Journal of Multimedia

ISSN 1796-2048

Volume 7, Number 3, June 2012

Contents

A Survey on Video-based Vehicle Behavior Analysis Algorithms

Jian Wu[1,2,3]
[1.]The Institute of Intelligent Information Processing and Application; [2.]Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, Suzhou 215006, China;
[3.]Jiangsu Yihe Traffic Engineering Co., Ltd., Suzhou 215002, China
Email: szjianwu@163.com

Zhi-ming Cui[1,2,3], Jian-ming Chen[1,2,3], Guang-ming Zhang[1,2]
[1.]The Institute of Intelligent Information Processing and Application; [2.]Provincial Key Laboratory for Computer Information Processing Technology, Soochow University, Suzhou 215006, China;
[3.]Jiangsu Yihe Traffic Engineering Co., Ltd., Suzhou 215002, China
Email: szzmcui@suda.edu.cn

***Abstract*—Analysis of the Vehicle Behavior is mainly to analyze and identify the vehicles' motion pattern, and describe it by the use of natural language. It is a considerable challenge to analyze and describe the vehicles' behavior in a complex scene. This paper first hackles the development history of the intelligent transportation system and analysis of vehicles' behavior, and then conducts an in-depth analysis of current situation of vehicle behavior analysis from the video processing, video analysis and video understanding, summarizes the achieved results and the key technical problems, and prospects the future development of vehicle behavior analysis.**

***Index Terms*—vehicle behavior analysis; vehicle detection; vehicle tracking; behavior understanding**

I. INTRODUCTION

Video analysis and recognition is to make the computer achieve the understanding of what the show is or what events occurred by running certain core algorithms through specific programs to extract content information or individual movement information in the video signal information. The most critical technology in this process is analyzing the video signal by the use of computer, so that we can extract some specific events occurs or the specific behaviors of the monitored target occurred in the video scene. Video analysis and recognition technology is a computer vision technology, and it belongs to one branch of artificial intelligence research which can extract and understand the contents of the video screen through digital image processing and video signal analysis. Video analysis and recognition technology has a wide range of applications, speaking in its contented definition, actually it has been used in the transport sector, and vehicle license plate recognition technology is a typical example. There are two ways to recognize license plate, one is by the use of captured image triggered by the sense coil, and another is by the use of captured images triggered by videos. For the latter, it means to make continued analysis of images from the video camera by the use of certain core algorithms and extract and then recognize the license plate number from the video.

Vehicles can be regarded as a special object in intelligent video analysis technology, behaviors of vehicles contain prohibition state and motion state, and the detection of vehicle behaviors has important application value. Analysis of vehicle behavior is to detect vehicles in monitored scene, analyze its direction and trajectory and analyze the correctness and rationality of vehicle behaviors according to the preset rules. Analysis of vehicle behavior is mainly to analyze and identify the movement pattern of vehicles and use natural language to describe it. The behavior patterns achieve automatic detection likely through video analysis and recognition technology are listed below:

(1) Parking due to violation/failure/accidents

Parking phenomenon appears in the driveway or the no parking area is extremely dangerous event or illegal behavior easily causes traffic jams, no matter it is parking due to vehicle breakdowns or illegal parking, we need timely processing. The management needs to know accident parking in time so that traffic is restored as soon as possible.

(2) Illegal left and right turns

It is not allowed to turn left or right in some crossing, otherwise it not only easily leads to traffic congestion and easily leads to traffic accidents cause loss of life and property.

(3) Illegal lane change

In some sections, it is an illegal act if high-speed motor vehicles change their lines freely; it easily leads to traffic accidents.

(4) Violation of traffic line

In traffic regulations, violation of traffic line is a serious illegal act and should be dealt with severely.

(5) Illegal retrograde

In traffic regulations, retrograde is a very serious violation of regulations, it is not only easily to cause traffic chaos, and is easily to cause traffic accidents, and

doi:10.4304/jmm.7.3.223-230

it should be detected in time, stopped in time and dealt with in time.

(6) Spilling objects

It is likely to cause traffic jam if moving vehicles spill objects inadvertently, especially when motor vehicles are high speed, the car in front spilling objects can easily lead to a serious car accident to the latter.

(7) Vehicle fire

Vehicle fire due to breakdown or other reasons is a very serious event; it can seriously block the traffic, there may even be the possibility of the spread of fire or explosion, it should be processed in the shortest time.

(8) Analysis of traffic congestion

Analyzing the traffic congestion on the roads or achieving optimal control of signal lights can help to ease traffic flow in time.

Vehicle Behavior Analysis has a wide range of applications and has become the research focus of security monitoring, intelligent transportation and other fields. This paper summarizes the vehicle behavior analysis algorithms current widely used based on video. It is organized as follows: Section II describes the history of video-based vehicle behavior analysis; Section III summarizes the most widely used vehicle behavior analysis algorithms based on video from the video processing, video analysis and video understanding the three levels; Section IV summarizes the existing key technical problems in video behavior analysis; Section V describes the evaluation methods of vehicle behavior analysis algorithm; Section VI summarizes this paper. The review of traffic incident detection algorithm is introduced in Literature [1,2], but not exhaustive or systematic. Currently the vehicle behavior analysis has become an independent research, it is necessary to conduct an independent, detailed and systemetic review for the behavior analysis algorithm.

II. History and Present Situation

The development of intelligent transportation systems can date back to the development and application of some new technologies of a series of traffic diversion system in Seventies and eighties in the 20th century. In 1991, the United States passed the "Surface Transportation Efficiency Act", and then IVHS research in the United States entered the operational phase of macro. In 1994, IVHS was renamed ITS in USA. After then, Europe, Japan also have joined in. After 40 years of development, United States, Europe, Japan become the three bases of the ITS. Currently, ITS study also has a large scale in other countries and regions, Such as Australia, Korea, Singapore, Hong Kong and so on. It can be said that a new ITS industry is being formed in the world and all kinds of projects are being carried out, whose scale and speed of development is staggering. The concept of ITS is gradually formed in the world, whose goal is "Protect the safety, increase efficiency, improve the environment and conservate energy". From the perspective of development of intelligent transportation systems, we count statistics of landmark events associated with intelligent transportation system occurred every year, and get Figure 1 which shows the overview of intelligent transport systems.

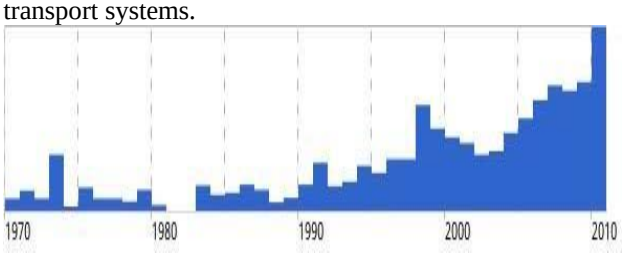


Figure 1. Development Overview of ITS

It is generally believed that a complete management system of ITS should be a combination of these various systems below: Traffic management system, public transportation systems, information service systems, surveillance systems, security systems and logistics management system. Among them, the vehicle behavior analysis is an important part in the intelligent transportation system, whose research originated in the 20th century, 70 years. Figure 2 shows the number of articles published on vehicle detection, vehicle tracking and traffic incident detection aspects in some important journals and conference from 2000 to 2010 until the first half (such as IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, International Journal of Computer Vision, IEEE International Conference on Intelligent Transportation Systems, IEEE Intelligent Vehicle Symposium, Computer Vision and Pattern Recognition,International Conference on Pattern Recognition , International Conference on Computer Vision, Europe Conference on Computer Vision and so on).

