as follow:

$$\omega_i=\frac{\exp(\beta_i)}{\sum_{k=1}^{M}\exp(\beta_k)} \tag{7}$$

$$\beta_i^n=\beta_i^o+\eta_\beta[g(i|X_t,\lambda^o)-\omega_i^o] \tag{8}$$

$$\mu_i^n=\mu_i^o+\eta g(i|X_t,\lambda^o)(\Sigma_i^o)^{-1}(X_t-\mu_i^o) \tag{9}$$

where in, β is the value of the Gabor's direction field; f is the frequency field value of Gabor; β_i^n and μ_i^n respectively are Gaussian envelope values on the two-axis direction in coordinates.

B. Bi-Narization

Refinement is a process of cutting the edge pixels until the line is a single pixel wide line, based on not influence the line's connectivity after the binarization of fingerprint. The dreaming refined skeleton of line should be in the ideal middle location of the original ridge line and maintaining the connectivity, topology and features of pattern lines are the same time.

Fingerprint image binarization processing refers to transfer the fingerprint image into fingerprint ridge structure of grayscale with black and white shown in Figure 2. If the sum of cut grey value is less than the sum of normal gray value, the black dot can be judged as that it locates in the fingerprint ridge line.

Figure 2. Binarization renderings

Processing formula of bi-narization method is as following:

$$(\Sigma_i^n)^{-1}=[1+\eta g(i|X_t,\lambda^o)](\Sigma_i^o)^{-1}-\eta g(i|X_t,\lambda^o)U_{t,i} \tag{10}$$

$$U_{t,i}=(\Sigma_i^o)^{-1}(X_t-\mu_i^o)(X_t-\mu_i^o)^T(\Sigma_i^o)^{-1} \tag{11}$$

where in, $U_{t,i}$ indicates the pixel gray value; X_t is the direction field; δ represents the determined lookup range; H is the sum of cut pixels value of the assessing pixels; V is the sum of normal pixels. If H> V, the pixel is located in the valley line, otherwise in the ridge line. Refining is

the main role of the fingerprint image, refining is removing the uneven lines in original binarization fingerprint image and the atypical feature information. Through deleting or keeping the way, the center line of the stripe is wide and unique, thereby the fingerprint image features are outstanding to improve the image processing speed. This thesis uses OPTA (One Pass Thinning Algorithm) to refine the fingerprint image.

C. Extracting Fingerprint Features

Fingerprint minutiae is the most common feature in artificial matching, because the probability of existing the branching point, end point is high and easy to be detected. And more importantly, they are sufficient to describe the uniqueness of the fingerprints. 8-neighbor method is used for extracting branch points and ends points of the fingerprint ridges. Assuming P is the target pixel point (pixel to be processed), the circumference of the adjacent 8:00 P1, P2,..., P8 are called 8 neighborhood points of point P.

For fully refined binary image there are only two kinds of conditions for gray value of pixel: 0 represents the background point gray; 1 is the grey value of ridge point. For any point P, set V1, V2,..., V8 are the gray value of point P's 8 neighborhood points; if P is the endpoint, its eight neighborhood points satisfies the following:

$$\omega_i^n = \frac{\exp(\beta_i^n)}{\sum_{k=1}^{M} \exp(\beta_k^n)} \tag{12}$$

If n is the branch point, its eight neighbor points satisfy the following;

$$\mu_i^n = \mu_i^o + \eta m_i \tag{13}$$

Finally, based on the direction of the edge features and detection of field features the detects false feature points of branch points and end points are extracted and removed.

Minutiae features of Bifurcation points and the endpoints can reflect the uniqueness of in the fingerprint image to some extent, but they can not reflect the correlation information between pixels. Especially for those fingerprint images which are hard to identify, in which the bifurcation point and endpoints are basically the same, making it impossible to distinguish, so a characteristic fingerprint associated with an earth system science is put forth. The extraction process of statistically associated with feature is as follows: First, the binarized fingerprint image is evenly divided into k × w small blocks; each block contains 10-20 pixels. After the division of the sub-block, the entire fingerprint image data can be seem as a series of data η (i), i=1, …, n, n=k×w, its which experimental semi-variogram values r (h) is shown in the following formula:

$$\sum_i^n = \sum_i^o + \eta(R_i - S_i) \tag{14}$$

(4) Extracting the statistical correlation characteristics. According to step (3) to determine the process variable, r (h) in the process is selected to reflect the feature value association information between image pixels; association features are the corresponding r(h) with distance one to six

When the traditional Occluded fingerprint recognition algorithm based on global features is used for fingerprint image recognition, since the Occluded fingerprint image lost parts of the texture, there is big difference from extracting its global features and extracting global features from complete fingerprint image, but some reserved parts are still the same with the complete images. Apparently in the recognition algorithm of Occluded fingerprint image, subset matching results based on local features are superior to global features, so a matching recognition method based on multi-subsets is designed. First, the entire pretreated image is evenly divided into k sub-images, and then no orientation features were extracted from each sub-image to form k feature subsets. Matching fingerprint images in fingerprint library are made the same process; each image is generated k feature subsets. Let Occluded fingerprint image to be identified as A, in which the number of complete image in the fingerprint database is B. Matching process of multi-subset is as follows: Determining the effective subset. Since after the division of Occluded image A the number of pixels in sub-sets formed by the Occluded parts is scarce, they are defined as the valid subset, which is not involved in match count. The subset image which its pixels are 80% is defined as valid subset; assuming a valid subset of the image to be recognized number is as m (m ≤ k). Matching the single sub-set. The first valid subset of the image A is respectively by-matched with all subsets of the image B; when its matching rate of a subset in image B exceeds a given threshold value, i.e. the valid subset is successfully matched; if all is not matched, the valid subset matching is unsuccessful. Matching multiple sub-set. The extracted m valid subsets are respectively repeat the process of matching the single subset. If the successfully matched valid subset exceeds a given threshold, the source of the Occluded image A is the same with images B with database.

IV. EXPERIMENT AND ANALYSIS

A. Experimental Platform and Standards

In order to verify the effectiveness and superiority of RA-MAMF, fingerprint data from FVC2004 DB1_B, DB2_B, DB3_B and DB4_B are usedto extract the 320 fingerprint images; each DB set contains 10 people; each people has 8 images. Three images are randomly selected from eight images of each people as the independent test samples, and the remaining is as the matching database. When the Occluded fingerprint recognition algorithm is verified, the fingerprint images in independent test samples are artificially make Occluded processing, and the Occluded range is from 5 % to 50 %. The whole algorithm platform is conducted on the Intel Pentium (R) 2.8GHz dual-core CPU, 4G memory, PC with windows 7 operating system, and all the algorithms are programming in Matlab 2012.

In order to make the results of RA-MAMF algorithm are more persuasive, three kinds of comparative

algorithms are chose: Tradition algorithm (only using the endpoint and the intersection of two minutiae features, and do not use the multi-subsets matching strategy); SM algorithm (only extracting two kinds of minutiae features, but using the multi-subset matching strategy); SCC algorithm (extracting the endpoint, intersection and the statistical associating features, but do not using the multi-subset matching strategy). For Occluded data sets, first three kinds of classic methods are selected to make comparing experiments. From the point of view of the effectiveness of RA-MAMF algorithm: local feature extraction method based on the micro characteristics [17] (2-SM, extracting peripheral point, branching point, isolated point, hole, intersection and short ridge); correlation characteristics based on triangle model [18] (3-SM); local features and global features of Gabor filter based on eight directions [19] (MSRC). The three kinds of feature extraction methods use multi-subset matching strategy; In addition, for Occluded fingerprint data, the Occluded fingerprint recognition model is built: Adaptive Generic Algorithm [20] (MSRC) and b Fingerprint Image Curve Changes [9] (FICC) are taken ad the reference model. Final evaluation criteria for each model use the fingerprint recognition rate, EER and average recognition time as three indicators.

d) Comparison of Occluded fingerprint image recognition

120 testing images are manually Occluded processed, and 5%-50% pixels are removed with the different shapes, different positions and multi-points, as shown in Figure 3.

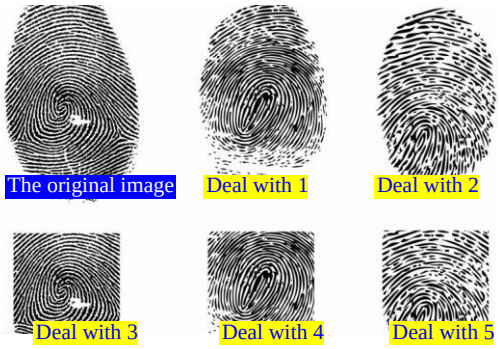


Figure 3. Different occluded processing

B. Results and Analyses

1) Results Comparison Of Recognition Algorithms

In order to validate the performance indicators of the proposed fusion algorithm in this chapter, the proposed algorithm and fusion algorithm based on entropy theory are compared with a single matching algorithm. On Database, in order to illustrate the problem of Occluded fingerprint identification, the fingerprint image in low-quality fingerprint image FVC2004 database are adopted to make simulation experiments with the proposed mechanism in this chapter. Firstly, three parameters measuring the efficiency of fingerprint identification are introduced. In fingerprint recognition system, False Match Rate (FMR), False Non-Match Rate (FNMR) and Equal Error Rate (ERR) are used to evaluate the performance of the algorithm.

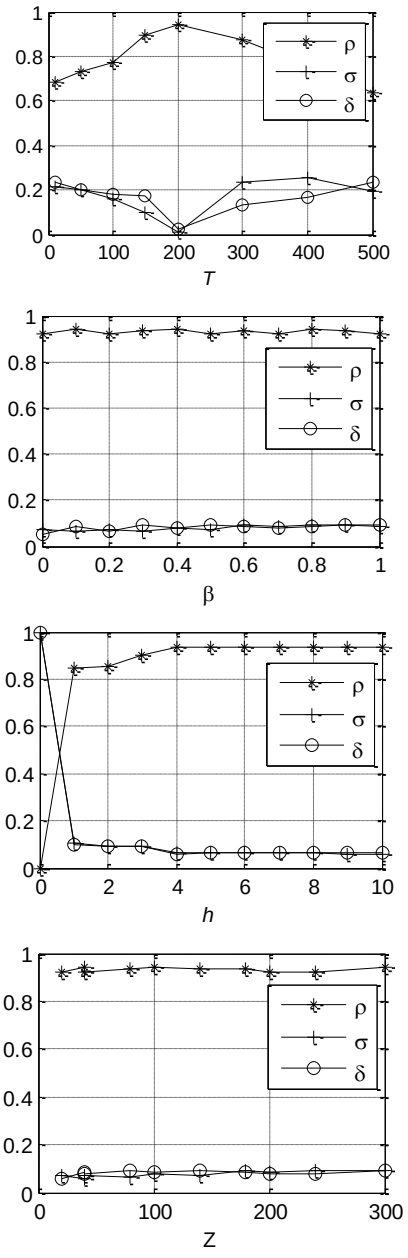


Figure 4. Contrast curve of several fingerprint matching algorithms

ROC curve of the proposed algorithm is always at the bottom of a single algorithm, which indicates that the proposed fusion mechanism can effectively identify the matching fingerprints and reduce the fingerprint matching errors. Second, the performance of fusion algorithm based on entropy theory is basically higher than fusion mechanism based on triangular mold operator, which indicates that the entropy theory-based algorithm can better adapt to the fusion of multi-class match operator. It can accurately reflect the real situation of different matching and improve the recognition performance. On the third, the proposed association features and multi-subset matching algorithm are associated with the quality of fingerprint, which indicate that when the quality of fingerprint is poor, the ROC curve of entropy-based algorithm is significantly lower than the curve of algorithm based on triangular mold operator; the

fusion method based on entropy theory can use the uncertainty to coordinate treatment for different matching algorithms. In the Occluded fingerprint matching identification, uncertainty is reduced and the comprehensive, accurate and reliable identification results are obtained. On the Fourth, the ROC curve of matching algorithm based on association features is lower than the algorithm based on minutiae, which indicates that the optimal solution can be found in Occluded fingerprint identification, i.e. the best match.

2) Results of Omparison of Match

To determine the optimal matching thresholds the single subset match" and "multi-subset match", the Occluded fingerprint images are used as objects and the thresholds are gradually sided (sliding range is between 0.5-1.0), and also the training sample is calculated to determine the identification accuracy shown in Figure 4. Score values of "single subset match" are defined as the correlation coefficient of two feature subsets.

$$m_i = \frac{1}{N}\sum_{t=1}^{N} g(i|X_t, \lambda^o)(\Sigma_i^o)^{-1}(X_t - \mu_i^o) \qquad (15)$$

where in, x is the feature subset of template fingerprint; y is the feature subset of fingerprint matching; r is the correlation coefficient.

The score of "Multi-subset match": the ratio of successfully matched subset to the effective subset.

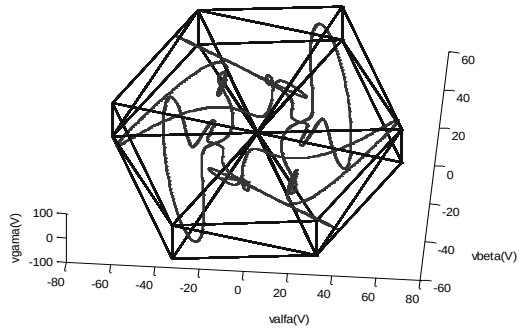


Figure 5. Accuracy comparison of different matching threshold

Can be seen from Figure 5, when the single subset Threshold is set as 1, which requires the feature subset of template is the same with the matching feature subset (correlation coefficient is 1), resulting in a low successful match rate; when the multi-subset Threshold is relatively high, the sample successfully matched rate is 0, so that the prediction accuracy is 0. When the single subset Threshold is set as 0.85 and the multi-subset threshold is set as 0.9, the right sentence rate of the training set is the highest. Therefore, the proposed single subset threshold and multi-subset match threshold are set as 0.78 and 0.86.

3) Contrast on the Complete Fingerprint Image Recognition

For the complete fingerprint image in FVC2004, the prediction accuracy of each algorithm model is shown in Table I. after analyzing the comparison in Table I, the following conclusions can be obtained:

(1) Tradition and SM algorithm only use the endpoints and the intersection of minutiae, and the average recognition rate is of 87.5%, which indicates that these two features have been distinguished the differences between the fingerprints to some extent, but its identification effect is not ideal.

(2) RA-MAMF and SCC algorithm use three kinds of feature framework recognition models containing statistical association features, and their recognition rate are higher than 92%, especially after adjustment of multi-subset matching; the recognition rate of RA-MAMF algorithm reaches 94.72%, and the results shows statistical association features can effectively reflect the correlation between regions of the fingerprint image, which make up minutiae defects only reflecting the structural changes, so that the e fingerprint image contains more comprehensive and credible information.

(3) The difference between SCC algorithm and RA-MAMF algorithm is the using of a multi-subset matching policy; the difference between Tradition and SM algorithm is the using of multi-subset matching strategy at the same time. As seen from Table I, when a fingerprint is completely recognized, multi-subset matching only slightly improves the recognition accuracy, but the running time relative increases and its value is not obvious.

Recognition results of different algorithms for Occluded fingerprint images are shown in Table II. From Table 2 it can be known:

(1) Compared with Tradition and SCC algorithm (not using the multi-subset matching strategy), recognition accuracy of RA-MAMF and SM algorithm (using the multi-subset match) is significantly higher.

(2) SCC algorithm is modeled based on three kinds of characteristics of the image, but the multi-subset matching strategy is not used; the recognition accuracy for complete fingerprint is of 92.17%, but recognition accuracy for Occluded fingerprint is only of 55.9%.

(3) SM algorithm uses two kinds of minutiae, but uses the multi-subset match strategy; when identify the Occluded fingerprint image the accuracy does not show significant decline.

(4) RA-MAMF algorithm has a very high recognition rate on complete fingerprint recognition and Occluded fingerprint recognition. Although using the multi-subset matching algorithm leads to a slight increase in running time, compared with the improvement of recognition accuracy the effects are not serious.

TABLE I. RECOGNITION ACCURACY OF DIFFERENT ALGORITHMS FOR OCCLUDED FINGERPRINT IMAGES

Model	Recognition rate (%)	Average running time (s)
RA-MAMF	93.76	1.67
Tradition	50.28	1.09
SM	84.01	1.58
SCC	55.9	1.12

TABLE II. RECOGNITION ACCURACY OF DIFFERENT ALGORITHMS FOR THE COMPLETE FINGERPRINT IMAGE

Model	Recognition rate %)	Average running time (s)
RA-MAMF	94.72	1.94
Tradition	89.36	1.12
SM	90.59	1.97
SCC	92.17	1.22

C. Recognition Performance

In order to further verify the efficiency of SCC association features for Occluded fingerprint image recognition, it compare with other three classic features extraction method for fingerprint image, and the results are shown in Table III, through the analyzing Table III it can be known: 2-SM model uses the typical local feature extraction method, which obtains the image' end point, a diverging point , an isolated point, hole, short ridge intersection of seven bridges minutiae, but the simple local characterization can not sufficiently express the information of fingerprint images, resulting in the low recognition accuracy; 3-SM extracts feature association features from the distance among three feature points lines and angles, reflecting association information of local area to some extent; compared with 2-SM method, the accuracy of 3-SM has been improved, but it do not get the relevant information and the final prediction accuracy is not very satisfactory; MSRC uses Gabor filter from eight directions to extract the local and global features of the fingerprint image; compared to the two reference models, the prediction accuracy is greatly improved, but the operator does not consider the feature association information of the image and recognition accuracy for Occluded fingerprint images is still weaker than RA-MAMF algorithm; RA-MAMF algorithm takes into account the local detail information of each subset and the global correlation information of the image at the same time; SCC association features through calculating the pixel differences between each sub-block with the same distance reflects pixel-related information of the semi-variogram functions, and finally based on different semi-variogram values multi-scale relevant information is achieved; and through multi-feature subset the importance of local features for Occluded fingerprint image is reinforced, and the independent testing accuracy is the highest in the participial model.

TABLE III. RECOGNITION ACCURACY OF DIFFERENT FEATURE EXTRACTION METHODS FOR OCCLUDED FINGERPRINT

Model	Recognition rate	Average running time (s)
RA-MAMF	93.45%	1.85
2-SM	80.96%	1.64
3-SM	82.62%	1.77
MSRC	86.33%	2.62

TABLE IV. RECOGNITION ACCURACY OF DIFFERENT OCCLUDED FINGERPRINT ALGORITHM

Recognition	Occluded degree			
algorithm	15%	25%	35%	50%
NN	76.8	77.6	78.9	79.0
SRC	83.4	77.8	65.9	41.2
BSRC(25*25)	93.9	89.8	78.2	62.6
BSRC(25*50)	92.0	88.5	73.4	57.7
BSRC(50*50)	85.5	81.3	68.4	45.0
MSRC	95.1	91.7	83.6	67.4
RA-MAMF	97.8	95.6	87.9	75.3

D. Recognition Performance Comparison of Occluded Fingerprint Algorithm

To compare the recognition performance of Occluded fingerprint recognition algorithm to the Occluded fingerprint image, 120 fingerprint images manually processed are divided into five levels according to the Defect Images (DI), the predict results of each Occluded fingerprint recognition algorithm are shown in Table IV.

MSRC algorithm is the optimal migration algorithm based on adaptive genetic algorithm; when it searches the appropriate benchmark matching feature points, matching problems of fingerprint features are partially solved, and there is some improvement in terms of accuracy; but when the degree of defect of fingerprints is between 48% $\leq$ DI <60%, the recognition accuracy is sharply decreased. Through analyzing the changes of fingerprint contour curves, FICC algorithm calculates the evolution of the fingerprint curve, and then through the discrete iterative methods, the Occluded fingerprint images are repaired. When the degree of defect is small, after image filled by the image recognition accuracy can be effectively improved; when the degree of defect is between 2% to 18%, the recognition accuracy of the test sample is up to 100%; but with the expand of the degree of defect, the effectiveness is greatly reduced, and the overall prediction accuracy is the lowest among the three reference models. The proposed RA-MAMF based on multi-subset match, and the prediction accuracy is the highest; with the increase of Occluded range, the recognition accuracy is decline slower than the reference models.

V. CONCLUSION

Recognition rate of mainstream fingerprint recognition algorithm is very low for Occluded fingerprint image. In order to solve this problem, this thesis proposes the algorithm based on association features and multi-subset matching (RA-MAMF). RA-MAMF algorithm improves the fingerprint recognition algorithm from the points of the integrity of feature extraction and multi-subset matching. And also the RA-MAMF algorithm is applied to the complete and Occluded fingerprint recognition, which has achieved excellent recognition accuracy; the RA-MAMF algorithm does not significantly increase the operating time, and this method effectively solves the low accuracy of traditional identifying the Occluded fingerprint image.

REFERENCES

[1] Wagner D, Reitmayr G. Real-time detection and tracking for augmented reality on mobile phones. *IEEE Trans. on Visualization and Computer Graphics*, 2010, 16(3) pp. 355-368.

[2] Kong Ww, Lei Yj, Lei Y, et al. Fusion Technique for Gray-scale Visible Light and Infrared Images Based on NSCT and IHS Transform. *IET Signal Processing*, 2011, 5(1) pp. 75~80.

[3] Zitova B, Flusser J. Image registration methods: a survey. Image and vision computing, 2003, 21(11) pp. 977- 1000.

[4] Rkodtook A, Makhanov S S. Selection of multiresolution rotationally invariant moments for image recognition. *Mathematics and Computers in Simulation*, 2009, 79(8) pp. 2458~2475.

[5] LOWE D G. Distinctive image features from scale invariant key points. *International Journal of Computer Vision*, 2004, 60(2) pp. 91-110.

[6] Yishuang Geng, Yadong Wan, Jie He, Kaveh Pahlavan, An Empirical Channel Model for the Effect of Human Body on Ray Tracing, *2013 IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, Britain Sep.* 2013 pp. 47-52.
[7] Mikolajczyk K, Schmid C. Scale & affine invariant interest point detectors. *Computer Vision*, 2004, 60(1) pp. 63-86.
[8] Lowe D. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2) pp. 91-110.
[9] Mikolajczyk K, Schmi C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2005, 27(10) pp. 1615 -1630.
[10] Zhu H Q, Liu M, Li Y. The RST invariant digital image watermarking using Radon transforms and complex moments. *Digital Signal Processing*, 2010, 20(6) pp. 1612~1628.
[11] Suk T, Flusser J. Affine moment invariants generated by graph method. *Pattern Recognition*, 2011, 44(9) pp. 2047-2056.
[12] Yishuang Geng, Jie He, Kaveh Pahlavan, Modeling the Effect of Human Body on TOA Based Indoor Human Tracking, *International Journal of Wireless Information Networks (IJWIN)* 20(4), 306-317, Dec. 2013
[13] Weston J, Watkins C. Support vector machines for multi-class pattern recognition: Proceedings of the seventh European symposium on artificial neural networks, 1999. Bruges: [s. n.], 1999, 4(6) pp. 219-224.
[14] Lowe D G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 2004, 60(2) pp. 91-110.
[15] Rekha J, Bhattacharya J, Majumder S. Hand Gesture Recognition for Sign Language: A New Hybrid Approach: proceeding of International Conference on Image Processing, *Computer Vision and Pattern Recognition*, IPCV, 2011. 2011, 11.
[16] Bay H, Ess A, Tuytelaars T, et al. SURF: Speeded up robust features. *Computer Vision Image Understand (CVIU)*, 2008, 110(3) pp. 346-359.
[17] JIANG Yu-Gang, NGO C W, YANG Jun. Towards optimal bag-of-features for object categorization and semantic video retrieval: Proceedings of the 6th ACM international conference on Image and video retrieval, 2007. *New York: Association for Computing Machinery (ACM)*, 2007 pp. 494-501.
[18] Arther D, Vassilvitskii S. k-means++: The advantages of careful seeding: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, 2007. *Pennsylvania: Society for Industrial and Applied Mathematics*, 2007 pp. 1027-1035.

Multi-View Distributed Video Coding Based on Discrete Cosine

Liu Guanqun and Ling Zi
HuNan Radio & TV University, Changsha, 410004, China

***Abstract*—To investigate the allocation scheme of the multi-view distributed video coding (DVC), the corresponding improvements are proposed correspondingly for traditional multi-view DVC. Traditional multi-view DVC (Wyner-Ziv DVC) encodes for all areas of Wyner-Ziv frame indiscriminately based on Turbo or LDPC. In this kind of encoding process, with regard to violent motor area, decoder can't decode violent motor area accurately and also send more solicited message to feedback channel, which lowers the code efficiency and decodes inaccurately for violent motor area, it causes a part of area distortion in the image. In this paper, a distributed video encryption algorithm is proposed which based on discrete cosine transform (DCT). The algorithm combines decision criteria of ROI to get violent motor area and non-violent motor area. For violent motor area, to extract low frequency coefficient of DCT as DCT-R algorithm to assist decoder end to decode, decoder utilizes low frequency coefficient of DCT which has already been decoded to carry on bi-directional movement evaluation. Simulation experiment tests and verifies the improved algorithm effectiveness of proposed multi-view DVC in this paper.**

***Index Terms*—Distributed Video Coding; Discrete Cosine Transform; Region of Interest; Multi-Perspective**

I. INTRODUCTION

In recent years, with the rapid development of wireless multimedia communication technology, wireless low power monitoring network, disposable camera, medical application, multi-perspective image acquisition and other new video applications occur unceasingly, its characteristic is a large number of resource(such as CPU, energy content of battery, communication broadband) is restricted in application [1].

Therefore, to adapt new demand of video application, it must require low-complexity encoder. Traditional video coding standard(such as MPEG, H.26) adopts technology of motion estimation and motion compensation , it causes encoding complexity is 5-10 times than decoding complexity [2], it's hard to satisfy the code low-complexity requirement of new video application. Slepian-Wolf and Wyner-Ziv theories of information theory indicate [3]: In encoding and decoding system, decoder end in the precondition of jointed decoding, encoder end carries on encoding independently or jointly for two related information resource [4], and achieving same distortion performance at last. Distributed video coding is based on these two theories, DVC adopts code rule which is completely different with traditional video encoding standard, encoder end adopts independent coding technique for the figures, videos which collect by wireless multimedia sensor networks, decoder end carries on decoded prediction by excavating interdependency between image frame and frame in video to achieve high efficiency encoding [6]. This kind of distributed video coding has characteristic of low-complexity encoding and good robustness, it's able to be the same with new video application (For example wireless camera, wireless low power consumption monitoring network, disposable camera, medical application, multi-perspective image collection et.). In some practical applications, video application has great capacities specially, so more and more experts at home and abroad start to follow with interest and deeply study distributed video coding program [7].

From the perspective of information theory: quantity of information and information redundancy capability form information source data, the main purpose of information source encoding is that to compress data, keep information quantity, and decrease information redundancy capability as far as possible [8]. In multi-angle of view, video data posses the overwhelming majority of multi-media information which needs to handle, therefore, the multi-media encoding technology is mainly to study video coding in multi-media data. Video sequence not only has spatial redundancy and time redundancy of one frame image which exists in two successive frames of images, but also posses certain coding redundancy and mind vision redundancy [9].

Multi-media information is necessary data in human communication, mainly including voices, images, videos etc. To make multi-media information carry on transmitting, saving to achieve the level of handling data in real time, bring great pressure to network transmission [10]. Thus, to reduce data volume of video stream during transmission, lighten pressure of network transmission, compression technology of multi-media information is very necessary [11].

As audios, images, videos etc multi-media data include large volumes of redundancy information, so all of them have certain compression space. Such as bit map format of image storage, whatever in row or column direction, interdependency between two pixels is very great, and then the whole image has certain redundancy. And video forms by several consecutive frames of images, not only with spatial redundancy of one frame of image and time redundancy of two successive frames of images, also has

doi:10.4304/jmm.9.7.918-925

some coding redundancy and mind vision redundancy. Compression technology of multi-media information is very necessary, reducing transmitting data volume of video stream, lightening network transmission capacity.

With the International standard appearance of H.261, H.263, H.264/AVC, MPEG-1, MPEG-2, MPEG-4, MPEG-7, a series of video compressions, video coding technology is tending to mature; it makes video compression and communication go into a new epoch [12]. The goals which H.26X series and MPEG-X series pursuit are same, transmitting codes as little as possible, the images which get by decoding can have higher quality. Favorable encoding technology is able to handle video, audio data in real time, improving efficiency of transmission channel, and then it can promise audio and video quality. Widely use of video conference, telemedicine, video education etc multi-media business, pushing forward the development of human communication. For the pixel value of different occurring probability in video sequence, if we adopt same length of code word to encode pixel value, it will cause certain Bit waste, which is coding redundancy. To solve this kind of redundancy, it adopts different length of code word to encode, short code word for pixel value of high frequency, long code word for pixel value of low frequency.

Distributed video coding constructs video coding program which based on coding theory of lossy information source of Wyner-Ziv. Slepian-Wolf and Wyner-Ziv theories in information theory show [13]: In encoding and decoding system, when decoder in the precondition of joint decoding, encoder carries out joint or individual encoding for two related information source, and achieves same distortion performance at last. DVC is based on these two theories, DVC adopts the coding rule which is different with the traditional video coding standard, encoder adopts technology of individual encoding and joint decoding for images and videos which WMSNs collects, encoder does not make prediction of frames, decoder predicts decoding to reach high frequency encoding and get more and more attentions and researches of scholars, by utilizing that excavating the interdependency method among frames of images in video.

Presently, typical distributed algorithm of encoding and decoding: Wyner-Ziv video coding [14] which proposed by Girod et al from Stanford University, video coding of Power-efficient Robust high-compression Syndrome-base Multimedia (PRISM)proposed by Ramchandran et al from University of California at Berkeley, video layered video coding of Wyner-Ziv proposed by Zixiang Xiong et al, bringing wavelet transform into distributed video coding program, it adapts to multi-perspective distributed video coding which has high requirement of video quality, state-free distributed video coding proposed by Sehgal et al and so on. Literature [15] based on turbo or LDPC that encode for all areas of Wyner-Ziv frames indiscriminately and uniformly, decoder can't decode violent motor area accurately and send more solicited messages to feedback channel, it lowers encoding frequency.

Although implementation methods of above several typical programs, these several distributed video coding has some characteristics in common: Encoder regards image frame as mutually independent information source in video stream, do not predict frame while encoding, encodes each frame of image independent which similar with traditional encoding among frames; Decoder passes through time domain interdependency among continuous frames of images, excavating interdependency between frame and frame of image to make a prediction, adopting motion evaluation method, using time domain interpolation for decoded frame of image to get side information, then utilizing this side information to go on decoding and rebuilding for present frame [16]. As DVC eliminates frame prediction of encoder, but puts it in decoder to carry out, so complexity of encoder could be very low, encoder design is very simple, decoder design is very complex, and this kind of design is contrary with traditional encoder mode [17].

Traditional Wyner-Ziv DVC is on the basis of LDPC, adopting uniform encoder mode for all the areas of Wyner-Ziv frames, decoder can't decode violent motor area accurately and send more solicited messages to feedback channel, it lowers encoding frequency [18]. Against this problem, firstly introducing the DVC theory in this article, and then on this basis, proposing an improved Wyner-Ziv DVC algorithm which is based on DVC algorithm of DCT-R: The first step is that encoder gets violent motor area and non-violent motor area through SAD judgment rule. The second step is to adopt interpolation methods based on DCT-R motion for violent motor area, extracting DCT coefficient of low frequency as DCT-R to assist decoder to decode, and carrying out entropy coding compression for DCT coefficient of low frequency, decoder utilizes decoded DCT coefficient of low frequency to carry on inter-frame bidirectional motion estimation. The third step, for non-violent motor area, do not extract DCT coefficient of low frequency, do not carry on compressing and sending of DCT coefficient of low frequency, adopting traditional motion interpolation algorithm to generate side information and achieving video transmission optimization at last.

The main contribution of this paper is that:

We simply introduced the DVC application background and presented research status.

From the ROI distinguishing algorithm of encoder and generating algorithm of DCT side information of decoder, we proposed new DCT-R distributed video coding program and utilized this algorithm to solve necessary improved problem for typical program.

Finally, the experiment test, performance test and comparison for proposed algorithms which are DVC, Wyner-Ziv video coding algorithm, H.263 intra-frame and inter-frame coding algorithm, traditional JPEG coding algorithm are carried out. The result of simulation experiment shows that proposed multi-view effectiveness can solve the problem preferably, which is decoder end can't accurately decode larger movement area in images,

improve video encoding efficiency and decoded image quality.

II. THEORY OF MULTI-PERSPECTIVE DVC

Theory of distributed information source coding developed from Slepian-Wolf theory of lossless compressed distributed coding, and Wyner-Ziv theory of loss distributed coding. Next, it was introduced in detail that Slepian-Wolf theory, Wyner-Ziv theory and typical distributed video coding program.

A. Distributed Video Coding

For distributed coding with two or more information source, supposing two individual, same distributed, infinite long random series, Slepian-Wolf gives its code rate limit, figure 1 shows limit diagrammatic sketch. In traditional entropy coding, the code rate could reach $R_X \geq H(X)$, $R_Y \geq H(Y)$, Such as area 1 in figure 1, X and Y are mutually independent encoding two signals, their entropy are $H(X)$ and $H(Y)$ respectively, corresponding code rate are R_X and R_Y, joint entropy of two signals shows by $H(X,Y)$, then relationship between total code rate and joint entropy $H(X,Y)$ is as formula 1:

$$R = R_X + R_Y \geq H(X) \tag{1}$$

For distributed coding of two information source, referential information only can use in decoder, Slepian-Wolf proves ultimate limit of distributed information source coding code rate, diagrammatic sketch is showed in figure 1. Area 2 in figure 1 gives compressed limit of Slepian-Wolf theory, utilizing interdependency between X and Y, encoding individually for two signals X and Y, its total code rate $R = R_X + R_Y$, also can reach the joint entropy $H(X,Y)$ of two signals, as below formula indicates:

$$\begin{aligned} R_X &\geq H(X \mid Y) \\ R_Y &\geq H(Y \mid X) \\ R_X + R_Y &\geq H(X) \end{aligned} \tag{2}$$

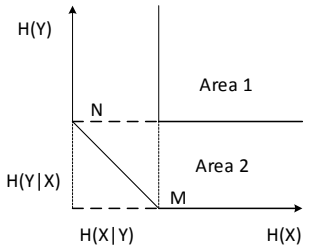


Figure 1. The source rate limit of Slepian-Wolf theory

In figure 1, point M shows code rate of information source X is $R_X = H(X)$, the code rate of information source Y could be $R_Y \geq H(Y \mid X)$ while compressing, it can decompress Y lossless; point N is in like manner, in information source Y, code rate is $R_Y = H(Y)$, when doing compression for information source X, only needs that its code rate is $R_X = H(X \mid Y)$, it's able to decompress X lossless.

On the basis of lossless coding theory of Slepian-Wolf, Wyner and Ziv proposed lossy distributed coding theory which utilizing side information to assist to decode. For side information which only can be used by decoder, Wyner and Ziv gave the code rate limit, established distortion theory of distributed lossy information source coding. Figure 2 is the block diagram of lossy encoding and decoding which side information Y done for information source X, X and Y represents two independent, same distributed random series which are related with statistics, information source X does not refer to side information Y while compressing, decoder needs to check side information Y for reference to decode for information source X. X' represents redevelopment value after A has been decoded, $R_{X|Y}^{WZ}(D)$ is distortion-rate function of Wyner-Ziv, which is distortion D has been given, the lower limit of bit rate distributed coding can get, in addition, encoder and decoder both can use distortion-rate function $R_{X|Y}(D)$ of side information, distortion D defines as $D = E[d(X, X')]$ in distortion-rate function. In the condition of $D > 0$, Wyner and Ziv proves that $R_{X|Y}^{WZ}(D) > R_{X|Y}(D)$.

When mean square error under the condition of distortion, relationship between information source X and Y could be showed in $Y = N + X$, N and Y both satisfy Gaussian distribution in independent case, in like manner, in the condition of $D > 0$, Wyner and Ziv prove it can get $R_{X|Y}^{WZ}(D) = R_{X|Y}(D)$. Distributed source encoding, distortion-rate function rate which decoder gets by using side information is same with distortion-rate function rate which encoder gets by using side information, which is distortion performance of these two schemes, are same.

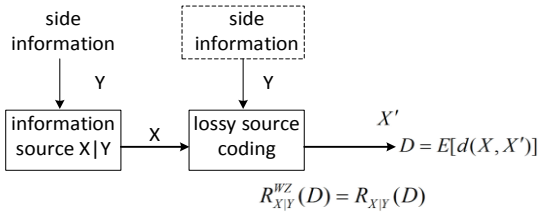


Figure 2. Lossy side decoding of message source X with relevant side information Y

B. Program of Layered Wyner-Ziv Video Coding

Scheme of layered Wyner-Ziv video coding is also a typical distributed video coding program, mainly to solve the problem of video information transmission while bandwidth is limited. Gradable encoding program makes video code rate possess the characteristics of adjustment, encoder only needs to encode video data once, it will be able to let decoder do decode for video stream in the light of video quality, spatial resolution or frame rate requirement, to satisfy different speed rate requirement in video communication. The basic layer is mainly to extract

video basic information in the program of gradable video coding, reinforcement is to extract video details and higher resolution ratio. Channel only need to make sure video data in basic layer transmits reliably and transmits moderate reinforcement data; it is able to solve the case which network switches from high speed to low speed suddenly to adapt all kinds of network transmission.

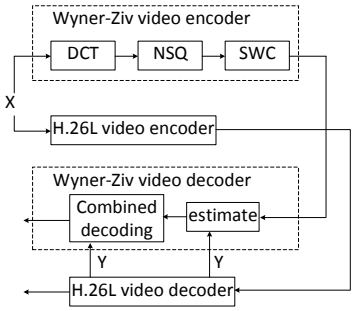


Figure 3. The structure diagram of layered Wyner-Ziv video coding

As figure 3 shows, layered Wyner-Ziv video coding which proposed by Zixiang Xiong. In this program, basic layer extracts video basic information to carry on traditional intra-frame encoded mode; enhanced layer extracts video details and higher resolution ratio to go on video encoding and decoding. The primary encoding process of Wyner-Ziv encoder is that carrying out DCT converting, nested scalar quantizing and SWC code based on LDPC. From the result of literature experiment, result of layered Wyner-Ziv video coding is basically same with MPEG-4/H.26L FGS. The important difference is encoder adopts individual encoded mode to decode for basic layer and enhanced layer, decoder utilizes basic layer to assist reinforcement to carry on joint encoding to get better code efficiency. Therefore, it has better error-resilience performance than MPEG-4/H.26L layered coding, to solve encoding and decoding error problems by error transmission.

As in the practical application, even though reinforcement data meets a loss during transmission, the video image which decoder utilized basic data to recover still can reach watching quality. Thus, layered distributed video coding is able to apply to different Qos network which supports differential service.

III. SCHEME OFTHE MULTI-VIEW DISTRIBUTED VIDEO CODING

Encoder of Multi-view Distributed Video Coding (MDVC) does not utilize complex encoder to encode, does not carry out data communication among visual angle, only utilizing perspective to carry on communication to get high quality decoding video in decoder, making it excel traditional multi-view coding. As MDVC monitoring system adopts several cameras to observe one and the same scene in many respects, so video data volume which needs to handle is really great, but the images which several cameras get at the same time should exist certain relationship in space. In the special MDVC area, side information not only gets from one and the same visual angle, but also can get from adjacent angles through spatial interdependency.

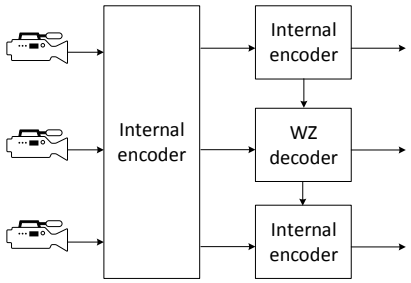


Figure 4. Block diagram of distributed multi view video coding

Figure 4 is a multi-view DVC block diagram, it consists of 3 cameras in total, so there exists three visual angles, the middle visual angle is primary one, WZ camera adopts independent coding, it needs to use self-view information and assistant view information to carry on joint decoding while decoding, which is MDVC mode; the other two adjacent visual views are assistant one which adopts H.264/ACV intra-frame encoding and decoding, decoder utilizes view convert to transform it into spatial side information for WZ decoder reference. From figure 1, it can be seen that side information in multi-view DVC program mix together from two part; One is inside visual angle, it adopts Motion Compensated Temporal Interpolation(MCTI); the other is inter-view, adopting Homography Compensated Inter- view Interpolation (HCII). Fused side information is getting by mixing together inter-view side information and inside one, in this case, the video quality which has been decodes is higher. Figure 5 is the decoding structure of MDVC, Cam2 is primary view, Cam1 and Canm3 are assistant one, in Cam2, frame WZ and I were got respectively by Slepian-Wolf decoder of DVC, H.264/AVC intra-frame encoding and decoding, frame C was got from H.264/AVC intra-frame encoding and decoding in Cam1 and Cam3.

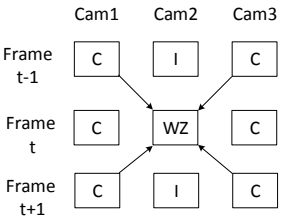


Figure 5. Decoding structure of distributed multi view video coding

IV. ALGORITHM DESIGN OF PROPOSED DCT-R DISTRIBUTED VIDEO CODING

Traditional multi-view DVC is on the basis of turbo or LDPC, encodes for all Wyner-Ziv frame areas indiscriminately, this kind of encoding handling, for motor area which is relative violent, decoder can't decode violent motor area accurately and send more solicited messages to feedback channel, it makes coding frequency lower, meanwhile, decoding for violent motor area is also not accurate, it causes part of areas distort in image. With regard to this problem, in this article, it proposed an improved Wyner-Ziv distributed coding method, getting violent and non-violent motor area by judgment rules of ROI. For violent motor area, extracting low frequency

coefficient of DCT as DCT-R algorithm to assist decoder to decode, decoder utilizes DCT low frequency coefficient which has been already decoded to evaluate bidirectional movement, adopting DCT motion interpolation to generate the best side information. But for non-violent motor area, do not extract DCT low frequency coefficient, do not compress and transmit DCT low frequency coefficient, adopts traditional motion interpolation method to generate side information and come true transmission optimization at last.