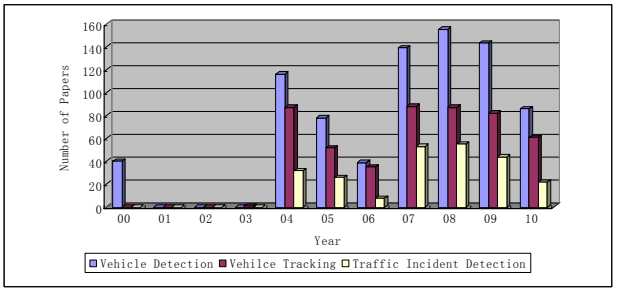


Figure 2. Relational research survey of vehicle bahavior analysis

Figure 2 shows that the number of articles on analysis of vehicle behavior in recent years has been significantly increased, which also reflects the development of researchs related to the vehicle behavior analysis. After 2004, analysis of vehicle behavior was in-depth study, and some research results were achieved. However, the current study is still focused on vehicle detection and vehicle tracking; few algorithms involve high-level understanding of complex behavior. If we applicate vehicles behavior analysis in real scenarios, there are still many serious problems.

III. Video-based Behavior Analysis

So far, many scholars have proposed a variety of algorithms of vehicle behavior analysis based on video.

As shown in Figure 3, the process of the algorithms is more or less the same and can be divided into there steps. First is ROI extraction and vehicle detection. Second is to get the static and dynamic parameters of the vehicles and establish descriptions of the vehicles. Thirdly, we should understand and describe the behavior of the vehicles by pattern matching and state estimating.

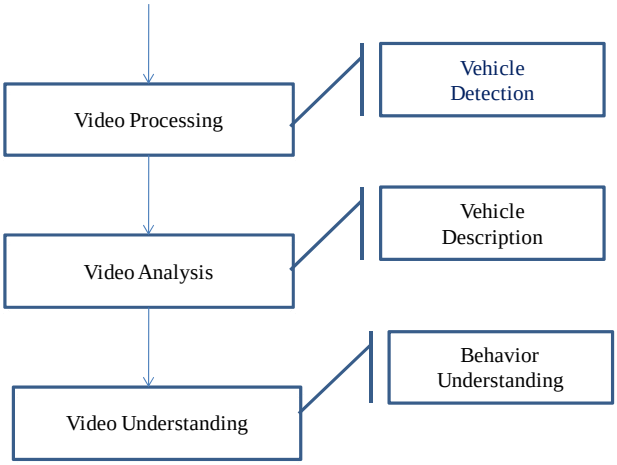


Figure 3. Hierarchy of vehicle behavior analysis

A. *Video Processing*

Vehicle detection is the core of the intelligent traffic and its purpose is to tell whether there is vehicle in the detected region. It can be attributed to the area of object detection. Moving object detection is to separating the varying region from the background of the image sequence. It is a major component of digital image processing. Effective segmentation of motion region plays a key role in the post processing, such as object segmentation and tracking. The mainly way detect the moving object is the following several methods, optical flow method [3,4], frame difference method [5-7], background subtraction method [8,9], expanding EM algorithm [10], scene change detection based on Morphology [11] and detection of energy motion algorithm [12]. Among the methods mentioned above, the first three algorithms are most commonly used.

(1) Optical flow method

Optical flow reflects the changes of the images caused by the movement in a certain time. It is a method to estimate the sport-field and merge the similar motion vector into a moving target. So calculating the optical flow is the prerequisite of the subsequent process. It mainly detects the moving area using the feature of vector flow of the moving object. Optical flow method can be divided into three categories, matching-based, frequency domain and the gradient method. The matching-based method consists of feature-based and region-based. There is some advantages of optical flow method that can be summarized in two points. First, it set the foreground moving object as the center without any priori scene information and can be used for the case of camera moving. Second, it measures the moving object based on the image pixels, for this reason it obtains accurate test results under the ideal case, pixels level or sub-pixel level. The disadvantage is that it is sensitive to noise and can't effectively deal with the issues such as target occlusion. On the other side, the high complexity of the algorithm is not conducive to a high degree of real-time applications.

(2) Frame difference method

Frame difference method is to get the pixel level difference between the two or three adjacent frames in a continuous sequence of images and extract the moving area in the image through the threshold. For example, VSAM [13] developed an algorithm that combines the adaptive background subtraction with the three frame difference, which can effectively detect the moving target from the background. Using this algorithm in the condition that the change of ambient luminance is little, we can hold that the scene is still where the pixel difference is small, while it is caused by the moving of the object where the pixel value vary widely and we can get the foreground moving object by subtracting the two images and threshold the result. For detecting moving object, frame difference method has strong adaptability for dynamic environment and is robustness to adapt the changing environment. But we can't get the entire foreground moving object, that is to say there are usually cavitations in the moving object.

(3) Background subtraction method

Background subtraction method is the most commonly used segmentation method. It is a technology that detects the moving regions using the difference of the current image and the background image. This method can detect the target accurately by subtraction when the image quality is good and the scene is stable. But the background is often more complex in reality. Be affect by so many destabilizing factors of noise, light and weather, that the simple method of reducing the background can't meet our needs. Usually, the adaptive threshold method largely eliminates the impact of light, but it is not useful for other factors. For this reason, the commonly used method is to update the background in real time by modeling of the background and get the closest background image for the current frame. Construction of dynamic background model has been being a heated thesis for the scholars domestic and foreign and among their results period average background is the simplest model, while mixture Gaussian model is one of the most popularly used model. There is lots of other dynamic background models including phone book model [14], the target centroid model [15] and the filter-based adaptive background model by Kalmann, Brandt and Kilger [16]. The background subtraction method is effective in moving target detecting when the camera is fixed and its greatest advantage is the ability to extract the entire moving target.

After the detection of moving objects, we should classify them. What we are interested in are the information of the moving vehicle, but the objects detected contain the pedestrians, bicycles, floating clouds, swinging tree branches, moving vehicles and other moving objects. To classify the moving objects is

necessary for the further analysis of the moving objects. We separate the vehicles from the moving objects by the appearance, shape and moving statistical properties of the moving object. The methods commonly used are shape-based classification, motion-based classification and co-occurrence matrix classification.

B. Video Analysis

Basing on the video processing, video analysis will describe the target features which are detected in video processing. The data of video is dynamic and its features are dynamic in the time dimension. Take full advantage of the dynamic features of the video image and we can further improve the retrieval accuracy of the video clip. The object of detecting and description in the traffic video is the vehicles and what we primary concern is how to measure the static and dynamic features and quantify the features so as to analyze the vehicles. The features used in description are as shown in Figure 4.

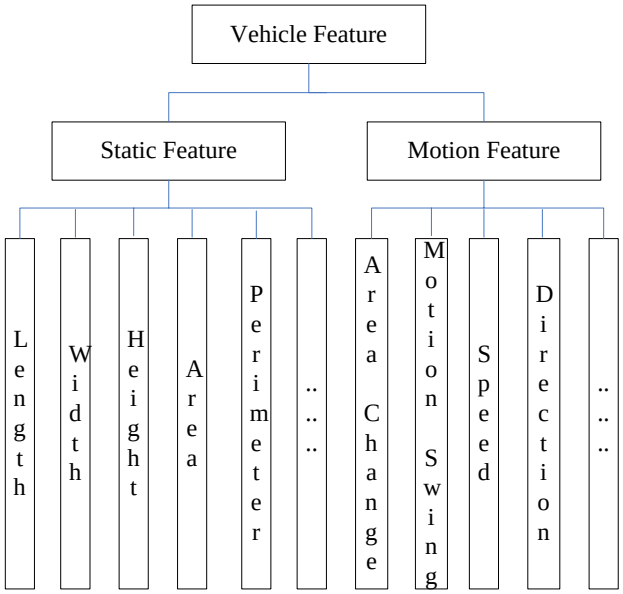


Figure 4. Features for vehicle description

The main features of the vehicles are divided into static features and dynamic features. The static features include brightness, color, texture, vehicle length, vehicle width, vehicle height, vehicle perimeter and other features while dynamic features include area change, speed, direction and other features. We have classified the moving objects and obtained the moving vehicles in vehicle processing. After the three-dimensional reconstruction and further action, to be assisted by the static information such as lane width, we can extract the some static features such as length, width, height, area of the vehicles.

For the dynamic features of the vehicles such as area change, speed, direction and other features, the main operation is tracking which belongs to the area of motion tracking. The commonly used motion tracking algorithm can be divided into several categories and we in this paper talks over region-based tracking, snake-based tracking, feature-based tracking, model-based tracking and multi-thread tracking.

(1) Region-based tracking

This algorithm is to extract the static features of the moving vehicle and match the features between the different frames, so as to track the vehicles. The most common method is sum of squared differences (SSD), kernel-based tracking algorithm [17,18] and affine image rectification algorithm [19,20], in which the most typical is mean-shift algorithm [21,22]. We can also use this algorithm to estimate the location of the target, combining with some forecasting methods such as linear prediction and Kalman prediction. It is frequently used in news video and teleconference which has little change. For example, Pahlavan use it for tracking the head and eyes. There are advantages of region-based tracking algorithm that it is intuitive, stable and accurate when the target is not under occlusion, while its disadvantage is that it is time-consuming especially when it searches globally and it results in loss of the target when there is change between the adjacent frames caused by occlusion or deformation of the target.

(2) Snake-based tracking

Proposed by Kass [23] and other scholars, snake is a algorithm with unique idea which is used in many application fields of image processing. It marks the initial contour manually at the border of the target, and defines the internal forces, external forces and constraint forces of the contour. And then make the contour moving to the interested target by minimizing the curve energy. Snake model is well suited for tracking the target with distortion, such as lips which is implemented by Kass and moving cells by Leymarie[24]. The shortcomings of Snake is that it require good initial contour and the hollow part is not available and there are solutions such as snack balloon model by Cohen[25]. The solutions for the shortcomings usually add external forces to the model.

(3) Feature-based tracking

The basic principle of the feature-based tracking is to use the information of the object's feature to identify the moving target. The general requirements for the features are that they uniquely represent the object and maintain stability even the object is moving. We assume the features as the unique stable features. Different from region-based tracking, it chooses one or several local features but not the whole vehicle as object. By contrast, the advantage of it is that even if some parts of the vehicle are under occlusion, we can track the vehicle accurately as long as there are any part is visible. However, there are shortcomings and difficulties that we should consider how to choose the unique stable features. Too little features results in high speed but wrong match, while too many features results in less efficient and low robustness.

(4) Model-based tracking

When there is a need to know the moving details of the object, we usually choose model-based tracking. We create a prior model of the moving objects first and the second is to search in the image sequence for tracking.

There are many ways to modeling the object and the common model is based on 2D/3D or based on view, where 2D means the model is established in two-dimensional space while 3D in three-dimensional space. 3D-based model is used in many algorithms in actual practice, especially for rigid moving target. For example, ZHU realize the tracking of the vehicles and ultimately get the speed and size of the vehicle by simplifying the 3D-model in the VISATRAM system [26]. View-based model describe the object making use of the collection of the images from different perspectives, to avoid three-dimensional space modeling of the target objects and complicated computation. For example, Matsumoto establish a "View Sequence" model [27] for the road the moving object is on using this method, which has a memory function and can identify the current environment through matching the current image and model.

(5) Multi-clue tracking

There are many deficiency and shortcomings in the tracking methods described above and the multi-clue tracking method combine the above method. It integrates the various information from different source, to improve the robustness of the tracking algorithm. But a major problem is how to handle the clues and make a reasonable combination. For example, Lu [28] and other scholars combine the edge, optical flow and shading information to optimize the model for tracking and realize the tracking of the hands under the changing condition. Its advantages are that it improves the tracking results using the multiple clues, which have complementary advantages, while there are many constraints about the tracked object and the object and environment assumed become special. Therefore, good algorithm should consider the compromise of the robustness and commonality.

We talked above about the advantages and disadvantages of the various tracking algorithm. Based on tracking, we can get the dynamic information of the moving object, such as area change, speed, direction and so on. These dynamic information are abstract and multi-dimensional and get ready for the next step.

C. Video Understanding

The ultimate goal of behavior analysis is to understand the behavior. Since the seventies of twenty century, people have developed a variety of traffic incident detection algorithm, which can be divided into pattern recognition based algorithm, statistic based algorithm, traffic flow model based algorithm and artificial intelligence based algorithm. Compared to other behavior detection, the vehicle behavior detection is more difficult, because the behavior of the vehicles is more complex and it is based on video processing and video analysis. Owing to the limitation of space, we will introduce the major classical algorithm.

(1) Pattern recognition based algorithm

Pattern recognition based algorithm is also known as comparison algorithm, which compares the measured values with the threshold and trigger the alarm system when the measured value exceed the threshold. The representative is the California algorithm [29] which was born in the late sixties of twenty century. Basing on road occupancy ratio of the adjacent sampling station, it detects the event and applied to the monitor center of Los Angeles highway. It requires three input variables. The first is the absolute difference of road occupancy parameters of two adjacent sampling stations. The second is the ratio of the difference and the road occupancy parameter of the upstream sampling station. The third is the ratio of the difference and the road occupancy parameter of the downstream sampling station. When the actual values of the input variables exceed the corresponding threshold, the alarm will send event alerts.

Among the improved California algorithm, California #7 and California #8 has better efficiency. The disadvantage of California algorithm is that the inputs are all about the road occupancy parameters without consider of flow rate and speed, while the advantage is that it seldom make mistake.

(2) Statistic based algorithm

Statistic based algorithm include the standard Normal Deviation method. SDN is proposed by the Texas Transportation Institute in seventies of twenty century, which is based on the statistic and analysis. It assumes that there is traffic incident when the traffic parameters change unexpected. In this method the predicted road occupancy value of the time of T is the arithmetic mean value of the former N values and the standard normal deviation of the time of T SND(t) is based on the standard deviation of the former N actual value and predicted value. If the value of SND(t) exceed the specified threshold, there may any incident occurred. The paper [30] detect the incident combining nonparametric regression and SND algorithm and gain good results.

(3) Traffic flow based model

Traffic flow based model [31] express the traffic incident making use of the complicated traffic parameters and predict whether there is a traffic incident or not by comparing the actual parameters with the predicted parameters. The representative algorithm is called Mc Master algorithm which is proposed by Mc Master university of Canada. Basing on catastrophe theory, it established a distribution mode of the traffic flow and road occupancy ratio, which is divided into four regions corresponding to four kinds of traffic state. If the measured parameter exceed some certain threshold and is in a specific range, we can hold that there may be traffic congestion and detect the traffic incident automatically by distinguish the congestion kind. It make use of both the traffic flow and road occupancy ratio, which excelled the California algorithm. But the threshold is difficult to set and unsuitable value will affect the results.