For some violent motor area in frame WZ of primary view, traditional multi-view DVC can't generate time side information accurately, and integrating time side information with spatial side information completely to generate fused side information to decode Wyner-Ziv frame in primary view. Regarding this problem, it was given a brand-new solution which is based on multi-view distributed coding algorithm of Macro block difference, in primary view, through judgment rules of ROI to get violent and non-violent motor area. For violent motor area, side information was generated by integrating time and spatial side information. The generated algorithm of time side information adopts DCT motor evaluation method; but for non-violent motor area, adopts traditional motor interpolate algorithm to generate time and spatial side information. DCT low frequency coefficient is used to assist decoder to decode. As DCT low frequency efficient includes important visual information of image, so it was chosen as assistant information to assist decoder to decode.

A. Judgment Rules

In the previous motor evaluation method based on hash code, encoder extracts all has bits of macro block of frame W, transmits it to decoder to assist side information to generate. However, in a same group of video sequence, only little part of image macro blocks to move greater, as motion vector of many image macro blocks is zero or smaller. For image macro block which is zero or smaller, it can get high quality side information by adopting motion compensated interpolation. Thus, in this chapter, adopted 8×8 macro block judgment based on SSH criterion at the encoder end, if it's greater than specific value, then doing ROI sign which is ROI macro block, or do not do any sign for target macro block. X_w represents present frame, X_p represents last key frame, so SSH criterion of this article is as below formula:

$$SSH = \sum_{(x,y)\in B_i} | X_w(x,y) - X_p(x,y) | \qquad (3)$$

In above formula, B_i represents a certain 8×8 macro block, x and y are pixel abscissa and ordinate respectively. Utilizing SSH value is greater or equal to a certain specific value which has been calculated, regarding that image macro block does greater movement, then doing ROI mark for targeted macro block. In this program, encoder end extracts DCT low frequency coefficient from ROI sign as DCTH to assist decoder end to decode, it makes encoder decode image macro block which has big movement more accurately.

B. Encoding and Decoding Framework of Distributed Video

Figure 6 is the whole process framework of DCT-R on the basis of DCT and ROI distributed video coding algorithm, in this article, it used violent motor macro block as ROI macro block, and exacted DCT low frequency coefficient of ROI macro block to assist decoder end to carry on motion evaluation to improve decoder end efficiency and quality of decoding image. Encoder end uses criterion of macro block SAD to judge 8x8 macro block for Wyner-Ziv frame, if greater or equal to certain specific value, then carrying out macro block ROI sign, extracting DCT low frequency coefficient as DCT-R algorithm to assist decoder end to decode, and doing compression of entropy coding for DCT low frequency coefficient; or carrying out LDPC encoding for unmarked macro block directly. Encoder end adopts traditional intra-frame coding for key frame to assist decoder end to decode Wyner-Ziv frame.

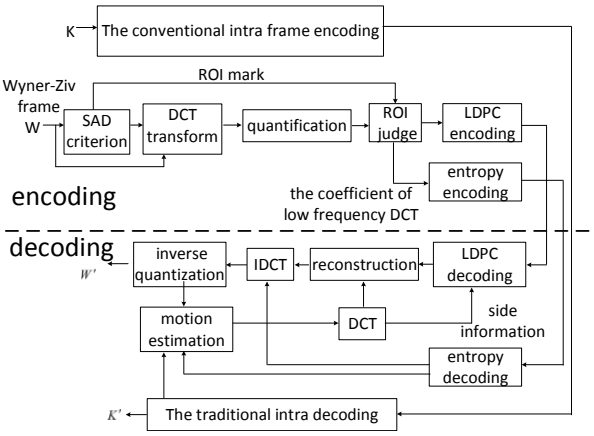


Figure 6. The process diagram of distributed video coding algorithm based on ROI and DCT-R DCT

If decoder end receives DCT low frequency coefficient, then will do bidirectional movement evaluation on the basis of DCT for targeted macro block; if decoder end doesn't receive DCT low frequency coefficient, then will adopt motor compensated interpolation method to generate side information for targeted macro block. Utilizing algorithm which proposed in this article to get the best side information to decode frame W, and then rebuilding, IDCT and inverse quantization to get decoding frame W'.

C. Side Information Generation

Motion vector strand or fall is vital for side information quality which decoder gets, reliability of side information which adopts weal motion vector interpolation to get decrease, it influences decoding image quality straightly. The generated method of side information which decoder end adopts can evaluate more real motion vector, the more accurate which side information gets, the more real which image has been decoded. In this article, to carry on interpolation

algorithm of bidirectional movement estimation based on DCT for ROI macro block, but for macro block which are not ROI, to carry on motion compensated interpolation algorithm to get the best side information. Supposing DCT inverse quantization of ROI is $H_{w[a,b]}$, $Z_{B[a,b]}$ represents DCT of all possible referential blocks, $[a,b]$ is the macro block position in frame, then matching criterion in DCT area is as below formula:

$$\arg\min\left\{\sum_{(x,y)\in S}|H_{w[a,b]}(x,y)-H_{B[a,b]}(x,y)|\right\} \quad (4)$$

In formula (4), x, y respectively represents abscissa and ordinate of DCT area, S represents DCT low frequency coefficient which macro block extracted. Utilizing criterion of formula 4 is able to find the best matching block for ROI, and then calculating its motion vector.

To carry on interpolation algorithm of bidirectional movement estimation for ROI macro block, to calculate B_R', which ROI macro block B_R side information.

First of all, decoder end utilizes DCT coefficient of B_R of received ROI macro block, utilizes formula (4) to calculate the best matching block B_R of previous key frame K_P and the best matching block B_P of latter key frame K_F.

And then inferring motion vector field $kv_1(x,y)$ which B_R relatives to previous matching block B_P, and motion vector field $kv_2(x,y)$ which B_R relative to previous matching block B_F.

The formula derivation for referential side information B_R' of B_R is as below:

$$\begin{aligned}B_R' = &\frac{|kv_2(x,y)|}{|kv_1(x,y)|+|kv_2(x,y)|}\\ &\cdot B_P(x+kvx_1(x,y), y+kvy_1(x,y))\\ &+\frac{|kv_1(x,y)|}{|kv_1(x,y)|+|kv_2(x,y)|}\\ &\cdot B_F(x+kvx_2(x,y), y+kvy_2(x,y))\end{aligned} \quad (5)$$

In formula (5), $kvx(x,y)$ and $kvy(x,y)$ respectively represents component product of motion vector field $kv(x,y)$ which ROI macro block relative to best matching block in x and y direction, finally, let referential side information of each macro block form into side information of a frame image to assist decoding frame W more accurately.

D. Side Information of Spatial Correlation

Homographic matrix using in formula (6), it is a 3×3 matrix, related views of the two stages is able to use coordinate system of this matrix to convert. Supposing selecting two nodes of V_1 and V_2, both of them monitor same area and adopt transparent motion model, formula (7) is mapping transformation which pixel point of frame V_1 towards pixel point of frame V_2, the 8 movement parameters in formula were got by gradient descent method.

$$\beta\begin{bmatrix}x_2\\1\\y_2\end{bmatrix}=\begin{bmatrix}a_0 & a_1 & a_2\\ a_0 & a_0 & a_0\\ a_0 & a_0 & a_0\end{bmatrix} \quad (6)$$

$$\begin{aligned}x_2 &= (a_4+a_0x_1+a_2\,y_1)/(a_6x_1+a_7\,y_1+1)\\ y_2 &= (a_5+a_3x_1+a_1\,y_1)/(a_6x_1+a_7\,y_1+1)\end{aligned} \quad (7)$$

In formula (6), $a_0, a_1, ..., a_7$ is the 8 movement parameters, β is ratio parameter.

(a) Camera decoding in the fourteen t frame(PSNR=16.53dB)

(b) Lena decoding in the fourteen frame(PSNR=16.42dB)

Figure 7. Lena and camera image which are decoded by the proposed algorithm

V. EXPERIMENT TESTS RESULT

In simulation experiment of this article, handling high frequency information of smooth motor area and overstated motor area of Wyner-Ziv frame, code rate of LDPC encoder uses 7/8 and adopts PEG method to generate its check code. After testing, analyzing and comparing for different compression rate so many times, it's ideal to take 64 as threshold value of ROI judgment rule, choose DC+8AC as DCT-R in DCT low frequency coefficient. The adopted GOP value is 2 during process of simulation experiment, choosing key frame as odd frame and utilizing JPEG standard to handle, Wyner-Ziv frame is odd frame, which is K-W-K-W sequence. In simulation experiment, it was proposed estimation algorithm of DCTH bidirectional movement, transform-domain Wyner-Ziv video coding algorithm, traditional JPEG coding algorithm, H.263 intra-frame coding, H.263 inter-frame coding, and made a

comparison of above five algorithms. H.263+ encoder end adopted TMN8, JPEG standard adopted Annex K [51]. Adopting camera and lena two standard sequences, image format is QCIF (176×144), encoding frame number is 100 frames, 30 frames each 30 seconds.

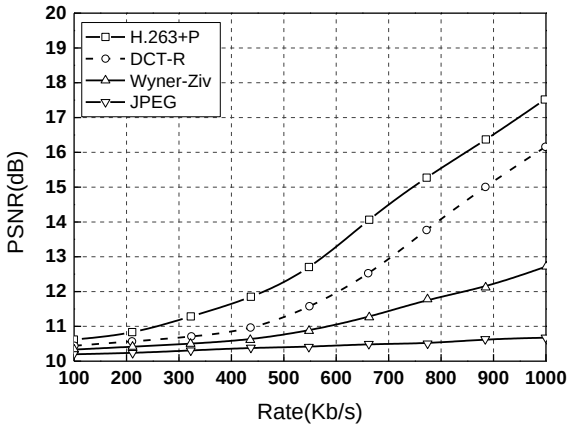
(a) Camera decoding in the fourteen t frame(PSNR=16.53dB)

(b) Lena decoding in the fourteen frame(PSNR=16.42dB)

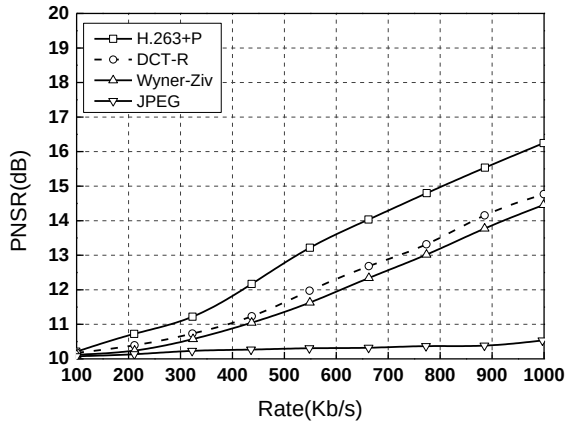
Figure 8. Lena and camera image which are decoded by the traditional DVC algorithm

Figure 7 and 8 are decoded images in simulation experiment of this chapter, adopting this chapter's algorithm and traditional DCV coding algorithm to carry on encoding and decoding for vide sequence, and make comparisons for image quality of video sequence after decoding. PSNR value of frame 21 in figure 7(a) foreman sequence is greater than PSNR value of frame 21 in figure 8(a) foreman; PSNR value of frame 21 in figure 7(b) salesman sequence is greater than PSNR value of frame 21 in figure 8(b) salesmen. According to experiment test and performance comparison, algorithm in this chapter is obviously excel than traditional DVC coding in same code rate, solving encoding problem of larger motion area in images preferably, the algorithm in this chapter not only decreases encoding efficiency of encoder end, but also increases decoded image quality.

Figure 9 is the performance comparison for above several coding, from there it was getting that encoding efficiency of Wyner-Ziv video algorithm is 2dB higher averagely than H.623+ intra-frame method, and the complexity is lower than H.263+ intra-frame, thus it can be seen how obviously the advantage is. However, PSNR value of H.263+ inter-frame is excelled than Wyner-Ziv, it is because that it utilizes a large amount of calculation of inter-frame prediction method while encoding. PSNR value which got by adopting algorithm proposed in this chapter is averagely 1.2dB higher than Wyner-Ziv encoding and decoding system, as the algorithm which proposed in this chapter can carry on more accurate motion evaluation for relative larger motion area and get the best side information.



(a) The Rate VS PSNR of image camera



(b) The Rate VS PSNR of image lena

Figure 9. The simulation results of video coding based on DCT hash of Wyner-Ziv

VI. CONCLUSION

In this chapter, first of all, simply introducing DVC application background and present research status, second, from ROI distinguishing algorithm of encoder end and generating algorithm of DCT side information of decoder end, it proposed new DCT-R distributed video coding program and utilized this algorithm to solve necessary improved problem for typical program, finally, carrying out experiment test, performance test and comparison for proposed algorithms which are DVC, Wyner-Ziv video coding algorithm, H.263 intra-frame and inter-frame coding algorithm, traditional JPEG coding algorithm. The result of simulation experiment shows that proposed multi-view effectiveness can solve the problem preferably, which is decoder end can't accurately decode larger movement area in images, improve video encoding efficiency and decoded image quality.

REFERENCES

[1] Akyildiz I F, Su W, Sankarasubramaniam Y. "A survey on sensor networks," *IEEE Communications Magazine*, vol. 8, no. 40, pp. 102-114, 2002.

[2] DeBardelaben J A. "Multimedia sensor networks for ISR applications," *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers. Pacific Grove, California: IEEE Press*, pp. 2009-2012, 2003.
[3] N. -M. Cheung, H. Wang and A. Ortega, "Sampling-based correlation estimation for distributed source coding under rate and complexity constraints," *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2122-2137, Nov. 2008.
[4] T. Maugey and B. Pesquet-Popescu, "Side information estimation and new symmetric schemes for multi-view distributed video coding," *Journal Of Visual Communication And Image Representation*, vol. 19, no. 8, pp. 589-599, Dec. 2008.
[5] M. Rezaei, I. Bouazizi and M. Gabbouj, "Joint video coding and statistical multiplexing for broadcasting over dvb-h channels," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1455-1464, Dec. 2008.
[6] H. Cui, X. Su and W. Shang, "On optimal media/video distribution in closed p2p-based iptv networks," Computer Networks, vol. 60, no. , pp. 217-232, Feb 26. 2014.
[7] K. -H. Choi, J. -Y. Lee, B. -W. Jeon and J. -K. Han, "Block boundary matching algorithm for generating side information in distributed video codec," *Optical Engineering*, vol. 52, no. 10, Oct. 2013.
[8] Y. Geng, J. Chen, K. Pahlavan, Motion detection using RF signals for the first responder in emergency operations: A PHASER project, 2013 IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), London, Britain Sep. 2013
[9] R. Li, Z. Gan, Z. Cui, M. Wu and X. Zhu, "Distributed adaptive compressed video sensing using smoothed projected landweber reconstruction," *China Communications*, vol. 10, no. 11, pp. 58-69, Nov. 2013.
[10] H. Liu, B. Song, H. Qin and Z. Qiu, "Dictionary learning based reconstruction for distributed compressed video sensing," *Journal Of Visual Communication And Image Representation*, vol. 24, no. 8, pp. 1232-1242, Nov. 2013.
[11] Y. Geng, J. He, K. Pahlavan, Modeling the Effect of Human Body on TOA Based Indoor Human Tracking, International Journal of Wireless Information Networks 20(4), 306-31
[12] T. Maugey, J. Gauthier, M. Cagnazzo and B. Pesquet-Popescu, "Evaluation of side information effectiveness in distributed video coding," IEEE Transactions on Circuits And Systems for Video Technology, vol. 23, no. 12, pp. 2116-2126, Dec. 2013.
[13] M. Narroschke, "Coding efficiency of the dct and dst in hybrid video coding," *Ieee Journal Of Selected Topics In Signal Processing*, vol. 7, no. 6, pp. 1062-1071, Dec. 2013.
[14] F. Verbist, N. Deligiannis, M. Jacobs, J. Barbarien, P. Schelkens, A. Munteanu and J. Cornelis, "Probabilistic motion-compensated prediction in distributed video coding," *Multimedia Tools And Applications*, vol. 66, no. 3, pp. 405-430, Oct. 2013.
[15] X. Du, V. Nguyen-Son, W. Cheng, T. Q. Duong and L. Shu, "Joint replication density and rate allocation optimization for vod systems over wireless mesh networks," *IEEE Transactions on Circuits And Systems for Video Technology*, vol. 23, no. 7, pp. 1260-1273, Jul. 2013.
[16] Z. Liu, A. Wang, B. Zeng, X. Zhang, H. Bai and Z. Li, "Distributed compressive video sensing with adaptive measurements based on structural similarity," *Chinese Journal Of Electronics*, vol. 22, no. 3, pp. 594-598, Jul. 2013.
[17] S. Milani and G. Calvagno, "Distributed video coding based on lossy syndromes generated in hybrid pixel/transform domain," *Signal Processing-Image Communication*, vol. 28, no. 6, pp. 553-568, Jul. 2013.
[18] L. Qing, X. He, X. Ou and R. Lv, "Distributed video coding based on multiple-source correlation model," *Applied Mathematics & Information Sciences*, vol. 7, no. 4, pp. 1609-1614, Jul. 2013.

Cache Performance Optimization for SoC Vedio Applications

Lei Li [1], Wei Zhang [1], HuiYao An [2], Xing Zhang [1], and HuaiQi Zhu [2]
1. School of Electronics Engineering and Computer Science, Peking University, Beijing, China
2. School of Software and Microelectronics, Peking University, Beijing, China
Email: mail_lilei@pku.edu.cn

***Abstract*—Chip Multiprocessors (CMPs) are adopted by industry to deal with the speed limit of the single-processor. But memory access has become the bottleneck of the performance, especially in multimedia applications. In this paper, a set of management policies is proposed to improve the cache performance for a SoC platform of video application. By analyzing the behavior of Vedio Engine, the memory-friendly writeback and efficient prefetch policies are adopted. The experiment platform is simulated by System C with ARM Cotex-A9 processor model. Experimental study shows that the performance can be improved by the proposed mechanism in contrast to the general cache without Last Level Cache (LLC): up to 18.87% Hit Rate increased, 10.62% MM Latency and 46.43% CPU Read Latency decreased for VENC/16way/64bytes; up to 52.1% Hit Rate increased, 11.43% MM Latency and 47.48% CPU Read Latency decreased for VDEC/16way/64bytes, but with only 8.62% and 4.23% Bandwidth increased respectively.**

***Index Terms*—CMP (Chip Multiprocessors); Memory Access; LLC (Last Level Cache); Writeback; Prefetch**

I. INTRODUCTION

In 2005, the computer industry adopted processor clock frequency as the primary indicator of processor performance. The change reflected the looming problem so that the frequency abounded by power dissipation, wire delay and the physical properties of CMOS transistors impeded the rate of the increase in clock speeds. Chip Multiprocessors (CMPs) are adopted by industry to deal with the speed limit of the single-processor. However, the "memory wall" problem is still an overwhelming bottleneck in current system performance [1] and over 50% clock time on processor is spent for memory access [2]. Cache optimization mechanism is the only efficient way to alleviate the latency. Considering the speed, the capacity and the increasing memory bandwidth requirement of the increased cores per chip, optimization proposals mainly focus on the last-level cache (LLC) [3, 4].

As the number of the integrated cores in Chip-Multiprocessor (CMP) designs continues to increase, the typically shared last-level cache memory gradually becomes a critical bottleneck for the performance. Ranging from the early Least Recently Used (LRU) replacement policy to the recent optimization policies [5-8], most of the work mainly focused on two points: when or how to replace (or write back) a block, and when or how to load (or prefetch) a new one.

Although statistical methodology has been proposed by researchers [9], the high level criterion to evaluate the performance is how exquisitely the replacement policy depicts the program locality. Work mentioned above is universal optimization policy to achieve great improvement of the performance. In practice, no policy is sufficient to satisfy each of the workload classifications. Furthermore, multimedia application is the most widely used and the roughest job for processors, especially for handheld devices. A new method for optimizing these applications was explored in this work on video application.

Our goal in this paper is to design a set of management policies to improve the performance of the cache for video application. The complete management policies include Memory-friendly writeback, multimedia-intensive prefetch and scheduler of last-level cache. Firstly, in modern memory systems, memory-write requests can cause significant performance loss by increasing the memory access latency for subsequent reading requests targeting the same device [7]. Memory-friendly writeback aims to use idle time to execute useful writeback command. Secondly, multimedia-intensive prefetch aims to execute useful prefetch by analyzing the behavior of video application. Finally, scheduler of last-level cache aims to manage all the commands (prefetch, writeback, critical reading and critical writing).

We evaluate our Management policies with System C on ARM Cotex-A9 processor model. Experimental study shows that the performance can be improved by the proposed mechanism in contrast to the general cache without Last Level Cache (LLC): up to 18.87% Hit Rate increased, 10.62% MM Latency and 46.43% CPU Read Latency decreased for VENC/16way/64bytes; up to 52.1% Hit Rate increased, 11.43% MM Latency and 47.48% CPU Read Latency decreased for VDEC/16way/64bytes, but with only 8.62% and 4.23% Bandwidth increased respectively.

II. RELATED WORK

To the best of our knowledge, little work about LLC in multimedia application has been done, especially for handheld devices. Restrictions on performance of the

doi:10.4304/jmm.9.7.926-933

processors for handheld devices often include, but not limited to [10] cost, power consumption, and functionality. Digital audio and video applications call for a larger amount of processes, compared with other widely used applications for handheld devices. The required process rate for compression ranges from 100 mega-operations per second (MOPS) to more than one tera-operations per second [11]. Considering the "memory wall" mentioned above, efforts which only focus on the increase of the speed performance for the processor's frequency, or highlighting multimedia dedicated processor can not solve all the problems. On the other hand, recent cache optimization proposals [5-8] don't work well enough for the multimedia application. Multimedia application has its own characteristics. Therefore the methods are discussed to improve cache performance on multimedia application. We discuss closely related work in prefetch and writeback.

A. Prefetch

There have been a handful of proposals for cache prefetch algorithms in the literature for the past few years. These proposed algorithms can be classified into several classes as follows.

A lot of previous DRAM scheduling policies were proposed to improve DRAM throughput in single-threaded [12, 16, 17], multithreaded [18, 19, 17], and stream-based [20, 21] systems. In addition, some recent works [22, 23, 24] have provided some methods for fair DRAM scheduling across different applications sharing the DRAM system. Some of these previous proposals [16, 25, 19, 22, 23, 24, and 17] don't mention how to treat prefetch requests and demand requests. Our management policies are based on these scheduling policies: they can be extended to adaptively prioritize demand and prefetch requests and to give useless prefetch requests up.

Other DRAM proposals use two different approaches to handle prefetch requests. Some proposals [12, 26, 27, and 28] prioritize demand requests over prefetch requests. Other proposals [21] treat prefetch requests the same as demand requests. Thus, these previous DRAM controller proposals handle prefetch requests strictly. Strict handling of prefetches can cause significant performance loss compared with adaptive prefetch handling. Our work improves upon these proposals by incorporating the effectiveness of prefetch into DRAM scheduling decisions.

Some previous works proposed execute the prefech command based on dynamic information. Our work is complementary to these proposals, which are described below.

Lee et al. [14] proposed a new low-cost memory controller, called Prefetch-Aware DRAM Controller (PADC), which aimed to maximize the benefit of the useful prefetches and minimize the harm caused by the useless prefetches. To accomplish the goal, PADC estimates the usefulness of prefetch requests and dynamically adapts to its scheduling and buffer management policies based on the estimates. In contrast, our mechanism adapts to the selected policy between demands and prefetches based on stream prediction, because our research is targeting video application. As a result, Lee's proposal can be combined with our prefetch scheduling policy to provide higher performance for multimedia application.

B. Writeback

Much previous work [30, 23, 22] does not take the writing interference problems into account. Eager writeback [34] is the first work which expands write resources by using the LLC to reduce write-induced interference. Eager writeback writes back dirty cache blocks in the least-recently-used (LRU) position of the LLC sets whenever the bus is idle instead of waiting for the block to be evicted to reduce the memory traffic. However, the scheduling window of eager writeback is still limited to the size of the writing buffer. Thus, the scheduling decision it makes is still far from optimal.

Stuecheli et al. [33] proposed a virtual writing queue (VWQ) technique. Their technique takes a fraction of the LRU positions in the LLC as the VWQ. Dirty cache blocks with high locality in the VWQ are written back in a batch, therefore improving writeback efficiency. The drawback of this technique is that it needs to search the dirty cache blocks in VWQ which hit in the same row when mapping to the DRAM. Although it uses the Cache Cleaner technique to help searching, it still consumes significant LLC power and searching time.

Wang et al. [7] proposed a rank idle time prediction driving LLC writeback technique. In contrast to previous work [33, 35] which does not exploit rank idle time, the technique allows the memory to service writing requests during the significant idle rank time. The technique can be used with LLC writeback scheduling techniques to improve the memory efficiency.

Based on the previous work, we propose the new writeback mechanism. When the memory is idle, the writeback command will be executed according to the data writeback command and Bank open/close information maintained, therefore the best time and schedule of writeback will be chosen.

III. MECHANISM

A. DRAM Access Background [14]

An SDRAM system consists of multiple banks which can be accessed independently. Each DRAM bank comprises rows and columns of DRAM cells. A row contains a fixed-size block of data (usually several Kbytes). Each bank has a row buffer (or sense amplifier), which caches the most recently accessed row in the DRAM bank. A DRAM access can be done only by reading (writing) data from (to) the row buffer using a column address.

There are three commands which need to be sequentially issued to a DRAM bank in order to access data:

1) a precharging command to precharge the row bitlines
2) an activating command to open a row into the row buffer with the row address

3) a read/write command to access the row buffer with the column address.

After the completion of an access, the DRAM controller can either keep the row open in the row buffer (open-row policy) or close the row buffer with a precharging command (closed-row policy). The latency of a memory access to a bank varies depending on the state of the row buffer and the address of the request as follows:

1. Row-hit: The row address of the memory access is the same as the address of the opened row. Data can be read from/written to the row buffer by a reading/writing command, therefore the total latency is only the reading/writing command latency.
2. Row-conflict: The row address of the memory access is different from the address of the opened row. The memory access needs a precharging, an activating, and a reading/writing command sequentially. The total latency is the sum of all three command latencies.
3. Row-closed: There is no valid data in the row buffer (i.e. closed). The access needs an activating command and then a reading/writing command. The total latency is the sum of these two command latencies.

Recent microprocessors employ hardware prefetch to hide long DRAM access latencies. If prefetch requests are accurate and fetch data early enough, prefetching can improve performance. Existing DRAM scheduling policies take two different approaches as to how to treat prefetch requests with respect to demand requests. Some policies regard a prefetch request to have the same priority as a demand request. This can significantly delay demand requests and cause degradation for performance, especially if prefetch requests are not accurate. Other policies always prioritize demand requests over prefetch requests so that data known-to-be-needed by the program instructions can be serviced earlier. One might think that it provides the best performance by eliminating the interference of prefetch requests with demand requests. However, such a rigid policy does not consider the non-uniform access latency of the DRAM system (row-hits vs. row-conflicts). A row-hit prefetch request can be serviced much more quickly than a row-conflict demand request.

DRAM access time is shortest in the case of a row-hit [15]. Therefore, a memory controller can try to maximize DRAM data throughput by maximizing the Hit Rate in the row buffer. Previous work [12] introduced the commonly-employed FR-FCFS (First Ready-First Come First Serve) policy which prioritized requests such that it services 1) row-hit requests first and 2) all else being equal, older requests first. This policy was shown to provide the best average performance in systems which do not employ hardware prefetching [12, 13]. However, this policy is not aware of the interaction and interference between demand and prefetch requests in the DRAM system, and therefore treats demand and prefetch requests equally.

Then the management and scheduling policies of LLC are proposed. They are based on the features of memory access in multimedia application. The efficient writeback command and the inefficient writeback command are differentiated and scheduled. Also the writing and reading commands are unified scheduling to maximize the DRAM efficiency and minimize the writeback overhead.

B. Logic Diagram of Last Level Cache Managment

In the management and scheduling policies of LLC, in order to schedule the prefech and writeback command more reasonably, it is necessary for the cache to get DRAMC bank open/close information as Figure1 shows. The logic diagram of the whole LLC management unit is shown in Fig. 1.

In the logic diagram, Prefech Manager, Writeback Collector and Scheduler are designed.

Prefetch Manager: The Hit Rate of prefetch is very critical for the performance of the memory access. So the prefetch manager is designed according to the feature of the video engine. It makes prediction according to the current behavior of the memory access, so that it can improve the Hit Rate of prefetch.

Writeback Collector: The time for writeback is an important decision for performance, so a Writeback Collector is designed. The collector maintains data writeback command and Bank open/close information, so the best time and schedule of writeback will be chosen.

Scheduler: The cooperation of the reading and writing operation will reduce the useless memory access and decrease the cost. Scheduler is designed to incorporate the prefetch command into the writeback command, and utilize the cycle of the Bank operation to improve the efficiency of every cycle.

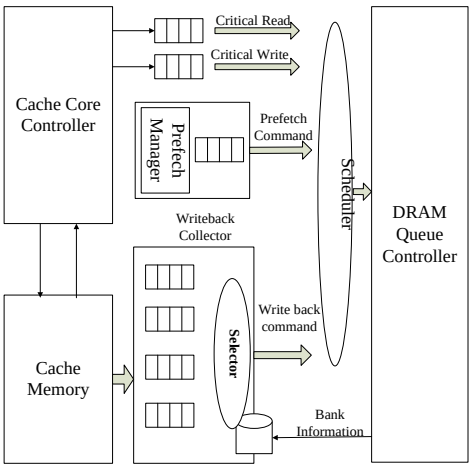


Figure 1. Logic diagram of last level cache managment

Behavior of the Video Engine on multimedia scenario has been analyzed when this replacement policy are designed. Two replacement methods are proposed in these multimedia applications.

C. Memory-Friendly Writeback

The sequence diagram of LLC WriteBack mechanism is shown in Fig. 2.

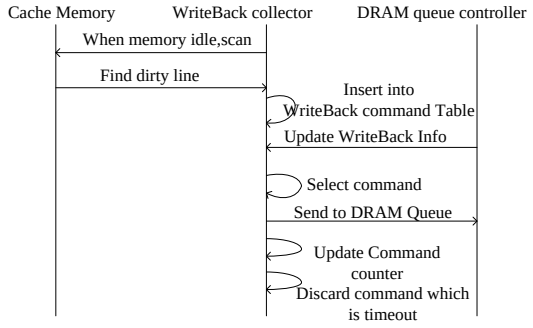


Figure 2. Sequence diagram of WriteBack mechanism

The LLC WriteBack data collector will scan the cache memory when the memory is idle. If dirty line is found, it will record this line in the table inside the collector with the index of DRAM's Row and Colum, and generate a writeback command. Besides, the collector will keep a counter for each pending writeback command. This counter will calculate the waiting time consumed in the buffer. When it hits a threshold once, this command will be discarded.

For each cycle, DRAM queue will update the bank open/close information for the LLC. Thus the collector could select the most appropriate writing command and send it to DRAM queue. The rule to select command is as follows:

1. Page hit command: According to the bank information provided by finding a page hits command. Because DRAM access time is the shortest in page hit case.
2. Page open command: If there is no page hit command, the collector will select a page open command.
3. The command which hits the waiting time threshold will be discarded. Because the selection will not try to interface the DRAM queue efficiency by issuing row-conflict command.

With the Memory-friendly writeback policy, the writeback operation will not bring extra cost of Bank open/close, and only will be executed when there is an appropriate open bank. In addition to avoiding the inefficiency writeback, during the period of waiting the scheduler, the same rewriting will be merged to reduce the actual times of writeback operation.

D. Prefetch Algorithm for Multimedia Intensive access

Prefetch modules issue prefetch command according to the prefetch prediction result. We use stream prediction style for prediction address. It means that the prediction address is the next 64 bytes or 128 bytes address.

For multimedia intensive scenario, the continuous address will appear with high possibility. However, because of cache resource limitation, it cannot hold all the prefetch data long time enough. It means that some prefetch data might not be used before it is written back. So only the address which will likely be accessed in near future should be prefetched.

There is a "prefetch manager" inside the prefetch module which records the cycle interval of the predicted command by each multimedia master. For each master, the manager will count the cycles between the continuous address for several times and record the maximum cycles. If the cycle value is smaller than a threshold, for the following reading command the prefetch module will issue prefetch command for the very master.

The procedure is as Fig. 3:

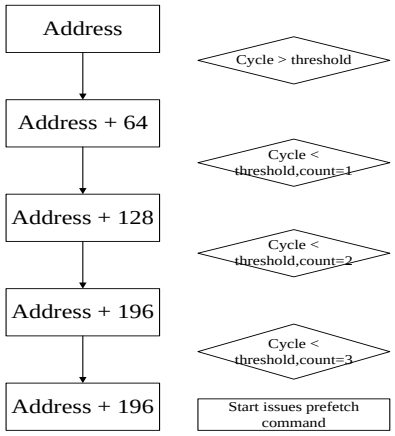


Figure 3. The prefetch algorithm for multimedia intensive

The matter in prefetch is that invalid prefetch will bring extra memory access. Some design mentioned above treats prefetch requests the same as demand requests, so invalid prefetch will influence the performance. To improve the performance and reduce the cost of prefeth, prefetch and writeback command are united scheduling to ensure the consecutiveness of reading/writing. The scheduler is described as follows.

E. Scheduler for LLC Manager

Scheduler of the LLC selects one command and sends it to DRAM queue manager. There are four kinds of commands under scheduler. Besides the writeback and prefetch command, there is critical writing and critical reading command. These two kinds of commands are issued by the cache core controller. Critical reading command is issued when a reading miss happened, and critical writing command is issued when writing miss happened.

The main purpose of the scheduler is to keep a balance between the latency and DRAM efficiency. The scheduler performs arbitration under following rules:

1. Continuous service of read or write until it hit the service count threshold.
2. Critical Read and Write have higher priority than prefetch read and writeback command.
3. Try to issue page-hit command for both read and write to maximize the DRAMC efficiency.

a) The writeback module has provided the most appropriate command for current DRAM page status.

b) Prefetch read command mostly is page-hit command to the original command.

With the selected policy of the scheduler, the time of page-hit is taken full advantage of and the writeback command will not be disturbed in continues reading command. Although Critical Reading/Writing command

which is the most influential for the performance will be executed continues, the page-hit writeback and prefech will be scheduled to maximum the efficiency.

IV. METHODOLOGY

A. System Model

To evaluate the improvement of CPU reading latency of LLC accurately, the ARM Cotex-A9 processor model provided by Carbon(TM) Design Systems was chosen. It was a highly accurate model with L1 Cache. We ran a segment of fake program which behaved like an image decoder. The designed bandwidth was 0.275GB/s. All modules of the testing platform were developed with systemC. The simulation frequency was 40 KHz, i.e., the processor's frequency was 40 KHz as seen from the real world. All modules of the platform except CPU worked at the frequency of 266MHz. The testing platform is shown in Fig. 4.

The System Cache in the figure was the one to be evaluated. The MT8320 External Memory Interface (EMI) connected to the cache was the DDR controller, whose total bandwidth was 8G and whose designed frequency was 266 MHz. The controller was a cycle-accurate systemC model. Each sub module exchanged data through systemC standard TLM interface. The interfaces were also cycle-accurate. After compilation, the testing platform could be used for simulation. The period needed for running one time was 27ms, which was also the period used for decoding one frame by the MMTG module. We used the standard Video Decoder Engine and the standard software running on CPU to decode the MPEG4 720P video file in the test. And we evaluated the improvement by observing indicators such as Bandwidth, Hit Rate, CPU Read Latency and MM Latency.

The test load was generated by Traffic Generator (TG in the figure). Traffic Generator issued the instructions of instruction-reading, instruction-writing and data-writing according to the configuration file. The time, address ranges and total bandwidth of these instructions were defined in the configuration file. The configurations and effects of the testing traffic generators are shown as follows.

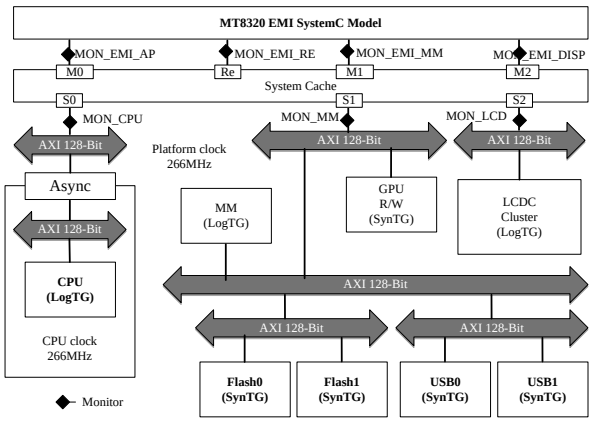


Figure 4. The architecture of simulation platform

B. Metrics

To measure the improvement of the performance for the whole system by LLC optimization, we choose Bandwidth, Hit Rate, MM Latency and CPU Read Latency as evaluating indicators.

1. Band Width: the system bandwidth describes the system overall spending.
2. Hit Rate: the Hit Rate of cache embodies DRAM's burden.
3. MM Latency: latency of multimedia encoding/decoding, key indicator for the video fluency.
4. CPU Read Latency: very important indicator, higher-performance CPU needs shorter total Data Path Latency; otherwise, CPU cannot get data and cannot reach the real performance.

We can evaluate the effect of the overall performance by incrementing bandwidth burden through the definition of Unit Bandwidth Hit Rate (Hit Rate/Bandwidth), Unit Bandwidth MM Latency (MM Latency/Band Width) and Unit Bandwidth CPU Read Latency (CPU Read Latency/Bandwidth).

V. EXPERIMENTAL EVALUATION

We analyze the experimental results on the platform in this section. To evaluate the results of LLC optimization accurately, we use test size of LLC 1MB, 2MB, 4MB and 8MB, respectively. As shown in Fig. 5(a) and (b), the test results of Bandwidth, Hit Rate, MM Latency and CPU Read Latency on platform with and without LLC optimization are respectively described. The experiment uses standard mpeg 720p video files. The light-gray part is with LLC optimization technique and the dark-gray part is without LLC optimization technique.

By comparing the two scenarios, we can conclude that the performance of cache has been improved greatly after optimization. As seen from the results, during encoding phase, bandwidth grows by an average of 8.26%, with 18.87% Hit Rate increased, 10.62% MM Latency and 46.43% CPU Read Latency decreased, respectively. During decoding phase, bandwidth grows by an average of 4.23%, with 52.1% Hit Rate increased, 11.43% MM Latency and 47.48% CPU Read Latency decreased, respectively.

There are two main reasons for the improvement:

1. Proposed prefech mechanism analyzes the way of memory access and the feature of memory address during video decoding phase, improves the Hit Rate and avoids too much invalid prefetch.
2. Memory-friendly writeback mechanism reduces the reading/writing operation switching frequency, improves the efficiency of cache loading and reduces the CPU Read Latency substantially.

The results shown in the figure can be analyzed respectively as follows. Firstly, overall bandwidth did not increase much and it means that the system overhead did not increase significantly. Secondly, Hit Rate increased substantially and it meant that the DRAM loading was reduced. Thirdly, CPU path delay decreased significantly. This was very meaningful because the high-performance

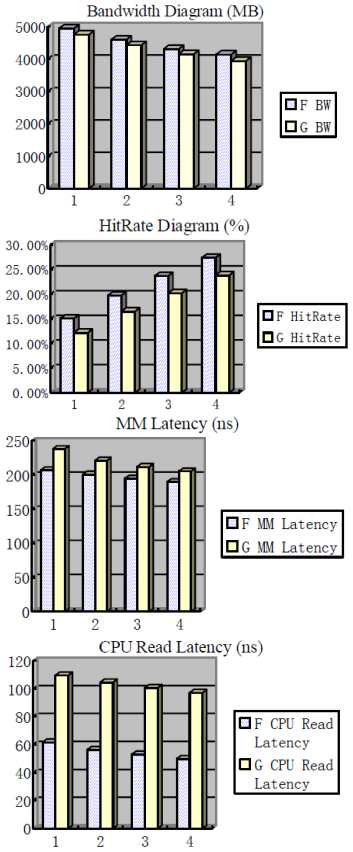


Figure 5. Experimental results (a) results for VENC/16way/64bytes(b)results for VDEC/16way/64bytes

CPU requested that the delay of entire Data Path must be low, otherwise without data the CPU would be idle, which claimed that several GHz are useless. Finally, Multimedia delay decreased and it meant that the movie would be played smoothly.

Fig. 6 describes performance enhancement per BW consumed when the size of LLC is 1MB, 2MB, 4MB and 8MB, respectively. As seen from the figure, bandwidth utilization is improved significantly by using the optimization technique.

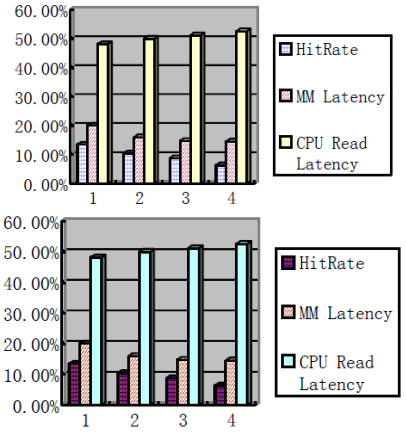


Figure 6. Performance Enhancement/BW(Encode) and Performance Enhancement/BW(Decode)

VI. CONCLUSIONS

In this paper, we have shown that existing Last Level Cache (LLC) optimization technique solves general memory access problems. Therefore this work mainly addresses video application on SoC, which is a common application in handheld devices.

According to the feature of the video application, we have proposed a set of LLC manage policies including memory-friendly writeback, prefetch algorithm for multimedia intensive access, and unified scheduling of the reading/writing command. Memory-friendly writeback mechanism chooses the most appreciating writeback command according to the Bank info of the current memory access, avoiding the inefficient row-conflict command. Prefetch algorithm improves the Hit Rate and avoids too much invalid prefetch. Unified scheduling of the reading/writing request insures the consecutiveness of the reading/writing request taking advantage of the time for the Page Hit, and improves the efficiency of cache loading.

These algorithms have been tested on the SoC simulation platform developed by systemC with ARM Cotex-A9 processor model. They are able to achieve significant Hit Rate improvement (18.87%), MM Latency reduction (10.62%) and CPU Read Latency reduction (46.43%) with only an average of 8.26% of bandwidth increased during encoding phase. Also, they achieve significant Hit Rate improvement (52.1%), MM Latency reduction (11.43%) and CPU Read Latency reduction (47.48%) with only an average of 4.23% of the bandwidth increased during decoding phase.

ACKNOWLEDGEMENT

This research is funded by the National Natural Science Foundation of China no. 61179029.