(4) Artificial intelligence based algorithm

Fuzzy logic. It was born in sixties in the twenty century and is realized in traffic video by D Teodorovic in 1994 [32]. Making use of the fuzzy logic, in the same year Chang and Wang deal with the abnormal traffic [33], define the fuzzy set of the traffic and distinguish whether the traffic state is a event or not. Because the concept of traffic flow is important but imprecise, to

detect the traffic incident with a fixed threshold is not well adapted to dynamic traffic state. The fuzzy logic makes up for it and has short-time performance. However, it is difficult to determine the fuzzy probability function and the detection rate and false-alarm rate will be ideal if we find the appropriate probability function.

Rough set theory. Z.Pawlak presented his classic paper Rough Sets in 1982 [34], which proclaimed the new birth of the Rough Sets. It is good at expressing the certain and uncertain knowledge, analyzing the inconsistent information and imitating analogue sorting. The paper [35] detects and sorts the traffic incident with the help of Rough Sets and improves the sorting accuracy by making use of cross-validation because the theory of Rough Sets is sensitive to the discrete attributes.

Artificial neural network algorithm. Neural network [36,37] is usually used for identifying the mode of the traffic when there is any incident. Because of the information gap, we can't obtain the accurate information of the incident and it's difficult to trace and analyze the reasons of the incident. The shortcomings of the algorithm are that there is no standard to determine the network structure and it needs a mass of training samples. It is liable to get the local minimum value and its ability to generalize the model is not strong because of its over-fitting.

Hidden Markov Model was born in about late sixties and early seventies in last century. Shunsuke Kamijo creates the traffic incident detection system based on HMM [38]. It studies the behaviors of each vehicle using HMM chain and identifies the current event using the output of tracking system. The results it identifies are collision, congestion and successful-pass. With single feature of the vehicle, this method is useful and practical in some situation. But when there are a lot of vehicles and pedestrians in the detected region, the tracking algorithm consumes too much computer source and results in system crash.

IV. Technical Difficulties

Since the 20th century, 70 and 80 years to now, behavior analysis have made great progress in the research at all levels. However, due to the complexity of the traffic scene, the vehicle behavior analysis is a very challenging research topic; errors existed in the results of each step will directly affect the analysis and description of the behavior. In order to improve the performance of behavior analysis algorithms, we need to solve many serious problems.

(1) Vehicle shadow

All monitored traffic scenes are in the outdoor, and illumination is very clear, so segmented motion regions often contain a shadow cast by vehicle. This not only leads to changes in the shape of the vehicle, but also increases the probability of the vehicles stick together. Therefore, to eliminate the shadow associated with moving objects is a key step for the extraction and tracking of vehicles. Shadow detection methods can be divided into two categories: model-based methods and feature-based approach. The problems of current algorithms: model-based approach requires prior knowledge and the computation is complex, so it is not suitable for real-time applications. Feature-based methods often have some limitations and rarely take the continuous nature of the shadow in space into account.

(2) Vehicles blocking

Block problem often appears between vehicles, which will affect the accuracy of the vehicle target detection, segmentation and positioning, and then affect the effectiveness of multi-target tracking algorithm. Therefore, it is of great significance to solve the mutual occlusion problem between vehicles in the intelligent traffic video surveillance, but it is also a more difficult problem in transport video processing to solve. In recent 10 years, scholars have put forward a number of ways to solve the problem of vehicles blocking, these methods can be basically divided into four categories: model method based on feature, model method based on three-dimensional, model method based on statistical and model method based on reasoning. However, these methods are studied for the partial occlusion and their requirement of scenario is high, so severe occlusion is difficult to deal with. Therefore, they are less practical, but also need further study.

(3) Semantic Expression of traffic accident

Traffic incident detection has become a research hotspot in the video surveillance. The researchers conducted a study on relevant issues, and have made some progress. However, because of traffic the system's strong nonlinearity, ambiguity and uncertainty, a variety of existing detection algorithms are more or less inadequate. Commonly used algorithms on the semantic expression of traffic incident can be divided into four catagories: the algorithm based on pattern recognition, the algorithm based on statistical techniques, the algorithm based on traffic flow model and the algorithm based on artificial intelligence. After the traffic exceptions occur on the highway, the affected objects are mainly vehicles driving on the line where exceptions occur, therefore, most algorithms basically has been widely used on the highway. Howere, vehicles on the urban road run on the highway in different station, the difficulty of detecting the vehicle behavior is more complex when compared to other detecting technologies.

V. Evaluation of The Algorithms

Vehicle behavior analysis is generally applied to real-time road monitoring in the fields like security surveillance, intelligent transportation and so on. Because of the requirements of real-time operation, which requires that the system can achieve real-time monitoring, we use the following three commonly used evaluations.

- Mean Time to Detect (MTD), which is defined as the average interval between the occurrence of an incident and the time when detection algorithm detects this incident.
- Detection Rate (DR), which is defined as the ratio of the number of detected traffic incidents and total number of events which reduces the traffic

capacity in a specified period time. It is expressed as a percentage.

- False Alarm Rate (FAR), which is defined as the ratio of the number of incidents of false positives and the total number of incidents detected with the algorithm in a certain period of time. It can be expressed as a percentage or the number of false positives in every period of time.

Efficient vehicle behavior analysis must have three basic characteristics: short MTD, high DR and low FAR. However, MTD, DR and FAR are interrelated, the improvement of one performance index often results in the degradation of others. So we should have a trade-off among those performance indexes according to actual situation.

Vehicle behavior analysis can also be applied to traffic video retrieval, whose performance can be judged by its efficiency and effectiveness. In general, its efficiency can be evaluated by how much work it needs when performing a query. Its effectiveness can be measured by Recall and Precision. Recall and Precision are relative reasonable indexes in measuring search results now.

$$Recall = \frac{number\ of\ searched\ relational\ traffic\ events}{number\ of\ relational\ traffic\ events\ of\ video} * 100\% \quad (1)$$

$$Precision = \frac{number\ of\ searched\ true\ traffic\ events}{number\ of\ searched\ relational\ traffic\ events} * 100\% \quad (2)$$

Both values are between 0 and 1. If the value is closer to 1, it means the Recall / Precision is higher. The former can measure the ability to detect relevant information of the retrieval system and searchers. The latter can measure the ability to reject non-relevant information of them. Take them together, and then we get the retrieval efficiency. Experimental results show that interdependence of the opposite exists between Recall and Precision. If we improve the Recall of output, its Precision will be reduced, and vice versa.

VI. CONCLUSION

Because of its complexity of the algorithm and the diversity of target behavior, the development of the study on Video Analysis and Recognition is undeniable to have been slow. Now foreign technology is more mature, but it could not meet the full satisfaction of the users. It has only a relatively low level of intelligence and can complete the automatic detection of some simple behaviors. Nowadays, the understanding of the vehicle behavior is still in the early development stage; however, the rapid development of vehicle behavior analysis will certainly promote the further development of behavior understanding. With the continuous development of video analysis technology, its application in the intelligent transportation industry is increasingly close to be realized. It will show its importance in the field of intelligent transportation and play a growing role.