REFERENCES

[1] Win. A. Wulf, Sally A. McKee. "Hitting the memory wall: Implications of the obvious," *ACM Sigarch Computer Architecture News*, vol. 23, pp. 20-24, 1995.

[2] J. Chang "Cooperative caching for chipmultiprocessors," *University of Wisconsin-Madison*, 2007.

[3] Jaehyuk Huh. Burger, D. Keckler, S.W. "Exploring the design space of future cmps," *in Proceedings of the 2001 International Conference on Parallel Architectures and Compilation Techniques*, 2001.

[4] Chun Liu, Anand Sivasubramaniam, M Kandemir. "Organizing the last line of defense before hitting the memory wall for cmp," *HPCA*, pp. 176-185, 2004.

[5] P. Stenström, K. De Bosschere, J. Albericio, R. Gran, P. Ibáñez, V. Viñals, J. M. Llabería. "ABS: A Low-Cost Adaptive Controller for Prefetching in a Banked Shared Last-Level Cache," *ACM Transactions on Architecture and Code Optimization*, vol. 8, pp. 1-20, 2012.

[6] Miao Zhou, Yu Du, Bruce R. Childers, Rami G. Melhem, Daniel Mossé. "Writeback-aware partitioning and replacement for last-level caches in phase change main memory systems," *ACM Transactions on Architecture and Code Optimization*, vol. 8, pp. 1-21, 2012.

[7] Zhe Wang, Samira Khan, Daniel Jimenez, "Rank idle time prediction driven last-level cache writeback," *in MSPC'12, Beijing, China*, 2012, pp. 21-30.

[8] Jeffrey Stuecheli, Dimitris Kaseridis, David Daly, Hillery C. Hunter, Lizy K. John. "The Virtual Write Queue: Coordinating DRAM and Last-Level Cache Policies," *in ISCA'10, Saint-Malo, France*, 2010.

[9] Parijat Dube, Michael Tsao, Li Zhang, Alan Bivens. "Performance modeling and characterization of large last level caches," *IEEE*, pp. 118-119, 2012.

[10] Jim Rasmusson, Harald Gustafsson, Tord Nilsson. "Multimedia in mobile phones—The ongoing revolution," *Ericsson Review*, vol. 2, pp. 98-108, 2004.

[11] I. Kuroda, T. Nishitani "Multimedia Processors," Proceedings of the IEEE, vol. 86, pp. 1203-1309, 1998.

[12] S Rixner, W. J. Dally, U. J. Kapasi, P. Mattson, and J. D.Owens. "Memory access scheduling". *In ISCA-27*, 2000

[13] S. A. McKee, W. A. Wulf, J. H. Aylor, M. H. Salinas, R. H. Klenke, S. I. Hong,and D. A. B. Weikle. Dynamic access ordering for streamed computations. *IEEE Transactions on Computers*, 49 pp. 1255–1271, Nov. 2000.

[14] Joo Lee Chang, "Prefetch-Aware DRAM Controllers", *In proceeding of: 41st Annual IEEE/ACM International Symposium on Microarchitecture (MICRO-41 2008), November* 8-12, 2008, Lake Como, Italy .

[15] Micron. Datasheet: 2Gb DDR3 SDRAM, MT41J512M4 - 64 Meg x 4 x 8 banks.

[16] W. Zurav leff and T. Robinbson. Controller for a synchronous DRAM that maximizes throughput by allowing memory requests and commands to be issued out of order. U. S. *Patent Number 5*,630,096, 1997.

[17] E. Ipek, O. Mutlu, J. Mart´ınez, and R. Caruana. Self-optimizing memory controllers: A reinforcement learning approach. *In ISCA-35*, 2008.

[18] C. Natarajan, B. Christenson, and F. Briggs. A study of performance impact of memory controller features in multi-processor server environment. *In WMPI*, 2004.

[19] Z. Zhu and Z. Zhang. A performance comparison of DRAM memory system optimizations for SMT processors. *In HPCA-11*, 2005.

[20] S. A. McKee, W. A. Wulf, J. H. Aylor, M. H. Salinas, R. H. Klenke, S. I. Hong,and D. A. B. Weikle. Dynamic access ordering for streamed computations. *IEEE Transactions on Computers*, 49 pp. 1255–1271, Nov. 2000.

[21] C. Zhang and S. A. McKee. Hardware-only stream prefetching and dynamic access ordering. *In ICS-14*, 2000.

[22] K. J. Nesbit, N. Aggarwal, J. Laudon, and J. E. Smith. Fair queuing memory systems. *In MICRO-39*, 2006.

[23] O. Mutlu and T. Moscibroda. Parallelism-aware batch scheduling: Enhancing both performance and fairness of shared DRAM systems. *In ISCA-35*, 2008.

[24] O. Mutlu and T. Moscibroda. Stall-time fair memory access scheduling for chip multiprocessors. *In MICRO-40*, 2007.

[25] O. Mutlu, H. Kim, D. N. Armstrong, and Y. N. Patt. Using the first-level caches as filters to reduce the pollution caused by speculative memory references. *Interna-tional Journal of Parallel Programming*, 33(5) pp. 529–559, October 2005.

[26] W.-F. Lin, S. K. Reinhardt, and D. Burger. Reducing DRAM latencies with an integrated memory hierarchy design. *In HPCA-7*, pp. 301–312, 2001.

[27] I. Hur and C. Lin. Memory prefetching using adaptive stream detection. *In MICRO-39*, 2006.

[28] S. Srinath, O. Mutlu, H. Kim, and Y. N. Patt. Feedback directed prefetching: Improving the performance and bandwidth-efficiency of hardware prefetchers. *In HPCA-13*, 2007.

[29] J. Shao and B. T. Davis. A burst scheduling access reordering mechanism. *In Proceedings of the 2007 IEEE 13th International Symposium on High Performance Computer Architecture*, pp. 285–294, Washington, DC, USA, 2007. IEEE Computer Society.

[30] K. Sudan, N. Chatterjee, D. Nellans, M. Awasthi, R. Balasubramonian, and A. Davis. Micro-pages: increasing DRAM efficiency with locality-aware data placement. *In Proceedings of the fifteenth edition of ASPLOS on Architectural support for programming languages and operating systems, ASPLOS '10,* pp. 219–230, New York, NY, USA, 2010. ACM.

[31] I. Hur and C. Lin. Adaptive history-based memory schedulers. *In Proceedings of the 37th annual IEEE/ACM International Symposium on Microarchitecture, MICRO 37*, pp. 343–354, Washington, DC, USA, 2004. IEEE Computer Society.

[32] Y. Xu, A. S. Agarwal, and B. T. Davis. Prediction in dynamic SDRAM controller policies. *In Proceedings of the 9th International Workshop on Embedded Computer Systems: Architectures, Modeling, and Simulation, SAMOS '09*, pp. 128–138, Berlin, Heidelberg, 2009. Springer-Verlag.

[33] J. Stuecheli , D. Kaseridis, D. Daly, H. C. Hunter, and L. K. John. The virtual write queue: coordinating DRAM and last-level cache policies. *In Proceedings of the 37th annual international symposium on Computer architecture, ISCA '10*, pp. 72–82, New York, NY, USA, 2010. ACM.

[34] H.-H. S. Lee, G. S. Tyson, and M. K. Farrens. Eager writeback - a technique for improving bandwidth utilization. *In Proceedings of the 33rd annual ACM/IEEE international symposium on Microarchitecture, MICRO '33,* pp. 11–21, New York, NY, USA, 2000. ACM.

[35] C. J. Lee, V. Narasiman, E. Ebrahimi, O. Mutlu, and Y. N. Patt. DRAM-aware last level cache writeback: Reducing

write-caused interference in memory system. *In HPS Technical Report*, TR-HPS-2010-002.

Lei Li was born in Jiangsu province, China in 1980. She is currently an Ph.D Candidate of the School of Electronics Engineering and Computer Science at Peking University, China. Her research interests include SoC Design and information security. (E-mail: mail_lilei@pku.edu.cn)

Wei Zhang was born in Beijing, China in 1973. She is currently an Ph.D Candidate of the School of Electronics Engizneering and Computer Science at Peking University, China. Her research interests include SoC Design and information security. (E-mail: mail_lilei@pku.edu.cn)

Huiyao An was born in Hunan province, China in 1972. He is currently an associate professor of the School of Electronics Engineering and Computer Science at Peking University, China. His research interests include wireless communications, mobile networks, and information security. (E-mail: anhy@ss.pku.edu.cn)

Discrete Point Cloud Registration using the 3D Normal Distribution Transformation based Newton Iteration

Fengjun Hu and Tiaojuan Ren
Institute of Information Technology, Zhejiang Shuren University, Hangzhou 310015, Zhejiang, China
Email: jainism@msn.com, 541776059@qq.com

Shengbo Shi
Institute of Information Technology, Zhejiang Shuren University, Hangzhou 310015, Zhejiang, China
Email: shirleyssb@gmail.com

***Abstract*—The technology of three-dimensional reconstruction based on visual sensor has become an important research aspect. Based on Newton iteration algorithm, the improved 3D normal distribution transformation algorithm" (NI-3DNDT) is put forward, aiming to fix the problem of discrete point cloud registration algorithm in poor astringency and being open to local optimum. The discrete 3d point cloud adopts one order and two order derivative of piecewise smooth functions on surface, divides the point cloud space into Cubic grids, and calculate corresponding value of the mean and covariance matrix. To downgrade algorithm complexity, the Gauss function approximation of the log likelihood function is introduced, the probability density function parameters of 3D normal distribution transformation algorithm is simplified, and the Hessian matrix and gradient vector is solved through translation, rotation relation and Jacobean matrix; to make sure algorithm is converged to one certain point after a small number of iterations, it proposes that Newton iterative algorithm step be improved by employing better line search. Finally, the algorithm is put on simulation experiment and compared with other ways, the result of which proves that the suggested algorithm is able to achieve better registration effect, and Improve accuracy and efficiency.**

***Index Terms*—Point Cloud Registration; NI-3DNDT; Newton Iteration; Line Search; Visual Sensor**

I. INTRODUCTION

With the development of computer-aided technology, reverse engineering technology has been widely applied. The point cloud technology uses visual sensor, structured light camera and other equipments to obtain the target data of point cloud [1, 2], and to incorporate point clouds of multiple perspectives into a complete point cloud in different perspectives of translation, rotation matrix of rigid body in space that solved with 3D point cloud. Point cloud registration technology is the core technology of the three-dimensional reconstruction, and it is also the hotspot and focus of research in the computer vision field such as virtual reality, simulation design, heritage digitization. In addition, a growing number of commercial companies (such as Microsoft and Intel) have launched low-cost point cloud acquisition equipment to develop the subversive model of human-computer interaction. The launch of more interactive design commercial products has opened up a new era.

There are three ways of point cloud registration: instrument registration, manual registration and automatic registration. Automatic registration refers to using algorithms or statistical regularity to compute relationship of translation, rotation matrix of two point clouds. Registration technology has two steps: 1) Pre-match: using a simple one-dimensional features, including curvature [3], normal [4], tagging method [5], etc., or a high-dimensional feature descriptions such as shape [6], contour curve [7], and image rotation [8]. The effect of pre-matching will affect the result of exact match directly; 2) exact match, the most commonly used method is ICP (iterative closest point) algorithm [9, 10].

Iterative closest point (ICP) algorithm based on the quaternary array is an iterative registration algorithm put forward by Bsel and Mckay in 1992 which firstly utilizes the Newton iteration method or other searching methods to search the corresponding closest point pair in two sets of point cloud, then does the iterative operation on the objective function based on Euclidean distance and finally completes the registration of the three dimensional point cloud data. Many productive results have been achieved with a lot of researches by the domestic and foreign scholars [11]. Almhdie [12] et al. proposed Comprehensive ICP algorithm. The process of looking for corresponding point registration had introduced the concept of search matrix, ensuring the only match by search distance matrix, making the convergence rate of Comprehensive ICP algorithm quicker and the registration results more accurate. Du [13] et al. proposed Affine ICP algorithm, which first estimated the initial position of point cloud by independent component analysis, then registered the point cloud by combining the affine transformation and original ICP. Besl and McKay

doi:10.4304/jmm.9.7.934-940

[14] proposed point-to-point registration, while this method did not consider the curved shape around the point, which was likely to fall into local minimum value or the relatively small domain of convergence. However, it needed more iterations to converge. Chen and Medioni [15] proposed point-to-plane registration to find out the minimum value of the tangent plane in the current point cloud and the reference point cloud. However, this method was difficult to converge when there was a big gesture initialization error. Pathak [10] proposed an improved ICP algorithm of point-to-plane, which derived uncertainty posture parameters through the least squares method. Chen [16] et al. proposed an automatic registration algorithm based 3D shape modeling. They present the Hong-Tan Iterative Closest Point registration algorithm. This improved ICP algorithm could achieve higher registration accuracy.

Sharp [17] proposed a kind of ICP registration algorithm using Invariant Features (ICPIF) that investigates the use of Euclidean invariant features in a generalization of ICP registration of range images and modified the definition of the distance between two points in the original ICP algorithm by linear weights of the straight-line distance and image eigenvalues. This method is more accurate in the searching of corresponding point pairs.

Du [18] et al. put forward a novel algorithm for affine registration of point sets based on ICP algorithm. It acquired the estimation of the initial position of point cloud through independent component analysis (ICA), then used a combination way of affine transformation and original ICP for registration. A preferable result of affine registration requires better initial registration parameters obtained from ICA. Experiment results show that this proposed algorithm possesses higher accuracy.

In view of the problems for the ICP algorithm, Biber and Straßer [19] first proposed the application of normal transform algorithm (NDT) to 2D registration data. This method was of high speed and accurate rate. Biber and Fleck [20] et al. proposed probabilistic framework algorithm based on NDT algorithm, which matched the data points and the probability distribution model. Magnusson [21, 22] proposed the registration that applied the NDT algorithm to 3D point cloud data, and verified that the relative ICP algorithm had higher registration accuracy and efficiency, but the convergence of this algorithm was unknown. The NDT-based 3D point cloud surface was expressed as the number of first derivative and second derivative of piecewise smooth function, and the smooth function could be optimized with numerical optimization algorithm, which could be quickly converged. Moreover, the point in reference registration point cloud was not directly involved in the match, unlike ICP algorithm, which needed a lot of time to search for the last collar. In the case of relative poor initial value, the registration results of 3DNDT algorithm which has improved the efficiency were more accurate.

II. 3D Normal Distribution Transform Algorithm

Frequency domain methods utilize the aliasing existing in each low resolution image to reconstruct a high Unlike the traditional algorithm which uses a single point model, the goal of 3D normal distribution transform algorithm is to find the maximum of the matched likelihood function between a point in the current point cloud and the reference point cloud, and it does not use the corresponding feature points to calculate and match in a registration process, so it is faster than the other algorithms. The first step of 3DNDT algorithm is to divide spatial point cloud into cubes grid, and one of the points in the cube is expressed as $U = \{\vec{u}_1, ..., \vec{u}_n\}$, and the mean of each cube:

$$\vec{q} = \frac{1}{n}\sum_i \vec{u}_i \tag{1}$$

Covariance matrix is expressed as:

$$\Sigma = \frac{1}{n-1}\sum_i (\vec{u}_i - \vec{q})(\vec{u}_i - \vec{q})^t \tag{2}$$

The probability density function of some point like $\psi'(a_l)$ in the cube can be expressed as normal distribution function:

$$p(\vec{u}_i) = \varphi \exp(-\frac{(\vec{u}_i - \vec{q})^t \Sigma^{-1} (\vec{u}_i - \vec{q})}{2}) \tag{3}$$

where in a_q and $\psi(a_l)$ represent the mean and the covariance matrix of the cube concluding the $\psi(a_t)$ point respectively. In order to find the gesture of current point cloud when the likelihood function has the maximum value, assume that the transformation function relationship of the space between current point cloud and the reference point cloud is $\psi'(a_l)$, a_s refers to a reference point that is transformed from the point $\psi(a_l)0$ in current point cloud with the relation of rotation and translation $\psi'(a_l)$. As the formula shows below, we can get the R and T relation $\psi'(a_t)$ when the maximum value of likelihood function is analyzed.

$$\zeta = \prod_{i=1}^{n} p(\Gamma(\vec{\gamma}, \vec{u}_i)) \tag{4}$$

Take the logarithm of the formula above, we can get:

$$\log \zeta = \sum_{i=1}^{n} \log(p(\Gamma(\vec{\gamma}, \vec{u}_i))) \tag{5}$$

$$p(\vec{v}_i) = p(\Gamma(\vec{\gamma}, \vec{u}_i)) \tag{6}$$

$$p(\vec{v}_i) = b_1 \exp(-\frac{(\vec{v}_i - \vec{q})^t \Sigma^{-1} (\vec{v}_i - \vec{q})}{2}) + b_2 p_{outlier} \tag{7}$$

b_1, b_2 and $p_{outlier}$ refer to the proportion of expected outliers.

$$\log(p(\vec{v}_i)) = \log(c_1 \exp(-\frac{(\vec{v}_i - \vec{q})^t \Sigma^{-1} (\vec{v}_i - \vec{q})}{2}) + c_2) \quad (8)$$

Likelihood function can be approximated by a Gaussian function:

$$\log(c_1 \exp(-\frac{v^2}{2\sigma^2}) + c_2) = d_1 \exp(-d_2 \frac{v^2}{2\sigma^2}) \quad (9)$$

From $v = 0$, $v = \sigma$ and $v = \infty$, we can get:

$$d_1 = \log(c_1 + c_2) \quad (10)$$

$$d_2 = -2\log(\frac{\log(c_1 \exp(-1/2) + c_2)}{d_1}) \quad (11)$$

The probability density function of 3DNDT algorithm can be expressed as the following equation approximated by Gaussian function:

$$p'(\vec{v}_i) = -d_1 \exp(-d_2 \frac{(\vec{v}_i - \vec{q})^t \Sigma^{-1} (\vec{v}_i - \vec{q})}{2}) \quad (12)$$

Score function $s(\vec{\gamma})$ can be identified as the sum of probability density function of grid Gaussian of all cubes:

$$s(\vec{\gamma}) = \sum_{i=1}^{n} p'(\vec{v}_i) \quad (13)$$

III. Newton Iterative Algorithm for Solving Parameter

A. Hessian Matrix and the Gradient Vector

Using Newton iterative algorithm to find the solution of parameter a_s in $s(\vec{\gamma})$, work out the formula, wherein H and $\vec{g}$ are the Hessian matrix and gradient vector of $s(\vec{\gamma})$, respectively. $\vec{\gamma} \leftarrow \vec{\gamma} + \Delta\vec{\gamma}$, $\Delta\vec{\gamma}$ is the increment of the current pose estimation in each iteration.

$$\vec{v}_k - \vec{q} = \Gamma(\vec{\gamma}, \vec{u}_k) - \vec{q} \quad (14)$$

$$\vec{u}_k' \equiv \Gamma(\vec{\gamma}, \vec{u}_k) - \vec{q} \quad (15)$$

The gradient vector $\vec{g}$ is identified as

$$\vec{g_i} = \frac{\partial s}{\partial \gamma_i} = \sum_{k=1}^{n} d_1 d_2 \vec{u}_k'^T \Sigma_k^{-1} \frac{\partial \vec{u}_k'}{\partial \gamma_i} \exp(\frac{-d_2}{2} \vec{u}_k'^T \Sigma_k^{-1} \vec{u}_k') \quad (16)$$

Hessian matrix is $H_{ij} = \frac{\partial^2 s}{\partial \gamma_i \partial \gamma_j}$

$$H_{ij} = \sum_{k=1}^{n} d_1 d_2 \exp(\frac{-d_2}{2} \vec{u}_k'^{\Gamma} \Sigma_k^{-1} \vec{u}_k') \vec{u}_k'^{\Gamma} (-d_2 (\vec{u}_k'^{\Gamma} \Sigma_k^{-1} \frac{\partial \vec{u}_k'}{\partial \gamma_i})(\vec{u}_k'^{\Gamma} \Sigma_k^{-1} \frac{\partial \vec{u}_k'}{\partial \gamma_j}) + \vec{u}_k'^{\Gamma} \Sigma_k^{-1} \frac{\partial^2 \vec{u}_k'}{\partial \gamma_i \partial \gamma_j} + \frac{\partial \vec{u}_k'^{\Gamma}}{\partial \gamma_j} \Sigma_k^{-1} \frac{\partial \vec{u}_k'}{\partial \gamma_i}) \quad (17)$$

R and T relation $\phi(0) = -s(\vec{\gamma})$ refers to a point in the current cloud point $a_l = 0$, which is transformed into a point $a_u = 0$, through Euler angles $\phi'(0) = -g\Delta\vec{\gamma}$:

$$\vec{\gamma} = [t_x, t_y, t_z, \alpha, \beta, \gamma]^T \quad (18)$$

$$\Gamma(\vec{\gamma}, \vec{u}_i) = \begin{bmatrix} \cos\beta\cos\gamma & -\cos\beta\sin\gamma & \sin\beta \\ \cos\alpha\sin\gamma + \sin\alpha\sin\beta\cos\gamma & \cos\alpha\cos\gamma - \sin\alpha\sin\beta\sin\gamma & -\sin\alpha\cos\beta \\ \sin\alpha\sin\gamma - \cos\alpha\sin\beta\cos\gamma & \cos\alpha\sin\beta\sin\gamma + \sin\alpha\cos\gamma & \cos\alpha\cos\beta \end{bmatrix} \vec{u}_i + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (19)$$

so $\frac{\partial \Gamma}{\partial \gamma_i}$ can be solved by the column vector of Jacobian matrix J:

$$J(\gamma_i) = \begin{bmatrix} \frac{\partial \Gamma_1}{\partial \gamma_1} & \frac{\partial \Gamma_1}{\partial \gamma_2} & \frac{\partial \Gamma_1}{\partial \gamma_3} & \frac{\partial \Gamma_1}{\partial \gamma_4} & \frac{\partial \Gamma_1}{\partial \gamma_5} & \frac{\partial \Gamma_1}{\partial \gamma_6} \\ \frac{\partial \Gamma_2}{\partial \gamma_1} & \frac{\partial \Gamma_2}{\partial \gamma_2} & \frac{\partial \Gamma_2}{\partial \gamma_3} & \frac{\partial \Gamma_2}{\partial \gamma_4} & \frac{\partial \Gamma_2}{\partial \gamma_5} & \frac{\partial \Gamma_2}{\partial \gamma_6} \\ \frac{\partial \Gamma_3}{\partial \gamma_1} & \frac{\partial \Gamma_3}{\partial \gamma_2} & \frac{\partial \Gamma_3}{\partial \gamma_3} & \frac{\partial \Gamma_3}{\partial \gamma_4} & \frac{\partial \Gamma_3}{\partial \gamma_5} & \frac{\partial \Gamma_3}{\partial \gamma_6} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & k_1 & k_2 \\ 0 & 1 & 0 & k_3 & k_4 & k_5 \\ 0 & 0 & 1 & k_6 & k_7 & k_8 \end{bmatrix} \quad (20)$$

$\frac{\partial^2 \Gamma}{\partial \gamma_i \partial \gamma_j}$ can be calculated by the following matrix:

$$H_{ij}=\begin{bmatrix} \frac{\partial^2\Gamma}{\partial\gamma_1^{\ 2}} & \frac{\partial^2\Gamma}{\partial\gamma_1\partial\gamma_2} & \cdots & \cdots & \cdots & \frac{\partial^2\Gamma}{\partial\gamma_1\partial\gamma_6} \\ \frac{\partial^2\Gamma}{\partial\gamma_2\partial\gamma_1} & \ddots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ \frac{\partial^2\Gamma}{\partial\gamma_6\partial\gamma_1} & \cdots & \cdots & \cdots & \cdots & \frac{\partial^2\Gamma}{\partial\gamma_6^{\ 2}} \end{bmatrix}=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & k_9 & k_{10} & k_{11} \\ 0 & 0 & 0 & k_{12} & k_{13} & k_{14} \\ 0 & 0 & 0 & k_{15} & k_{16} & k_{17} \end{bmatrix} \tag{21}$$

B. Step Update Based on Linear Search

For step update problems of Newton iteration algorithm, we can use linear search algorithm. When the linear search algorithm satisfies the sufficient descent and curvature conditions, with a given continuously differentiable function ϕ, incoming 3DNDT algorithm fraction equation into the formula: $\phi(a_t)=-s(\vec{\gamma})$, $\phi^{'}(a_t)=-g\Delta\vec{\gamma}$.

For a continuously differentiable function f, and the down trend p of one point x of the f, we identify the function as:

$$\phi(a)\equiv f(x+ap) \tag{22}$$

$f\in[-\infty,\infty)$, $a>0$, so the function $\psi(a)$ can be identified as:

$$\psi(a)\equiv\phi(a)-\phi(0)-\mu\phi^{'}(0)a \tag{23}$$

$$\psi^{'}(a)=\phi^{'}(a)-\mu\phi^{'}(0) \tag{24}$$

For a search range $I[a_l,a_u]$, there is $a_t\in[a_l,a_u]$, for the update search interval $I^{'}[a_l^{+},a_u^{+}]$, exists $a_t^{+}\in[a_l^{+},a_u^{+}]$, so there exists:

$$\psi(a_t)=\phi(a_t)-\phi(0)-\mu\phi^{'}(0)a_t \tag{25}$$

$$\psi^{'}(a_t)=\phi^{'}(a_t)-\mu\phi^{'}(0) \tag{26}$$

Definition a_c is the minimum of cubic interpolation of $\psi(a_l)$, $\psi(a_t)$, $\psi^{'}(a_l)$ and $\psi^{'}(a_t)$, a_q is the minimum of cubic interpolation of $\psi(a_l)$, $\psi(a_t)$ and $\psi^{'}(a_l)$, a_s is the minimum of cubic interpolation of $\psi(a_l)$, $\psi^{'}(a_l)$ and $\psi^{'}(a_t)$. The renewal of a_t^{+} can be divided into the following four kinds of circumstances:

1) When $\psi(a_t)>\psi(a_l)$:

$$a_t^{+}=\begin{cases} a_c, if\ |a_c-a_l|<|a_q-a_l| \\ 0.5(a_c+a_q), otherwise \end{cases} \tag{27}$$

2) When $\psi(a_t)\le\psi(a_l)$ and $\psi^{'}(a_l)\psi^{'}(a_t)<0$:

$$a_t^{+}=\begin{cases} a_c, if\ |a_c-a_t|\ge|a_s-a_t| \\ a_s, otherwise \end{cases} \tag{28}$$

3) When $\psi(a_t)\le\psi(a_l)$, $\psi^{'}(a_l)\psi^{'}(a_t)\ge0$ and $|\psi^{'}(a_t)|\le|\psi^{'}(a_l)|$:

$$a_t^{'+}=\begin{cases} a_c, if\ |a_c-a_t|\ge|a_s-a_t| \\ a_s, otherwise \end{cases} \tag{29}$$

$$a_t^{+}=\begin{cases} \min\{a_t+0.66(a_u-a_t),a_t^{'+}\}, if\ a_t>a_l \\ \max\{a_t+0.66(a_u-a_t),a_t^{'+}\}, otherwise \end{cases} \tag{30}$$

4) When $\psi(a_t)\le\psi(a_l)$, $\psi^{'}(a_l)\psi^{'}(a_t)\ge0$ and $|\psi^{'}(a_t)|>|\psi^{'}(a_l)|$, then a_t^{+} is the minimum of cubic interpolation of $\psi(a_u)$, $\psi(a_t)$, $\psi^{'}(a_u)$ and $\psi^{'}(a_t)$.

According to the theory and methods of optimization, a_c, a_q and a_s can be defined respectively as:

$$a_c=a_l+\frac{(a_t-a_l)(w-\psi^{'}(a_l)-z)}{\psi^{'}(a_t)-\psi^{'}(a_l)+2w} \tag{31}$$

$$a_q=a_l-\frac{0.5\psi^{'}(a_l)(a_l-a_t)}{\psi^{'}(a_l)-(\psi(a_l)-\psi(a_t))/(a_l-a_t)} \tag{32}$$

$$a_s=a_l-\frac{(a_l-a_t)(w-\psi^{'}(a_l)-z)}{\psi^{'}(a_l)(\psi^{'}(a_l)-\psi^{'}(a_t))} \tag{33}$$

$$w=\sqrt{z^2-\psi^{'}(a_t)\psi^{'}(a_l)} \tag{34}$$

$$z=\frac{3(\psi(a_t)-\psi(a_l)}{a_t-a_l}-\psi^{'}(a_t)-\psi^{'}(a_l) \tag{35}$$

It is concluded, $\vec{\gamma}\leftarrow\vec{\gamma}+\Delta\vec{\gamma}a_t^{+}$, the renewal of algorithm is determined by the following conditions in search interval:

1) When $\psi(a_t)>\psi(a_l)$, then $a_l^{+}=a_l$, $a_u^{+}=a_t$, the algorithm continue to search.

2) When $\psi(a_t)\le\psi(a_l)$, $\psi^{'}(a_t)(a_l-a_t)>0$, then $a_l^{+}=a_t, a_u^{+}=a_u$, the algorithm continue to search.

3) When $\psi(a_t) \le \psi(a_l)$, $\psi'(a_t)(a_l - a_t) < 0$, then $a_l^+ = a_t, a_u^+ = a_l$, the algorithm continue to search.

4) Under other circumstances, the algorithm terminates the search.

IV. Experimental Results and Analysis

Programs in this paper are all written in C + + language, and run in the Intel Core i7-3820QM 2.70GHZ, 8.0GB of memory. As shown in figure 1, the experiment adopts different types of point cloud, and it contains the scan data of 360-degree in different perspectives in the same room. There are 110,000 data points in each point cloud. (The data were downloaded from http://kos.informatik.uni-osnabrueck.de/3Dscans)

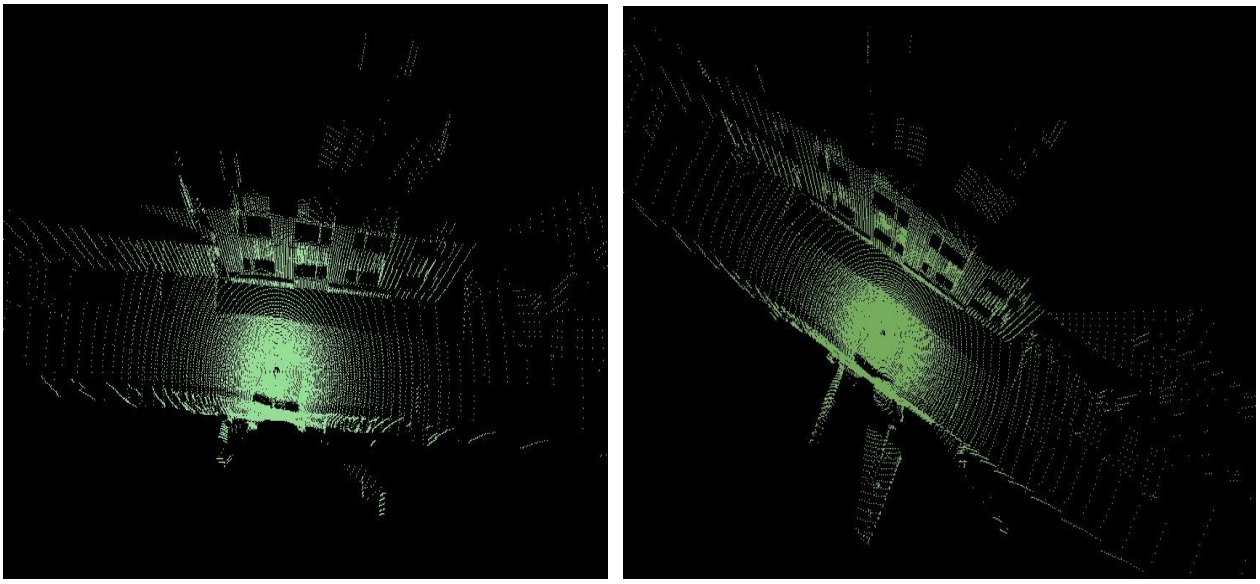
A. Example 1

Take the algorithm in Doctor Magnusson's paper to test in point cloud a1 and a2: The resolution of NI-3DNDT is $r = 1$, $p_{outlier} = 0.55$, $c_1 = 10*(1 - p_{outlier}) = 4.5$, $c_2 = \frac{p_{outlier}}{r^3} = 0.55$.

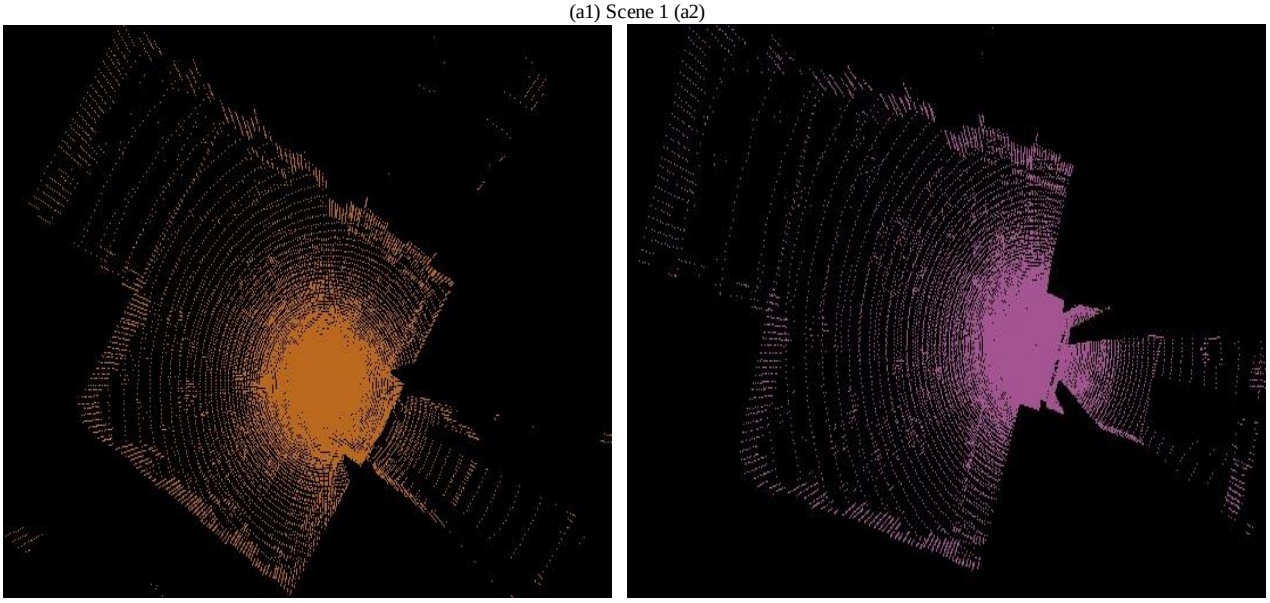
According to the Newton iterative initialization parameter $\phi(0) = -s(\vec{\gamma})$, $a_l = 0$, $a_u = 0$ and $\phi'(0) = -g\Delta\vec{\gamma}$, the translation, rotation matrix can be calculated:

0.793983	–0.447965	0.0110171	2.27322
0.447847	0.794034	0.0116674	0.0730304
–0.0150726	–0.00549647	0.999871	0.0523041
0	0	0	1

To test the improved algorithm NI-3DNDT which is proposed in this paper, the translation, rotation matrix can be calculated:



(a1) Scene 1 (a2)



(a3) Scene 2 (a4)

Figure 1. Different types of 3D point cloud data

0.893912	–0.448124	0.0103027	2.48265
0.448013	0.893954	0.0114286	0.0755997
–0.0143315	–0.0056004	0.999882	0.0527146
0	0	0	1

The error distance of the two kinds of detection algorithm are 0.17511 and 0.14543 respectively, and the running time of them are 1350ms and 1160ms respectively. Therefore, with the detection algorithms in this paper, we can get better results. Besides, it has a high convergence speed and accuracy registration. Figure 2 and figure 3 are the comparison of 3DNDT algorithm and the algorithm which is proposed in this paper in registration time and errors when the three-dimensional spatial resolution from 0.1 to 4.5. As shown in figure 2, when the resolution range is from 0.7 to 1.5, NI-3DNDT obtains a good effect of registration, registration algorithm convergence rate is also significantly faster than the improved algorithm (about 50ms). As shown in figure 3, detection time of these two kinds of registration algorithm increases with the improvement of detection resolution.

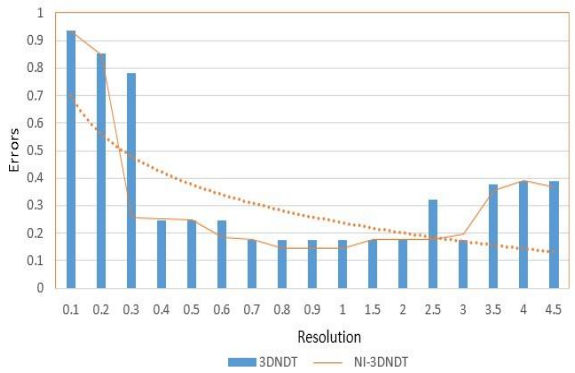


Figure 2. Algorithm detects errors comparison chart based on Scene one

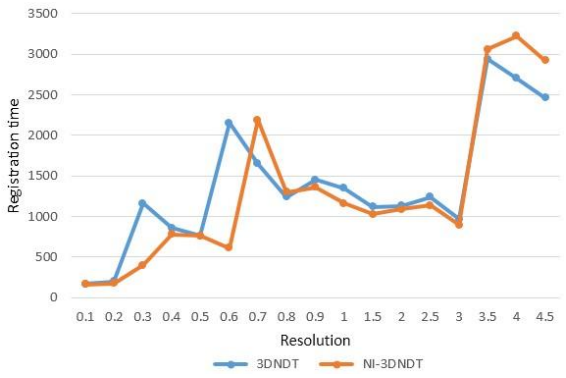


Figure 3. Algorithm matching time comparison chart based on Scene one

B. Example 2

To test the two registration algorithm in another scenario, put the algorithm of Doctor Magnusson's paper in point cloud a3 and a4: The resolution of 3DNDT is $r = 0.6$, other experimental conditions are the same as the experimental one, and the translation, rotation matrix can be calculated:

0.881566	–0.191123	0.0006638	0.1313989
0.191124	0.881564	–0.001689	–0.347889
–0.000329	–0.001783	0.999998	0.0006301
0	0	0	1

To test the improved algorithm NI-3DNDT which is proposed in this paper, the translation, rotation matrix can be calculated:

0.984268	–0.176681	–0.000009	–0.032258
0.176681	0.984276	–0.001173	–0.325616
0.000296	0.0011387	0.999989	–0.000732
0	0	0	1

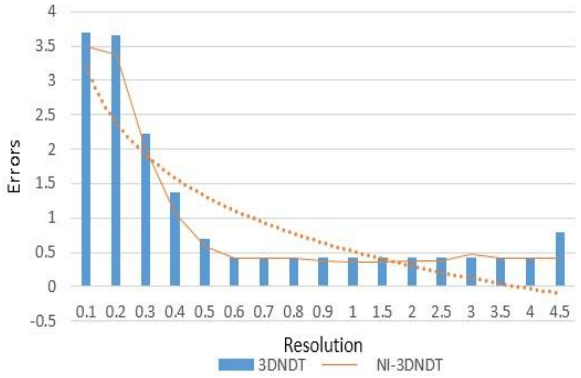


Figure 4. Algorithm detects errors comparison chart based on Scene two

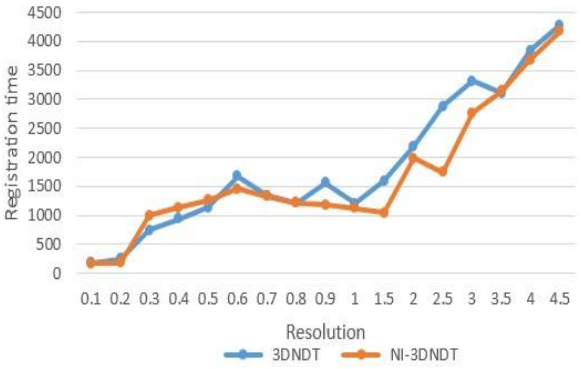


Figure 5. Algorithm matching time comparison chart based on Scene two

Under the circumstance that the resolution is $r = 0.6$, the error distance of the two kinds of detection algorithm are 0.418381 and 0.349911 respectively, and their running time are 1210 ms and 1120 ms respectively. Registration algorithm proposed in this paper is better. As shown in figure 4, when the resolution range is from 0.6 to 1.0, NI-3DNDT obtains a good effect of registration. Registration algorithm convergence rate, about 140ms, is also significantly faster than the improved algorithm. Compared with scenario one, NI-3DNDT algorithm is still obtained good effect of registration when the resolution is extreme. Therefore, we should set different detection resolutions and step lengths according to the type of different scenarios.

V. CONCLUSION

3DNDT algorithm, as a new algorithm of point cloud registration, is put forward in recent years. And it has attracted more and more attention from the researchers because of its high matching precision and convergence speed. NI-3DNDT algorithm proposed in this paper is simplified in the terms of the parameters of the probability density function and improved in view of the linear search algorithm. Besides, it updates step length of Newton iteration algorithm dynamically, accelerates the convergence speed of Newton iteration algorithm, and improves the matching precision. Experiments show that the algorithm has more effective stability on the point cloud in different kinds of scenarios. The algorithm has a good performance under the condition of appropriate spatial resolution, and it has certain advantages in the matching accuracy and matching speed compared with existed algorithms. Therefore this algorithm has higher application value and theoretical reference value.

ACKNOWLEDGEMENT

This work was supported by Natural Science Foundation of Zhejiang province, China (Grant No. LY13F010013), the Department of Science and Technology of Zhejiang Province (Grant No. 2014C31065), and the National Natural Science Foundation of China (Grant No. 51175473).

REFERENCE

[1] Chen S Y, Li Y F, Zhang J W. Vision Processing for Real time 3D Data Acquisition Based on Coded Structured Light. *IEEE Transactions on Image Processing*, 2008, 17(2) pp. 167-176.

[2] Chen S Y, Zhang J H, Li Y F, A Hierarchical Model Incorporating Segmented Regions and Pixel Descriptors for Video Background Subtraction. *IEEE Transactions on Industrial Informatics*, 2012, 8(1) pp. 118-227.

[3] Wang S, Wang Y, Jin M, et al, Conformal geometry and its applications on 3D shape matching, recognition, and stitching. *Medical Engineering & Physics*, 2007, 29(7) pp. 1209-1220.

[4] Kalle Å, Johan K, Olof E, Anders E, Fredrik K. Automatic feature point correspondences and shape analysis with missing data and outliers using MDL. *Image Analysis*, 2007, 4522(5) pp. 21-30.

[5] Luo X, Zhong Y X, Li R J. Data registration in 3D scanning systems. *Journal of Tsinghua University (Science and Technique)*, 2004, 44(8) pp. 1104-1106.

[6] Frome A, Huber D, Kolluri R, et al. Recognizing objects in range data using regional point descriptors // *In Proceedings of the European Conference on Computer Vision, Prague, Czech Republic*, 2004 pp. 224-237.

[7] Yang R, Allen P. Registering integrating and building cad models from range data //*In: IEEE International Conference on Robotics and Automat ion, Leuven, Belgium*, 1998 pp. 3115-3120.

[8] Johnson A, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(5) pp. 674-686.

[9] Herrmann M, Otesteanu M, Otto M. Fast and robust point cloud matching based on EM-ICP prepositioning // *10th International Symposium on Electronics and Telecommunications, Timisoara, Rumania*, 2012 pp. 99-103.

[10] Pathak K, Birk A, Vaškevičius N, Poppinga J. Fast Registration Based on Noisy Planes With Unknown Correspondences for 3-D Mapping. *IEEE Transactions on Robotics*, 2010, 26(3) pp. 424-441.

[11] Xin W, Pu J X. Point cloud integration base on distances between points and their neighborhood centroids. *Journal of Image and Graphics*, 2011, 16(5) pp. 886-891.

[12] Almhdie A, Léger C, Deriche M, et al. 3D registration using a new implementation of the ICP algorithm based on a comprehensive lookup matrix: Application to medical imaging. *Pattern Recognition Letters*, 2007, 28(12) pp. 1523 –1533.

[13] Du S Y, Zheng N N, Meng G F, Yuan Z J. Affine registration of point sets using ICP and ICA. *IEEE Signal Processing Letters*, 2008, 15 pp. 689-692

[14] Besl P J, McKay N D. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, 14(2) pp. 239-256.

[15] Chen Y, Medioni G. Object modelling by registration of multiple range images. *Image and vision Computing*, 1992, 10(3) pp. 145-155.

[16] Chen J, Wu X J, Michael Y W, et al. 3D shape modeling using a self-developed hand-held 3D laser scanner and an efficient HT-ICP point cloud registration algorithm. *Optics and Laser Technology*, 2013, 45(1) pp. 414-423.

[17] Sharp G. C, Lee S. W, Wehe D. K. ICP registration using invariant features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(1) pp. 90-102.

[18] Du S. Y, Zheng N. N, Meng G. F, Yuan Z. J. Affine Registration of Point Sets Using ICP and ICA. *IEEE Signal Processing Letters*, 2008, 15 pp. 689-692.

[19] Biber P, Straßer W. The normal distributions transform: A new approach to laser scan matching //*In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Las Vegas, USA*, 2003 pp. 2743–2748.

[20] Biber P, Fleck S, Strasser W. A Probabilistic Framework for Robust and Accurate Matching of Point Clouds. *Lecture Notes in Computer Science*, 2004, 3175 pp. 480-487.

[21] Magnusson M, Andreasson H, Nuchter A, Lilienthal A J. Appearance-Based Loop Detection from 3D Laser Data Using the Normal Distributions Transform //*IEEE International Conference on Robotics and Automation, Osaka, Japan*, 2009 pp. 23-28.

[22] Magnusson M. The Three-Dimensional Normal-Distributions Transform-an Efficient Representation for Registration, Surface Analysis, and Loop Detection. Sweden: *Orebro University*, 2009.

Dr. Fengjun Hu is a PhD student in control theory and control engineering from Zhejiang University of technology. Currently, he is a researcher at Zhejiang Shuren University. His major research interests include intelligent robot and machine vision. He has published nearly ten papers in related journals indexed by Ei Compendex. He has been supported by Zhejiang Department of Education research project and Zhejiang Province excellent Youths Research Plan of Colleges and Universities.

Salient Region Detection based on Global Contrast and Object-Biased Gaussian Refinement

Xing Zhang and Xiaolin Zhang
Key Laboratory of Wireless Sensor Network & Communication, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China

***Abstract*—In this study, we address the problem of salient region detection. Recently, salient region detection methods with histogram-based contrast and region-based contrast have given promising results. Center bias is a hypothetical characteristic in human vision system which is applied in many existing salient region detection methods. In this paper, we propose an object-biased Gaussian model to refine the histogram-based contrast method and region contrast method. The proposed algorithm is simple, efficient, and produces full-resolution, high-quality saliency maps. We extensively evaluated our algorithm using traditional salient object detection benchmark, as well as a more challenging co-saliency object detection benchmark. Experimental results demonstrate that the proposed algorithm outperforms the original global contrast methods and other existing salient object detection methods.**

***Index Terms*—Visual Attention; Color Contrast; Saliency; Gaussian Refinement**

I. INTRODUCTION

With the development of technology in recent years, the data volume of image and video is growing rapidly, and the demand for applications based on image and video is growing, too. It is a hotspot in computer vision to extract effective information from large amount of image and video fast and accurately. The principle of human visual system could help to solve the problem. Human visual system is one of the most complex and efficient systems in the world. And Visual attention is the key mechanism to guide attention allocation and visual cognition in the human visual system. The ability of image and video data processing in computer could be improved by imitating the visual attention mechanism. Visual attention detection is an effective method to solve difficult issues in computer vision.

Saliency denotes the visual uniqueness, rarity or surprise and is related to many attributes of image like color, gradient, edges, and boundaries. Different facts of science tried to find out the principle of this phenomenon. Psychologists have studied visual attention related human behaviors like change blindness [1], inattentional blindness [2] and attentional blink [3]. Neurophysiologists have shown that the neurons could effectively accommodate themselves to represent salient objects better [4]. Computational neuroscientists have tried to simulate and explain the attention behavior by building realistic neural network models [5, 6]. Enlightened by these studies, robotists and computer vision scientists tried to solve and implement visual saliency in computer systems.

According to the work by Treisman and Gelade [7], Koch and Ullman [8], and Itti et al. [9], the visual attention in human system could be divided into two stages: scene-driven bottom-up (BU) saliency extraction, which is mainly based on the feature of a certain visual scene; expectation-driven top-down(TD) saliency extraction, which is determined by cognitive phenomena like knowledge, expectations, reward, and current goals. Usually, the bottom-up saliency extraction is fast, pre-attentive; while the top-down saliency extraction is slower and task dependent.

We focus on bottom-up data driven saliency detection via global contrast. The definition of salient regions that capture human's visual attention is a combination of the insights from neuroscience, biology, computer vision, and other fields. Based on the bionic model of visual attention provided by Koch and Ullman [10], the saliency is defined by Itti et al. [9] with center difference of multi-scale image characteristic. The meaning of salient region detection in computer system lies in the ability to allocate computational resources for image analysis in priority. Recent researches have shown that the bottom-up stimuli-driven attention models [8, 11, 12] could be effectively applied to many application scenarios, including image segmentation [13-15], resizing [16], and object recognition [17].

Cheng et al. [18] proposed two global contrast methods to detect salient regions, which is histogram-based contrast method (HC) and region-based contrast method (RC). HC and RC have shown to give promising results. Based on HC and RC method, we propose an object-biased Gaussian refinement method to improve the performance of HC and RC. Center bias is a hypothetical characteristic in human vision system which is applied in many existing salient region detection methods [19]. But most of the methods adopt the center coordinate of the input image as the center of center bias. We calculate the center of the salient object and make a refinement to highlight the salient regions and downplay backgrounds. The exemplary effect of our method is shown in Fig. 1.

The meaning of the object-biased Gaussian refinement method lies in the reduction of unnecessary details, small features, local features or errors in the background of a certain saliency map while keeping salient regions with

doi:10.4304/jmm.9.7.941-947

its high saliency value. As shown in the third row of Fig.1, after the refinement, the salient regions retain the same, but the background is more clean and simple.

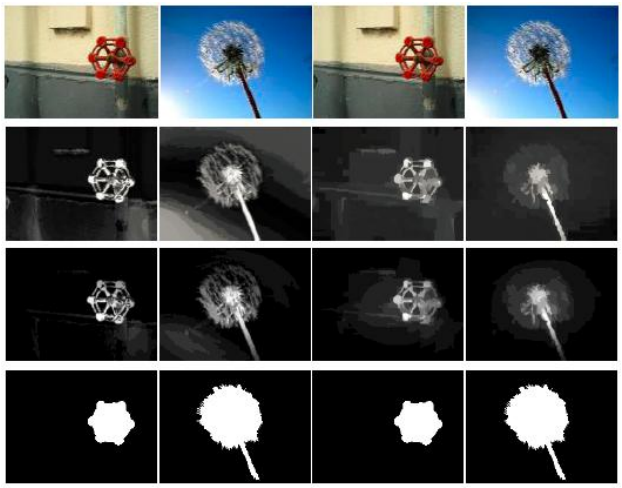


Figure 1. Exemplary results of our methods. The first row are original images, the second row are HC and RC results, respectively, the third row are object-biased Gaussian refined results HC* and RC*, respectively, the last row are ground truth

In the evaluation, we adopt the benchmark provided by Achanta et al. [20], which contains 1000 images with binary ground truth in the form of accurate human-marked labels for salient regions. In order to show the effectiveness of HC* and RC*, we adopt another challenging benchmark provided by Li and Ngan [21], which contains 210 images with binary ground truth. This benchmark is used for co-saliency detection, which means 210 images are divided into 105 image pairs and within each image pair, the salient objects are similar but the backgrounds are different.

The remainder of this paper is organized as follows. In Section 2, related works of saliency detection are reviewed. In Section 3, histogram-based contrast method (HC) and region-based contrast method (RC) are presented. In Section 4, the object-biased Gaussian refinement method is proposed. Experimental evaluations and discussions are presented in Section 5, and Section 6 is the conclusion.

II. Related Work

According to the definition of saliency, bottom-up visual attention (saliency) could be broadly classified into three categories [22]: fixation prediction, salient object detection, and objectness proposal generation.

Fixation prediction models are designed to imitate human eye fixation function [23, 24]. Based on the highly influential biologically inspired early representation model introduced by Koch and Ullman [10], Itti et al. [9] defined image saliency using central surrounded differences across multi-scale image features. To extend this method, Ma and Zhang [25] introduce a fuzzy growing method, and a graph-based visual saliency detection algorithm is proposed by Harel et al. [26]. Hou and Zhang [27] determine saliency based on phase spectrum of the Fourier transform of an image. Zhang et al. [28] determine salient image regions using information theory. However, as shown by Cheng et al. [18], the final saliency maps are generally blurry, and the small, local features are often over emphasized, which limit these approaches' applications to the areas such as segmentation, detection, etc.

Salient object detection models aim at detecting and segmenting the most salient objects in a scene [19, 29]. Achanta et al. [20] proposed a frequency-tuned approach to detect salient object. Goferman et al. [30] adopt a patch based approach. Cheng et al. [18] proposed a global contrast approach to localize and segment salient objects. Rahtu et al. [31] propose a saliency detection algorithm by measuring the center-surround contrast of a sliding window over the entire image. Borji and Itti [32] computed global saliency inverse proportionally to the probability of a patch appearing in the entire scene. Perazzi et al. [33] adopted a filter frame work to produce high quality saliency map. We combine global contrast method HC and RC with object-biased Gaussian refinement to generate high performance saliency maps. The corresponding methods are called HC* and RC*.

Objectness proposal generation methods produce category-independent proposals to point out all the objects in a certain image [34-36]. Generating rough segmentations as objectness proposals has shown to be an effective approach [35, 37], while the processing time unacceptable. Alexe et al. [34] adopted a cue integration model to improve the performance. Zhang et al. [38] took oriented gradient feature within cascaded ranking SVM method to generate the proposal. Uijlings et al. [36] improved the performance by a selective search approach. Cheng et al. [22] proposed a simple ,intuitive and fast method, resizing the window and using the norm of the gradients as a simple 64D feature.

The pre-attentive bottom-up saliency detection could also be broadly classified into local and global schemes [18, 39]. Local contrast based methods investigate the rarity of image regions with respect to (small) local neighborhoods [25]. The saliency is determined by dissimilarities at pixel-level. Multi-scale Difference of Gaussians [11] or histogram analysis [12] may be taken. On the contrary, global contrast based methods evaluate saliency of an image region using its contrast with respect to the entire image [20]. Specifically, Goferman et al. [30] adopt a patch based approach. Wang et al. [40] compute saliency over the whole image relative to a large dictionary of images. Liu et al. [12] measure center-surrounded histograms over windows of various sizes and aspect ratios in a sliding window manner, and learn the combination weights relative to other saliency cues. Cheng et al. [18] propose a region contrast-based method to model global contrast. Cheng et al. [41] also propose an algorithm taking the advantage of soft image abstraction. Perazzi et al. [33] decompose an image into perceptually homogeneous elements, which abstract unnecessary details, to improve salient object detection.

The evaluation results in label [12] indicate that contrast based themes using individual feature are difficult to assess. Our work adopts two kinds of contrast

methods and refines the saliency maps, aiming at highlighting the salient regions and downplaying the background separately.

III. GLOBAL CONTRAST METHOD HC AND RC

In this section, histogram-based contrast method (HC) and region-based contrast method (RC) are discussed in details and exemplary results are given.

A. Histogram-based Contrast Method

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image [42]. As proposed in Cheng et al. [18], the saliency of a pixel is determined by color global contrast. The saliency value of a pixel I_k in image is defined as,

$$S(I_k) = \sum_{\forall I_i \in I} D(I_k, I_i) \tag{1}$$

where $D(I_k, I_i)$ is the color distance metric between pixels I_k and . The color is computed in the L∗a∗b∗ color space. Expand Eq.1, is represented as,

$$S(I_k) = D(I_k, I_1) + \cdots + D(I_k, I_N) \tag{2}$$

where *N* is the number of pixels in image *I*. It's obviously that the saliency value of the pixels with the same color is the same. Hence, group the pixels with the same color c_j, and rearrange Eq. 2, we get,

$$S(I_k) = S(c_l) = \sum_{j=1}^{n} f_i D(c_l, c_j) \tag{3}$$

where c_l is the color value of pixel , n is the number of distinct pixel colors, and f_i is the probability of pixel color in image *I*. Exemplary results are given in Fig. 2.

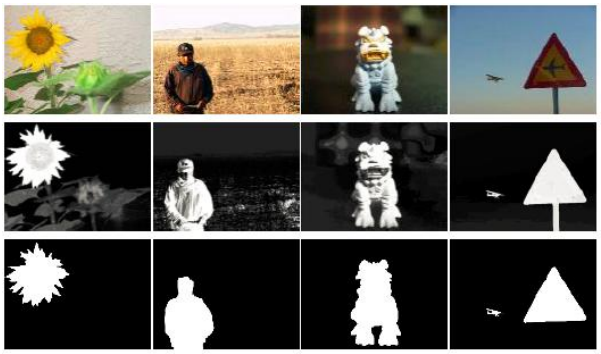


Figure 2. Exemplary results of HC. The first row are original images, the second row are HC results, the third row are ground truth

B. Region-based Contrast Method

Human pays more attention to those image regions that contrast strongly with their surroundings [43]. The image is first segmented into regions via graph-based image segmentation [44]. Then generate histogram of each region. The saliency of region r_k is determined by the color global contrast to other regions.

$$S(r_k) = \sum_{r_k \neq r_i} \omega(r_i) D_r(r_k, r_i) \tag{4}$$

where $\omega(r_i)$ is the weighted value of region . In our work, is represented as the point number of region r_i . $D_r(\cdot,\cdot)$ denotes the color distance of two regions. Give region r_1 and r_2 the color distance is defined as,

$$D_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_{1,i}) f(c_{2,j}) D(c_{1,i}, c_{2,j}) \tag{5}$$

where $f(c_{k,i})$ is the probability of the i-th color $c_{k,i}$, among all n_k colors in the k-th region Exemplary results are given in Fig. 3.

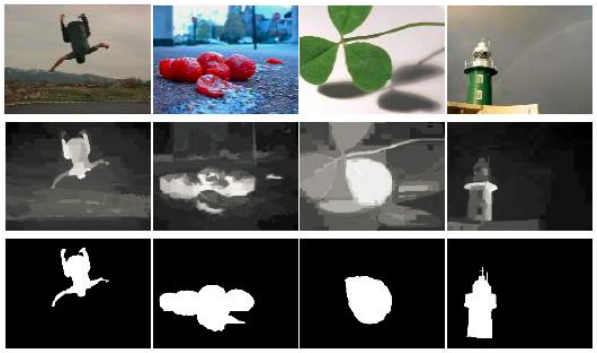


Figure 3. Exemplary results of RC. The first row are original images, the second row are RC results, the third row are ground truth

IV. OBJECT-BIASED GAUSSIAN REFINEMENT METHOD

As shown in Fig. 1, Fig. 2 and Fig. 3, the salient object are closed to the ground truth while in the background, there are lot of unnecessary details. HC generates pixels level saliency map and the same color has the same saliency. RC measures the saliency scale of a single region. So how to highlight the salient object and downplay the background is a valuable issue to be addressed.

Ideally, colors belonging to the background will be distributed over the entire image exhibiting a high spatial variance, whereas foreground objects are generally more compact [30]. We treat the unnecessary details in the background as noise and take the advantage of Gaussian filter to downplay the noise.

The first step is to calculate the center coordinate of the salient object. With this coordinate as the center of the Gaussian model, we could refine the saliency map and generate saliency maps with highly differentiated salient object and background.

Borji et al. [19] point that there is center bias in some of the saliency detection benchmark. Center bias is also a visual phenomenon in human visual system. It is applied in several salient region detection methods, and could be formulated by Gaussian model,

$$G(z) = \exp\left[-\left(\frac{(x_z - \mu_x)^2}{2\sigma_x^2} + \frac{(y_z - \mu_y)^2}{2\sigma_y^2} \right) \right] \tag{6}$$

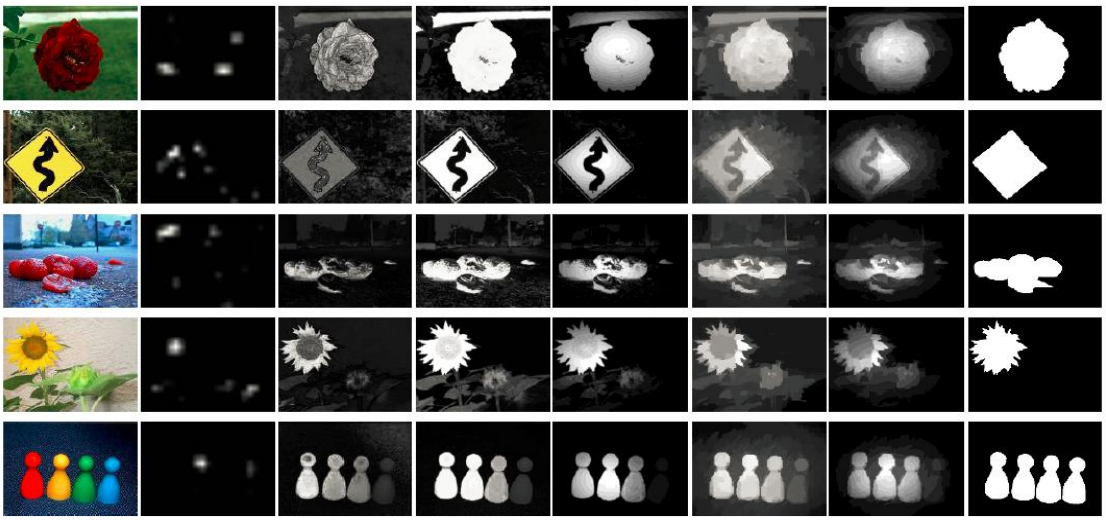


Figure 4. Visual comparison of saliency maps on benchmark I. HC* is the refined results of HC method, RC* is the refined results of RC method

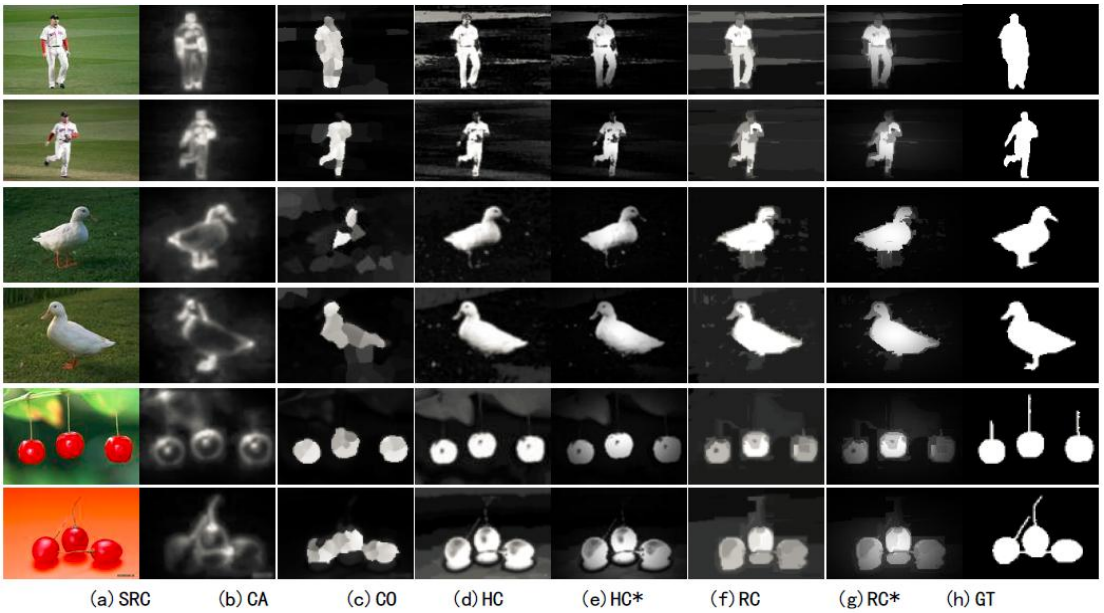


Figure 5. Visual comparison of saliency maps on benchmark II. HC* is the refined results of HC method, RC* is the refined results of RC method

where, (μ_x, μ_y) is the center of the Gaussian model. The center of the Gaussian model is set to be the center of the input image in several applications [45-47]. Although it provides promising results, but the center should be determined by the content of the image.

In this work, we adopt the Gaussian model to refine the saliency maps generated by HC and RC methods. And the center of the Gaussian model is generated by calculating the center of the object, which is defined as,

$$\begin{cases} x_0 = \sum_i \frac{S(i)}{\sum_j S(j)} x_i \\ y_0 = \sum_i \frac{S(i)}{\sum_j S(j)} y_i \end{cases} \quad (7)$$

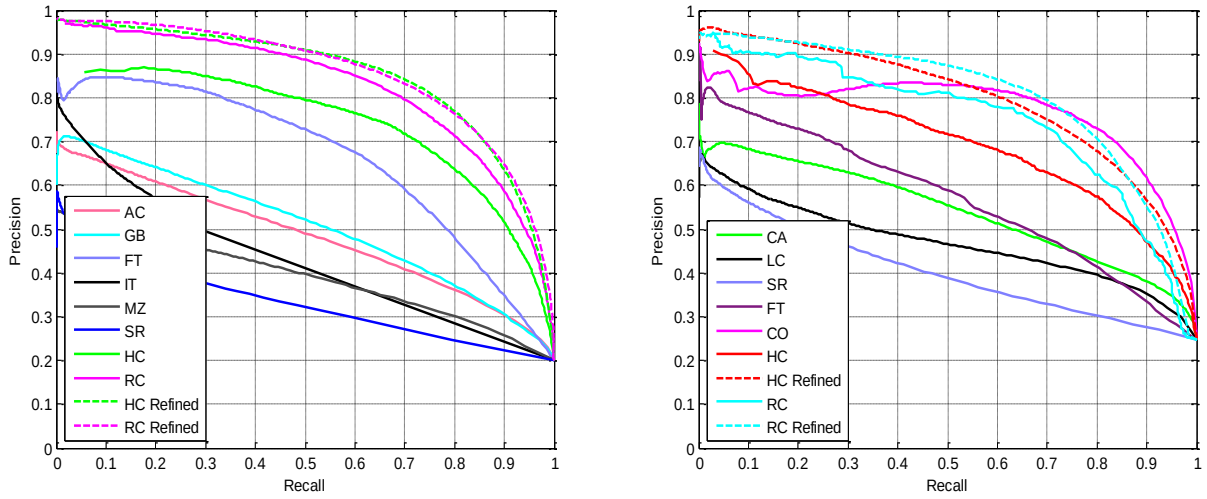
where, (x_i, y_i) is the center of each region. $\sigma_x = 0.25 \times H$, $\sigma_y = 0.25 \times W$, H and W are the height and width of the input image separately. With Eq. 6 and Eq. 7, the saliency map could be refined as,

$$S_{refined}(z) = G(z) * S(z) \quad (8)$$

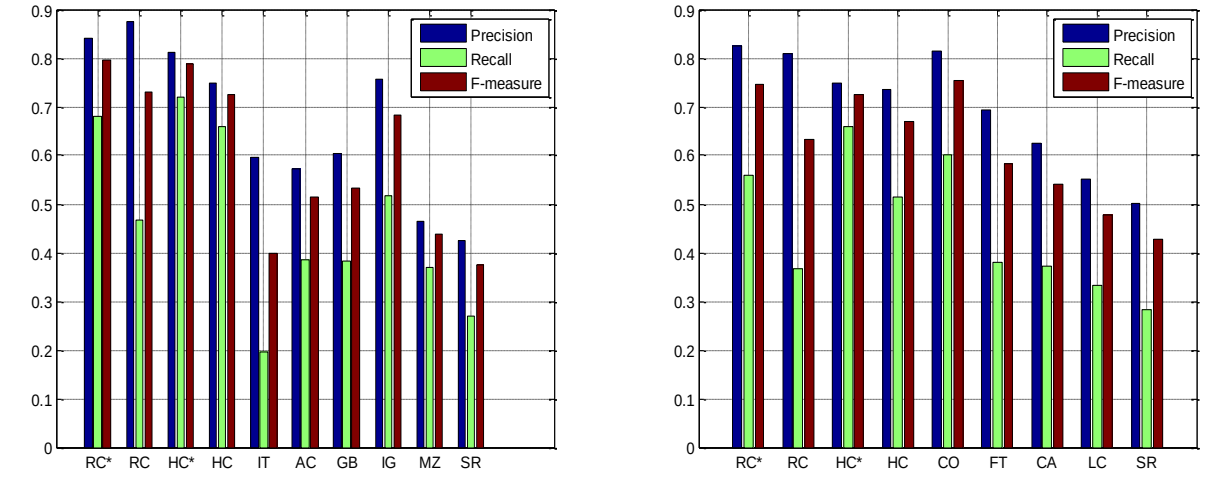
With Eq. 8, we could generate the refined saliency maps, exemplary results of refined HC and RC methods are shown in Fig. 1.

V. Experimental Results and Discussion

The evaluation of our method contains three parts. First, a visual comparison of HC* and RC* results with other existing methods are given. Second, we adopt precision and recall curve to evaluate our results with other existing results. Third, we evaluate our results via F-measure.



(a) Fixed Thresholding



(b) Adaptive Thresholding

Figure 6. Statistical comparison with 11 other saliency detection methods using benchmark I (Left) and benchmark II (Right). (a) and (b) are average precision and recall curves by fixed/ adaptive thresholding separately

A. Evaluation Benchmarks

The results of our methods are evaluated on two benchmarks. Benchmark *I* is provided by Achanta et al. [20], which contains 1000 images with binary ground truth in the form of accurate human-marked labels for salient regions. Benchmark *II* is more challenging. It is provided by Li and Ngan [21], which contains 210 images with binary ground truth. Benchmark *II* is used for co-saliency detection, which means 210 images are divided into 105 image pairs and within each image pair, the salient objects are similar but the backgrounds are different. Our method is compared with FT [20], GB [26], AC [48], IT [49], MZ [25], CA [30], LC [50], SR [27], CO [21].

B. Visual Comparison

The visual comparison of saliency maps are given in Fig. 4 and Fig. 5 in order to explore the effect of object biased Gaussian refinement directly. In Fig. 4, benchmark *I* is adopted and in Fig. 5, benchmark *II* is adopted. The results of refined HC and RC method are also given.

As shown in the comparison, especially in the comparison of HC and HC*, RC and RC*, after the object-biased Gaussian refinement, the background regions are effective downplayed and the salient regions keeps their saliency. Through the visual comparison, the proposed HC* and RC* also show better performance than other approaches.

C. Precision and Recall Curve

According to [12, 18, 20], we evaluate our method by calculating its *precision and recall* rate. *Precision* measures the pixels that are correctly assigned in percentage, and *Recall* measures the salient pixels that are correctly detected in percentage. First we adopt fixed threshold in the range of [0.255] to generate corresponding binary saliency maps. Each binary saliency map is assigned to a *precision and recall* value with comparison to the ground truth mask. The resulting

precision and recall curve with fixed thresholding is shown in Fig. 6(a).

In the second evaluation, we adopt the image depend threshold proposed by [20], defined as twice the mean saliency of the image:

$$T_a = \frac{2}{W \times H} \sum_{x=1}^{W} \sum_{y=1}^{H} S(x, y) \tag{9}$$

where, W and H are the width and height of the saliency map S and $S(x, y)$ is the saliency value in saliency map S within the position . The resulting precision and recall curve with adaptive thresholding is shown in Fig. 6(b).

As shown in Fig. 6(a), our object-biased Gaussian refinement method could effectively improve the precision of HC and RC in both benchmark *I* and *II*. And the improvement of HC is greater than that of RC. This is because the HC method focuses on histogram and the same color has the same saliency, which means there will be some small regions with high saliency in the background. The salient regions are not that centralized. While the object-biased Gaussian refinement method could effective downplay this kind of regions.

As shown in Fig. 6(b), the *precision and recall* of HC are improved in both benchmarks, the *recall* of RC are improved too, while in benchmark *I*, the precision is reduced and *recall* is improved. The performance of the refinement method will be evaluated by *F-measure* which is the combination of *precision and recall*.

D. F-measure

The weighted harmonic measure or *F-measure* is also taken to evaluate our method, which is defined as:

$$F_\beta = \frac{(1+\beta^2) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \tag{10}$$

Similar to [18, 20, 33] , we set $\beta^2 = 0.3$ to weigh *precision* more than *recall*.

The resulting *F-measure* is shown in Fig.6(b). The *F-measure* of HC* and RC* are both improved significantly, which means the performance of the saliency maps could be effective improved via our object-biased Gaussian refinement method.

VI. CONCLUSIONS

In this paper, we propose an object-biased Gaussian refinement method to improve the global contrast method. HC* and RC* are proposed to shown the effect. Experiment based on two benchmarks are given, including visual comparison, *precision and recall curve* and *F-measure* evaluation. The results illustrate that, our object-biased Gaussian refinement method could effectively improve the performance existing saliency maps generated by global contrast methods.

ACKNOWLEDGMENT

This work is supported by Chinese Academy of Science Strategic Pioneer Research and Development Program (Project Number: XDA06020800) and the Plan of Action on Science and Technology Innovation, Science and Technology Commission of Shanghai Municipality (Project Number: 13511507503).

REFERENCES

[1] D. T. Levin, and D. J. Simons, "Failure to detect changes to attended objects in motion pictures," *Psychonomic Bulletin & Review*, vol. 4, no. 4, pp. 501-506, 1997.

[2] D. J. Simons, and C. F. Chabris, "Gorillas in our midst: Sustained inattentional blindness for dynamic events," *Perception-London*, vol. 28, no. 9, pp. 1059-1074, 1999.

[3] J. E. Raymond, K. L. Shapiro, and K. M. Arnell, "Temporary suppression of visual processing in an RSVP task: An attentional blink?," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 18, no. 3, pp. 849, 1992.

[4] S. Treue, "Neural correlates of attention in primate visual cortex," *Trends in neurosciences*, vol. 24, no. 5, pp. 295-300, 2001.

[5] E. T. Rolls, and G. Deco, "Attention in natural scenes: neurophysiological and computational bases," *Neural networks*, vol. 19, no. 9, pp. 1383-1394, 2006.

[6] G. A. Carpenter, and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Computer vision, graphics, and image processing*, vol. 37, no. 1, pp. 54-115, 1987.

[7] A. M. Treisman, and G. Gelade, "A feature-integration theory of attention," *Cognitive psychology*, vol. 12, no. 1, pp. 97-136, 1980.

[8] C. Koch, and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Matters of Intelligence*, pp. 115-141: Springer, 1987.

[9] L. Itti, "Models of bottom-up and top-down visual attention," California Institute of Technology, 2000.

[10] K. Koch, J. McLean, R. Segev, M. A. Freed, M. J. Berry Ii, V. Balasubramanian, and P. Sterling, "How Much the Eye Tells the Brain," *Current Biology*, vol. 16, no. 14, pp. 1428-1434, 7/25/, 2006.

[11] L. Itti, and P. F. Baldi, "Bayesian surprise attracts human attention." pp. 547-554.

[12] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 2, pp. 353-367, 2011.

[13] J. Han, K. N. Ngan, M. Li, and H.-J. Zhang, "Unsupervised extraction of visual attention objects in color images," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 1, pp. 141-145, 2006.

[14] B. C. Ko, and J.-Y. Nam, "Object-of-interest image segmentation based on human attention and semantic region clustering," *JOSA A*, vol. 23, no. 10, pp. 2462-2470, 2006.

[15] M. Donoser, M. Urschler, M. Hirzer, and H. Bischof, "Saliency driven total variation segmentation." pp. 817-824.

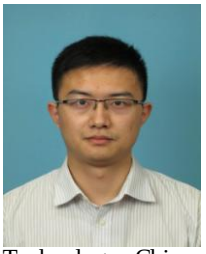
[16] S. Avidan, and A. Shamir, "Seam carving for content-aware image resizing." p. 10.

[17] U. Rutishauser, D. Walther, C. Koch, and P. Perona, "Is bottom-up attention useful for object recognition?." pp. II-37-II-44 Vol. 2.

[18] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection." pp. 409-416.

[19] A. Borji, D. N. Sihite, and L. Itti, "Salient object detection: A benchmark," *Computer Vision–ECCV 2012*, pp. 414-429: Springer, 2012.

[20] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection." pp. 1597-1604.
[21] H. Li, and K. N. Ngan, "A co-saliency model of image pairs," *Image Processing, IEEE Transactions on*, vol. 20, no. 12, pp. 3365-3375, 2011.
[22] M.-M. Cheng, Z. Zhang, W.-Y. Lin, and P. Torr, "BING: Binarized normed gradients for objectness estimation at 300fps."
[23] A. Borji, D. N. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study," *Image Processing, IEEE Transactions on*, vol. 22, no. 1, pp. 55-69, 2013.
[24] T. Judd, F. Durand, and A. Torralba, "A benchmark of computational models of saliency to predict human fixations," 2012.
[25] Y.-F. Ma, and H.-J. Zhang, "Contrast-based image attention analysis by using fuzzy growing." pp. 374-381.
[26] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency." pp. 545-552.
[27] X. Hou, and L. Zhang, "Saliency detection: A spectral residual approach." pp. 1-8.
[28] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, "SUN: A Bayesian framework for saliency using natural statistics," *Journal of Vision*, vol. 8, no. 7, 2008.
[29] Y. Li, X. Hou, C. Koch, J. Rehg, and A. Yuille, "The secrets of salient object segmentation."
[30] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 10, pp. 1915-1926, 2012.
[31] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos," *Computer Vision – ECCV 2010*, pp. 366-379: Springer, 2010.
[32] A. Borji, and L. Itti, "Exploiting local and global patch rarities for saliency detection." pp. 478-485.
[33] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection." pp. 733-740.
[34] B. Alexe, T. Deselaers, and V. Ferrari, "Measuring the objectness of image windows," 2012.
[35] I. Endres, and D. Hoiem, "Category independent object proposals," *Computer Vision–ECCV 2010*, pp. 575-588: Springer, 2010.
[36] J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders, "Selective search for object recognition," *International journal of computer vision*, vol. 104, no. 2, pp. 154-171, 2013.
[37] J. Carreira, and C. Sminchisescu, "Cpmc: Automatic object segmentation using constrained parametric min-cuts," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 7, pp. 1312-1328, 2012.
[38] Z. Zhang, J. Warrell, and P. H. Torr, "Proposal generation for object detection using cascaded ranking SVMs." pp. 1497-1504.
[39] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S.-M. Hu, "Salient object detection and segmentation," *Image*, vol. 2, no. 3, pp. 9, 2011.
[40] M. Wang, J. Konrad, P. Ishwar, K. Jing, and H. Rowley, "Image saliency: From intrinsic to extrinsic context." pp. 417-424.
[41] M.-M. C. J. W. Wen, Y. L. S. Z. V. Vineet, and N. Crook, "Efficient Salient Region Detection with Soft Image Abstraction."
[42] E. Sutton, "Histograms and the Zone System," *Illustrated Photography. Accessed at http://www. illustratedphotography. com/photography-tips/basic/contrast*.
[43] W. Einhäuser, and P. König, "Does luminance-contrast contribute to a saliency map for overt visual attention?," *European Journal of Neuroscience*, vol. 17, no. 5, pp. 1089-1097, 2003.
[44] P. F. Felzenszwalb, and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167-181, 2004.
[45] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, "Automatic salient object segmentation based on context and shape prior." p. 7.
[46] X. Shen, and Y. Wu, "A unified approach to salient object detection via low rank matrix recovery." pp. 853-860.
[47] L. Duan, C. Wu, J. Miao, L. Qing, and Y. Fu, "Visual saliency detection by spatially weighted dissimilarity." pp. 473-480.
[48] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, "Salient region detection and segmentation," *Computer Vision Systems*, pp. 66-75: Springer, 2008.
[49] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 11, pp. 1254-1259, 1998.
[50] Y. Zhai, and M. Shah, "Visual attention detection in video sequences using spatiotemporal cues." pp. 815-824.



Xing Zhang was born in Tangshan, Hebei, China, in 1986. He received the B.S. degree from School of Information Science and Engineering, Southeast University, Nanjing, China, in 2009. He is currently working towards the Ph.D. degree in information and communication engineering in Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China. His research interests include visual attention, bionic vision, SLAM.

Xiaolin Zhang received the B.E. degree in Power Systems from Northeast China Institute of Power Engineering, China, in 1985, and the M.S.E. and Ph.D. degrees from Yokohama National University, Yokohama, Japan, in 1989 and 1995, respectively. From 1989 to 1992, he served as assistant professor in Intelligent Mechanical Laboratory at Yokohama National University, Japan. From 1995 to 2003, He served as assistant professor at Tokyo Medical and Dental University, Japan. Since 2003, he has been associate professor in Precision and Intelligence Laboratory at Tokyo Institute of Technology, Japan. Presently, he is visiting professor at Tokyo Institute of Technology, Japan, and professor in the Internet of Things System Technology Laboratory, Shanghai Institute of Microsystem And Information Technology, Chinese Academy of Sciences, Shanghai, China. His research interests include robot vision, artificial intelligence, medical engineering, virtual reality, image recognition.

Multi-scale and Multi-feature Segmentation of High Resolution Remote Sensing Image

Li Zhao and Hao Fang
School of Computer Science, Wuhan Donghu University, Wuhan, 430212, China

***Abstract*—With the development of the remote sensing technology, high resolution remote sensing images widely penetrates into the common people's life. Traditional medium or low resolution image processing method based on pixel doesn't meet people's requirement any more. In view of it, this paper puts forward a high resolution remote sensing image segmentation method based on the traditional watershed algorithm with multiple scales and characteristics and introduces two new concepts-wrong waterlogging basin and merging array. It firstly improves the immersion process of traditional algorithm, then optimizes the wrong water logging basin and the merging array of basin, adopts the eight immersion neighbors avoiding the wrong division based on four neighbors immersion, optimizes the multilayer immersion process, segments the high resolution remote sensing image with different scales and different features, and finally combines the similar regions. The simulation experiments show that the improved watershed algorithm not only keeps the accurate segmentation effect of traditional algorithm, but also has much higher efficiency.**

***Index Terms*—Remote Sensing Image Segmentation; Multi-Scale and Multi-Feature; Wrong Waterlogging Basin Optimization; Immersion Process Optimization; Improved Watershed Algorithm**

I. INTRODUCTION

Remote sensing image is the major carrier of the remote sensing information. Traditional remote sensing information is extracted mainly through the visual interpretation by experts of medium or low resolution remote sensing image or through computer technology based on pixel [1]. Image segmentation is a key step from image processing to image analysis and understanding so that the segmentation resolution will influence the resolution of subsequent analysis and understanding, and then influence the human behaviour. Therefore, it is of great importance to study the image segmentation, especially the segmentation of high resolution remote sensing image [2].

Many algorithms have emerged in the area of remote sensing image segmentation in recent ten years. In 2005, Guo et al. improved the segmentation algorithm with maximum between-cluster variance, adjusting the profile curve of histogram to choose the optimal segmentation threshold according to the pixel distribution in the image histogram, and effectively setting apart the cross part of object and background [3]. In 2010, De R.L. et al. studied the watershed algorithm and proposed to add the edge information into the watershed so as to detect the boundary of the target object, which improved the boundary location accuracy [4]. In 2011, Chen et al. put forward an improved multi-threshold segmentation method based on simulated annealing to find out multi optimal threshold to segment the remote sensing image [5]. Feng et al. combined the mathematical morphology and graph theory and provided a new segmentation algorithm, firstly splitting out many micro-region with watershed transform and then segmenting these regions with Normalized Cut method from a global perspective which avoids the over segmentation in watershed [6]. Parvathi K and Prakasa P B S provided a pyramid watershed segmentation algorithm based on independent wavelet transform in 2009 [7]. At the same year, Mohamad M A and Ahmad N used fuzzy C-means clustering and self-organizing map to segment the remote sensing image with complex and changing geographical conditions [8]. Li X.F. and other researchers made a breakthrough on the segmentation efficiency by putting forward a fast remote sensing image segmentation method. They adopted a C-V model and wavelet transform which doubled the segmentation speed under the condition of same segmentation accuracy [9]. In 2010, Huang et al. improved the fuzzy entropy algorithm from its noise immunity combining the local information [10]. Liu put forward a new type of remote sensing image segmentation algorithm combining the wavelet transform and Otsu, well segmenting the image in noise [11].