ACKNOWLEDGEMENT

This research was partially supported by the Natural Science Foundation of China under grant No. 60970015, the 2009 Special Guiding Fund Project of Jiangsu Modern Service Industry (Software Industry) under grant No. [2009]332-64, the Project of Jiangsu Key Laboratory of Computer Information Processing Technology under grant No. KJS0924, the 2010 Program for Postgraduates Research Innovation in University of Jiangsu Province under grant No. CX10B_041Z, the Applied Basic Research Project (Industry) of Suzhou City under grant No. SYJG0927 and No. SYG201032, and the Beforehand Research Foundation of Soochow University.

REFERENCES

[1] Zhang Jinglei, Wang Xiaoyuan. Research Progress of Traffic Incident Automatic Detection Algorithms[J]. Journal of Wuhan University of Technology(Transportatioon Science & Engineering), 2005, 29(2):215-218.

[2] Tang Shu-ming, Wang Kun-feng, Li Yuan-tao. Survey of Vision-based Automatic Incident Detection Technology[J]. Journal of Highway and Transportation Research and Development, 2006, 23(8):116-121.

[3] A Verri, S Uras, E De Micheli. Motion Segmentation from Optical Flow[C]. In Proceedings of the 5th Alvey Vision Conference, 1989:209-214.

[4] J Barron, D Fleet, S Beauchemin. Performance of Optical Flow Techniques[J]. International Journal of Computer Vision, 1994, 12(1):42-77.

[5] A J Lipton, H Fujiyoshi, R S Patil. Moving Target Classification and Tracking from Real-time Video[C]. In Proceedings of IEEE Workshop on Applications of Compute Vision,1998:8-14.

[6] A Neri, S Colonnese, G Russo, et al. Automatic Moving Object and Background Separation[J]. Signal Processing, 1998, 66(2):219-232.

[7] F Moscheni, S Bhattacharee, M Kunt. Spatial Temporal Segmentation Based on Region Merging[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1998, 20(9):897-915.

[8] N Ohta. A Statistical Approach to Background Subtraction for Surveliance Systems[C]. In Preceedings of IEEE international Conference on Computer Vision, 2001, Vol.2:481-486.

[9] A Mital, N Paragios. Motion-based Background Subtraction Using Adaptive Kernel Density Estimation[C]. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2004,Vol.2:302-309.

[10] G McLach Ian, T Krishnan. The EM Algorithm and Extensions[M]. John Wiley& Sons,Inc.,1997.

[11] E Stringa. Morphological Change Detection Algorithms for Surveillance Applications[C]. In Proceedings of British Machine Vision Conference, 2000:402-411.

[12] R P Wildes. A Measure of Motion Salience for Surveillance Applications[C]. In Proceedings of International Conference on Image Processing, 1998:183-187.

[13] Fujiyoshi H, Kanade T. VSAM:Video Surveillance and Monitoring Project[J]. Kyokai Joho Imeji Zasshi/Journal of the Institute of Image Information and Television Engineers, 2003, 57(9):1068-1072.

[14] Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, LarryDavis. Real-time foreground-background segmentation using codebook model[J]. Real-Time Imageing, 2005, 11(3):172-185.

[15] Mustafa Oral, Umut Deniz. Centre of mass model-A novel approach to background modeling for segmentation of moving objects[J]. Image and Vision Computing, 2007, 25(8):1365-1376.
[16] Kilger M. A shadow handler in a video-based real-time traffic monitoring system.in:Processings[C]. IEEE Workshop on Applications of Computer Vision, 1992:11-18.
[17] Patras L., Hendriks E.A., Lagendijk R.L. Video segmentation by MAP labeling of watershed segments[J]. IEEE Trans. On Pattern Analysis and Machine Intelligence, 2001,23(3):326-332.
[18] Markandey V., Reid A., Wang S. Motion estimation for moving target detection[J]. IEEE Trans. on Aerospace and Electronic Systems, 1996,32(3):866-874.
[19] Nickels K., Hutchinson S. Estimating uncertainty in SSD-based feature tracking[J]. Image and Vision Computing. 2002,20(1):47-58.
[20] Haritaoglu I., Harwood D., Davis L.S. Ghost. A human body part labeling system using silhouettes[C]. In Proc. Of the 14th Int'1. Conf. on pattern recognition, 1998, Vol.1:77-82.
[21] Anderson M. Tracking methods in computer vision[D]. In Dept. of Numerical Analysis and Computing science, KTH, Stockholm, Sweden, 1994.
[22] Zhang Z. Token tracking in a cluttered scene[J]. Image and Vision Computing, 1994, 12(2):110-120.
[23] Kass M., Witkin A., Terzopoulous D. Snakes: Active contour models[J]. International Journal of Computer Vision, 1987, 1(4):321-331.
[24] Leymarie F., Levine M.D. Tracking deformable objects in the plane using an active contour model[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence,1993,15(6):617-624.
[25] Cohen L.D. Note on active contour models and ballons[J]. Computer vision, Graphics and Image Processing: Image Understanding. 1991, 53(2):211-218.
[26] Zhu Z., Xu G., Yang B., Shi D., Lin X. VISATRAM: A real-time vision system for automatic traffic monitoring[J]. Image and Vision Computing, 2000, 18(10):781-794.
[27] Matsumoto Y., Sakai K., Inaba M., Inoue H. View-based approach to robot navigation[C]. In Proc. Of IEEE/RSJ Int'l. Conf. on Intelligent Robots and Systems, Takamatsu, Japan, 2000, Vol.3:1702-1708.
[28] Lu S., Metaxas D., Samaras D., Oliensis J. Using multiplue cues for hand tracking and model refinement[C]. In: Proc. of IEEE Conf. on computer vision and pattern recognition, 2003,Vol.2:443-450.
[29] Dia H., and Rose G. Development and Evaluation of Neural Network Freeway Incident Detection Models Using Field Data[J]. Transp. Res. C, Emerg. Technol., 1997, 5(5): 313-331.
[30] Tang S., Gao H. Traffic-Incident Detection-Algorithm Based on Nonparametric Regression[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(1):38-42.
[31] Barak Fishbain, Ianir Ideses, David Mahalel, Leonid Yaroslavsky. Real-time vision-based traffic flow measurements and incident detection[C]. Real-Time Image and Video Processing 2009.
[32] D Teodorovic. Fuzzy Sets Theory Applications in Traffic and Transportation[J]. European Journal of Operational Research, 1994, 74(3):379-390.
[33] Chang, Wang. Improved Freeway Incident Detection Using Fuzzy Set Theory[R]. Transportation Research Record No. 1453, Intelligent Transportation Systems: Evaluation, Driver Behavior, and Artificial Intelligence:75-82.
[34] Z. Pawlak. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356.
[35] Shuyan Chen, Wei Wang, Gaofeng Qu. Traffic Incident Detection Based on Rough Sets Approach[C]. Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, 2007:3734-3739.
[36] Dougherty M. A Review of Neural Networks Applied to Transport[J]. Transp. Res. C, Emerg. Technol., 1995, 3(4):247-260.
[37] Seung-Heon Lee, Jin-Woo Choi, and Nam-Kwan Hong. Development of Incident Detection Model Using Neuro-Fuzzy Algorithm[C]. In Proceedings of the Fouth Annual ACIS, International Conference on Computer and Infromation Science, 2005:364-368.
[38] Shunsuke Kamijo, Yasuyuki Matsushita, Katsushi Ikeuchi, and Masao Sakauchi. Traffic Monitoring and Accident Detection at Intersections[J]. IEEE Transactions on Intelligent Transportation Systems, 2000, 1(2):108-118.