With the development of computer programming language, object-oriented thought was introduced into the remote sensing image segmentation and developed quickly around the world. Especially when a German company Definiens developed an object-oriented image analysis software eCognition, this kind of image processing method was firstly brought to practical use. In 2010, Wang Z W et al. put forward an adaptive object-oriented segmentation algorithm, involving five steps to realize the segmentation of remote sensing images—K-means clustering, segmentation initialization, seed point selection, region growing and merging. Huang Liang made a second development of eCognition software and presented a new segmentation method based on Canny operator. Nakagawa proposed the varying threshold value method, including new concepts, two dimensional grey level histogram and two dimensional maximum fuzzy entropy etc., effectively using the grey information and space information together with fast algorithm to deal with the images that cannot be processed by traditional

doi:10.4304/jmm.9.7.948-954

threshold value segmentation methods. With the development of image segmentation method, the objects of image segmentation become more and more complicated from two dimensional to three dimensional image, from grey images to color images and multispectral images, from static analysis to dynamic tracking, and so on.

However, due to the difficulty of the remote sensing image segmentation itself, all of these segmentation algorithms don't make a huge progress and form a common theory and evaluation standard.

In order to ensure the truth of the image segmentation and avoid the distortion, this paper put forward a multi-scale and multi-feature high resolution remote sensing image segmentation method. It firstly improved the immersion process of traditional algorithm, then optimizes the wrong pondering basin and the array merger of the basin, and segments high resolution remote sensing images in different scales and different features. Theoretical study and experiment results show that this scheme is feasible.

II. STANDARD WATERSHED ALGORITHM

The mathematical model of the simulated immersion algorithm is: Suppose $I(p)$ represents two dimensional grey image function, where p is the pixel point, $D_I \subset Z^2$ is the range of I, $h_{\min}$ is the minimum in the D_I, and $h_{\max}$ is the maximum in the D_I.

Definition 1: $T_h(I)$ is the threshold function, shown as follows:

$$T_h(I) = \{p, I(p) \le h\} \quad (1)$$

Definition 2: If $C(M)$ is basin, M is the minimum grey in the basin, then $C_h(M)$ represents the subset of the $C(M)$ with the following expression:

$$C_h(M) = \{p \in C(M), I(p) \le h\} = C(M) I T_h(I) \quad (2)$$

Definition 3: geodesic distance, if A is a simple connected area, the path P between two points x and y are completely included in, as shown in fig.1, then the geodesic distance is defined as:

$$d_A(x, y) = \inf\{I(P)\} \quad (3)$$

Definition 4: Geodesic influence zone is indicated as $iz_A(B_i)$ where $B_1, B_2, B_3, K\, B_k$ is the connected region included in A, then $iz_A(B_i)$ is the pixel set of the geodesic distance from A to B_i below $B_j (j \in 1:k, j \ne i)$, defined as:

$$iz_A(B_i) = \{p \in A, \forall j \in 1:k, j \ne i, d_A(p, B_i) < d_A(p, B_j)\} \quad (4)$$

Geodesic affected area is shown in fig. 2, $\delta_n(F_k(x,y))2(F_k) = \delta_n(F_k) - \varepsilon_n(F_k)(\delta_n(F_k))2nh+1_{\max}$, the pixels out of any geodesic affected area forms the skeleton of the influence zones, denoted as $SKIZ_A(B)$, with the following expression:

$$SKIZ_A(B) = (A \cap IZ_A(B))^c, IZ_A(B) = \underset{i \in [l:k]}{Y}{}^{iZ_A(B_i)} \quad (5)$$

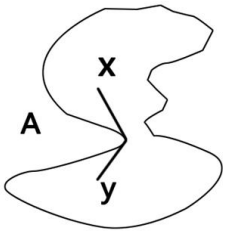


Figure 1. Geodesic distance map

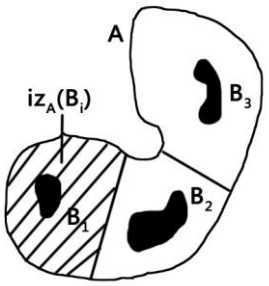


Figure 2. Geodesic affected area

In order to simulate the immersion process, we take the set $T_{h\min}(l)$ to represent the set of the pixel that is flooded at first by the "water", namely the set of the lowest points in the basin and the initial point of our recursion operation. Here, we suppose $X_{h\min} = T_{h\min}(l)$ where $X_{h\min}$ is composed of the pixel points with smallest grey values in the image. It is seen that $X_{h\min} \subset T_{h\min+1}(l)$, if Y is set as the connection part of $T_{h\min+1}(l)$, then there are three possibilities about the relationship between Y and $X_{h\min}$:

- If $Y \cap X_{h\min} = \varnothing$, then it is the bottom of the new basin in the image. For $\forall p \in Y$, from $p \in X_{h\min}$, it is easy to get $I(p) \ge X_{h\min} + 1$; from $p \in Y$, it is easy to get $I(p) \le X_{h\min} + 1$. Therefore, the grey value of all the points around Y that doesn't belong to the $T_{h\min+1}(l)$ is larger than $X_{h\min} + 1$, and we can think Y as a new basin bottom that is pierced and water begin to flood it;
- (2) If $Y \cap X_{h\min} + 1 = \varnothing$ and they are connected, Y is formed by the pixel that has a grey value less than or equal to $X_{h\min} + 1$ in the basin base of $Y \cap X_{h\min}$, namely $Y = C_{h\min+1}(Y \cap X_{h\min})$;
- (3) If $Y \cap X_{h\min} \ne \varnothing$ and they are not connected, then Y is thought to include the basin different from that in the I, named $Z_1, Z_2, K\, Z_k$. Z_i is one of them, and then the basin taking the Z_i as the bottom can be written as following expression with geodesic affected zone, as shown in fig. 3:

$$C_{h\min+1}(Z_i) = iz_Y(Z_i) \quad (6)$$

Based on three possibilities, the third step of the recursion algorithm can be defined as: For different grey value,

Definition 5: Waterlogging basin and watershed, and the waterlogging basin of grey image can be obtained with following steps:

$$X_{h\min} = T_{h\min}(I) \quad (7)$$

$$\forall h \in [h_{\min}, h_{\min} - 1], X_{h+1} = \min_{h+1} \cup IZ_{Th+I(i)}(X_h) \quad (8)$$

Figure 3. Three kinds of diagrams

III. SPEED OPTIMIZATION OF WATERSHED ALGORITHM

A. Improved Immersion Process

Definition 6: Suppose point P_{cur} belongs to the labelled water basin L_{label}. In the improved watershed algorithm, if there is no label point in the neighbourhood of point P_{cur}, then the labelled point P_{cur} is the new minimum area which forms a new water basin. It is apparently impossible to label two different basins with a point P_{cur}, thus this new waterlogging basin is called the wrong water basin.

In fast watershed algorithm, when point P_{cur} belongs to the L_{label}, if P_{cur} is close to the water basin L_{label}, then P_{cur} is merged into L_{label}, otherwise, there are three kinds of spatial relationship between them as shown in fig. 4:

- In fig. 4(a), if water basin L_{label} is on top of or left to the point P_{cur}, following the order from top to bottom and from left to right, point P_{cur} will surely be merged into L_{label}.
- In fig. 4(b), if L_{label} is at the right or bottom of the point P_{cur}, namely P_{cur} will be scanned firstly, from definition (1) it is known that P_{cur} will be labelled as wrong waterlogging basin. Because we have supposed that P_{cur} belongs to L_{label}, learning from the watershed algorithm, there must have a connected path that doesn't be climbed between P_{cur} and L_{label}.
- In fig. 4(c), the situation is similar to (b), if B point belongs to water basin, but due to the discontinuity between B and L_{label}, then B is wrongly labelled as the new minimum zone, thereby a wrong basin is formed around point B. And then P_{cur} will be merged into this wrong basin during sequential scanning.

Among them, (2) and (3) will produce much wrong water basins. We list the identification methods and correction methods of wrong water basins following.

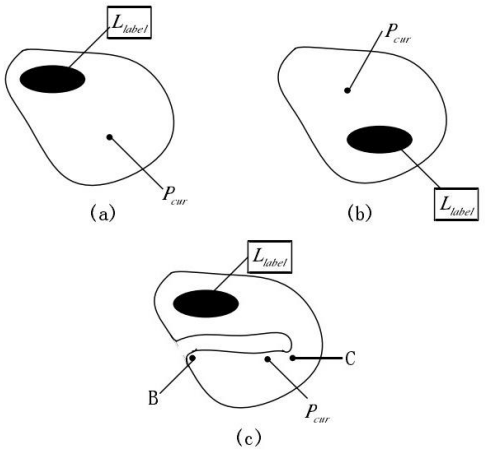


Figure 4. Spatial relationship diagram point and neighbourhood water basins

B. The Optimization of Wrong Waterlogging Basin

After sorting, these pixel points are immersed in a gradient order from low to high. Therefore, before dealing with current gradient order h, all the pixel points lower than h are immersed, namely they are labelled. The number of formed water basin is noted as $N_{\min}$, and a new labelled water basin is marked as $N_{\min}++$. The number which is $\le N_{\min}$ is right, otherwise, it requires to judge, if it is wrong water basin, then it is corrected.

In 4(c), the neighbourhood of point C has two labelled basins, N_{old} and N_{new}, if $N_{old} < N_{new}$, then the relationship and corresponding treatment can be listed into three kinds as follows:

- If $N_{old} \le N_{\min}$ and $N_{new} \le N_{\min}$, it indicates that N_{old} and N_{new} are right labelled, and point C is the watershed.
- If $N_{old} \le N_{\min}$ and $N_{new} > N_{\min}$, it indicates that N_{old} is labelled right, N_{new} is a new labelled water basin, and when they encounter at the point C, they are connected and should be merged into one water basin.

From definition 6, N_{new} is the wrong water basin and should be merged into N_{old}, then N_{new} will be changed by N_{old}. This process is called as the correction process of the wrong water basin.

According to the watershed segmentation principle, if P_{cur} belongs to a certain water basin, then it is thought to be labelled into the wrong water basin. The immersion process always has the similar points like C, thus the correction is necessary.

- If $N_{old} > N_{\min}$ and $N_{new} > N_{\min}$, it indicates that, both of these basin is new labelled water basin, when they meet at the point C where they are connected, because $N_{new} > N_{old}$, from definition 6, N_{new} is wrong water basin, then N_{new} will be merged into the N_{old} to be corrected.

C. Array Merging Optimization of Waterlogging Basin

The correction process of the wrong water basin is a process that adjusts the label of each pixel point to the label of its neighbourhood water basin. In Fig. 5(a), each label number represents a water basin, 6 is firstly merged into 4, and then 4 is merged into 1, which means the label numbers of all the pixel points in basin 6 are corrected to 4 and then 1. The label number in 6 are corrected twice. According to definition 6, the immersion process will produce much wrong water basin. Therefore, if we correct the wrong water basin in time when checking a wrong one and change its label number, the merging efficiency will significantly low.

However, in the improved watershed algorithm, it doesn't correct the wrong water basin timely but uses a merging array to record these water basins. The step is as follows:

In the merging array $a[i] = j$, i is the label number of the wrong water basin, j is the label number of neighbourhood water basin, and equation means merging the i into j. After immersing all the gradient order, according to the record of the merging array, the final label number of the wrong water basin before merging is obtained.

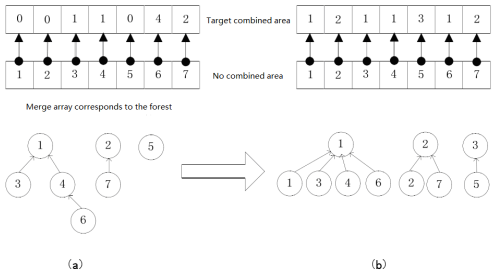


Figure 5. Figure merge array merger process

In Fig. 5(b), 6 is merged into 1 one-time while it requires twice to merge 6 into 1 in Fig. 5(a). The merging process is shown in fig. 5:

- Using merging array to record the correction process of each wrong water basin, as shown in fig. 5(a).
- Obtaining the final merging basin of each wrong water basin, and merging them according to the record condition of merging array.
- Refreshing the label number of all the pixel points in each water basin. The label number of the wrong water basin is useless after merging, and should be compressed to make the label number continuous in the zone after segmentation by watershed. As shown in fig. 5(b), the label number of 5 is changed to 3 so as to be continuous to the 1and 2.

IV. Immersion Optimization of Watershed Algorithm

A. Immersion Neighboring Optimization

The improved watershed algorithm mentioned above only searches four neighboring points, up, down, left and right point during the pixel immersion process, ignoring four points in the direction of opposite angles. In fig. 7, point P_{cur} only searches the 1, 2, 3, and 4 which are not labeled while the point C of direction of opposite angles has been labeled into the waterlogging basin N. Actually, point P_{cur} is close to N, but in the immersion process of four regions, point C is seen as the non –neighboring point, and P_{cur} will not combined into N, that is, P_{cur} is in the wrong basin.

This wrong waterlogging basin is not the one defined by the fast watershed algorithm. It cannot be corrected.

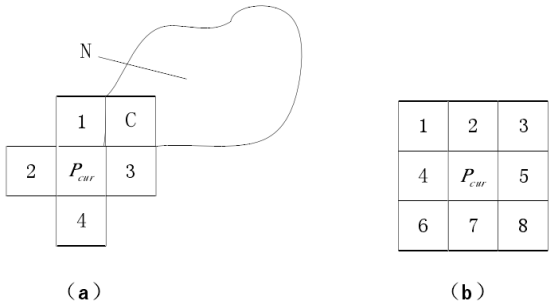


Figure 6. The current domain with eight points neighbors neighborhood graph

It is seen that the immersion based on four neighbors will divide the neighboring and similar points into different regions so as to generate wrong watershed algorithm and divide one region in opposite anlges to two regions. All these contribute to the increase of segmentation region numbers.

When the four neighbors mode are changed to eight neighbors mode, each point in eight neighbors of current point is searched in a number order in fig. 7. The point P_{cur} be divided into the waterlogging basin correctly as long as the neighbor of P_{cur} exists the labeled points, which avoids the wrong immersion partition based on four neighbors.

B. Immersion Hierachical Optimization

In standard watershed algorithm, the immersion process is usually the layer-by layer immersion that immerses a pixel point of a gradient one time. This way will divide the neighboring regions with many similar characteristics into large number of small regions. To improve the over-segmentation in the immersion process, this paper put forward the multilayer immersion based on the improved watershed algorithm.

The layer by layer immersion of standard watershed algorithm will split out two waterlogging basins A and

B, and W is the watershed line between them. In the double layer immersion, A and B will be combined to a waterlogging basin, which means two similar neighbors are merged into one.

Suppose N_h is the number of pixels of layer h through layer by layer immersion. In multilayer immersion, if the layer height is set as Num, then N_h is defined as,

$$N_h = \{p \in D \mid h \le I(p) \le h + Num\}, Num = 0,1,... \quad (9)$$

In (9), $Num = 0$ means the layer by layer immersion, $Num = 1$ means double layers immersion of h and $h+1$, and $Num = n$ means the immersion of layer s h, $h+1$, … $h+n$.

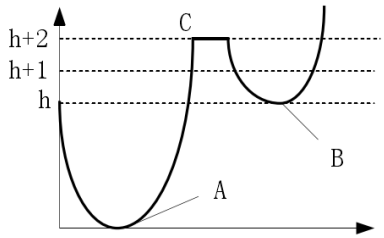


Figure 7. Immersion process analysis

In fig. 8, after layer by layer immersion, B must be labeled as a new minimum region, but when the layer height is three or more, B will be regulated to the region of A.

According to the improved watershed principle, the gradient level of neighbors of B is higher than that of B in layer by layer immersion, thus B must be labeled as the minimum region and C forms the watershed of A and B. The difference between single layer immersion and multilayer immersion is determined by the labeled region number after immersion which is noted as $N_{\min}$. In fig. 8, if the labeled region number of last immersion before B is $N_{\min}$, then the value of B must be larger than $N_{\min}$, and this is the new waterlogging basin. If the immersion process has a condition of C that have two different neighbors. According to definition 1, B is the wrong waterlogging basin. When C is scanned, wrong basin is collected, namely B is combined into old basin A.

V. Merging Optimization of the Remote Sensing Image After Segmentation

A. Color Space Transformation

The transformation from remote sensing image to XYZ color space means mapping the pixel points (R,G,B) into (X,Y,Z).

$$\begin{cases} X = \text{var\_}R \cdot 0.4124 + \text{var\_}G \cdot 0.3576 + \text{var\_}B \cdot 0.1805 \\ Y = \text{var\_}R \cdot 0.2126 + \text{var\_}G \cdot 0.7152 + \text{var\_}B \cdot 0.0722 \\ Z = \text{var\_}R \cdot 0.0193 + \text{var\_}G \cdot 0.1192 + \text{var\_}B \cdot 0.9505 \end{cases} \quad (10)$$

In (10), var_R is transformed through R,

$$\text{var\_}R = \begin{cases} ((R/255 + 0.055)/1.055)^{2.4} \cdot 100, R > 255 \cdot 0.04045 \\ 12.92 \cdot R/255 \cdot 100, R \le 255 \cdot 0.04045 \end{cases} \quad (11)$$

The situation is same to var_G and var_B.

The transformation from XYZ to LUV color space is shown as follows,

$$L = \begin{cases} 116 \cdot (Y/Y_n)^{1/3} - 16, (Y/Y_n) > (6/29)^3 \\ (29/3)^3 (Y/Y_n), (Y/Y_n) \le (6/29)^3 \end{cases} \quad (12)$$

$$U = 13 \cdot L \cdot (u' - u'_n) \quad (13)$$

$$V = 13 \cdot L \cdot (v' - v'_n) \quad (14)$$

In (14), u' and v' is,

$$u' = (4 \cdot X)/(X + 15 \cdot Y + 3 \cdot Z) \quad (15)$$

$$v' = (9 \cdot Y)/(X + 15 \cdot Y + 3 \cdot Z) \quad (16)$$

Y_n, u'_n and v'_n is the white color point at the center of the chromaticity coordinates, respectively valued by, $Y_n = 100$, $u'_n = 0.2009$ and $v'_n = 0.4610$.

B. Similar Region Merging of Segemented Images

There still exists over-segmentation after the image segmentation with improved fast watershed algorithm. Due to the uniformity of the inner pixel points in each waterlogging basin, the neighboring basins have a certain of similarity or a larger diversity. Therefore, an appropriate region merging rules can merge the regions with higher similarity so as to solve the over-segmentation problems of the watershed. This paper combines the mean chroma value information of the remote sensing image and gradient information of two sides to calculate the neighboring region similarity.

When the mean chroma value is close, the chromacity of two sides of the watershed must be as little different as possible to merge.

The equation of the neighboring region similarity is determined,

$$D(i,j) = C(i,j) \cdot E(i,j) \quad (17)$$

In (17), $C(i,j)$ represents the chroma similarity, defined as,

$$C(i,j) = \sqrt{(L_i - L_j)^2 + (U_i - U_j)^2 + (V_i - V_j)^2} \quad (18)$$

Here, L_i, U_i and V_i is the mean value of the three components $L,UandV$ of all the pixels in region i while L_j, U_j and V_j is the mean value in region j.

$E(i,j)$ in (17) indicates the similarity of the gradient mean value, defined as follows,

$$E(i,j) = \mid \eta_i - \eta_j \mid \quad (19)$$

η_i and η_j nis the gradient mean value of the pixels at the sides of the watershed between two neighboring regions. The more apparent edges the watershed has, the more η_i and η_j differs, that is, the higher $E(i, j)$ will be.

C. Merging Optimization Based on Spatial Pattern Clustering

The merging method of the similar regions above ignores the space region information, which will contribute to the wrong region segmentation. Therefore, this paper utilizes an algorithm based on fuzzy clustering of spatial pattern to realize the merging optimization.

R_i is defined as the region i, and the distance between pixel x_k and clustering center v_i is defined as,

$$d(x_k, v_i) = (d_{ik}^E / N + d_{ik}^S)^2 \tag{20}$$

where d_{ik}^E is the Euclidean distance between them ($d_{ik}^E = \| x_k - v_i \|_2$, and $|...\|_2$ is two norm), d_{ik}^S is the spatial distance, N is grey level, $N = 256$.

d_{ik}^S is determined by the membership of neighboring pixel with following expression. Before calculating the space distance, the membership of each pixel must be determined.

$$d_{ik}^S = 1 - \frac{\sum_{t \in \eta_s} u_{rt} \beta_t}{\sum_{c=1}^{c} \sum_{t \in \eta_s} u_{rt} \beta_t} \tag{21}$$

where c is the expected cluster number; β_t is the contribution factor of neighboring pixel; u_{rt} is the membership of pixel S to the r cluster V_r, $0 \le u_{rt} \le 1$.

According to (21), we acquire the distance $d(x_k, v_i)$ from pixel to cluster center, then the objective function of cluster method based on spatial pattern is described as follows,

$$J_m(U,V) = \sum_{k=1}^{m} \sum_{i=1}^{c} u_{ik}^m d^2(x_k, v_i) \tag{22}$$

Here, $\sum_{k=1}^{c} u_{ik} = 1, 1 \le k \le n$. n is the sample number of cluster space; c is the type of cluster; u_{ik} is the membership of pixel x_k; m is the weighted factor of membership; $d(x_k, v_i)$ is the distance between pixel x_k and cluster center v_i. Therefore, the membership function based on these distance features is obtained in following expression.

$$u_{ik} = \frac{[d(x_k, v_i)]^{-\frac{2}{m-1}}}{\sum_{j=1}^{c} [d(x_k, v_i)]^{-\frac{2}{m-1}}} \tag{23}$$

VI. SIMULATION EXPERIMENT

In this paper, all the experiments were conducted on the computer PC P4 T2310, with1.86Gmemory, 2GRAM, Inte182865G graphics card. The experiment environment is MAT-LAB6. 5, the image 256 * 256, grayscale256 and clustering number 2. Both of the traditional watershed algorithm and improved watershed algorithm are used and compared for image segmentation in two groups. The experimental results are as follows:

TABLE I. TRADITIONAL WATERSHED ALGORITHM AND IMPROVED WATERSHED ALGORITHM COMPARISON

Image Name	Algorithm	Sort by time	Immersion time	Total time	Split area
Fig1	Traditional algorithms	15ms	235 ms	250 ms	7200
	Improved Algorithm	15ms	15 ms	30 ms	7202
Fig 2	Traditional algorithms	16ms	188 ms	204 ms	6831
	Improved Algorithm	16ms	15 ms	31 ms	6832
Fig 3	Traditional algorithms	15ms	297 ms	312 ms	4292
	Improved Algorithm	15ms	15 ms	30 ms	4292

In Table 1, the sorting time of two algorithms are exactly the same, but the immersion time is largely different. The immersion time of improved watershed algorithm on the three images is only 15ms, but the immersion time of the traditional algorithm is up to about 200ms, and fig. 3 is even close to 300ms. The total time is the sum of the sorting time and immersion time, and the total time of improved watershed algorithm is 30ms while that of the traditional algorithm is about 300ms. The number of split area of the two algorithms are basically the same: Fig. 3 has the same number, fig. 1 differs by two, and fig. 2 differs by one.

Therefore, it is concluded that improved watershed algorithm not only maintains accurate segmentation results of the traditional algorithm, but also has much higher efficiency. In addition, the immersion time of improved watershed algorithm is proportional to the image size.

Then we applied the improved watershed algorithm into high resolution remote sensing image. The original image is shown in fig. 9. The segmented result with standard watershed algorithm is shown in fig. 10. Figure 11 shows the result of our improved watershed algorithm.

Figure 8. Original image

Figure 9. Standard watershed segmentation algorithm results

Figure 10. Improved watershed segmentation algorithm results

From these three figures, it is evident that our imporve watershed algorihtm has a more preicious segmentation result on high resoultion remote sensing images than standard watershed algorithms.

VII. CONCLUSIONS

Image segmentation is one of the key parts in the computer vision and image processing. This paper put forward a multi-scale and multi-feature high resolution remote sensing image segmentation method, improved the identification method and correction method of the wrong waterlogging basin in the watershed algorithm, divided the merging conditions of the neighbouring waterlogging basins with different labels into three kinds, and finally realized and tested this algorithm. Experiment results show that this algorithm makes the most use of the texture information and multistate information and obtains good segmentation effect, which provides favourable foundation for the subsequent remote sensing image information extraction.

REFERENCES

[1] Li Nan, Huo Hong, Fang Tao, "Adatpitve multiscale Intergrated Remote Sensing Image Segmetation Method", *Computer Engineering*, vol.38, pp. 208-210, 2012.

[2] Cui Bingge, Meng Aoxiang, "Fast Remote Sensing Image Segmentation Algorithm Based on Nearest Neighbor Direct Graph ", *Computer Science*, vol.40, pp. 274-278, 2013.

[3] Guo Jianxing, Liu Songlin, Ni Li, "An improved Image Segmentation Algorithm Based on the Otsu Method", *Chinese Journal of Scientific Instrument*, vol. 26, pp. 665-666, 2005.

[4] De R L,Gui F Z, Zhao C W et al., "An Edge Embedded Marker-Based Watershed Algorithm for High Spatial Resolution Remote Sensing Image Segmentation", *IEEE Transactions on Image Processing*, vol. 19, pp. 2781-2787, 2010.

[5] Chen Ming, Wang Xingfeng, "Improved Multithreshold Remote Sensing Image Segmentation Based on Simulated Annealing ", *Information System Engineering*, vol. 12, pp. 73-75, 2011.

[6] Feng Lin, Sun Tao, "Image Segmentation Method Based on Watershed Transformation and Graph Theory", *Chinese Journal of Scientific Instrument*, vol. 29, pp. 649-653, 2008.

[7] Parvathi K, Prakasa R B S , Mariya D M et al, "Pyramidal Watershed Segmentation Algorithm for High-Resolution Remote Sensing Images Using Discrete Wavelet Transforms ", *Discrete dynamics in nature and society*, vol. 2, pp. 1-11, 2009.

[8] Mohamad M A, Ahmad N., "Satellite Image Segmentation Using Self-Organizing Maps and Fuzzy C-Means", *IEEE International Symposium on Signal Processing and Information Technology*, pp. 398-402, 2009.

[9] Li Xiaofeng, Zhang Shuqing, Liu Qiang, "Fast Segmentation Method of High Resoultion Remote Sensing Image", *Journal of Infrared and Millimeter Waves*, vol. 28 , pp.146-150, 2009.

[10] Huang N. N., He Di, Yang Jie, "A new algorithm for remote sensing image segmentation based on the combination of the improved fuzzy entropy and local information", *Laser Journal*, vol. 31, pp. 20-22, 2010.

[11] Liu Wenjing, Jia Zhenhong, "Noisy Remote Sense Image Segmentaiton Algorithm Based on Wavelet Packet and Otsu", *Computer Engineering*, vol. 37, pp. 203-204, 2011.

A Fast Fractal Coding Method for Image with Primary Additional Errors

Shuai Liu
College of Computer Science, Inner Mongolia University, Hohhot, China
School of Physical Science and Technology, Inner Mongolia University, Hohhot, China
*Corresponding author, Email: cs_liushuai@imu.edu.cn

Mengxi Liu, Qi Jia, Lingyun Qi, and Haipeng Li
College of Computer Science, Inner Mongolia University, Hohhot, China
Email: wn_fu@sohu.com

***Abstract*—Today, in the multimedia encoding technology, fractal image coding is an effective coding method without resolution. The effectiveness is because of the high compressing ratio of fractal image coding. But the computational complexity of this coding method is so high that it needs long encoding time. In this paper, a novel fast fractal coding method is constructed to decrease the coding time by the capture of primary additional error values. This method is a universal algorithm, which is independent of image types. First, we abstract the additional error values from classic image coding. Then, we present a method to abstract the primary error values with a given rule of weight. Moreover, the encoding and decoding processes are reformed to store the primary additional error values. Finally, experimental results shows the improved fractal image coding method has higher compressing ratio and better effectiveness (signal to noise ratio) than the classic algorithm.**

***Index Terms*—Fractal Coding; Image Coding; Primary Additional Error; Compressing Ratio; Signal to Noise Ratio**

I. INTRODUCTION

Today, multimedia is used everywhere in the human society. But one bottleneck of multimedia improvement is that its large size needs more space to apply. In order to decrease the size of multimedia, many encoding method are presented. In this way, image coding, which is basis of multimedia coding, is a highlight in this domain.

Nowadays, there are many image encoding methods, such as discrete cosine transform (DCT) [1], Huffman code [2], wavelet image coding [3], etc. Also, there are many international standards, which have been presented by these coding methods, such as BIG, JPEG, H.263, MPEG, etc.

However, the basic ideas of these image coding methods are similar so that the compressing ratios are also similarly [4, 5]. So we need a novel thinking to code images with higher compressing ratio because of the larger images. In this way, fractal image coding is created by the self-similar in nature.

In real world, the geometrical form can be classified as two kinds. One is regular and smooth, which can be described by traditional geometry. Contrarily, the other is rough and anomalistic, which can't be described by traditional geometry. Besides, the natural objects usually have rough and anomalistic forms. So a novel subject is created to research in this domain, which is called fractal geometry [6, 7]. Admittedly, there are much natural scenery is fractal, coastline, mountain's shape, stream, tree, lightning, etc. Nowadays, when fractal theory is combined with computer technology, it becomes an interdisciplinary and nonlinear subject. Fractal images coding is such a technology in this subject, which depends the fractal geometrical form in the images.

Fractal image coding technology is based on the local self-similar of natural images. It uses contractive affine transformation (CAT) to iterate a created image to the coded image. In fractal coding technology, we only need to store the quantization parameters of CAT, whose size is much smaller than the original image. In this case, it can reach the image compression with higher compressing ratio.

After encoding, decoding is also a novel fast iterating process. The existence and uniqueness of the iteration are proved by Banach fixed point theorem. It means that we can use any image to iterate the original image, and the original image is limit of the iteration. So the main problem in fractal coding is to find the best approximation of CAT.

So we find that the fractal coding is a finite distortion method. It means that we can decrease coding time when we decrease part of quality of the coding image. However, this method can be applied when the real-time is required without quality. But in many domains, both real-time and qualities of images are required. So researchers have to find faster fractal coding methods.

In this paper, we decrease coding time by used larger element of the coding set. In other word, the search space changes to small. In additional, we exact primary errors to store in iterating space in order to compensate decoding image. In this case, the fast fractal coding method both has better coding time and coding quality.

The remainder of the paper is organized as follows. We achieve the additional error value from classic image

doi:10.4304/jmm.9.7.955-962

coding, and abstract the primary error value in Section 2. Then, we present the reformed coding process and structure in Section 3. It needs to be reformed because of the storage of primary errors. Moreover, experimental results are presented and analyzed in Section 4. Finally, Section 5 summarizes the whole paper.

II. Related Works

First, in year 1988, Barnsley and Jacquin used iterated function system (IFS) and Recurrent Iterated Function System (RIFS) to code some images [8]. They found that the highest compressing ratio is more than 10000:1. In year 1992, Jacquin achieved self-adaptive fractal coding method in computer [9]. This is the symbol of fractal coding, and means the generation of fractal coding. Meantime, Monro and Dudbridge presented another fractal coding method by used fractal block [10]. Then, Bedford et al extended research of Jacquin, and presented a fractal coding method for monochrome images [11]. Later, Kim and Park presented a coding method of still image [12]. They reach the method by fractal approximation of the image. Kim et al used fractal coding into video sequence [13]. Their research focused on the mapping and non-contractive interframe mapping in the video. Chang and Kuo presented an iteration-free fractal image coding method [14]. Their work is based on the designed domain pool.

After year 2000, due to the requirement of multimedia, especially the images, fractal image coding developed more rapidly than before. Many researchers studied many methods to increase the coding rate and decrease the coding time. In 2002, Li et al used fuzzy image metric into fractal coding [15]. Lai etc presented a fast fractal image coding with kick-out and zero contrast conditions [16]. Belloulata used a non-iterative block clustering to code subbands [17]. Wang etc researched into no-search fractal image coding with a modified gray level transform and fitting plane [18, 19]. His team also paid attention to fractal coding with wavelet transform [20-21]. Meanwhile, Lu etc studied Huber fractal coding with fitting plane [22]. Bhayani and Thanushkodi compared fractal coding methods of medical image compression [23]. Later, there were also some coding methods with fractals. Our team reached some results in this area. We researched fractal properties in k-M set and use it into facial capture [24-25].

Fractal image coding has many advantages, such as high compressing ratio and independence with resolution. However, the coding has high complexity of computation and long coding time. It is because fractal coding needs to find the best matching block for every input sub-block in a large matching set. Admittedly, the result of best matching block can be found by global search of whole matching set. But the computational cost is too high to apply. Though there are many methods are presented to solve this problem, these methods are all for special application (type of image). So in this paper, a universal method is presented to fit all image types when the image can be transformed to a matrix.

III. Primary Additional Error Value

First, we present the steps of a fractal image coding method in Fig. 1.

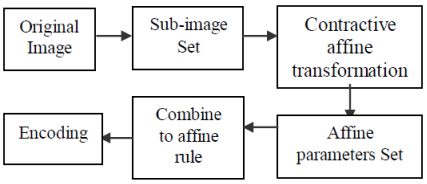


Figure 1. Processes of fractal image coding

Then, we present fractal image coding method step by step in Alg. 1.

Algorithm 1. Fractal Coding Method

Input. Original Image I(n×n), where n is the size of image I; c is the size of each element in coding set (c|n); f is the size of each element in affine set ($f < c$).

Output. Encoding file F.

Step 1.

To find a partition I_r of I ($r = 1 \ldots d, d = \left(\frac{n}{c}\right)^2$), where $\forall i,j|i \neq j \rightarrow I_i \cap I_j = \emptyset$ and $\cup_{1 \leq i \leq d} I_i = I$ are all true. The size of I_r is c.

Step 2.

To divide I to D_k. The size of D_k is f.

Step 3.

To select the best affine transformation from all D_k for each I_r. Then, to store the affine transforming table as encoding file to output.

Alg. 1 finished.

In order to corresponding to Alg. 1, we have decoding method in Alg. 2.

Algorithm 2. Fractal Decoding Method

Input. Encoding file F, which contains s, o, (x, y), direction where s is scaling of luminance, o is offset of luminance, (x, y) is affine starting position, direction has eight values (1-8) and denotes the types of equilong transformations; iteration time T; c and f are same to Alg.1.

Output. Decoding image D.

Step 1.

For each rectangular area Di with size f as a decoding area, to get the corresponding s and o as the affine transforming parameters of this area.

Step 2.

For every decoding area Di, to get random image Ri with size c as affine image, then affine mapping it to an image with size f by corresponding parameters s and o.

Step 3.

To collage all rectangular areas to an image D'. Then, let R_i = corresponding part of D', iterating step 1-3 until iterating time =T.

Step 4.

To output D' as D.

Alg. 2 finished.

Admittedly, there exist errors between D and I for any affine transformation. It is because that Eq. 2 is applied to instead of Eq. 1 in real application.

$$\frac{1}{B^2}\min_j\{\min_{s,o\in R,|s|<1}\left\|R_i - \left(s \cdot D_j + o \cdot E\right)\right\|^2\} \quad (1)$$

$$\min_j\{\min_{s,o\in R}\|R_i - (s \cdot D_j + o \cdot 1)\|\} = \min_{s,o\in R}\left\|R_i - \left(s_i \cdot D_{m_i} + o_i \cdot E\right)\right\| \quad (2)$$

In Eqs. 1-2, B is the grey level, E is the identity matrix, $\| \|$ is the vector norm which is usually 2-norm in fractal coding. R_i is each affine sub- image, D_j is each pattern sub-image, mi is serial number of the best D_j, si is scaling of R_i and o_i is offset of R_i in affine mapping.

In this way, we can extract additional error $\Delta_{n\times n}$ from Eq. 3.

$$\Delta_{n\times n} = I_{n\times n} - \bigcup_{1\le j\le d} D_j \quad (3)$$

For a general nature image, the value zero in Δ is usually a little. But many values are small. Realistically, to ensure the quality of the image by visual angle, we don't interest in those small values in Δ. Meanwhile, in order to drop the blocking effect in the decoding image, we have to store those error values Δp and Δq that Δp locates the edge e_p of D_p, Δq locates the edge e_q of D_q and e_p=e_q. In this case, we have an evaluation standard to extract the valuable errors from Eq. 4. Then we use and named them "additional error values".

$$w_{i,j} = \Delta_{i,j} + \sum_{k=1}^{5} u_k \cdot \frac{1}{k-1} \sum_{t_1+t_2=k, t_1,t_2\ge 0} \left(\left|w_{i,j} - w_{i-t_1,j+t_2}\right|\right) \quad (4)$$

In Eq. 4, u_k is penalty point which is smaller when k is larger. In our method, we don't think that there is blocking effect when k>5. Then, with the sequence of all w_{ij} by their values, we extract the primary additional error values. The number of the primary additional error values is d/2, which can be stored in the encoding file with same size. Then, we present a fast fractal coding method in the following Section.

IV. A Fast Fractal Coding Method with Primary Additional Error Value

Since we have extracted the additional error values, the quality of the decoding image is high enough. However, the coding time is still large. So we can use a higher R and D with same c/f to decrease coding time. Then we can compare the fast coding method to the classic one.

Assuming that variable parameters of the classic method are c and f for a static image with n size, we use c1=2c and f1=2f as the parameters of ours. Then, we can compute the compressing ratio of the two methods by Eqs. 5-6, and the coding time of the two methods by Eqs. 7-8 when assuming the time of one affine mapping is t.

In these equations,

$$CR_{classic} = \frac{n^2}{6\cdot\left(\frac{n}{f}\right)^2} = \frac{f^2}{6} \quad (5)$$

$$CR_{fast} = \frac{n^2}{6\cdot\left(\frac{n}{2f}\right)^2 + 8d^2} = \frac{f^2}{8} = \frac{3}{4} CR_{classic} \quad (6)$$

$$T_{classic} = t\cdot\left(\frac{n}{c}\right)^2\left(\frac{n}{f}\right)^2 \quad (7)$$

$$T_{fast} = t\cdot\left(\frac{n}{2c}\right)^2\left(\frac{n}{2f}\right)^2 = \frac{1}{16} T_{classic} \quad (8)$$

In this case, we know that we have both higher compressing ratio and smaller compressing time with the improved algorithm. Then, we present the encoding and decoding algorithms in Algs. 3-4.

Algorithm 3. Fast Fractal Coding Method
Input. Same to Alg. 1.
Output. Same to Alg.1
Step 1.
Encoding F by Alg.1 with parameters c_1 and f_1.
Step 2.
Extracting primary additional error values with Eq. 4 for Δ. The number of primary additional errors is $8d^2$. Its coding rule likes $P = \{p_{xy}\}$, where (x,y) is the position of errors and pxy is the value of the error.
Step 3.
Attaching $V = \{v_{xy}\}$ to F where $v_{xy} = x\cdot 2^{16} + y\cdot 2^8 + p_{xy}$ after rewrite original position (e_x, e_y) to $e = e_x\cdot 2^8 + e_y$. Then, F with addition V is the output.
Alg. 3 finished.

Algorithm 4. Fast Fractal Decoding Method
Input. F with additional V.
Output. Same to Alg. 2.
Step 1.
Extracting $P = \{p_{xy}\}$ where $p_{xy} = v_{xy} \bmod 2^8$, $y = \left(\frac{v_{xy}-p_{xy}}{2^8}\right) \bmod 2^8$, $x = \frac{v_{xy}-p_{xy}-y\cdot 2^8}{2^{16}}$. Extracting (ex, ey) with $e_y = e \bmod 2^8$ and $e_x = \frac{e-e_y}{2^8}$.
Step 2.
Decoding D with same process of Alg. 2.
Step 3.
Add P to D by used $d_{xy} = d_{xy} + p_{xy}$. D is the output.
Alg. 4 finished.

Because the scale of F is $F_c = (n/2c)^2 = d/4$, we have that $d/2 = 2F_c$. It means we have two additional dimensions to attach these values. In additional, we rewrite (e_x, e_y) to e, which economizes one dimension. Thus, we drop value of the 6th dimension. So the total scale doesn't increase.

Furthermore, in the decoding period, we spend additional time to extract V and e. But these are only basic computation, which spends only a little computational time.

V. Experiment and Analysis

In this paper, we use four images as the examples. In these examples, the 1st is classic "Lena", the 2nd is named "building", the 3rd is named "bride", and the 4th is named "windmill". They are all reformed to greyscale images with size 256×256 and presented in Fig. 2. In Fig. 2, we name these figures in following. Fig. 2a is named as "Lena", Fig. 2b is named as "building", Fig. 2c is named as "bride", and Fig. 2d is named as "windmill".

Figure 2. Original images

Figure 3. Decoding Images with both classic and improved method in image "Lena" (by six iterated time)

Figure 4. Decoding Images with both classic and improved method in image "building"

In our experiments, c=8 and f=4. So c_1=16 and f_1=8. Figs. 3-6 show the comparisons of the decoding images of these two methods. Fig. 3 shows the comparison of the decoding images of image "Lena". Fig. 4 shows the comparison of image "building". Fig. 5 shows the comparison of image "bride". Fig. 6 shows the comparison of image "windmill". In the sub-figure "a" of each figure, we decode the image by the classic method. In the sub-figure "b" of each figure, we decode the image by the improved method.

Figure 5. Decoding Images with both classic and improved method in image "bride" (by six iterated time)

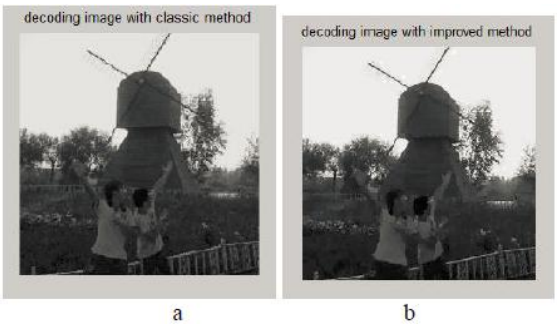


Figure 6. Decoding Images with both classic and improved method in image "windmill" (by six iterated time)

In fact, we can't judge which decoding image is better by used human vision. So we compute PSNR to measure the quality of the two decoding image from Eq. 9.

$$\mathrm{PSNR} = 10 \cdot \log_{10} \frac{n^2 \cdot 255^2}{\sum_{i=1}^{n^2} (I_i - D_i)^2} \tag{9}$$

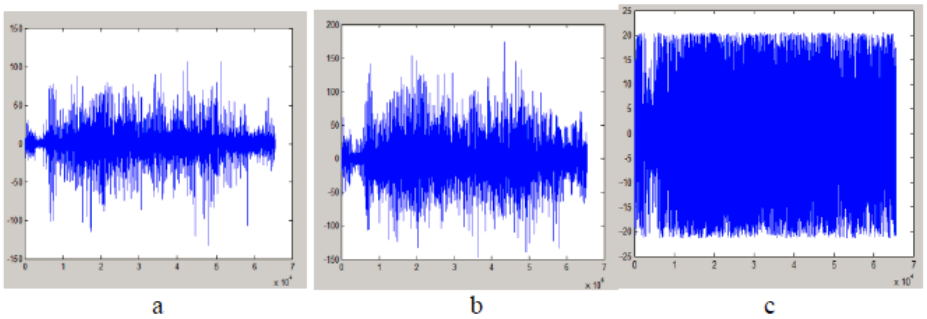


Figure 7. Comparison of Additional error values with both the classic, improved method (without primary additional error values), and improved method (with primary additional error values) in the image "Lena"

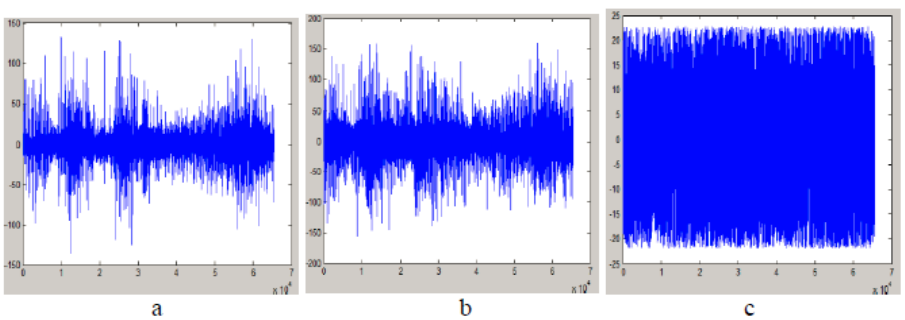
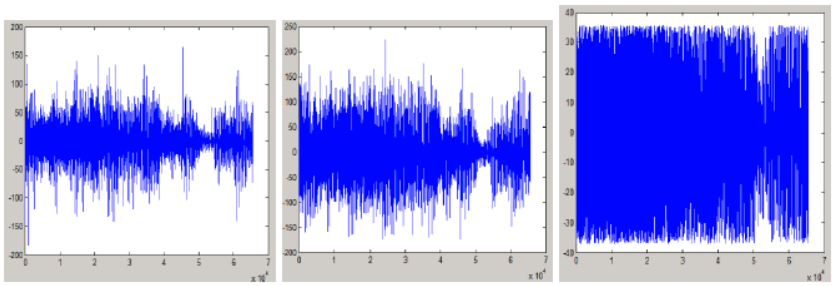
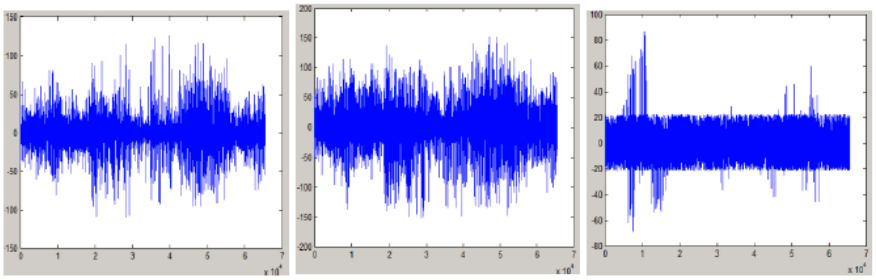


Figure 8. Comparison of Additional error values with both the classic, improved method (without primary additional error values), and improved method (with primary additional error values) in the image "building"



a b c

Figure 9. Comparison of Additional error values with both the classic, improved method (without primary additional error values), and improved method (with primary additional error values) in the image "bride"



a b c

Figure 10. Comparison of Additional error values with both the classic, improved method (without primary additional error values), and improved method (with primary additional error values) in the image "windmill"

Figure 11. Decoding images with each iterated time (1-5) for image "Lena"

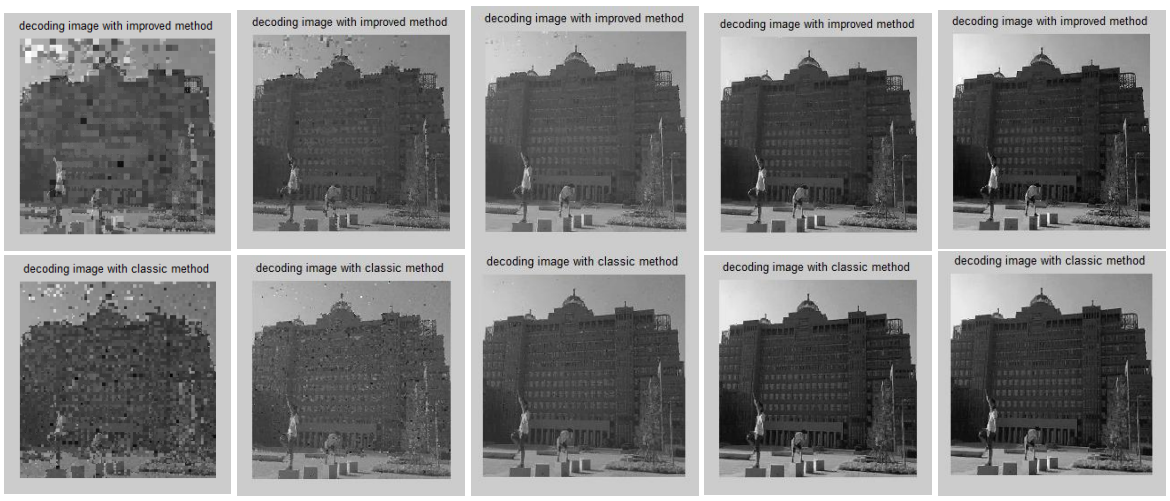


Figure 12. Decoding images with each iterated time (1-5) for image "building"

Then, we provide Figs. 7-10 to show the comparisons of additional error values between the classic method and the improved method in these four images. Fig. 7 shows the comparison of the additional error values with these two methods in image "Lena". Fig. 8 shows the comparison of image "building". Fig. 9 shows the comparison of image "bride". Fig. 10 shows the comparison of image "windmill".

Similar to Figs. 3-6, in Figs. 7-10, the sub-figure "a" of each figure shows the additional error values with the classic method, the sub-figure "b" of each figure shows the additional error values of the improved method without primary additional error values, and the sub-figure "c" of each figure shows the additional error values of the improved method with primary additional error values. First, we know that the sizes of the matrix are same because the number of points is 256×256=65536.

In these figures, we have that the additional error values of classic method are lower than the improved method without primary additional error values. It is because that the size of divided blocks in classic method is smaller than the size in improved method (the mean of value is 50~80 and maximum value is 150~200 in the classic method, the mean of value is 80~100 and maximum value is 200~250 in the improved method without primary additional error values).

TABLE I. EXPERIMENTAL RESULTS OF CLASSIC FRACTAL CODING

Image	Coding Time (s)	Decoding Time (s)	PSNR (dB)	Compressing ratio
Lena	1003.496	1.327	30.0933	2.67
Building	1212.999	1.032	29.0667	2.67
Bride	1228.219	1.014	25.9097	2.67
Windmill	1235.142	1.402	28.7373	2.67

However, we also find that the additional error values change to 25~40 per point when the primary additional error values are added. It means that the primary additional error values increase the compressing effect indeed.

TABLE II. EXPERIMENTAL RESULTS OF FAST FRACTAL CODING

Image	Coding Time (s)	Decoding Time (s)	PSNR (dB)	Compressing ratio
Lena	74.134	0.799	31.5946	3.59
Building	80.497	0.973	30.4684	3.59
Bride	81.135	0.924	26.8855	3.59
Windmill	81.775	0.856	31.7594	3.59

Figure 13. Decoding images with each iterated time (1-5) for image "bride"

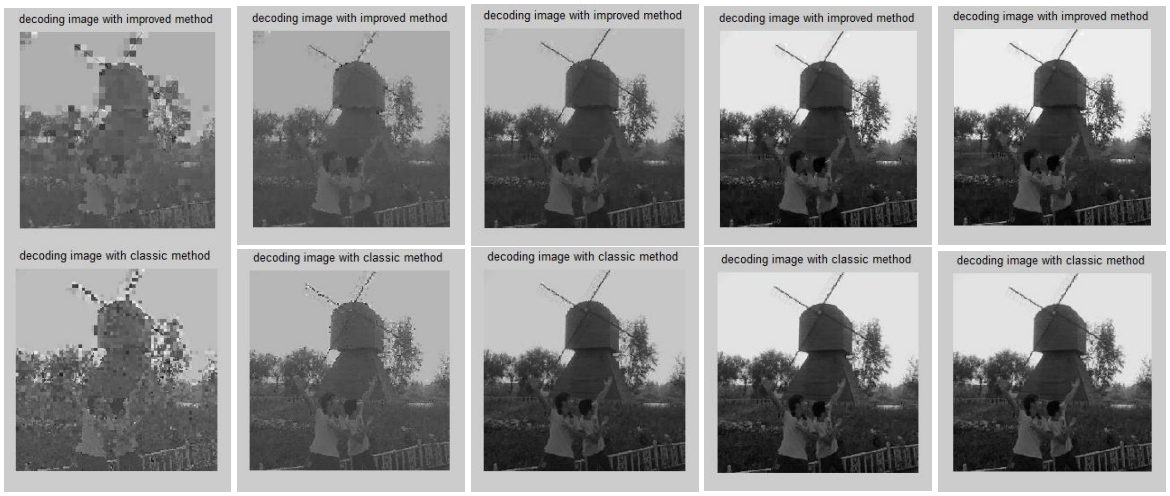


Figure 14. Decoding images with each iterated time (1-5) for image "windmill"

In additional, we have the results of the classic and improved algorithms in tables 1-2. The two tables show the coding time, decoding time, PSNR and compressing ratio of classic algorithm of these two algorithms. In table 1, we find that the coding time is very long (>1000s) when it reaches the well PSNR (>25). Though it has short decoding time, it can not be used in the real world because of the long coding waste. But in table 2, we find both the encoding and decoding time are all well enough to use in application.

Then, we compute the ratio of the two methods and find they are all nearby 16. Concretely, ratio of Lena is 13.54 and others are all between 15 and 16. It validates our conclusion of coding time in Eqs. 5-8. Furthermore, the ratio of "compressing ratio" of the two is 0.744. It also validates our conclusion of compressing ratio in Eqs. 5-8.

In additional, we show the decoding image of the four original images for each iterated time 1-5 in Figs. 11-14. The upper sub-images of each figure are all decoding with the improved algorithm, and the lower sub-images are all with the classic algorithm. In these figures, we have that the decoding image of improved algorithm is better than the classic algorithm when the iterated time is more than three.

VI. Conclusion

In this paper, we present a fast fractal coding method with primary component analysis of additional error value. We extract the primary additional error value and store them into encoding file. We also improve structure of the encoding file in order to store the error value. Moreover, we change the decoding rule of the improved coding method to add the error values to the decoding image. Finally, we use some images to experiment. The experimental results show that the fast fractal coding has higher coding speed and better "signal to noise ratio".

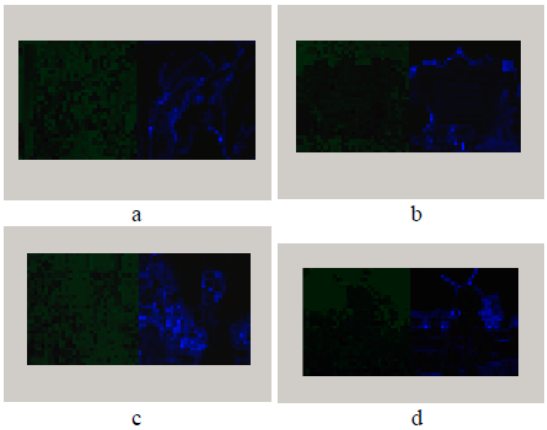


Figure 15. The encoding images of the four original images

In future, our work will pay attention to microcosmic coding rule. We project that we will construct a novel storage rule to decrease space of additional error value. We also project that we will construct a rule of indemnity by fuzzy set. Another area we want to study is to use parallel environment to increase coding time [26-28].

Another work will be the encryption with fractal coding. In Figs. 15, we have presented the encoded images of these four images. Fig. 15a is corresponding to "Lena", Fig. 15b is corresponding to "building", Fig. 15c is corresponding to "bride", and Fig. 15d is corresponding to "windmill". In Fig. 15, we find that the encoding image is so confusing to the original image. So we will use this technology into encryption in future.

Acknowledgment

This work is supported by Grants Programs of Higher-level talents of Inner Mongolia University [No. 125126, 115117], Innovation Training Program of Graduates of Inner Mongolia University [201315226], Scientific projects of higher school of Inner Mongolia [No. NJZY13004], National Natural Science Foundation of China [No.61261019, 61262082].

The authors wish to thank the anonymous reviewers for their helpful comments in reviewing this paper.

References

[1] Rao K R, Yip P. Discrete cosine transform: algorithms, advantages, applications. *Academic press, Boston*, 1990.
[2] Huffman D A. A method for the construction of minimum redundancy codes. *Proceedings of the IRE (IEEE)*, 40(9) (1952) 1098-1101.
[3] Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients. Signal Processing, *IEEE Transactions on*, 41(12) (1993) 3445-3462.
[4] M. Egmont-Petersena, D. de Ridderb and H. Handelsc, Image processing with neural networks -a review, *Pattern Recognition* 35 (2002) 2279–2301
[5] M. Vetterli, Wavelets and Filter Banks: Theory and Design, IEEE Trans. *Signal Processing*, 40(9)(1992) 2207-2232
[6] Mandelbrot BB, The Fractal Geometry of Nature, *Freeman W H, San Francisco*, 1982.
[7] K Falconer. Fractal Geometry: Mathematical Foundations and Applications, *Second Edition*. *John Wiley @ Sons, Inc*, 2003.
[8] M. F. Barnsley and A. E. Jacquin, Application of recurrent iterated function systems to images, in Proceedings of the SPIE, *Visual Communications and Image Processing*, 1001 (1988) 122-131
[9] A. E. Jacquin, Image coding based on a fractal theory of iterated contractive image transformations, *IEEE Trans. Image Process*. 1 (1) (1992) 18-30
[10] Monro D M, Dudbridge F. Fractal block coding of images. *Electronics letters*, 28(11) (1992) 1053-1055.
[11] Bedford T, Dekking F M, Breeuwer M, etc. Fractal coding of monochrome images. Signal processing: *Image communication*, 6(5) (1994) 405-419.
[12] Kim I K, Park R H. Still image coding based on vector quantization and fractal approximation. *Image Processing, IEEE Transactions on*, 15(4) (1996) 587-597.
[13] Kim C S, Kim R C, Lee S U. Fractal coding of video sequence using circular prediction mapping and noncontractive interframe mapping. *Image Processing, IEEE Transactions on*, 7(4) (1998) 601-605.
[14] Chang H T, Kuo C J. Iteration-free fractal image coding based on efficient domain pool design. *Image Processing, IEEE Transactions on*, 9(3) (2000) 329-339.
[15] Li J, Chen G, Chi Z. A fuzzy image metric with application to fractal coding. Image Processing, *IEEE Transactions on*, 11(6) (2002) 636-643.

[16] Lai C M, Lam K M, Siu W C. A fast fractal image coding based on kick-out and zero contrast conditions. *Image Processing, IEEE Transactions on*, 12(11) (2003) 1398-1403.
[17] Belloulata K. Fast fractal coding of subbands using a non-iterative block clustering. *Journal of Visual Communication and Image Representation*, 16(1) (2005) 55-67.
[18] Wang X Y, Wang S G. An improved no-search fractal image coding method based on a modified gray-level transform. *Computers & Graphics*, 32(4) (2008) 445-450.
[19] Wang X Y, Wang Y X, Yun J J. An improved no-search fractal image coding method based on a fitting plane. *Image and Vision Computing*, 28(8) (2010) 1303-1308.
[20] Zhang Y, Wang X. Fractal compression coding based on wavelet transform with diamond search. *Nonlinear Analysis: Real World Applications*, 13(1) (2012) 106-112.
[21] Wang X Y, Zhang D D. Discrete wavelet transform-based simple range classification strategies for fractal image coding. *Nonlinear Dynamics*, 75(3) (2014) 439-448.
[22] Lu J, Ye Z, Zou Y. Huber fractal image coding based on a fitting plane. *Image Processing, IEEE Transactions on*, 22(1) (2013) 134-145.
[23] Bhavani S, Thanushkodi K G. Comparison of fractal coding methods for medical image compression. *IET Image Processing*, 7(7) (2013) 686-693.
[24] Liu S, Cheng X, Lan C, etc, Fractal property of generalized M-set with rational number exponent. *Applied Mathematics and Computation*, 220 (2013), 668-675.
[25] S Liu, W Fu, W Zhao, etc. A Novel Fusion Method by Static and Moving Facial Capture. *Mathematical Problems in Engineering*, (2013) doi:10.1155/2013/503924.
[26] Liu M, Liu S, Fu W, etc, Distributional Escape Time Algorithm based on Generalized Fractal Sets in Cloud Environment. *Chinese Journal of Electronics* (In press).
[27] S Liu, W Fu, H Deng, etc. Distributional Fractal Creating Algorithm in Parallel Environment, *International Journal of Distributed Sensor Networks*, (2013) doi:/10.1155/2013/281707.
[28] G Yang, S Liu, "Distributed Cooperative Algorithm for Set with Negative Integer by Fractal Symmetrical Property," *International Journal of Distributed Sensor Networks*, (2014) doi:10.1155/2014/398583.

Stain Detection in Video with Background Restructured

Jia Wang and Xin Li
Experimental Instrument Center, Dalian Polytechnic University, Dalian, China
Email: wangjia@dlpu.edu.cn

Shuai Liu*
College of Computer Science, Inner Mongolia University, Hohhot, China
*Corresponding author, Email: cs_liushuai@imu.edu.cn

***Abstract*—Nowadays, multimedia, especially video, is used widely in many domains. But camera lens are easily stained because they are always outside, such as in traffic and monitoring. Admittedly, the video from these cameras is out of action because it cannot provide clear frames. In this paper, we provide a detecting method from video with background restructure. This method is presented to detect stains in video. Firstly, the first video frame is used as a pseudo-background because original cameras are clear with no stains. Then, background is constructed by the pseudo-background and the continuous frames which are from the video. Moreover, with training in the frames, we restructure background by dropped moving objects in it. Finally, after background restructured, we detect stains in camera lenses by the restructured background, and provide positions of stained area. Experimental results show its robustness and practicability.**

***Index Terms*—Background Restructured; Background Robustness; Stain Detection; Video Stain; Practicability**

I. INTRODUCTION

Nowadays, with development of information technology, network cameras are used widely in many areas. Today, it becomes an important application in many domains, e.g. military affairs, national defense and so on [1-2]. It uses background and current frame to detect stains on cameras. Nowadays, we always use background subtraction method to detect camera stains [3]. But a major problem in this method is that the camera must recognize the stability of all extra objects in video. It is because that moving objects and stains must be separated. So we must track trajectory of all moving object, and then gain stains in a video. So in this paper, we restructure background with a method fusing background subtraction method and temporal difference method, and then track trajectory and detect stains.

When we use background subtraction method to catch moving objects, we must construct a pixel model of background at first. In this paper, we use temporal difference method to restructure the background by comparing differences in continuous three frames [4]. In fact, its computation amount is small for background subtraction method and its accuracy is well enough for temporal difference method. It has robustness with transformations of illumination, noise and other nature characters.

The threshold of differences in continuous frames is key matter of temporal difference method. Moreover, a moving object is detected when differences are larger than the threshold. This method is robustness with transformations of illumination and shadow. But the threshold needs to select well, otherwise, there will be large void or miss object when threshold is too large and large noise when threshold is too small. However, this method cannot treat those objects which stay at a position for a long time, and it do not process well when movement is not uniform [5-6].

Nowadays, there are lots of study of moving tracking and capture in many areas [7-9]. But there is no suitable way to solve stains in videos. So in this paper, we put forward a new tracking and detection method without origin background. In our opinion, we detect stains in cameras by avoided moving objects with or without movement. So in our method, we use temporal difference method to restructure background at first. Then we detect stains by using background and moving trajectory.

In the following sections, we present our work in the following order.

At first, in Section 2, we present our method and model, which using background restructuring method by temporal difference and mathematical morphology method.

Then, in Section 3, we further track trajectory by using background restructured, and then we find those static trajectories and separate stains where they were moved before.

Moreover, in Section 4, we use experimental result to validate robustness and practicability of our method.

Finally, in Section 5, we summarize this paper.

II. BACKGROUND RESTRUCTURED

A. Temporal Difference Method

In this paper, we use temporal difference method to detect points of background by Eqs. 1-2 and inequation.3. Firstly, we create a function G_{ij} to compute difference for

doi:10.4304/jmm.9.7.963-969

two continuous frames X_{i-1} and X_i. Then we use another function Bi,j to show if there are moving objects in these frames. Next, when applying In(x) as pixel gray of nth fore-frame and Tn(x) as threshold, we consider pixels z is moving when it fits Inequation.3 and not when contrary. So when the moving objects moves between frames, we can train to find all background points. Flow chart of background training is shown in Fig. 1.

$$G_{i,i-1}(x,y) = |X_i(x,y) - X_{i-1}(x,y)| \quad (1)$$

$$B_{i,i-1}(x,y) = \begin{cases} 1, & \text{if } G_{i,i-1}(x,y) \geq \text{threshold} \\ 0, & \text{Otherwise} \end{cases} \quad (2)$$

$$\begin{aligned} |I_n(x) - I_{n-1}(x)| &> T_n(x) \\ |I_n(x) - I_{n-2}(x)| &> T_n(x) \end{aligned} \quad (3)$$

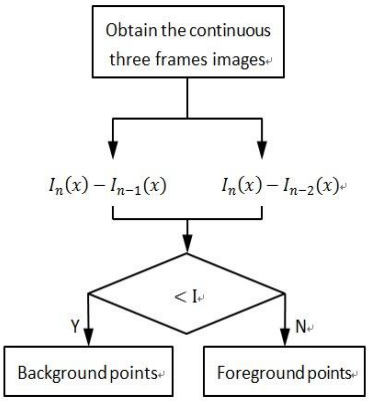


Figure 1. Restructure background with training

B. Restructuring Method of Background

When an object moves, there are some shaded points of background displayed. With our restructuring method, we find these static points and put them into a new frame. This new frame can be called restructuring background. Then, we train frames of a video some time to fill the restructuring background. Admittedly, points of background are changed when there are changes of illumination, wind (mainly outside) and other elements.

As we known, the movement and environment are usually unpredictable. For example, intense movement can bring sharp or fuzzy border and strong illumination and elastic deformation also shows negative detection. It means that this problem is also in the background we restructured. So we need to use some methods to compensate the background we restructured.

In fact, we need to modify its structure when we take clear and complete background. In our work, we use mathematical morphology to modify the background. Specifically, we use erosion to fill small hole and dilation to remove isolated noise. To use them in special order, we get opening and closing operation derived from erosion and dilation.

When we call B_z as translation of structure element B, we use Eq. 4 to give the dilation. Form Eq. 4, we have that dilation X_d is a set construct with all points, which makes B_z and X are not empty set. Furthermore, with this expanded process, the exchanged result expands background and fills empty holes in X_d. So we use Eq. 4 to reduce empty holes in background and initialize it to a connected domain.

$$X_d = X \oplus B_z \quad (4)$$

Oppositely, erosion is such a set construct with all points z, which makes B_z is a subset of X. We use E.q.5 to express it. Erosion is a process of shrink transform to remove boundary points. The exchange result X_e expands empty holes and shrinks background. In this way, we use erosion to remove isolated noise point actively.

$$X_e = X \ominus B_z \quad (5)$$

Generally, both erosion and dilation are irreversible operations. It means that the result always differs to X when we transform X by using Eqs. 4-5 step by step. Then, when we use Eq. 4 at first, then use Eq. 5, we have the result X_o in Eq. 6. In contrary, when we use Eq. 5 at first, then use Eq. 4 behind, we can have another result X_c in Eq. 7.

In Eq. 6, we find that this presented novel morphological transformation is an opening operation because X_o is constructed by the set of B_z, which is translation of B. Otherwise, it always smooth border of the exist background, and it removes small sharp or isolated points. Furthermore, it sharpens angles, disconnects limited gaps and removes thin tips.

Then, with same analysis, we know X_c is an closing operation from Eq. 7. It transforms X by dilation firstly and erosion secondly. In this case, X_c is a complementary intersection set of all translations B_z, and it outside X. We use this morphological transformation to smooth border, remove limited gaps and long thin blanks like opening operation. Moreover, it removes small holes and fills boundary ruptures.

$$X_o = (X \ominus B_z) \oplus B_z = \cup \{B_z^c : B_z \subset X^c\} \quad (6)$$

$$X_c = (X \oplus B_z) \ominus B_z = \cup \{B_z^c : B_z \subset X^c\} \quad (7)$$

So, in the next step, we execute several transformations by Eqs. 6-7 to reach continuous background. Then, with both temporal difference method and morphological transformations, we restructure the background by trained several frame teams.

III. STAINS DETECTION AND TRAJECTORY TRACKING METHOD

With background restructured, we can detect stains in monitoring videos. In fact, the background is static in monitoring cameras. So we research in differences of frames with fixed distance. Then, we compare static area without background and then use morphological transformations to process. We detect a stained area and send administrator messages when area of one static region is more than a threshold.

But this process has shortcomings. For example, there is a video for a part of a classroom, and original background is an empty desk. Then someone puts a book on it. This will lead to a wrong detection because the book is detected as a stain. So we should use trajectory tracking method in our method [10-11]. We track

trajectories of all moving objects with their differences are less than a threshold in continuous frames.

We use E.q.8 as basic shift formula [12].

$$\mathrm{M_h}(x) = \frac{1}{k}\sum_{x_i \in S_h}(x_i - x) \quad (8)$$

where

$\mathrm{S_h}(x) \equiv \{y: (y-x)^T(y-x) \le h^2\}$

It is known that a point is more effective of statistic characteristics estimation x when it is nearer to x. So we use distance into Eq. 8, and give an improved tracking method with Eq. 9.

$$\mathrm{M(x)} = \frac{\sum_{i=1}^{n} G_H(x_i - x)w(x_i)(x_i - x)}{\sum_{i=1}^{n} G_H(x_i - x)w(x_i)} \quad (9)$$

Then, to set them in kernel function we have E.q.10 to describe the trajectory, and Eq. 11 must be a maximum likehood of Eq. 10. So we use Eq. 12-13 to give the final tracking method by simplified formulas.

$$\widehat{q_u} = C\sum_{i=1}^{n} k\left(\left\|\frac{x_i^s - x_0}{h}\right\|^2\right)\delta[b(x_i^s) - u] \quad (10)$$

$$\widehat{p_u}(y) = C_h\sum_{i=1}^{n_h} k\left(\left\|\frac{x_i^s - y}{h}\right\|^2\right)\delta[b(x_i^s) - u] \quad (11)$$

$$\mathrm{w_i} = \begin{cases} \sum_{u=1}^{m}\delta[b(x_i) - u]\sqrt{\frac{q_u}{p_u(y_0)}} & when\ i = 1 \\ 0 & otherwise \end{cases} \quad (12)$$

$$\rho[\mathrm{p(y)}, \mathrm{q}] = \begin{cases} \frac{1}{2}\left\{C_h\sum_{i=1}^{n} w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\sum_{u=1}^{m}\sqrt{p(y_0)q_u}\right\} \\ \qquad when\ i = 1 \\ \frac{1}{2}\sum_{u=1}^{m}\sqrt{p(y_0)q_u} \qquad when\ i = 0 \end{cases}$$

In this way, we research features of the stains, and find a difference area is a stain where it is a static object which differs to background and it is appeared suddenly. In other words, this static object is not appeared in previous frames sequence. Therefore, this kind of objects is probably a stain. Admittedly, camera movement is also detected as a stain. In fact, this step is possible because our final objective is to find all cases that makes monitor cannot photo/film what we need.

The process of the novel method is shown in Fig. 2.

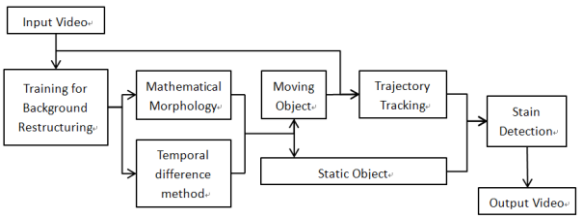
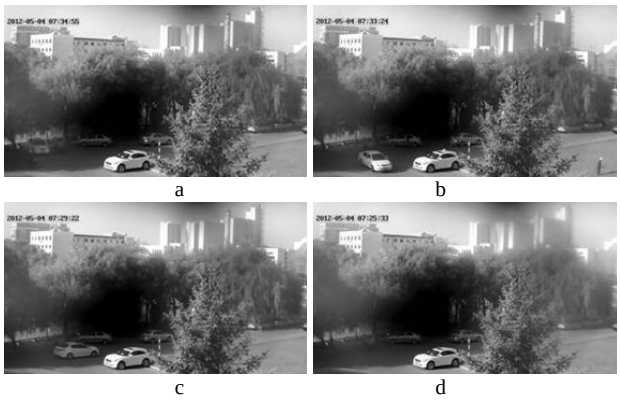


Figure 2. Process of the novel stains detection method

IV. Experiment and Analysis

The first experiment is moving object capture from a video which is filmed outer of my office window by a camera. The background is full of noises because of wind and illumination. We show some images caught from video in following figures. Upper left of each image is filmed time. The form is yyyy-mm-dd hh-mm-ss. At first, we give our restructured background in Fig.3. We train it for a short time that there are some static cars are detected as background. It is because they do not move after we film our video. Then, we stain the camera lens in Fig.4-7. We use soil on the camera to simulate stains on it. Moreover, we use the stained camera to film another video and have some frames from it, which are shown in Fig. 4-7 by time sequence.

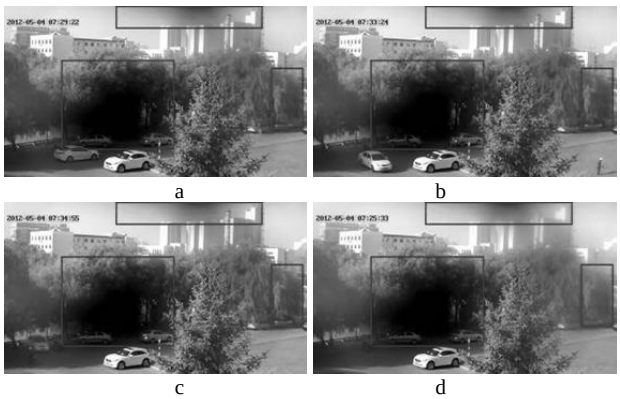
Figure 3. Original background outside



a b c d

Figure 4. Stained video frames

In this way, we use the stains detection method to process these frames of the video and the results are shown in Fig. 5.



a b c d

Figure 5. Stains detection of stained video frames

Then, we use some video fragment from websites to validate our method [13-15]. Unfortunately, we do not find videos with stains. So we use a stain randomized algorithm to create stains for these videos, and then add these stains into these videos separately. Original frames of these videos are shown in Figs. 6-8.

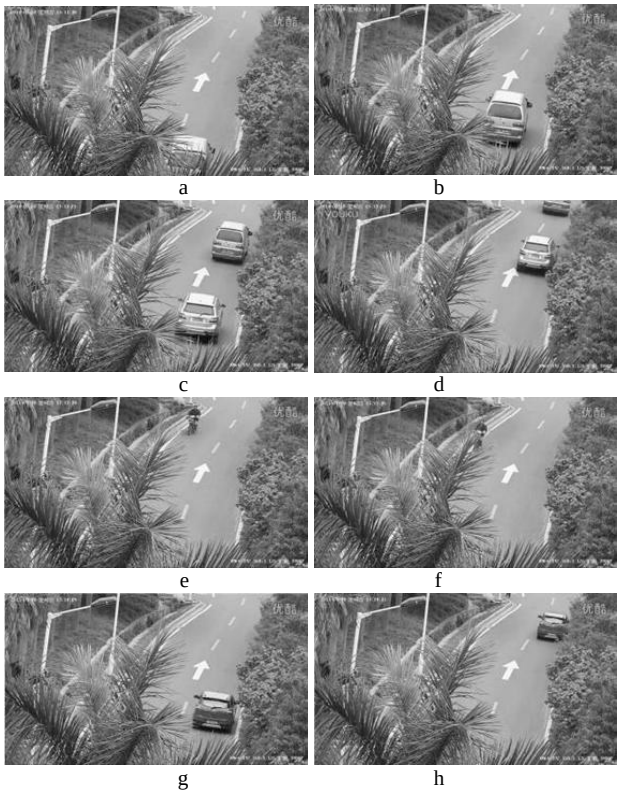


Figure 6. The first monitoring video fragments

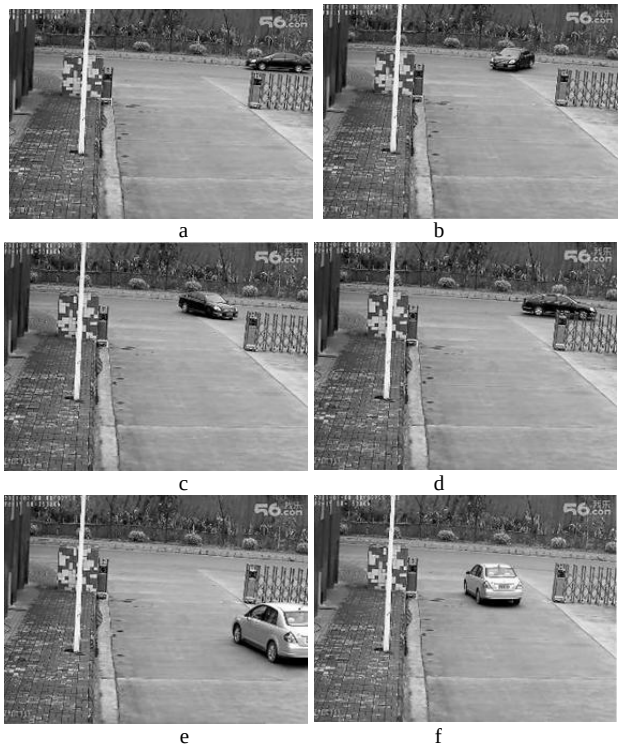


Figure 7. The second monitoring video fragments

Figure 8. The third monitoring video fragments

The corresponding stained frames are shown in Fig. 9-11.

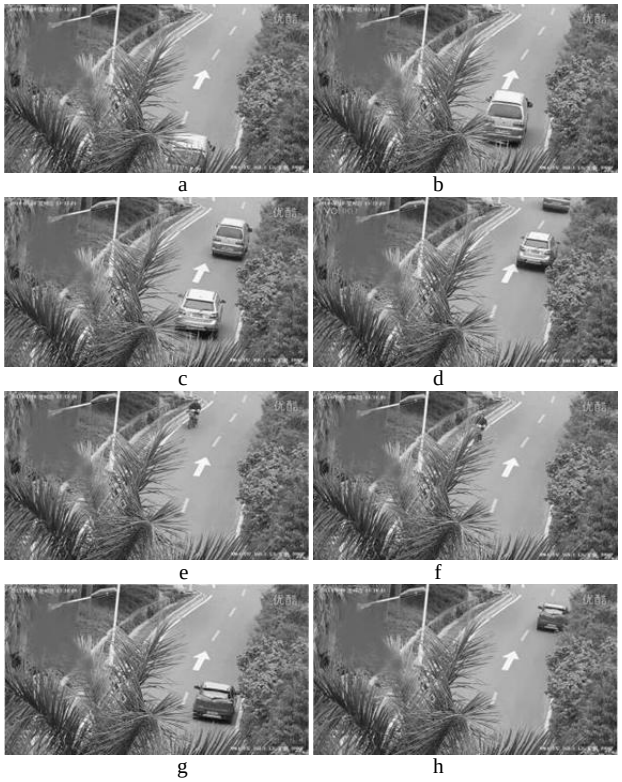
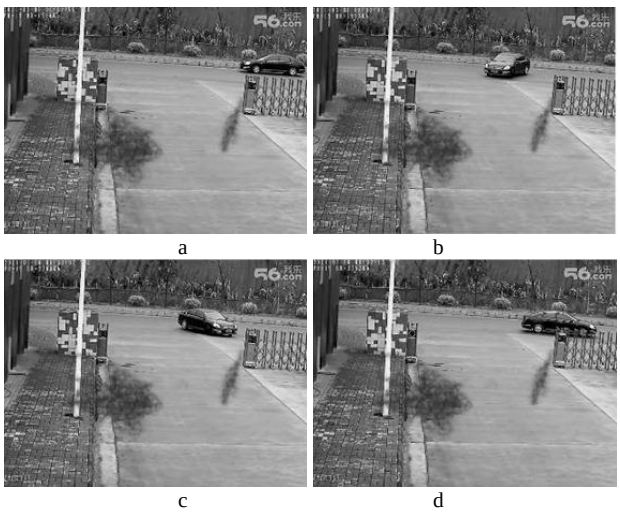


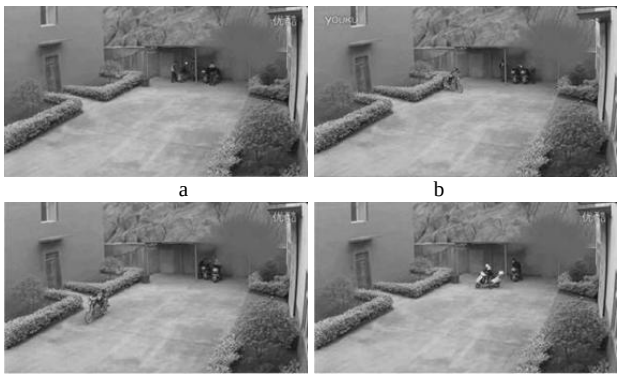
Figure 9. The first stained monitoring video fragment

Moreover, the stains we created are shown in Fig. 12. Fig. 12a is the stain for the first monitoring video, Fig. 12b is the stain for the second monitoring video, and Fig. 12c is the stain for the third monitoring video.



e f

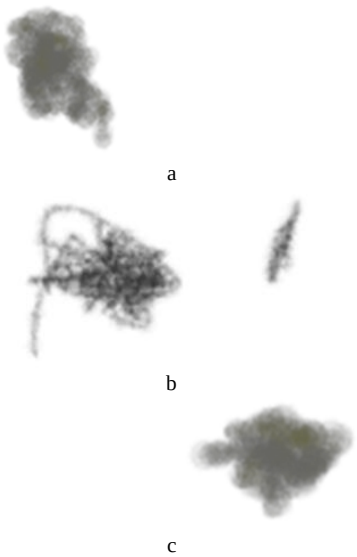
Figure 10. The second stained monitoring video fragment



a b

c d

Figure 11. The third stained monitoring video fragment



a

b

c

Figure 12. Three stains

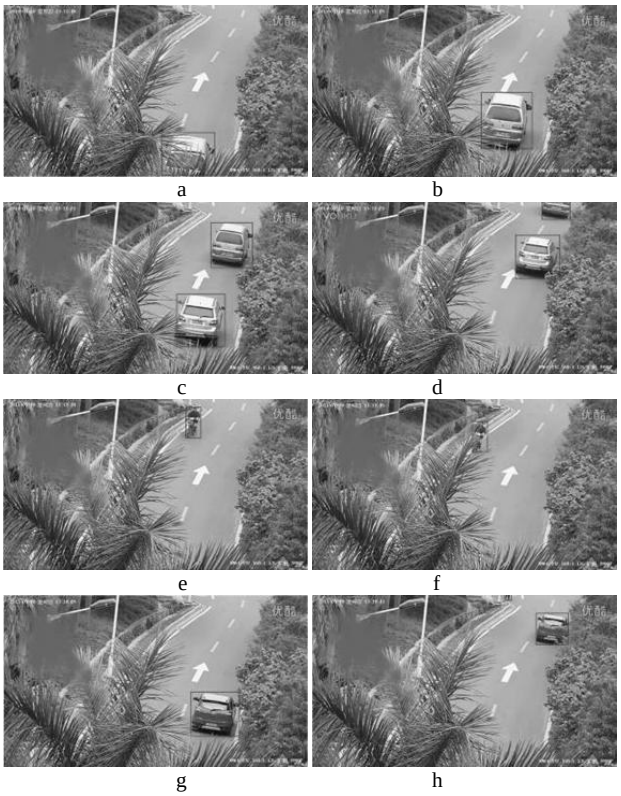
Figure 13. The first background

Then, we use our stains detection method in these videos, and get the background of all videos in Figs. 13-15. Meantime, I have to say that we cannot process displayed time on top left because the time changed tinily that we cannot detect time as moving or static object. We can only treat them as background.

Figure 14. The second background

Figure 15. The third background

Meanwhile, we reach the moving trajectories of these three videos. They are shown in Figs. 16-18.



a b

c d

e f

g h

Figure 16. Trajectory tracking in the first stained monitoring video fragment

Figure 17. Trajectory tracking in the second stained monitoring video fragment

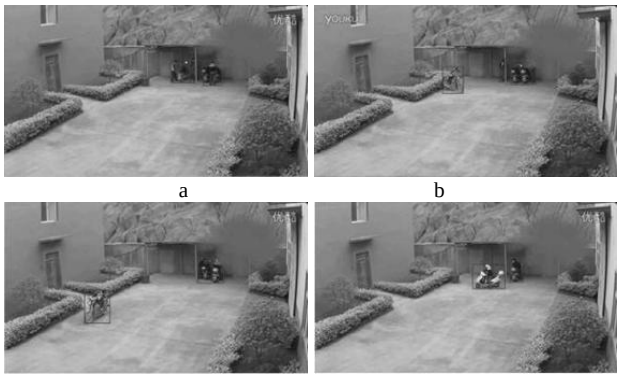
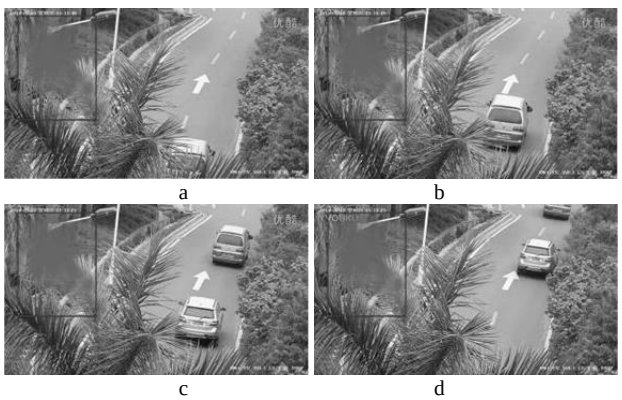


Figure 18. Trajectory tracking in the third stained monitoring video fragment

Finally, we reach the stains of these three videos. They are shown in Figs. 19-21.