Jian Wu was born in Nantong on the 29th April, 1979, and got master degree in the field of computer application technology from Soochow university, Suzhou city, China in 2004. The main research direction is computer vision, image processing and pattern recognition. He works as a teacher in the same college after his master graduation. Now he is pursuing the doctoral degree. He was awarded the Third Prize of 2007 Suzhou City Science and Technology Progress and the 2008-2009 Soochow University Graduate Scholarship Model.

Zhi-ming Cui was born in Shanghai on the 4th July, 1961. Professor, PhD Candidate Supervisor. The main research direction is deep web and video mining.

Jian-ming Chen was born in Suzhou on February, 1960. Associate professor, Master Supervisor. The main research direction is intelligent information processing and software engineering.

Guang-ming Zhang was born in Suzhou on the 10th Februry, 1981. PhD Candidate, his main research direction is image processing and video retrieving.

Geometrically Invariant Watermarking Scheme Based on Local Feature Points

Li Jing
School of Computer and Information Engineering,
Henan University of Economics and Law, Zhengzhou, China
Email: jingli776@126.com

Xiaowen Zhang
Department of Computer Science,
University College Cork, College Road, Cork, Ireland
Email: skyof45 @yahoo.com.cn

***Abstract*—Based on local invariant feature points and cross ratio principle, this paper presents a feature-point-based image watermarking scheme. It is robust to geometric attacks and some signal processes. It extracts local invariant feature points from the image using the improved scale invariant feature transform algorithm. Utilizing these points as vertexes it constructs some quadrilaterals to be as local feature regions. Watermark is inserted these local feature regions repeatedly. In order to get stable local regions it adjusts the number and distribution of extracted feature points. In every chosen local feature region it decides locations to embed watermark bits based on the cross ratio of four collinear points, the cross ratio is invariant to projective transformation. Watermark bits are embedded by quantization modulation, in which the quantization step value is computed with the given PSNR. Experimental results show that the proposed method can strongly fight more geometrical attacks and the compound attacks of geometrical ones.**

***Index Terms*—digital watermark, invariant feature points, the cross ratio of four collinear points, geometrical attacks, quantization modulation**

I. INTRODUCTION

Geometric attacks means watermarked image is transformed by some geometrical distortions, such as rotation, scaling, translation (RST), random bend attack (RBA), and shearing etc. Geometric attacks induce synchronization errors between the original and the extracted watermark during the detection process, and the inserted watermark can not be detected correctly although it is still present in watermarked image. Up to now a few algorithms have presented the topic of how to achieve robustness against geometric attacks, which can be approximately classified into the following categories: watermarking based on invariant transform[1,2], watermarking based on synchronization templates[3,4], watermarking based on invariant moments[5,6], and watermarking based on feature points[7,8,9,10]. Compared with other kinds of algorithms, watermarking based on feature points has better performance against geometric attacks. It is not only robust to RST attack and local geometric attacks (for example RBA and cropping), but also to combination of more geometric attacks. For this reason watermarking based on feature points has been a researching focus in recent years.

Feature-point-based watermarking uses local invariant feature points to decided local feature regions, watermark is inserted these local regions repeatedly. The watermark inserted positions combined with these invariant feature points, in this way keeping watermarking synchronization. The stability of feature points to geometric transform and signal processes decides the robustness of watermarking system. Bas et al. [7] used Harris corner detecting method to extract feature points from image. Tang and Hang [8] extract feature points by Mexican hat wavelet scale interaction. Wang et al. [9] use Harris-Laplace method and Lee et al.[10] use Scale Invariant Feature Transform (SIFT) method. The stability of extracted feature points by these methods is very different. Harris corner points are robust to rotation, but not to scaling and affine transform. Feature points extracted by Mexican hat wavelet scale interaction are robust to signal processes, but not to geometric transform. Compared to these methods, feature points from Harris-Laplace and SIFT have better robustness [11].

Other important factors of Feature-point-based watermarking are the selection of local feature regions and the watermark inserting and extracting method. Many algorithms [9, 10] choose circle patches centering local feature points to insert watermark. They are robust to rotation and scaling, but not to shearing of affine transform. When the watermarked image is transformed by shearing, part of inserted watermark bits will beyond the detected circle patches and watermark can not be extracted correctly. Tang and Hang [8] used normalized circle patches to resist affine transform. However, the normalization is sensitive to the image contents used, so the robustness of these patches will decrease when the image is distorted. Up to now these algorithms only

doi:10.4304/jmm.7.3.231-238

concern affine transform attack, dot not involved projective transform yet.

This paper utilizes improved SIFT algorithm and invariance of cross-ratio in projective transform, proposed a watermarking scheme. The proposed scheme belongs to feature-point-based watermarking, is robust to geometric attack. It extracts stable feature points using the improved SIFT algorithm, and optimizes the distribution of feature points. Then choose qualified quadrilateral regions to insert watermark. Every region is decided by four feature points. The watermark is inserted into every region repeatedly. The location to insert watermark bit is based on invariance of cross-ratio, and using quantization modulation method to insert. Experimental results show the proposed scheme is not only robust to general signal processes, but also to many geometrical distortions, such as affine transform, projective transform, cropping and their combination.

The rest of this paper is organized as follows. Section 2 describes the feature points extracting and adjusting method used in the proposed scheme. Section 3 reviews cross-ratio theory of four collinear points and given two lemmas, which is the theory basis of watermark insert method. In Section 4, we present watermark inserting process. Section 5 covers the details of the watermark detection procedure. Simulation results in Section 6 will show the performance of our scheme. Finally, Section 7 concludes this presentation.

II. EXTRACTING FEATURE POINTS AND OPTIMIZING THEIR DISTRIBUTION

This paper utilizes SIFT algorithm to extract feature points. SIFT feature is extracted by considering local image properties and is invariant to rotation, scaling, translation, and partial illumination changing. And also robust to a substantial range of affine distortion, change in 3D viewpoint, addition of noise[12].

The basic idea of the SIFT is to extract features through a staged filtering that identifies stable points in the scale space. It first selects candidates for features by searching for peaks in the scale space of the difference-of-Gaussians (DoG) function, then localizes each feature using measures of its stability. The candidate locations that have a low contrast or are poorly localized along edges are removed by measuring the stability of each feature using a 2×2 Hessian matrix H as follows:

$$stablity = \frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$

$$\text{Where } H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \qquad (1)$$

Here r is the ratio of the largest to the smallest eigenvalue and is used to control stability. They use r=10. The quantities D_{xx}, D_{xy}, and D_{yy} are the derivatives of the scale space images.

The next work is to assign orientations based on local image gradient directions, at last give a descriptor for every SIFT feature point. Our scheme only use SIFT points to decide quadrilateral regions when inserting and detecting watermark. So we abnegate constructing feature vector process and save computing time.

SIFT algorithm generates large numbers of features points that densely cover the image. A typical image of size 500×500 pixels will give rise to about 2000 stable features points, and these feature points densely covering structure complex regions of image. And the feature points used in watermark inserting and extracting should be stable and be distributed equally. So we improved SIFT algorithm to get stable feature points with optimized distribution, removing those feature points that are susceptible to watermark attacks.

Combined our work, we change the procedure of SIFT algorithm, and the following is our improving measures.