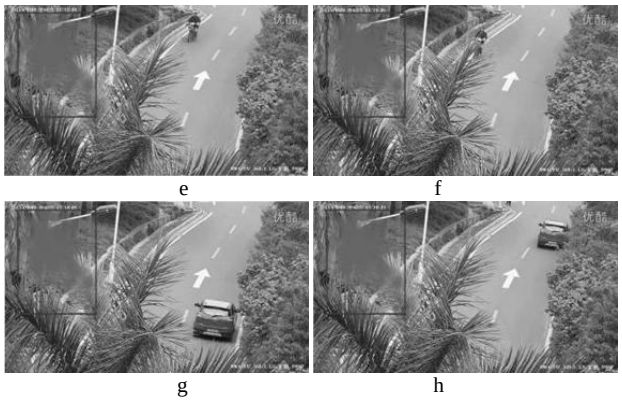


Figure 19. Stain detection in the first stained monitoring video fragment

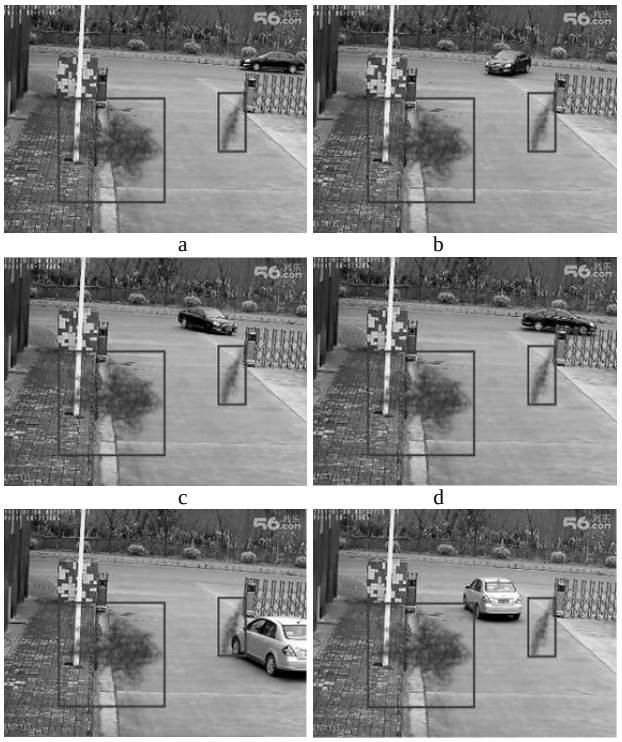


Figure 20. Stain detection in the second stained monitoring video fragment

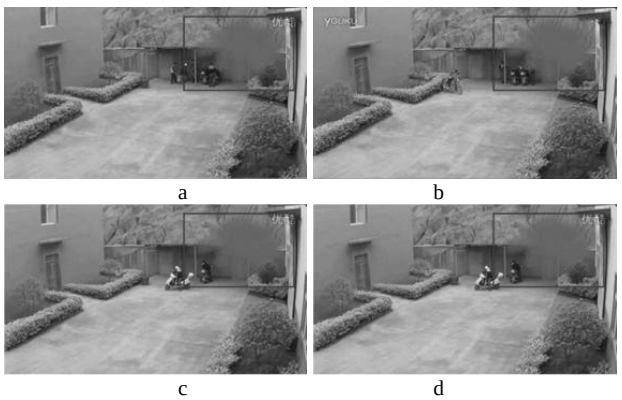


Figure 21. Stain detection in the third stained monitoring video fragment

V. CONCLUSION

This paper puts forward a stain detection method in

network monitoring video with background restructured. In this paper, we restructure background by using both temporal difference method and morphological transformations. Then with the trained background, all moving trajectory and static trajectory are detected. Moreover, with stain detection method, we separate static trajectory to stains and static objects with movement. This judgment is given by detecting if these objects were moved before in the video. This method is antinoise when we set a threshold to separate moving fragments and noise. Admittedly, these noises may be created by any condition, e.g. illumination and wind. Finally, we process experiments to validate our method. Experiment results show its positive computational complexity, accuracy and robustness.

ACKNOWLEDGMENT

This work is supported by grants Programs of Higher-level talents of Inner Mongolia University [No. 125126, 115117], Innovation Training Program of Graduates of Inner Mongolia University in year 2014, Scientific projects of higher school of Inner Mongolia [No. NJZY13004], National Natural Science Foundation of China [No: 61261019, 61262082].

The authors wish to thank the anonymous reviewers for their helpful comments in reviewing this paper.

REFERENCES

[1] Q Du, B Jiang, Y Tang, et al. Design and Implementation of the Embedded Based Web Camera System. *Journal of Software*, 2012, 7(11), 2560-2566

[2] I Halliday, A A Griffin, A T Blackwell. Detailed data for 259 fireballs from the Canadian camera network and inferences concerning the influx of large meteoroids. *Meteoritics & Planetary Science 31*, 185-217 (1996)

[3] S Gupt, O Masound, R F K Martin, N P Papanikolopoulos. Detection and classification for vehicles. *Intelligent Transportation Systems*. 2002, 3(1):37-47

[4] X Wang, Y Zhang, H Dong, et al. Algorithm Research on Three-Frame Difference for Detection of Moving Target. *Journal of Shenyang Ligong University*. 2011, 30(6):82-85

[5] W Fu, Z Xu, S Liu, et al. The capture of moving object in video image. *Journal of Multimedia*, 2011, 6(6):518-525

[6] Yubin Zhan, Jianping Yin, Xinwang Liu. A Convergent Solution to Matrix Bidirectional Projection Based Feature Extraction with Application to Face Recognition. *International Journal of Computational Intelligence Systems*, 2011, 4(5):863-873

[7] J Park, S H Oh, Y S Lee. Network security camera system and its application for consumer electronics in ubiquitous environment. *Multimedia Tools and Applications*, 2013, DOI 10.1007/s11042-013-1442-8

[8] J Yu, W Zhang. Study on Agricultural Condition Monitoring and Diagnosing of Integrated Platform Based on the Internet of Things. *IFIP Advances in Information and Communication Technology* 2013, 392:244-250

[9] Z Tu, P Bhattacharya. Game-theoretic surveillance over arbitrary floor plan using a video camera network. *Signal, Image and Video Processing*, 2013, 7(4):705-721

[10] S Liu, W Fu, W Zhao, et al. A Novel Fusion Method by Static and Moving Facial Capture. *Mathematical Problems in Engineering*, 2013, doi:10.1155/2013/503924.

[11] S Liu, X Cheng, W Fu, et al. Numeric characteristics of generalized M-set with its asymptote. *Applied Mathematics and Computation*, 2014, 243: 767–774

[12] J A Balista, M N Soriano, C A Saloma. Compact time-independent pattern representation of entire human gait cycle for tracking of gait irregularities. *Pattern Recognition Letters*, 2010, 31(1): 20-27.

[13] http://v.youku.com/v_showid_XNTAwMDI1MzUy.html

[14] http://www.56.com/u44v_NjEwMDQ2MjA.html

[15] http://v.youku.com/v_showid_XNTU5NTI3NDQ4.html

Recommendation System Based on Fuzzy Cognitive Map

Wei Liu and Linzhi Gao
Beijing University of Posts and Telecommunications, Beijing China
Email: westllife@163.com, twhlw@163.com

***Abstract*—With the increase of data volume and visitor volume, the website faces great challenge in the environment of network. How to know the users' requirements rapidly and effectively and recommend the required information to the user becomes the research direction of all websites. The researchers of recommendation system propose a series of recommendation system models and algorithms for the user. The common challenge faced by these algorithms is how to judge the user intention and recommend the relevant content by little user action. The paper proposes the user situation awareness and information recommendation system based on fuzzy clustering analysis and fuzzy cognitive maps, and verifies the validity of the algorithm by the application to recommendation site of academic thesis.**

***Index Terms*—Recommendation System; Fuzzy Cognitive Map; Fuzzy Cluster**

I. INTRODUCTION

Recommendation system started from electronic commerce, which mainly solves the problems of information overload caused by rapid expansion of Internet scale and application range [1, 2]. According to [3], recommendation system uses electronic commerce website to provide merchandise news and suggestion to help the user buy products and simulate salesmen to help the client to complete purchase process. With the coming of big data era, the mainstream search engine and data index service can't meet increasing individualized information requirements of network user. And realizing individualized recommendation for every user has become the research focus of data mining technology, data fusion and e-commerce.

Recommendation algorithm is the core of recommendation system. The mainstream recommendation algorithms include content recommendation, collaborative filtering recommendation and knowledge recommendation. The database structure established by most algorithms is a series of set structure of traditional mathematics. The correlations between data, and the mapping relation between data and category is too simple and qualitative, especially when it relates to the concept of fuzzy semantic word, the correlation between data lacks of scientific description. For example, some recommendation systems realize recommendation according to the match of user retrieval word and the established key lexicon, and the systems make recommendation unsuccessful because of vocabulary limitation of key lexicon and training deviation. On the other hand, with the deep intersection and combination between subjects, there are many words appearing in many fields, which is difficult to complete effective and reliable recommendation behavior by keyword match, the reason for which is the correlation between words, especially the quantitative correlation is not emphasized and described fully.

Introducing the idea of fuzzy logic into recommendation system can effectively solve the problems. Literature [5, 6, 7] introduces some application of fuzzy clustering analysis algorithm. Although many recommendation algorithms use fuzzy clustering analysis and achieves some results, but most of them use collaborative filtering idea [4]. And the key is to classify the users according to the user behavior and predict the preference of the present users according to the interest model of similar users, which causes the following problems.

1. It needs to acquire enough user information and user behavior to realize effective recommendation.

2. The information structure in database is simple and mechanical, and lacks of description based on fuzzy logic.

Therefore, the paper establishes a recommendation system of academic paper based on fuzzy clustering analysis and fuzzy cognitive maps, uses fuzzy clustering analysis to quantitatively describe the correlation between lexical semantics. The algorithm of analyzing user interest uses fuzzy cognitive mapto figure out qualitative distribution of user interest and quantitative interest degree according to much less user information compared with these systems mentioned above. Fuzzy cognitive mapconverts the original tree structure between discipline classification and key words into mesh structure, describes the correlation between interdisciplinary and the map of a key word corresponding to many discipline classifications, and provides iterative operation function for reasoning and states evolution, for realizing effective recommendation.

The implementation process of the system is as follows. Firstly, key words are selected for discipline classification, and the frequency of each key word is for statistics. Then, AFS fuzzy clustering analysis algorithm is used to process the key words and discipline classification in database, which gets fuzzy description of each classification. And the quantitative correlation value

doi:10.4304/jmm.9.7.970-976

between discipline classifications and the correlation value between each key word an each discipline classification is achieved according to fuzzy description, which establishes net data structure. Lastly, according to input information of the user, the related discipline classification is screened out in net data structure. Fuzzy cognitive map is used to compute the screened key words. And the papers with key words in database are recommended to the user.

Compared with the existing recommendation systems based on collaborative filtering algorithm, the recommendation system based on fuzzy clustering analysis and fuzzy cognitive map established in the paper has the following innovation points. Firstly, fuzzy clustering analysis introduces the view of information entropy theory. The information content and information structure in database received quantitative description, and the tree data structure is converted into net data structure, which prepares for effective recommendation. Secondly, the model firstly applies fuzzy cognitive map to recommendation system. Fuzzy cognitive map has evident advantage for processing data structure with quantitative association, so the method is suitable for net data structure of the model based on fuzzy clustering analysis. Thirdly, most algorithm processes of the model are completed in previous preparation and feedback training process. The algorithms executed by the recommendation system from input information to recommendation results are very easy, which can ensure to use fewer resources to get recommendation results in a shorter time.

II. FUZZY CLUSTERING ANALYSIS

The paper uses fuzzy logic clustering algorithm to realize the above fuzzy clustering process. Compared with the previous fuzzy clustering algorithm, the algorithm integrates information entropy theory.

Fuzzy logic clustering algorithm firstly needs to collect the relevant data sample to acquire fuzzy description of samples, the objective of which is to expand a sample to be a concept of set to evaluate the correlation between samples.

And the sample data is divided into two conditions.

1. A data sample is a discipline classification of the bottom. The property of sample is a series of relevant key words, and the attribute value is the frequency of key words.

2. A data sample is a key word. The property of sample is a series of relevant discipline classification, and the attribute value is the frequency of key words in classifications.

Fuzzy description of samples is the description with a certain threshold on a sample based on fuzzy set theory. Ideal fuzzy description on a sample can make the fuzzy description of similar sample attaching to the sample great, and make the fuzzy description of non-similar sample attaching to the sample small.

And a threshold is needed to classify the samples according to the degree of the sample attaching to fuzzy description. The original sample set is converted into the set of classification, and the same method is used to make weighting fuzzy description on every classification. And the original samples are used as the elements, and the classification is used as the set to analyze and evaluate the subjection degree, which can determine the classification of the sample.

Literature [8, 9, 10 and 11] introduces AFS fuzzy clustering analysis algorithm integrating information entropy theory. As FCM to be established consists of a series of key words and discipline classifications, and the frequency of key words can be achieved by statistical samples, which meets the characteristics of the clustering algorithm. So the paper applies the algorithm to clustering analysis on discipline classification and key words, for realizing effective recommendation. 9 definitions are used to make mathematical description on the algorithm, as follows.

Definition 2.1: ς is an object of data X and ς corresponds to a relationship R_ς of X, ($R_\varsigma \subseteq X \times X$), and $(x,y) \in R_\varsigma \Leftrightarrow x$ attaches to ς. To a certain degree, and the degree attaching to ς should be greater than or equal to the degree of y attaching to ς, $(x,y) \in X$.

Definition 2.2: R is the relationship of data set X. For $(x,y) \in X$, $x \neq y$, R meets:

If $(x,y) \in R$, $(x,x) \notin R$;

If $(x,x) \notin R$, and $(y,y) \in R$, $(y,x) \in R$;

If $(x,y) \in R$, and $(y,z) \in R$, $(x,z) \in R$;

If $(x,y) \in R$, and $(y,y) \in R$, $(x,y) \in R$, or $(y,x) \in R$.

R is the weak intention. The objects corresponding to weak intention is simple object, conversely, it is called complicated object.

Definition 2.3: m is simple fuzzy object of X, $\rho_m : X \to R^+ = [0,\infty)$. If ρ_m meets the following conditions, ρ_m is the attaching density function of simple fuzzy object m.

$\rho_m(x) = 0 \Leftrightarrow (x,x) \notin R_m, x \in X$;

$(x,y) \in \Rightarrow \rho_m(x) \geq \rho_m(y), x, y \in X$.

Definition 2.4: M is the sample fuzzy object set of data set X, and 2^m is the power of M, $\tau : X \times X \to 2^M$. If it meets the condition AX1, AX2, (M, τ, X) is AFS structure.

AX1: For any $(x_1, x_2) \in X \times X, \tau(x_1, x_2) \subseteq \tau(x_1, x_1)$;

AX2: For any $(x_1, x_2)(x_2, x_3) \in X \times X$,

$\tau(x_1, x_2) \cap \tau(x_2, x_3) \subseteq \tau(x_1, x_3)$.

X is called data set, M is called object set and τ is called structure.

Definition 2.5: X and M is set, (M, τ, X) is a AFS structure. $B \subseteq X, A \subseteq X$, and the define symbol is

$\bar{A}(B) = \{y \mid y \in X, \tau(x,y) \supseteq A, \text{ for any } x, x \in B\}$.

$X \subseteq R^n$ is data set, ρ_m is the attaching density function of simple fuzzy object m, $\overline{\{m\}}(x)$ is given by:

$$L_m(x) = \frac{\sum_{x \in \overline{\{m\}}(x)} \rho_m(x)}{\sum_{x \in X} \rho_m(x)} \quad (1)$$

It is the degree of data x attaching to object m.

Definition 2.6: M is the set of simple fuzzy objects of data set X. For fuzzy object $A \subseteq M$, the subjection function of fuzzy object A is defined as follows, for $x \in X$,

$$\mu_A(x) = \min_{\alpha \in A}(L_a(x)) \in [0,1] \quad (2)$$

Definition 2.7: X is data set, and M is simple fuzzy object set of X, $A \subseteq M$. And information entropy function of subjection degree is:

$$E(A) = -\sum_{x \in X}\left(\mu_A(x)\log_2(\mu_A(x))\right) \quad (3)$$

Smaller E(A) means that the subjection degree of all data belonging to object A is closer to two sides of closed interval [0,1], and dividing data according to fuzzy object A is clearer.

Definition 2.8: X is data set, M is the set of simple fuzzy object of X, $A \subseteq M$, and n is the number of data. The distribution function D of subjection degree is:

$$D(A) = -\left(\frac{(\sum_{x \in X} \mu_A(x))}{n}\right)\log_2\left(\frac{(\sum_{x \in X} \mu_A(x))}{n}\right) \quad (4)$$

Smaller D(A) means that the subjection degree of all data belonging to object A is closer to one side of closed interval [0,1]. When it is close to the right side, it means that fuzzy object A and data set X are similar.

Definition 2.9: M is the set of simple fuzzy objects of X, $A \subseteq \omega$, ω is the weight set of simple fuzzy object of A. (A, ω) is weighted fuzzy object, in which $A \subseteq \omega$, $\omega = \{\omega_m \mid m \in A, 0 \le \omega_m \le 1, \sum_{m \in A} \omega_m = 1\}$.

|A| is the number of elements of A. The weighted membership function $\overline{\mu}$ of data x attaching to weighted fuzzy objects (A, ω) is as follows. For any $x \in X$,

$$\overline{\mu}_{(A,\omega)}(x) = \sum \omega_m L_m(x) \quad (5)$$

The following is description on algorithm implementation process.

Firstly, the key word library needs to be established. The resource of the key words used in the paper is not Keyword of the paper but is achieved by refinement from the collected original paper. And the concrete steps are as follows.

Step 1: 500 samples are collected for each text to train, which establishes key word library. And the concrete method is to match the words in 500 samples, and order the words according to the frequency of each word in all samples, for which the problem is that the appearance frequency of some prepositions and auxiliary words is high.

Step 2: The solution is that 500 text samples are collected, and the number of samples is distributed averagely. 500 samples receive matching exercise, and the words with high appearance frequency are found out. And the words are the prepositions and auxiliary words in each category.

For the order of Step 1, the words achieved in step 2 are eliminated, and the key words of each category which are achieved finally are ordered according to appearance frequency.

Each data sample $x_i \in X$ achieves fuzzy description of sample with the following step.

Step1:

$\eta_{x_i} \leftarrow \{m_{i,j} = \arg\max\{\mu_{x_i,1}(x_i), \mu_{x_i,2}(x_i), ..., \mu_{x_i,r_i}(x_i)\}, i = 1,2,...,s; 1 \le j \le r\}$

η_i is the simple fuzzy object set with the maximal subjection degree taken from r_i simple fuzzy objects.

Step2: $E(\eta_{x_i}) = -\sum_{x \in X}\left(\mu_{\eta_{x_i}}(x)\log_2(\mu_{\eta_{k_i}}(x))\right)$.

Step3:

$$D(\eta_{x_i}) = -\left(\frac{(\sum_{x \in X} \mu_{\eta_{x_i}}(x))}{n}\right)\log_2\left(\frac{(\sum_{x \in X} \mu_{\eta_{x_i}}(x))}{n}\right).$$

Step4: $V_{post} \leftarrow V_{pre} \leftarrow E(\eta_{x_i}) / D(\eta_{x_i})$, V_{post}, V_{pre} on behalf of the minimum evaluation indexes of two rounds before and after a While loop.

Step5: While $V_{pre} \ge V_{post}$, do:

* $\xi_{x_i} \leftarrow \eta_{x_i}$.

* $V_{pre} \ge V_{post}$.

*For each $m \in \eta_{x_i}$, do:

$E(\eta_{x_i}^m) \leftarrow -\sum_{x \in X}(\mu_{\eta_{x_i}^m}(x)\log_2(\mu_{\eta_{x_i}^m}(x)))$.

$$D(\eta_{x_i}^m) = -\left(\frac{(\sum_{x \in X} \mu_{\eta_{x_i}^m}(x))}{n}\right)\log_2\left(\frac{(\sum_{x \in X} \mu_{\eta_{x_i}^m}(x))}{n}\right).$$

$V_m \leftarrow E(\eta_{x_i}^m) / D(\eta_{x_i}^m)$.

* $m \leftarrow \arg\min\{V_m \mid m \in \eta_{x_i}\}$.

* $V_{post} \leftarrow V_m$.

* $\eta_{x_i} \leftarrow \eta_{x_i} - \{m\}$.

Return ξ_{x_i}.

Based on the above process, the fuzzy description clustering analysis process is as follows.

Step1: Establishing equivalent relation matrix of correlation between samples.

* $m_{i,j} \leftarrow \min\{\mu_{\{\xi_{x_i},\xi_{x_j}\}}(x_i), \mu_{\{\xi_{x_i},\xi_{x_j}\}}(x_j)\}$.

* $M \leftarrow (m_{ij})$.

* $Q \leftarrow M^t$ finds the integer t to make $(M^t)^2 = M^t$, M^t is the equivalent relation matrix.

Step2: Determining the initial classification $\overline{C_1}$, $\overline{C_2}$, …, $\overline{C_c}$ of the initial threshold α.

* $Q=(q_{ij})$, if $q_{ij}\geq\alpha$, $q_{ij}=1$, which means x_i, x_j belongs to the same classification, which achieves the initial classification $\overline{C_1}$, $\overline{C_2}$, …, $\overline{C_c}$.

Step3: Acquiring weighted fuzzy description $\xi_{C_i}=(A_{C_i},\omega_{C_i})$ of each classification.

* $A_{C_i}\leftarrow\{m\mid m\in\xi_x, x\in\overline{C_i}\}$.

* $\omega_{C_i}\leftarrow\{\omega_m\mid m\in A_{C_i}\}$,

$$\omega_m\leftarrow\frac{|\overline{C_i^m}|}{\sum_{x\in\overline{C_i}}\xi_x},$$

in which $|\overline{C_i^m}|=|\{x\mid x\in\overline{C_i}, m\in\xi_x\}|$,
|.| is the number of elements.

$\omega_m\in\omega_{C_i}$ is the weight of $m\in A_{C_i}$.

* $\xi_{C_i}\leftarrow(A_{C_i},\omega_{C_i})$.

Step4: determining the final classification $C_1, C_2, ..., C_c$ of the corresponding threshold.

For each $x\in X$,

$q\leftarrow\arg\max\{\overline{\mu}_{\xi_{C_i}}(x)\}$, $x\in C_q$.

The weighted membership is calculated according to:

$\overline{\mu}_{(A,\omega)}(x)=\sum\omega_m L_m(x)$.

Step5: return $C_1, C_2, ..., C_c$ and $\xi_{C_1}, \xi_{C_2}, ..., \xi_{C_c}$.

After clustering process, it realizes more rational, and is suitable for discipline classification of fuzzy cognitive mapmodeling.

Figure 1 takes the key word situational awareness as an example to describe the relationship comparison of key words of discipline classification before and after cluster. We can see from the figure that after cluster, the description of key words connects with more relevant disciplines by quantitative relationship of administrative subordination, which makes the key words become the representative key factor of interdisciplinary, and combines many discipline classifications which have great difference in tree diagram. Next, the paper establishes quantitative correlation between discipline classifications based on the same key word by achieving causal weight.

III. Establishing Fuzzy Cognitive Mapto Recommender

Fuzzy cognitive map came from the concept of cognitive map proposed in Structure of Decision: The Cognitive map of Political Elites proposed by Axelrod in 1976. And Bart Kosko of University of Southern California firstly systematically proposed fuzzy cognitive maps. For the concept, FCM is the combination of fuzzy logic and concept mapping. Fuzzy logic is from fuzzy set theory and is a theory that process is similar to reasoning process. And it can be considered to be the application of a fuzzy set theory. FCM uses pre-definite knowledge or constructs the causal relationship between concepts. It plays the role of processing symbolic knowledge, so FCM is applicable to software knowledge. Fuzzy cognitive map have the characteristics and advantages of neural network and Hidden Markov Chain, which overcomes the fixed defect of probability model [12].

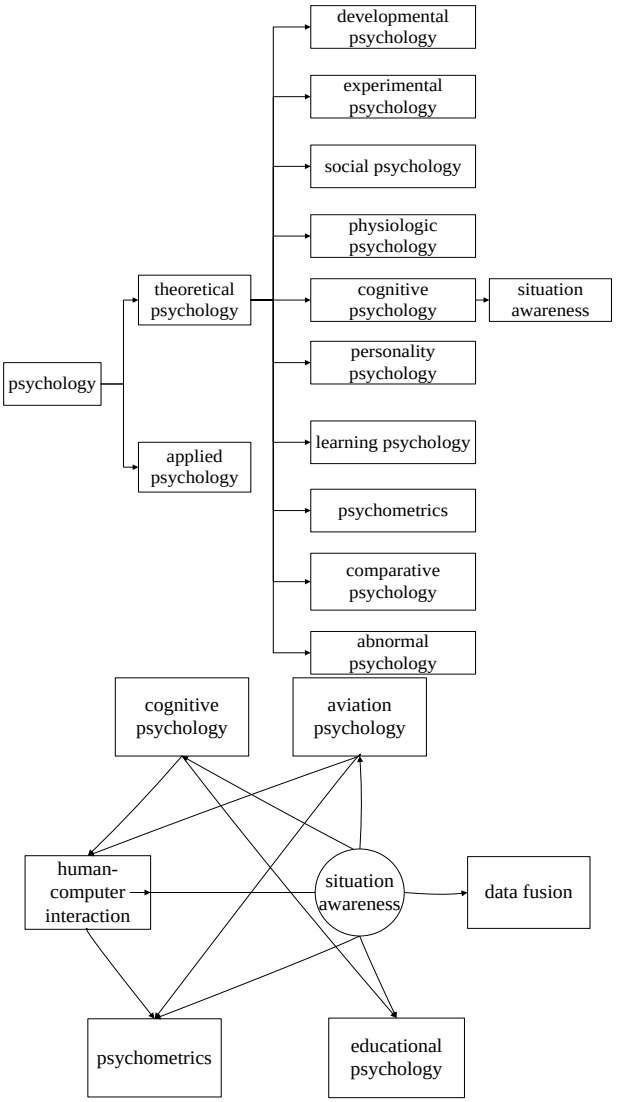


Figure 1. Relationship comparison between discipline classification and key words before and after cluster

According to [13, 14, 15], FCM provides an adaptive structure, which can give qualitative reasoning and quantitative elementary evaluation for the level or state of a complicated system. The core of FCM is a diagram structure consisting of many concepts connected by causal relationships. And the causal relationship is characterized by the numerical value between the interval [-1,+1]. -1 means reverse causal relationship, +1 means forwarded causal relationship, and 0 means there is no causal relationship.

In fact, the fuzzy clustering analysis process is to prepare for constructing the nodes of fuzzy cognitive maps. We extract the concept with the same level from the network structure with quantitative correlation after clustering to be as the nodes of fuzzy cognitive maps.

The causal weight between nodes is achieved from positive and negative correlation of appearance frequency after the key words represented by two nodes receive cluster. We use spearman correlation coefficient to

represent causal weight. And the concrete calculation process is as follows.

Spearman correlation coefficient is Pearson correlation coefficient between classification variables. For the sample with the capacity , n original data X_i, Y_i is converted into graded data x_i, y_i.

The original data is assigned a grade according to the average descending position in overall data. According to the appearance frequency of key words in text samples, descending sort is made, and the grade is position. The key words with the sample appearance frequency take the mean of the position values.

The difference of levels between two variables is $d_i = x_i - y_i$,

and ρ is:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{6}$$

$\omega_{ij} = \rho$. After getting all ω_{ij}, we can use a N-order matrix to express, which is called node correlation matrix.

Figure 2 is a FCM example of cognitive domain.

Incidence matrix of Figure 2:

$$\Omega = \begin{bmatrix} 0 & 0.8 & 0.2 & 0.3 & -0.7 & 0 & 0 \\ -0.9 & 0 & 0.75 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0.6 & 0.6 & 0.4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.9 & 0.3 & 0 & 0.5 & 0 \\ 0 & -0.8 & 0 & 0.2 & 0.5 & 0 & 0.7 \\ -0.7 & 0 & 0 & 0.4 & 0 & 0.7 & 0 \end{bmatrix}$$

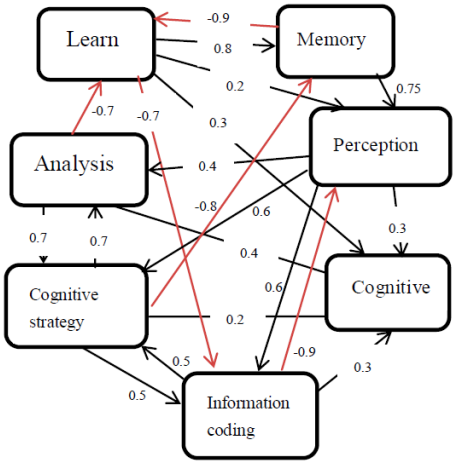


Figure 2. Key word FCM relating to cognition

If the initial state of concept node of the system is known, the state value of any concept node at any time can convert function by FCM.

$$A_i^{t+1} = f\left(A_i^t + \sum_{j=1, j\neq i}^{N} A_j^t \omega_{ij}\right) \tag{7}$$

$C = (c_1, c_2, ..., c_7)$ is supposed to be the set of system concept nodes, and A_i^t means the state value of the i concept node at time t, and ω_{ji} means the weight of the j concept node on the causality influence of the i concept node. And f is a threshold function. In the experiment, the threshold function is a three-value step function.

$$f(x) = \begin{cases} -1, x \le -0.5 \\ 0, -0.5 < x < 0.5 \\ 1, x \ge 0.5 \end{cases} \tag{8}$$

When the user inputs the retrieval words, if the retrieval words are in the established keyword library, the recommendation system can capture the retrieval words. If the retrieval words are not in the lexicon, some words with similar semantics provided by research engine are used to be as the recommendation engine of inputting retrieval words. If the retrieval words are used as key words in many classifications, the state value of the lexical node is 1, and that of other nodes is 0 as the initial vector for iterative operation until FCM system is stable. The words represented by the activated nodes are used as key words to filter the papers and recommend them to the users. For example, when the user inputs the word, study, the research engine generally considers that learning from someone is closer to retrieval words than cognition. But in the recommendation of the paper, the system thinks that the user inputs an initial state in FCM of Figure 2.

$A^{t=0}[1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$

$A^{t=1} = f([1 \quad 0.8 \quad 0.2 \quad 0.3 \quad -0.7 \quad 0 \quad 0])$

$= [1 \quad 1 \quad 0 \quad 0 \quad -1 \quad 0 \quad 0]$

$A^{t=2} = f([0.1 \quad 1.8 \quad 1.9 \quad 0 \quad -1.7 \quad -0.5 \quad 0])$

$= [0 \quad 1 \quad 1 \quad 0 \quad -1 \quad 0 \quad 0]$

$A^{t=3} = f([-0.9 \quad 1 \quad 2.6 \quad 0 \quad -0.4 \quad 0.1 \quad 0.4])$

$= [-1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]$

$A^{t=4} = f([-1.9 \quad 0.2 \quad 1.6 \quad 0 \quad 1.3 \quad 0.6 \quad 0.4])$

$= [-1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 1 \quad 0]$

$A^{t=5} = f([-1 \quad -1.6 \quad -0.1 \quad 0.5 \quad 2.8 \quad 2.1 \quad 1.1])$

$= [-1 \quad -1 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1]$

$A^{t=6} = f([-0.8 \quad -2.6 \quad -1.9 \quad 1.6 \quad 2.2 \quad 2.2 \quad 1.3])$

$= [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1]$

$A^{t=7} = f([-0.8 \quad -2.6 \quad -1.9 \quad 1.6 \quad 2.2 \quad 2.2 \quad 1.3])$

$= [-1 \quad -1 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1]$

The system enters stable state. And the activated nodes are cognition, information coding, cognitive strategy and information analysis. When the recommendation system selects words in paper database, the paper with four words is recommended to the user. If there is no meeting the demand, it is ordered according to the stable weighting value of four words, and the non-important words are deleted.

The most evident characteristic of fuzzy cognitive map is the hierarchy and flexibility of the structure. Every node of FCM image can form a FCM sub-image, which

means that fuzzy cognitive map can form a multi-layer structure and can establish correlation between different layers of classifications.

We can see that recommendation system based on fuzzy cognitive map can implement effective recommendation with little or incomplete user information. Multilateral layer of fuzzy cognitive map reduces the number of nodes of each layer, and avoids too complicated associated matrix, which can make iterative process of the system too difficult to be stable.

IV. Recommendation System Structure

Recommendation system includes two modules, search engine and recommendation engine.

Research engine uses open-source retrieval engine, Sphinx. The user inputs the key words in input interface of retrieval engine, and then retrieves the engine and returns the retrieval results. We expect the user can make some direct feedback according to retrieval results to represent the satisfaction degree. The purpose of the above figure using user feedback module is to meet the requirements of the users. We provide port in the page of retrieval results for the user to feedback, such as Bookmark and Google it function of RecULike recommendation system. If the user has no feedback, we can implement recommendation according to retrieval words inputted by the user.

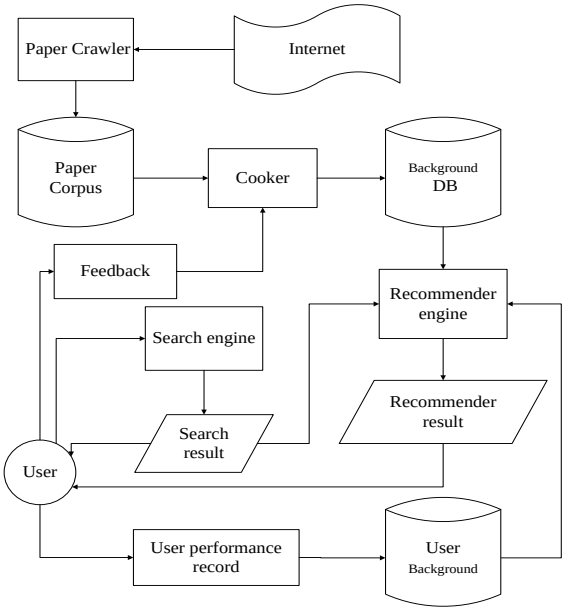


Figure 3. Structure of recommendation system

The recommendation engine is shown in the top-right figure. The background data generation module can process the original material of the paper collected by crawler (many academic databases are downloaded with charge, but the information of them is public. And the information is enough for our system). Background data generator also processes the information feedback by user feedback module and paper classification. The background database is the back-end storage of recommendation system. And the stored information includes the general satisfaction degree of the user on a paper, discipline classification and key words list after clustering, correlation and causal weight data. When the system operates initially, we don't know if the general satisfaction degree of the user on a paper. With the user using the system and feedback, we collect the information and update the database. The storage content of the background database depends on the requirements of recommendation algorithm. If using a database for implementation is too confused, the information can be stored by many independent libraries. Lastly, recommendation engine can figure out and predict the paper liked by the users and recommend it according to the information and user background in background database.

In the recommendation system, search results and recommendation results are independently shown in user interface. When the user processes the key words input by the user, it needs to update regularly, and too old information can't timely reflect the requirements of the user. If the users have no interest in the content, and keywords are stilled stored, it may mislead the recommendation system to produce wrong recommendation [16].

In order to verify the recommendation effect, we make experimental comparison on Fishgirl system and RecUlike system in the paper. 100 are tested and are demanded to write the search terms to be input and the expected recommendation results. After comparing the recommendation results of two recommendation systems, 7-level grading mechanism is used to grade the satisfactory degree of 5 dimensions. The tendency of the average absolute error of two systems increasing with testing amount is shown in Figure 4.

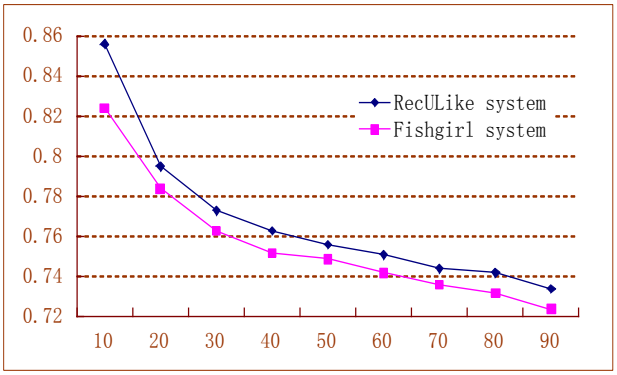


Figure 4. Recommendation results MAE comparison between Fishgirl system and RecUlik system

V. Conclusion

Based on fuzzy cluster and fuzzy cognitive maps, the paper proposes a recommendation algorithm, and constructs a academic paper recommendation system based on recommendation algorithm. It is different from the collaborative filtering recommendation algorithms based on fuzzy logic, AFS fuzzy clustering algorithm of the algorithm focuses on clustering and analyzing background data, and does not make clustering analysis on the user interest after the user inputting the information, which not only greatly reduces the workload

of recommendation engine after the user inputs information, and improves the operation speed, but also reduces the operation load of the system.

Fuzzy cognitive mapis a dynamic adaptive structure integrating neural network and Hidden Markov Chain, and can achieve lots of quantitative data with simple and rapid matrix operation in initial state with little information. Fuzzy cognitive mapfocus on the correlation of elements, and can convert the correlation between nodes into the parameters of the network, which realizes automation of decision process. The characteristics make the fuzzy cognitive mapapply to recommendation algorithm.

The algorithm provides new idea and example for applying fuzzy logic theory to recommendation system. The future research focuses on analyzing the background data of user and integrating the advantages of collaborative filtering algorithm into the recommendation system based on fuzzy cognitive maps.

REFERENCES

[1] Jose Jesus Castro-Schez, Raul Miguel, David Vallejo, Lorenzo Manuel López-López (2011): 'A highly adaptive recommender system based on fuzzy logic for B2C e-commerce portals', *Expert Systems with Applications 38* (2011) 2441–2454

[2] Y. Xin, H. Steck, 'Multi-value probabilistic matrix factorization for IP-TV recommendations', *in: ACM Conference on Recommender Systems (RecSys)*, 2011.

[3] M. A. Tayebi, M. Jamali, M. Ester, U. Glasser, R. Frank, CrimeWalker: 'a recommendation model for suspect investigation', *in: RecSys '11 Proceedings of the fifth ACM conference on Recommender systems*, 2011.

[4] Jesus Serrano-Guerrero, Enrique Herrera-Viedma, Jose A. Olivas Andres Cerezo, Francisco P. Romero: 'A google wave-based fuzzy recommender system to disseminate information in University Digital Libraries 2. 0', *Information Sciences* 181 (2011) 1503–1516

[5] J. Senthilnath, S. N. Omkar, V. Mani, 'Clustering using firefly algorithm: performance study', *Swarm and Evolutionary Computation*, 1 (3) (2011), pp. 164–171

[6] Qiu Baozhi, Zheng Zhijie, Mesh- based clustering algorithm based on local density and dynamic generation, *Computer Engineering and Design*, 2010, 31(2) pp. 385—387.

[7] A. Hatamlou, S. Abdullah, H. Nezamabadi-pour'A combined approach for clustering based on K-means and gravitational search algorithms', *Swarm and Evolutionary Computation*, 6 (2012), pp. 47–52

[8] Liu X D. The fuzzy sets and systems based on AFS Structure. Elalgebra and EII alge-bra. *Fuzzy Sets and Systems*. 1998. 95 pp. 179-188.

[9] Liu X D. The fuzzy theory based on AFS algebras and AFS structure. *Journal of Mathematical Analysis and Applications*. 1998, 217 pp. 459-478.

[10] Lin X D. \\rang W, Cha T Y. The fuzzy clustering analysis based on AFS theory. *IEEE Transactionon Systems, Man. and CvbcrncLicb-partB: Cybernetics*. 2005, 35(5) pp. 1013-1027.

[11] Zhang Y J, Liang D Q. Tong S C. On AFS algebra -part I. *Information Sciences*, 2004. 167 pp. 263-286.

[12] Glykas, M. Glykas, 'Fuzzy cognitive maps: Advances in theory, methodologies, tools and applications'. *Studies in fuzziness and soft computing*, Vol. 247Springer-Verlag (2010)

[13] G. A. Papakostas, D. E. Koulouriotis, A. S. Polydoros, V. D. Tourassis, 'Towards Hebbian learning of Fuzzy Cognitive mapin pattern classification problems' *Expert Systems with Applications*, Volume 39, Issue 12, 15 September 2012, Pages 10620-10629

[14] Ghaderi et al., S. F. Ghaderi, A. Azadeh, B. P. Nokhandan, E. Fathi, 'Behavioral simulation and optimization of generation companies in electricity markets by fuzzy cognitive map'. *Expert Systems with Applications*, 39 (5) (2012), pp. 4635–4646

[15] Kun Chang Leea, Habin Leeb, Namho Leea, Jaehoon Limb, 'An agent-based fuzzy cognitive map approach to the strategic marketing planning for industrial firms', *Industrial Marketing Management*. Volume 42, Issue 4, May 2013, pp. 552–563

[16] Tintarev and Masthoff, N. Tintarev, J. Masthoff 'Evaluating the effectiveness of explanations for recommender systems—methodological issues and empirical studies on the impact of personalization' *User Model. User-Adapt. Interact.*, 22 (4–5) (2012), pp. 399–439

Algorithm Optimization of Resources Scheduling Based on Cloud Computing

Zhongli Liu, Hangjun Zhou, Sha Fu, and Chaoqun Liu
Department of Information Management, Hunan University of Finance and Economics, Chang Sha, Hu Nan, 410205, P. R. of China

Huizhi Liang
School of Information Technology, Queensland University of Technology, Brisbane, Australia

***Abstract*—In some service areas, such as cloud computing of virtual technology, the specific virtual technology's level configuration processing power is limited and it has defects.[1] Under normal circumstances, the traditional cloud computing resource sharing configuration can not be effectively applied, for it has a time difference and several machines can not be connected to the network by a broadband link and work simultaneously, making the virtual technology particularly difficult in the efficiency and effectiveness of the implementation. This paper presents a cloud computing of ant colony and genetic combination of resources virtual scheduling optimization, which is applied to multiple machines physical layer simultaneously, at the same time it can optimize the allocation of resources to maximization. Experimental results show that the model algorithm is more convergent and more practical than traditional methods.**

***Index Terms*—Multi-processor; Objective Function; Cloud Utility; Node**

I. INTRODUCTION

The rise of cloud computing is gradually changing the entire computer industry and academia, the way which uses computer as the utility tool and software as a service brings great changes to the computer software and hardware design [1]. Cloud computing is an information service mode, which can distribute user tasks to a resources pool form by a large number of physical or virtual organizations, allows users to access of required information service, cloud computing resource scheduling mainly includes five features, as Figure 1.

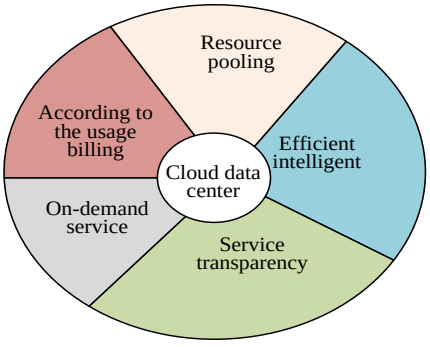


Figure 1. Feature of cloud computing resource scheduling

Since the existing traditional cloud resources data center does not have a global two-way choice transaction, has not fully considered the multidimensional cloud resource's scheduling and communication energy between cloud tasks. Therefore, cloud resource providers of the existing cloud data center has low income, multidimensional cloud resource utilization is low and the cloud data center energy consumption is high. Cloud users are service users at the mean time they can gain income by providing resources [2-6]. An increasing number of available cloud resources added to the cloud data center, thus alternative cloud service resources is increasing. Cloud service users can obtain more cost-effective service, and cloud resource providers can also obtain greater benefits. Buyers need to pay fees for using the service. In the continuous developing process of cloud computing applications, the user scale of the managing resources and access in cloud resources platform will further expand. The access of vast resources information, users' resource request with high concurrency and vast resource update bring huge challenges for resource management systems, which makes the resource efficient management and allocation become one of the key issues that must be addressed [7-11].

Cloud resource scheduling computing is a hot research topic at home and abroad, it plays an important role in promoting the development of information. Hawei Tang, Wei Lin, etc., proposed pricing mechanism based on the consumption of resources thinking, which can put cloud resource providers and users together, allowing users to run computing tasks in the public cloud [12-13]. Wu Pinyun and others proposed a secondary positioning pricing mechanism based on dynamic auction mechanism to solve the problem of resource allocation [14-18]. An important element of the cloud service secondary pricing mechanism is that cloud service providers can ensure a reasonable return and efficient computing resource allocation [19-25]. Service providers have two tasks: first, run-time monitoring and allocation of resources to the cloud user. After determining the amount of resources allocated to the previous task, the cloud service provider will sell the remaining resources to the newly arrived cloud user tasks. After the auction quoting phase, the user

doi:10.4304/jmm.9.7.977-984

submits their offer to cloud service providers, cloud service providers to collect all the quotes and then determine the price. Second, they propose a theoretical framework to solve the problem of capacity allocation under the cloud computing framework. Under the direct decision-making rules, secondary positioning price mechanism can ensure effective ability distribution and to make cloud service providers obtain greater benefits, but the efficiency of the system is not high and the cloud service provider's income is relatively lower. And the cloud computing resource scheduling is traditional, which does not solve the fundamental problem of information storage services fundamentally. In the way of building and distributing software, SOA (service-oriented software architecture) and SAAS are increasingly becoming the core way of software building and distribution under cloud computing software mode. And for the infrastructure supporting the running underlying computing services, the challenges are focused on two areas: system software and hardware, in which, adaptive resource allocation is the biggest challenge of the system software, and how to allocate the resources is the core problem of this challenge. According to peak load to static allocate the resources will cause serious waste of resources. According to statistics, in the current data center the utilization rate of CPU resources is average between 5% to 20%, and the utilization rate of memory resources is not more than 60%. Clearly, the current resource allocation model has a high degree of resource waste, which increases the user's total cost ownership (TCO), while static allocation of resources based on the average will decrease the service quality under the situation of peak load and ultimately it will affect potential profit of cloud users. And Tniee Marry, Harry Li, etc. proposed a new scheduling model of cloud computing, to increase rental income of cloud services. In this new model, service providers and users follow a service level agreement. The scheduling model allows re-allocation of cloud resources after the initial distribution. The focus of the model is to provide fair trade between users and resource providers, while optimizing revenue it also can increase trading service volume, and maintain a reasonable price considering the cloud market. Based on a true cloud computing system, it designed cloud resource scheduling algorithm based on service level agreements. The cloud resource scheduling algorithms based service level agreements can provide liaison between service providers and users and timely provide services for users.

Introduction of virtualization technology has brought the following benefits: the more fine-grained resource sharing, security isolation and high reliability. Through the sharing multiplexing technology with the virtual machine as the granularity, the virtualization platform can effectively make the servers consolidation to improve resource utilization, and ultimately improve the overall resource utilization of cloud computing infrastructure. But now the virtualization technology focuses on resource optimization within a single physical server, and through resources multiplexing technology the resource utilization of a single server is improved. For the global resource optimization, because of the lack of virtualization supporting technology of accessing remote resources, studies on global optimization conduct more research from the perspective of the operating system, such as distributed shared memory systems (DSM) and network memory, etc. These works will need to modify the operating system source code or insert the corresponding driver to use remote resources. As for deploying the cloud data centers with multiple operating system platforms it increases the cost of deployment. Overall, the use of virtualization technology has two major challenges for building the cloud data centers of operating system transparent: (1) how to build a more flexible and reliable global resource space and reliable global resource space, and to establish the framework of resource flows to achieve resource optimization a crossing physical machine border; (2) how to dynamically identify the idle resources to provide flow resource space for the resource optimization.

This paper mainly makes the extensive and innovative work in the following areas:

(1) For mapping defects of traditional virtual machine, this paper proposes the cloud computing resource scheduling model. The cloud resource scheduling maximizes network utilization of effective mechanism, to achieve cloud computing resource sharing for traditional physical identification and distribution levels, as well as network's broadband connection link, improving user's satisfaction of resource sharing. It is better than traditional model.

(2) In order to further validate the new model of cloud computing resources' effectiveness of the algorithm in physical computing is stronger than that of traditional one, this paper conducts simulation experiments on cloud computing model 's all levels of ant colony and agenetic combination's optimized algorithm. The results show that new cloud computing resources with virtual technology has better sharing and convergence, and it can effectively solve issues such as virtual resource's optimal allocation and sharing.

II. RESOURCES SCHEDULING MODEL WITH HIGH UTILIZATION

Resources scheduling model with high utilization is shown in Figure 2.

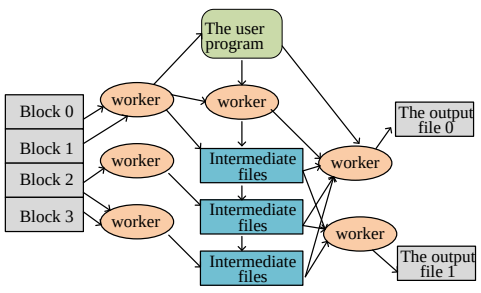


Figure 2. Highly resource model

Highly cloud resources mainly uses centralized scheduling mode, so there is a master scheduler, all

scheduling requests must be submitted to the master scheduler, the master scheduler decides how to allocate resources in response to user requests and so on. The model is completed by Master, Worker and User. Among them, Master acts as the centralized system's central scheduler and it is responsible for receiving user's request, the data segmentation, task scheduling and balance control, etc.; Worker is responsible for receiving the tasks sent by Master and carrying out specific tasks, such as data processing, communications and computing, saving the computing data and resulting data in appropriate storage location.

High resource scheduling utilization's problem can be described as:

$$\min f(x_1, x_2, ..., x_n)$$
$$ca_j = \overline{ca}_i + \frac{\sum_{i=1}^{m} dev_{i,j}}{M-1} \quad (1)$$

where in: a_j and b_j is an integer and they are the bounds of variable x_j; n is the number of variables x_j.

Defined delt $a_j = (b_j - a_j)/(l_j - 1)$, l_j is the number of values of the variables x_j, $j = 1, 2, ..., n$. Feasible solution space is shown in Figure 3.

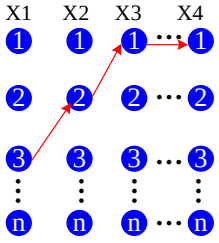


Figure 3. Feasible solution space

In Figure 3, x_j has l_i nodes, and each variable in a node of a column, which can form a solution $(x_1, x_2, ..., x_n)$ of the optimization problem. Node number of each column 1 ~ l_j is defined as the code of the selected nodes of the j_{th} variable node. A set of solution is as follows:

$$(x_1, x_2, ..., x_n) = (a_1 + (m_1 - 1)\bullet delta_1, a_2, + (m_2 - 1)\bullet delta_2, ..., a_n + (m_n - 1)\bullet delta_n) \quad (2)$$

n variables constitute the decision problem with n-level; the j_{th} level (the j_{th} variables) has l_j nodes; at the beginning m ants are on the first level; the j_{th} level select the probability of the i_{th} node.

$$diff(c_i, c_j) = \frac{cv_i . cv_j}{|cv_i| . |cv_j|} \quad (3)$$

τ_{ij} is the intensity of attracting of the cv_i node on the j-level. Update equation is as follows:

$$m_{ij}^{(1)} = [\int_s \phi_i \phi_j dA] I_3, m_{ij}^{(2)} = \int_s \varphi_i \varphi_j dA;$$
$$l_{ij}^{(1)} = \begin{cases} 2\int_s [\phi_i \varnothing \varphi_j - n(\nabla s \varphi_j)^T \nabla s \varphi_j] dA, \text{ for SDF} \\ \int_s [\phi_i [\varnothing \varphi_j + 2n(H^2 - K)\varphi_j] - n(\nabla s \phi_i)^T \nabla s \varphi_j] dA, \text{ for WF} \end{cases} \quad (4)$$
$$l_{ij}^{(2)} = -\frac{1}{2}\int_s [\varphi_i (\varnothing \phi_j)^T - (\nabla s \varphi_i)^T \nabla s \phi_j n^T] dA$$

where in: ρ represents the attenuation coefficient of intensity; generally 0 67 ~ 0.89; Q is the positive number; f is the value of objective function.

III. Cloud Resource Scheduling Model

*N*M* cloud resource scheduling model (referred to as the CRS model) is consists of n resources and m users; the cloud resource scheduling problem can be described as a four-tuple, *M=(R, U, Of, Θ)*. Wherein, R is the resource collection composed of n resources; U is users collection composed of m users; Of presents the scheduling objective function of cloud scheduling system; Θ indicates the scheduling optimization algorithm. Its specific characteristics are described as follows:

Cloud collection $R = \{r_1, r_2, ..., r_n\}$ is consisted with n resources and the i_{th} resource $r_i = \{r_i(1), r_i(2), ..., r_i(t)\}$, $0 \le i \le n$, which contains t characteristic attributes with different dimensions; $r_i(k)$, $0 \le k \le t$ i_{th} are the r_i characteristic attribute of the i_{th} resource r_i. In a service period T, the available service time of resources r_i obeys the Poisson distribution and each resource is independent, namely:

$$\forall r_i \in R, \bigcap_{i=0}^{n} r_i = \Phi \quad (5)$$

The available services rate vector of cloud collection R $s_r = \{s_{r1}, s_{r2}, ..., s_{rn}\}$, where in, $s_{ri} (0 \le i \le n)$ is the available service rate for the resource u_j.

Cloud user set $U = \{u_1, u_2, ..., u_m\}$ is constituted by m users, and at any time user u_j only has one task ts_j, and the ts_j task $ts_{jk} = \{ts_{jk}(1), ts_{jk}(2), ..., ts_{jk}(h)\}$, $0 \le j \le m$ of the user u_j contains the characteristic properties with h different dimensions. The task of the cloud user set U generates the rate vector $t_r = \{tr_1, tr_2, ..., tr_m\}$, where in ts_j, $\forall ts_j \in U$ u_j is the task generation rate of the ts_j user u_j, and each task is independent.

$$\forall ts_j \in U, \bigcap_{j=0}^{m} ts_j = \Phi \quad (6)$$

$r_i \in R$, $ts_j \in U$; if the resource r_i can satisfy the scheduling requirements of task ts_j i.e. $\Theta(r_i, ts_j) > 0$, the task ts_j can be assigned to resource r_i to perform. Normally there are several resources r_i which can meet

mission ts_j. In order to achieve effective load balancing system, the task ts_j can be scheduled to resource r_i with the probability p_{ij}, and then the cloud system is the probability matrix.

$$P_{n\times m}=(p_{ij})_{n\times m}, P_{ij}\geq 0, \sum_{i=0}^{n} p_{ij}=1 \tag{7}$$

For the resource scheduling matrix $\begin{cases}\dfrac{\partial x}{\partial t}=-2\Delta_s Hn,\ S(0)=S_0\\ \partial S(t)=\Gamma\end{cases}$ of cloud system, if $\Theta(r_i,ts_j)>0, x_{ij}=1$, $x_{ij}=1$. Otherwise, $x_{ij}=0$. Meanwhile, in order to ensure the task ts_j is successfully completed, there is a need to meet constraints of formula (8); in order to ensure resource r_i can accept the assignment within the capabilities of the service, there is a need to meet constraint condition of formula (9); in order to ensure cloud systems is in a stable environment, there is a need to satisfy the constraints of formula (10)

$$\sum_{i=0}^{n} p_{ij}\bullet x_{ij}=1 \tag{8}$$

$$\begin{cases}\dfrac{\partial x}{\partial t}=-\Delta_s^2 x,\ S(0)=S_0\\ \partial S(t)=\Gamma\end{cases} \tag{9}$$

$$\sum_{i=0}^{n}\sum_{j=0}^{m} p_{ij}\bullet x_{ij}\bullet tr_j\leq \sum_{i=0}^{n} sr_i \tag{10}$$

IV. Optimal Algorithm Integration of Resource Model Combined with Genetic and Ant Colony

A. System Features

Because the virtual resource has the characteristics like heterogeneous, distribution and dynamic in cloud environment, it is important to conduct a relatively comprehensive distributed scheduling algorithm simulation environment. Open source cloud computing simulation platform Cloudsim develop the function library in the discrete event simulation package Simjava, which inherites the programming model of Gridsim and supports the research and development of cloud computing. Cloudsim is advantage for accelerating the algorithm and method of the cloud computing and development of norms, which is an ideal simulation platform. Virtual resource collection: Suppose there are k virtual resources in cloud environment, and the virtual resources is defined as follows:

$$W(S):=\int_S R^2 d\sigma \tag{11}$$

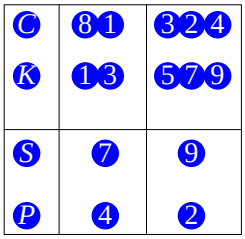
Expected execution saves the completion time matrix: the required time of resources $d\sigma$ executing task R^2 under the no load conditions, denoted as ETC_{ij}. All value of ETC_{ij} elements make up the ETC matrix.

ETC matrix corresponding to c_1 tasks and g_1 virtual resource are given as follows:

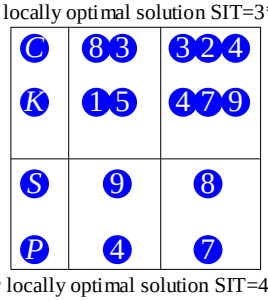
$$(c_1, g_1, r_1)\bullet(c_2, g_2, r_2)=(c_1\oplus c_2, g_1\odot\sigma^{r_1}(g_2), r_1+r_2)$$

B. Algorithm's Optimization

In the existing resource optimization in cloud computing, cloud resources users' needs is always greater than cloud resources' actual demand. In the past, there were kinds of algorithms' cloud resource scheduling model in the traditional cloud computing resource scheduling, but most algorithms are unable to play its full role more of virtual resource space, on the effect of optimized iterative method for solution, traditional general methods will at a certain extent depend on the batch rules used by the algorithm, resulting in a bottleneck to optimizing algorithm's performance. Thus this paper presents cloud resource scheduling model of combining both genetic and ant colony algorithm, which can effectively overcome this shortcoming. There are both advantages and shortcomings of this algorithm, both are global optimization algorithm which are unsure on probability. It starts from the initial population, utilizing factor's crossover and mutation to make the population can quickly get the optimal solution of iterative function. In the optimization process, it does not depend on the objective function of the problem itself, without any restrictions of end conditions, and seeks its own optimal solution under uncertain limited conditions. Its status before and after optimal solution are schematically shown in Figure 4.



(a) Before locally optimal solution SIT=3*8+4*5=41



(b) After locally optimal solution SIT=4*8+5=45

Figure 4. Schematic diagram of iterative optimization algorithm for solution

In the cloud computing model, resource scheduling model combining genetic and ant colony is an optimized mathematical model. In this model, cloud storage users can submit resource requirements to cloud computing resources, asking it to provide services. By using his model, cloud resources can be stored and computed, and

the function of the corresponding time can be gained. Algorithm is as follows:

At the initialization of $S(0)$, place N ants randomly in the cloud resource scheduling N, while initialize information $S(0)=S_0$ as a constant value. Ants and genetic decide the direction in accordance with the intensity of the cloud resources on the path, the following formula represents the probability of T transferring from cloud resources B to cloud resources U for virtual storage while initializing ants.

$$\begin{cases}\dfrac{\partial x}{\partial t}=-\Delta_s^2 x,\ S(0)=S_0\\ \partial S(t)=\Gamma\end{cases} \tag{12}$$

Among them, $\partial S(t)$ represents the intensity of cloud resources at the moment T, $\Delta_s^2 x$ 7 indicates the type of information stored, which is defined in terms of the structural characteristics of the studied question of resources. For TSPI problem is defined as U, representing that the shorter the distance speed between the memory used, the greater the probability of being selected. Y represents the relative weight of information resource's optimization scheduling, U is the candidate list of a combination of ants and genetic, i.e. the untapped cloud resource integration. When the access to all cloud resources is finished, stored information of all paths through can be iteratively updated according to formula (13):

$$\begin{cases}\int_s \dfrac{\partial x}{\partial t}\phi dA+2\int_s[\phi\varnothing R-n(\nabla_s\phi)^T\nabla_s R]dA=0,\\ \forall\phi\in R^I(S)\\ \int sR\varphi dA-\dfrac{1}{2}\int_s tr(\varnothing x)\varphi dA=0,\forall\varphi\in R^I(S)\end{cases} \tag{13}$$

Among them, p is the evaporation coefficient, v represents the information strength of stored cloud resources released in path K by the ant and genetic in this cycle Y, which can be defined as:

$$\begin{cases}\dfrac{\partial x}{\partial t}=-2\Delta_s Hn,\ S(0)=S_0\\ \partial S(t)=\Gamma\end{cases} \tag{14}$$

C represents the path length of cloud resources storage established by combined a of ant and genetic. It can be seen that the better the path built by ant, the better resource pheromone can be got by each side. Set T = T +1, and then repeat the above process of building and updating information pheromone, until it reaches the termination condition of iterative algorithm.

Ant Colony and genetic optimal algorithm's process is simplified as follows:

Setting the initial point - $X^{(1)}$ is a wholly-zero matrix, that is, any element $x_{ij}=0$, $\lambda^{(1)}=[1,1,...]^T$, $\mu^{(1)}=[\omega_1,\omega_2,...,\omega_n]^T$, Tolerance $\varepsilon>0, k=1$.

Calculating the sub-gradient: $g^{(k)}$

Take λ,μ into Lagrangian function L; let $\dfrac{\partial L}{\partial x}=0$, after solving equations, get the target $x^{(k)}$ with the k iterations is obtained.

If $\| g(k)\|\le\varepsilon$, stop the iteration, otherwise seek step $s^{(k)}=\dfrac{\overline{L}-L(x^{(k)},\lambda^{(k)},\mu^{(k)})}{\| g^{(k)}\|^2}$.

Using Equation (11) to update the Lagrangian operator and solve the $\lambda^{(k+1)}$ and $\mu^{(k+1)}$. $k=k+1$, turn to step 2.

Where in, each element in the matrix represents the expected time of execution of the i_{th} task on the j_{th} virtual resource.

Algorithm steps:

Setting the initial point - $X^{(1)}$ is a wholly-zero matrix, that is, any element $x_{ij}=0$, $\lambda^{(1)}=[1,1,...]^T$, $\mu^{(1)}=[\omega_1,\omega_2,...,\omega_n]^T$, Tolerance $\varepsilon>0, k=1$.

Calculating the sub-gradient: $g^{(k)}$

Take λ,μ into Lagrangian function L; let $\dfrac{\partial L}{\partial x}=0$, after solving equations, get the target $x^{(k)}$ with the k iterations is obtained.

If $\| g(k)\|\le\varepsilon$, stop the iteration, otherwise seek step $s^{(k)}=\dfrac{\overline{L}-L(x^{(k)},\lambda^{(k)},\mu^{(k)})}{\| g^{(k)}\|^2}$.

Using Equation (11) to update the Lagrangian operator and solve the $\lambda^{(k+1)}$ and $\mu^{(k+1)}$. $k=k+1$, turn to step 2.

According to the need setting the allowable error ε the different optimal solution precision is obtained. In the algorithm the estimated optimal target L is used to estimate through the heuristic search algorithm, it can be get by directly using matlab, LINGO software. In CUM simplified gradient algorithm; after adjusted by step 4, the effect of the results of this bias can be reduced.

C. Settings of Algorithm's Parameter

In the ant colony and genetic's optimization algorithm, a large number of simulation experiment the algorithm of target parameters are made, and finally this paper proposes to set common parameters of the algorithm, as shown in Table 1.

TABLE I. OPTIMIZED PARAMETERS' SETTING OF COLONY AND GENETIC ALGORITHM

Parameters	α	β	U	S	s_{ij}
Genetic algorithm	22	m	0.6	4	$\min/\sum_i^u 0$
Ant colony algorithm	22	m	0.6	5	$\min/\sum_i^{u\backslash o} 0$
genetic's optimization algorithm	24	m	0.2	2	$\min/\sum_i^{o/ii} 0$

In Table 1, respectively, three parameters settings of the ant colony algorithm, the genetic algorithm and the optimized algorithm combined with ant colony and

genetic are set. It is not difficult to be seen from Table 1 that the optimized algorithm combined with ant colony and genetic is able to obtain optimal solution of cloud resource scheduling, and also it has a distinct advantage in the algorithm parameters' setup. The key parameters that affect the performance of the ant colony and genetic algorithm optimization is the number of ants gene M. When the number of ants increases, algorithm can get the optimal solution of optimized cloud resources, but meanwhile the computational time will increase sharply, thus it becomes very important to require a balance of computational efficiency in this two algorithm's combined optimization. But with the number of ants up to 3", the single colony algorithm's solution will not be a great improvement, on the contrary it will add burden to the time for solving, extending the time rate of calculation. Because when the ant parameters are set as fixed 3, in order to achieve balance and maximum optimization in terms of time and quality, making the corresponding objective function value will increase and at a certain promoting release of cloud resources information sharing. But when the value is set as R is L, so that virtual storage of cloud resource scheduling increases as the number of tasks and physical increase and thus seek the optimal solution. Thus relatively better objective function solution through iterative solution can be gained when algorithm's parameters are from 3 to 5.

V. Experimental Simulation and Analysis

A. Comparison Between Algorithm's Cloud Resources Rate

In order to verify the capacity of cloud computing scheduling model is faster and stronger than that of traditional one in resource sharing, storage and utilization, this paper analyzes and compares cloud resources' utilization rate with several sets of experimental data, to further prove of the correctness of the experiment. The specific experiments are analyzed as follows:

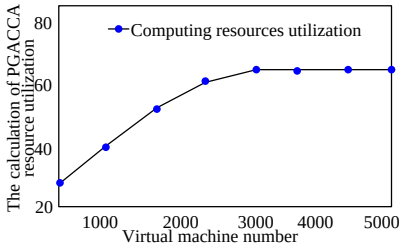


Figure 5. Cloud computing resource utilization

In Figure 5, the ordinate is the number of virtual technology of resource scheduling, the red line is cloud computing resource's utilization growth rate of greedy algorithm, the blue one indicates cloud computing resource's utilization growth rate of ant colony algorithm, while yellow one is cloud computing resource's utilization growth rate of parallel genetic algorithm. From Figure 5, we can see that in the experiment of three different resource scheduling algorithm, as the number of virtual technology increases, namely when it creases from 120 to 6000, the corresponding growth rate will increase, from 37.4% to 60.5%, and the average also increases to 51.2%, showing that the utilization of cloud computing resource scheduling has a significant increase. Meanwhile in the experiment, from the three coordinate curves with different colors, it can also be seen that when the curvilinear coordinates of different algorithms increase sharply then reach a steady state gradually. This shows that the proposed algorithm's performance is the best and most stable in cloud computing resources's utilization. Because the algorithm combines both the advantages of the ant colony and genetic algorithm, so the calculation of interest rates of this two combined cloud computing resource scheduling algorithm is significantly higher than that of greedy algorithm. Similarly, with the gradual increase of the number of virtual technology, fragmentary resources also increase, which leads to resources' increasing.

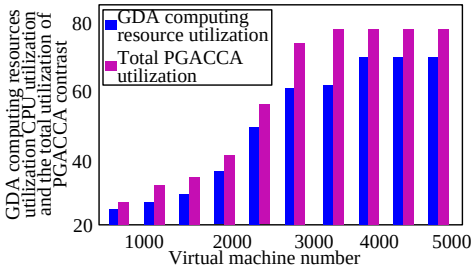


Figure 6. Comparison of several different algorithms' resources virtual technology

From the experimental data in Figure 6, it can be seen that abscissa is number of resources scheduling's virtual technology, and the vertical coordinate is the trend of genetic and ant colony's combined algorithm's utilization rate of cloud resources. Through this 10 groups of experimental data with different trends, it can be found that as the number of virtual technology increases, namely when it creases from 120 to 6000, the corresponding growth rate will increase, from 37.4% to 60.5%, and the average also increases to 51.2%, showing that the utilization of cloud computing resource scheduling has a significant increase. From the four coordinate curves with different color of different algorithms, it also shows the blue line's trend. The curvilinear coordinate first increases sharply then reaches a steady state gradually. This shows that the proposed algorithm's performance is the best and most stable in cloud computing resources utilization, and it has a obvious advantage. Because the algorithm combines both the advantages of the ant colony and genetic algorithm, so the calculation of interest rates of this two combined cloud computing resource scheduling algorithm is significantly higher than that of greedy algorithm. Similarly, with the gradual increase of the number of virtual technology, its computational resource's storage, usage, bandwidth, etc. will also increases as the calculating number increases, thereby increasing the magnitude of its resources, and further shows that the resource utilization of the algorithm is the most reasonable. It can fully meet the needs of cloud

resources' users and the characteristics of cloud computing data center's system architecture.

B. Algorithm Convergence

Assuming L is the optimal solution of resource scheduling model, L is a function value of the algorithm. During the optimization of ant colony and genetic algorithm's combination, a brown factor u of the algorithm will be gained after k +1 times' iteration, and the factor u will be closer to the u after N iterations, expressed as follows:

$$\| u^*-u^{(k+1)} \| < \| u^*-u^{(k)} \|, \forall k \quad (15)$$

According to the comparison of nature of the algorithm's Brown factor u , it can be found that:

$$g^{(k)}(u^*-u^{(k)}) \geq L^*-L^{(k)}$$

So,

$$2g^{(k)}(u^*-u^{(k)})-(L^*-L^{(k)})>0$$

Then

$$-s^{(k)}(-2g^{(k)}(u^*-u^{(k)})+L^*-L^{(k)})<0,$$

$$\| u^*-u^{(k+1)} \|^2 < \| u^*-u^{(k)} \|^2$$

And then,

After the proving of equation (15), So the optimized algorithm combining ant colony and genetic algorithm is convergent.

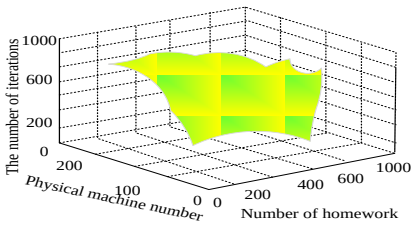


Figure 7. Surface chart of scale of the problem and the number of iterations

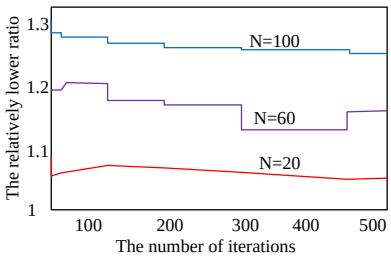


Figure 8. Iteration times of algorithm

In order to verify the convergence of the algorithm, this paper conducts comparison between the number of physical machines and tasks of cloud resource scheduling. Figure 7 shows the iterative relationship trend of ant colony and genetic algorithm. Figure 8 shows the number of iterations of the algorithm. In the experiment, it can be seen from Figure 8 that after the resources user input a question of tasks with size M and N physical machines, the combined algorithm optimizes the function iterative, the complexity of the algorithm is $o(mn)$. As can be seen from the comparison of experimental data, when the number of tasks is 40, the number of physical machines is 40, its number of iterations of the algorithm is 154. When the number of tasks increases, its value as 250 and the number of physical machines is 150, the times of iterations of cloud resource scheduling algorithm reach 498. Then further set a larger scale. when the number of tasks is up to 1200, the number of physical machines is 400, then the times of iterations of resource scheduling algorithm's virtual technology can be 750, and it is the biggest number. Experiments further show that: in different size, resource users enter a different question, the algorithms' iterative function's value can all get the optimal solution, indicating that the algorithm has convergence.

C. Application of the Model

VM model can be applied to the cloud environment with a larger scale. Figure 9 shows the application mode.

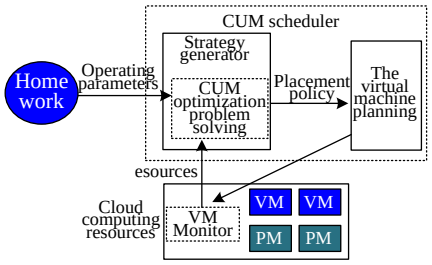


Figure 9. Application scenario of VM in the cloud

In Figure 9, VM global scheduler real-time receive the computing resource usage sent by the virtual machine monitor (VM Monitor). According to the job and remaining resources to determine the maximized effectiveness of the virtual machine placement policy, create the virtual machine and assign jobs to the virtual machine. Compare with the job scheduler in traditional clusters and grid computing environment, firstly VM scheduler determines the virtual machine placement strategy, namely the instance of how many instances of physical computing resources into a virtual machine, and then assign the jobs to a virtual machine for execution. In a real cloud environment, the jobs preprocessing module can be increased in the front of VM scheduler. Through the user bids and the user's preference the job level is divided, and it is set from the high and low. The value of willing to pay wi are set with different value. After grading the jobs are submitted to VM scheduler for scheduling. VM model uses the utility maximization as the optimization objective and this optimization approach has achieved good results in the computer network bandwidth allocation. The carried business of the computer network has the following characteristics: at the same time there are a lot of business reached for service, and the quality of service requirements are different, and in the peak the network resources are basic full load. So when evolving cloud computing, users' demand for computing resources are as common as the demand on the network, and the scheduling advantage of VM model becomes more apparent.

VI. CONCLUSION

In order to overcome the defects of cloud computing resources' scheduling and optimization, this paper proposes resources scheduling algorithm's optimization based on cloud computing. By comparing the experimental curves and the optimized solution process of ant colony algorithm, genetic algorithm and combined algorithm of ant colony and genetic, the experimental results show that: the optimal algorithm combined with ant colony and genetic enables the storage, computation and operation of cloud resources to reach the optimal solution, meeting the user's need for using cloud resources, and it further proves that the algorithm is convergent.

ACKNOWLEDGMENT

This research work is supported by the Hunan Science and Technology Project (No. 2012FJ6011) and the Construct Program of the Key Discipline in Hunan Province, China.

REFERENCES

[1] Xin Huang, Xiao Ma, Bangdao Chen, Andrew Markham, Qinghua Wang, Andrew William Roscoe. Human Interactive Secure ID Management in Body Sensor Networks. *Journal of Networks*, Vol 7, No 9 (2012), 1400-1406

[2] Ramesh Raskar, Immersive planar display using roughly aligned projectors. *New York, USA, Proceedings of IEEE Virtual Reality*, 2000.

[3] Chen Y, Clark D W.' Automatic alignment of high-resolution multi-projector displays using an un-calibrated camera, *IEEE Visualization*, 2000.

[4] Jannick P. Rolland and Kevin P. Thompson, Freeform optics: Evolution? No, revolution SPIE Newsroom 10. 1117/2. 1201207. 004309, pp. l-3, 2012.

[5] Vu Q H, Ooi B C, Rinard M, et al. Histogram-based global load balancing in structured peer-to-peer systems. Knowledge and Data Engineering, *IEEE Transactions on*, 2009, 21(4) pp. 595-608.

[6] Hsiao H C, Liao H, Chen S T, et al. Load Balance with Imperfect Information in Structured Peer-to-Peer Systems. Parallel and Distributed Systems, *IEEE Transactions on*, 2011, 22(4) pp. 634-649.

[7] Jones P, EASTLAKE D E. US secure hash algorithm 1 (SHA1). 2001 pp. 902-910

[8] Jelasity M, Voulgaris S, Guerraoui R, et al. Gossip-based peer sampling. ACM Transactions on Computer Systems (TOCS), 2007, 25(3) pp. 8.

[9] Stoica I, Morris R, Liben-Nowell D, et al. Chord: a scalable peer-to-peer lookup protocol for internet applications. *Networking, IEEE/ACM Transactions on*, 2012, 11(1) pp. 17-32.

[10] Kevin P. Thompson, Pablo Benitez and Jannick P. Holland, Freeform Optical Surfaces - Report from OSA's First Incubator Meeting, *OPN Optics & Photonics News*, pp. 32-37, 2012.

[11] J H. Z. Guo, Q. C. Chen, C. Xin, et al. "A Length-variable Feature Code Based Fuzzy Duplicates Elimination Approach for Large Scale Chinese WebPages", *Journal of Software*, 2012, 7(11), pp. 262-269

[12] K. Atual, J. Weissman, "ViDeDup: An Application-Aware Framework for Video De-duplication", *In Proc. of the 3rd USENIX conference on Hot topics in storage and file systems*, 2011, pp. 7-7

[13] Xiao-bo Li, Quan Zhou, "A lossless data hiding transmission method for satellite remote sensing image based on histogram modification, " *Journal of Astronautics*, vol. 34, no. 5, pp. 686- 692, May, 2013.

[14] SHENG L, GANG W, STONES DS, et al. T-code: 3 Erasure Longest Lowest-Density MDS Codes. *IEEE Journal on Selected Areas in Communications*, 2010, 28(2) pp. 289-296.

[15] N. P. Ramaiah, "De-duplication of Photograph Images Using Histogram Refinement", *Recent Advances in Intelligent Computational Systems*, 2011, pp. 391-395

[16] A. V. Sreedhanya and K. P. Soman, "Secrecy of cryptography with compressed sensing," *International Conference on Advances in Computing and Communications*, pp. 207-210, 2012.

[17] C. Wengert, M. Douze, H. Jegou. "Bag-of-colors for Improved Image Search", *In Proc. of the 19th ACM international conference on Multimedia*, 2011, pp. 1437-1440

[18] CHEN T W., SHEN G W., XU B H., et al. H-Code: A Hybird MDS Array Code to Optimize Partial Stripe Writes in RAID-6. // *Proceedings of the 2011 IEEE International Parallel & Distributed Processing Symposium, Anchorage, Alaska, USA, IEEE* Press, 2011 pp. 782-793.

[19] J. Tian, "Reversible watermarking using a difference expansion, " *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.

[20] L. An, X. Gao, C. Deng, and F. Ji, "Reversible watermarking based on statistical quantity histogram," *in Proc. Adv. Multimedia Inform. Process, LNCS 5879*. 2009, pp. 1300-1305.

[21] FENG G. -L., DENG R., BAO F., SHEN J. -C.. New Efficient MDS Array Codes for RAID, Part I: Reed Solomon Like Codes for Tolerating Three Disk Failures. *IEEE Trans. Computers*, 2005, 54(9) pp. 1071-1080.

[22] X. Zhang, Y. Ren, G. Feng and Z. Qian, "Compressing encrypted image using compressive sensing," *2011 7th International Conference on Intelligent Information Hiding and Multimedias Signal Processing (IIH-MSP)*, pp. 222-225, 2011.

[23] XIANG H L, JI W S. Summary of Research for Erasure Code in Storage System. *Journal of Compute Research and Development*, 2012, 49(1) pp. 1-11.

[24] P. Lu, Z. Xu, X. Lu and X. Liu, "Digital image information encryption based on Compressive Sensing and double random-phase encoding technique," *Optik-International Journal for Light and Electron Optics*, vol. 124, pp. 2514-2518, 2013.

[25] Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 2008, 51(1) pp. 107-113.

Instructions for Authors

Manuscript Submission

All paper submissions will be handled electronically in EDAS via the JMM Submission Page (URL: http://edas.info/newPaper.php?c=7325). After login EDAS, you will first register the paper. Afterwards, you will be able to add authors and submit the manuscript (file). If you do not have an EDAS account, you can obtain one. Along with the submission, Authors should select up to 3 topics from the EDICS (URL: http://www.academypublisher.com/jmm/jmmedics.html), and clearly state them during the registration of the submission.

JMM invites original, previously unpublished, research papers, review, survey and tutorial papers, application papers, plus case studies, short research notes and letters, on both applied and theoretical aspects. Submission implies that the manuscript has not been published previously, and is not currently submitted for publication elsewhere. Submission also implies that the corresponding author has consent of all authors. Upon acceptance for publication transfer of copyright will be made to Academy Publisher, article submission implies author agreement with this policy. Manuscripts should be written in English. Paper submissions are accepted only in PDF. Other formats will not be accepted. Papers should be formatted into A4-size (8.27" x 11.69") pages, with main text of 10-point Times New Roman, in single-spaced two-column format. Authors are advised to follow the format of the final version at this stage. All the papers, except survey, should ideally not exceed 12,000 words (14 pages) in length. Whenever applicable, submissions must include the following elements: title, authors, affiliations, contacts, abstract, index terms, introduction, main text, conclusions, appendixes, acknowledgement, references, and biographies.

Conference Version

Submissions previously published in conference proceedings are eligible for consideration provided that the author informs the Editors at the time of submission and that the submission has undergone substantial revision. In the new submission, authors are required to cite the previous publication and very clearly indicate how the new submission offers substantively novel or different contributions beyond those of the previously published work. The appropriate way to indicate that your paper has been revised substantially is for the new paper to have a new title. Author should supply a copy of the previous version to the Editor, and provide a brief description of the differences between the submitted manuscript and the previous version.

If the authors provide a previously published conference submission, Editors will check the submission to determine whether there has been sufficient new material added to warrant publication in the Journal. The Academy Publisher's guidelines are that the submission should contain a significant amount of new material, that is, material that has not been published elsewhere. New results are not required; however, the submission should contain expansions of key ideas, examples, elaborations, and so on, of the conference submission. The paper submitting to the journal should differ from the previously published material by at least 30 percent.

Review Process

Submissions are accepted for review with the understanding that the same work has been neither submitted to, nor published in, another publication. Concurrent submission to other publications will result in immediate rejection of the submission.

All manuscripts will be subject to a well established, fair, unbiased peer review and refereeing procedure, and are considered on the basis of their significance, novelty and usefulness to the Journals readership. The reviewing structure will always ensure the anonymity of the referees. The review output will be one of the following decisions: Accept, Accept with minor changes, Accept with major changes, or Reject.

The review process may take approximately three months to be completed. Should authors be requested by the editor to revise the text, the revised version should be submitted within three months for a major revision or one month for a minor revision. Authors who need more time are kindly requested to contact the Editor. The Editor reserves the right to reject a paper if it does not meet the aims and scope of the journal, it is not technically sound, it is not revised satisfactorily, or if it is inadequate in presentation.

Revised and Final Version Submission

Revised version should follow the same requirements as for the final version to format the paper, plus a short summary about the modifications authors have made and author's response to reviewer's comments.

Authors are requested to use the Academy Publisher Journal Style for preparing the final camera-ready version. A template in PDF and an MS word template can be downloaded from the web site. Authors are requested to strictly follow the guidelines specified in the templates. Only PDF format is acceptable. The PDF document should be sent as an open file, i.e. without any data protection. Authors should submit their paper electronically through email to the Journal's submission address. Please always refer to the paper ID in the submissions and any further enquiries.

Please do not use the Adobe Acrobat PDFWriter to generate the PDF file. Use the Adobe Acrobat Distiller instead, which is contained in the same package as the Acrobat PDFWriter. Make sure that you have used Type 1 or True Type Fonts (check with the Acrobat Reader or Acrobat Writer by clicking on File>Document Properties>Fonts to see the list of fonts and their type used in the PDF document).

Copyright

Submission of your paper to this journal implies that the paper is not under submission for publication elsewhere. Material which has been previously copyrighted, published, or accepted for publication will not be considered for publication in this journal. Submission of a manuscript is interpreted as a statement of certification that no part of the manuscript is copyrighted by any other publisher nor is under review by any other formal publication.

Submitted papers are assumed to contain no proprietary material unprotected by patent or patent application; responsibility for technical content and for protection of proprietary material rests solely with the author(s) and their organizations and is not the responsibility of the Academy Publisher or its editorial staff. The main author is responsible for ensuring that the article has been seen and approved by all the other authors. It is the responsibility of the author to obtain all necessary copyright release permissions for the use of any copyrighted materials in the manuscript prior to the submission. More information about permission request can be found at the web site.

Authors are asked to sign a warranty and copyright agreement upon acceptance of their manuscript, before the manuscript can be published. The Copyright Transfer Agreement can be downloaded from the web site.

Publication Charges and Re-print

The author's company or institution will be requested to pay a flat publication fee of EUR 360 for an accepted manuscript regardless of the length of the paper. The page charges are mandatory. Authors are entitled to a 30% discount on the journal, which is EUR 100 per copy. Reprints of the paper can be ordered with a price of EUR 100 per 20 copies. An allowance of 50% discount may be granted for individuals without a host institution and from less developed countries, upon application. Such application however will be handled case by case.

More information is available on the web site at http://www.academypublisher.com/jmm/authorguide.html.

(Contents Continued from Back Cover)