- Using a Gaussian filter to blur the original image before feature points extraction. By this preprocess, it can reduce the interference of noise and increase the robustness of extracted feature points.
- In SIFT algorithm we leave out the step of representing feature descriptor that is 128 element feature vector for each feature. Only remain feature points detector, which reduces largely calculating time without destroying our scheme.
- In order to improve the robustness of feature points, properly increase the sampling frequency and depth in scale direction at scale space. With shorter sampling distance and larger sampling depth, the extracted feature points are more robust to scale changing.
- Set a threshold Th, when searching local extremum in 3×3×3 region of scale space, DoG function $D(x,y,\sigma)$ should meet the following formula:

$$\{D(x,y,\sigma) \,||\, D(x,y,\sigma)| > Th\} \qquad (2)$$

In our experiences we set Th=0.05

The improving method mentioned above can strengthen the robustness of extracted feature points, but their distribution is still not equal. The distribution of local feature points is related to the performance of watermarking systems. In other words, the distance between adjacent feature points should be determined carefully. In order to fit our work, their number and distribution need to be adjusted. The following is adjusting method.

Detect feature points with the improved SIFT algorithm firstly. In the local circle regions centered detected feature points with the radius R, if the DoG fuuction $D(x,y,\sigma)$ value of circle center is the local extremum of the circle region, we keep the feature points, otherwise remove the points. In this way, remained feature point is the most robust point in the local circle region. The distribution of remained feature points is comparatively even. And the distance of two feature points can keep between R and $2R$. Fig. 1 (a) is the extracted feature points using improved SIFT algorithm, and Fig. 1 (b) is the distribution adjusted result by the above mentioned method.

The scale s of features point is related to the scaling factor of the Gaussian function in the scale space, and varies with the local image characteristic. So the local

circle radius R is accordant with the scale of the selected point, namely $R = r \times s$. r is the adjusting parameter of the distance between feature points. The value of r can be used to control the number of feature points. With lager r, the distance between feature points is larger, the number of remained feature points will be smaller, and the local region to insert watermark will be larger. On the contrary, r is smaller, the distance between feature points will be shorter, remain more feature points, and the local region to insert watermark will be smaller.

Smaller local region to insert watermark can be robust to local attacks, but can not contain more watermark bits. So setting the value of r should take into account both robustness and watermark capacitance. In our experience we set r=5.

(a) extracted feature points (b) distribution adjusting result

Figure 1. Extracted feature points and their adjust result

III. CROSS RATIO THEORY OF FOUR COLLINEAR POINTS

As seen in Fig. 2, *A*, *B*, *C*, *D* are four collinear points they all locate at line m_1 . Choose a direction along the line as the positive direction (so that distances measured in the opposite direction are treated as negative). The cross-ratio r_{ABCD} of these four points in the given order is defined to be:

$$r_{ABCD} = \overline{AC} \times \overline{BD} \Big/ \overline{BC} \times \overline{AD} \tag{3}$$

Where $\overline{PQ}$ denotes the signed distance between two points P and Q for some choice of orientation.

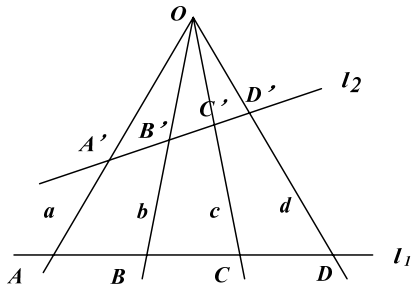


Figure 2. The sketch map of cross-ratio

It is a standard result that the cross ratio of four collinear points does not change under a transform. Affine transform is the special form of projective transform. The principle of projective transform is useful to affine transform, so the cross ratio of four collinear points also keep invariant under affine transform. The parallel lines keep parallel after being operated by affine transform, but do not keep parallel operated by projective transform. And line is still line after being operated by projective transform

From cross ratio is projective invariant, we can get the following two lemmas:

Lemma 1: Suppose three points of four collinear ones have been known, and the cross ratio of the four collinear points is also given, the forth point is certain.

Proof: Suppose *A*, *B*, *C* and *D* are distinct collinear points, as shown in Fig. 2. Now know *A*, *B*, *C* are certain, and the cross ratio $r_{ABCD} = \frac{CA}{CB} : \frac{DA}{DB}$ is also certain. If the forth point *D* is not certain, then segment *BD* is not certain, from which can deduce the cross ratio $r_{ABCD} = \frac{CA}{CB} : \frac{DA}{DB}$ is not certain. That is in contradiction with the known facts. So the forth point *D* is certain.

Lemma 2: Suppose four lines *a*, *b*, *c*, *d* cross the same point *O*, intersect line l_1 at four points *A*, *B*, *C*, *D*, and intersect line l_2 at four points *A′*, *B′*, *C′*, *D′*, as shown in Fig.2. According to the duality principle of projective geometry, the following equation can be got.

$$r_{ABCD} = r_{abcd} = r_{A'B'C'D'} \tag{4}$$

Proof: According to the duality principle of projective geometry, if r_{ABCD} is certain, then the cross ratio r_{abcd} of line *a*, *b*, *c* and *d* is certain. The definition of r_{abcd} is following:

$$r_{abcd} = \frac{\sin(ca)}{\sin(cb)} : \frac{\sin(da)}{\sin(db)} \tag{5}$$

Where $\sin(ca)$ is the sine of the directed angle formed by line *c* and *a*.

The areas of triangle *COA*, *COB*, *DOA* and *DOB* can be calculated by the following formula:

$$\begin{cases} S(COA) = \frac{1}{2} h \cdot CA = \frac{1}{2} OC \cdot OA \cdot \sin(ca) \\ S(COB) = \frac{1}{2} h \cdot CB = \frac{1}{2} OC \cdot OB \cdot \sin(cb) \\ S(DOA) = \frac{1}{2} h \cdot DA = \frac{1}{2} OD \cdot OA \cdot \sin(da) \\ S(DOB) = \frac{1}{2} h \cdot DB = \frac{1}{2} OD \cdot OB \cdot \sin(db) \end{cases} \tag{6}$$

Where h is the high from point O to line l_1 .

From the above four area computing formula, the following equation can be deduced:

$$\begin{aligned} r_{ABCD} &= \frac{CA}{CB} : \frac{DA}{DB} = \frac{\sin(ca) \cdot OA}{\sin(cb) \cdot OB} : \frac{\sin(da) \cdot OA}{\sin(db) \cdot OB} \\ &= \frac{\sin(ca)}{\sin(cb)} : \frac{\sin(ba)}{\sin(bd)} = r_{abcd} \end{aligned} \tag{7}$$

Similarly we can deduce $r_{A'B'C'D'} = r_{abcd}$

So $r_{ABCD} = r_{abcd} = r_{A'B'C'D'}$

Based on the Lemma 1 and Lemma 2, we choose the watermark bits inserting points in the local quadrilateral

region decided by feature points, which guarantees watermark is invariant to projective transformation.

IV. WATERMARK INSERTING

The whole process of watermark inserting can be seen in Fig. 3, firstly extract feature points from carrier image with the improved SIFT algorithm given in part II，and adjust the number and distribution of feature points using the method in part II. Then choose the proper

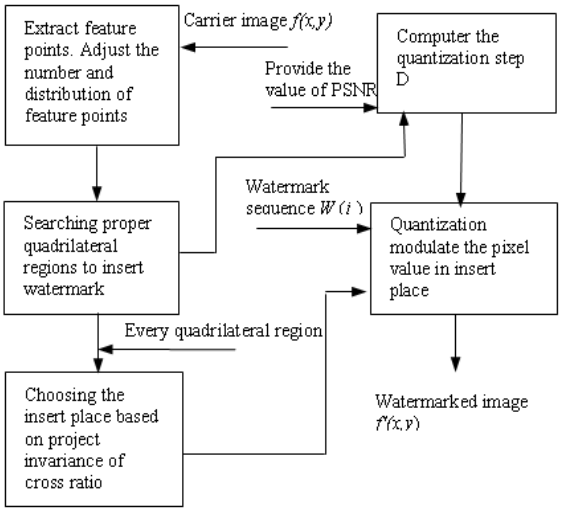


Figure 3. Watermark inserting process

quadrilateral as local feature regions to insert watermark and get watermarked image.

A. Choose Local Feature Regions

This paper uses the local quadrilateral regions to insert watermark, and the apexes of quadrilateral are the chosen feature points. Use the improved SIFT algorithm to extract feature points, adjust the number and distribution and get a set of feature points $P=\{p_i, i=1,\cdots,N\}$. And the local feature regions are decided by the feature points in set P.

The method to choose local feature regions is as the following: regard feature points $p_i, i=1,\cdots,N$ as center, in a round region with radium $4R$ searching three feature points, and construct a quadrilateral with point p_i. The chosen quadrilateral should be convex quadrilateral that is close to square. To meet this requirement, the internal angle of chosen quadrilateral should be smaller than 100 degree, the length of two adjacent sides should be similar, and their difference can not be larger than 30 pixels.

And the chosen quadrilateral regions maybe overlap each other. If two regions overlap (not include the two regions with same point and side), choose the one approaching square, which will help watermark inserting and detecting.

B. Choose the Inserting Locations for Watermark Bits

In every local feature region, decide the inserting location for watermark bit based on the projective invariance of cross ratio. Suppose $P_1P_2P_3P_4$ standing for a local feature region, as shown in Fig. 4, line P_1P_3 intersects line P_2P_4 at point O

If extremity points of two segments P_1, P_2, P_3, P_4 are invariant to projective transformation, their cross point O keeps invariance to projective transformation too. Because projective transformation is linear transformation in two dimensions space, line transformed by projective is still line. Two segment keep invariance, and their cross point keep invariance. According to

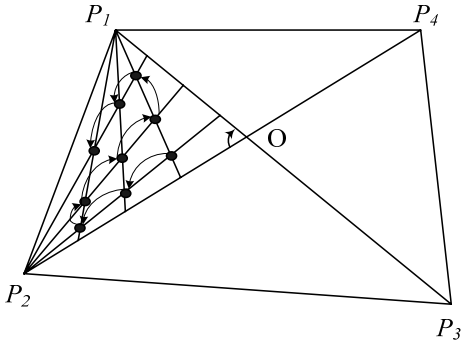


Figure 4. The sketch map of choosing watermark bits inserting locations

Lemma 1 in part III, given three points of four collinear points and the cross ratio of the four collinear points, the forth point is certain.

Given n cross ratios in advance, locate n points in every segment, and control $n/2$ points distribute one side of cross point O, another n/2 points distribute the other side, as shown in Fig. 4. In the triangleP_1P_2O, draw lines between point P_1 and the points on segment P_2O, draw lines between point P_2 and the points on segment P_1O. According to Lemma 2 in part III the cross points of these lines keep invariance to projective transformation.

We choose the cross points (marked black dot in Fig. 4) as the inserting location for watermark bits, which will be robust projective transformation.

In order to keep the correct order of watermark bits when watermarked image destroyed by geometrical attacks, the watermark bit inserting location should arrange in a certain order. We regulate arranging the inserting location from the original side along clockwise of the internal angle at cross point O. Arranging order is along the arrowhead orientation in Fig. 4.

Use the same way to choose the watermark inserting location in the other three triangles.

C. Watermark Inserting Method

We adopt quantization modulation to modify the pixel value at chosen watermark bit inserting location. The pixel values at N watermark inserting locations are supposed to be $I_j(x,y)$ j=1,2,…,N , and watermark is a {0,1}sequence, supposed to be $W(j)$ j=1,2,…,N. $I_j(x,y)$ and $W(j)$ are one to one correspondence. If $W(j)$ corresponding $I_j(x,y)$ is 0, modulate $I_j(x,y)$ to center of even interval; If $W(j)$ corresponding $I_j(x,y)$ is 1, modulate $I_j(x,y)$ to center of odd interval.

Describe the quantization step as D, and the quantization interval λ is defined as:

$$\lambda = \lfloor I(x,y)/D \rceil \quad (8)$$

Where $\lfloor \bullet \rceil$ mean rounding a number to the nearest integer.

In order to insert watermark bits, The pixel value of $I_f(x,y)$ is modified by (9)

$$\begin{cases} I_W(x,y) = (\lambda - 0.5) \times D \\ \quad if\ (\lambda + W(i)) \bmod 2 = 1 \\ I_W(x,y) = (\lambda + 0.5) \times D \\ \quad if\ (\lambda + W(i)) \bmod 2 = 0 \end{cases} \quad (9)$$

In order to strengthen the robustness of watermarking system, the pixels in a 3×3 region centered the chosen inserting point are also modified. If the center inserting point is modified to correspond 0, the 8 pixels around is modified to correspond 0. If the center inserting point is modified to correspond 1, the 8 pixels around it is also modified to correspond 1.

In modifying the pixels with (9), how to choose the quantization step D is a crucial problem. The step D is a key balance factor between robustness and imperception of watermark. If choose large value for step D, watermark will have good robustness, but not guarantee imperception. If choose small value for step D, watermark will have good imperception, but not guarantee robustness. In many former reported paper, the step D was adjusted through repeated experimentation, to meet the requirement for Peak Signal-to-Noise (PSNR) of watermarked image. In this paper we deduce the quantitative relation between PSNR and step D. Based on the relation, compute the step D with PSNR directly, not need repeated experimentation.

The following is PSNR definition.

$$PSNR(I, I') = 10 \log_{10} \frac{255^2}{MSE} \quad (10)$$

$$MSE = \frac{1}{M} \sum_{i=1}^{N} (I_i(x,y) - I_i'(x,y))^2 \quad (11)$$

Where MSE is mean square error, I is original image, I' is watermarked image. $I(x,y)$ denotes the value of pixel (x,y), M is pixel number in all local feature region to insert watermark, and N is pixel number used to insert watermark bits. In many reported paper, M means the size of image and PSNR stands for the imperception of the whole image. In this paper watermark is inserted local feature regions, not the whole image, if set N to be size of the whole image, PSNR will be small even local region changed largely. In this case, watermark has been seen and PSNR can not express the quality of watermarked image.

In local feature region the watermark inserting ratio is $\rho = N / M$, the relation of step D, inserting ratio ρ and mean square error MSE can be deduced.

We can learn from (9) that quantization errors distribute the interval $[-D, D]$ equably. So the mathematical expectation of the square quantization errors is (12).

$$E(I(x,y) - I'(x,y))^2 = \frac{(2D)^2}{12} = \frac{D^2}{3} \quad (12)$$

And the mathematical expectation of the mean square error MSE is (13).

$$E(MSE) = \frac{1}{M} \sum_{i=1}^{N} E(I(x,y) - I'(x,y))^2 = \frac{N}{M} \cdot \frac{D^2}{3} = \frac{1}{3} \rho D^2 \quad (13)$$

So we can deduce the relation of step D, inserting ratio ρ and mean square error MSE

$$D = 255 \cdot \left[\frac{1}{3} \rho \cdot 10^{\frac{PSNR}{10}} \right]^{-0.5} \quad (14)$$

V. WATERMARK DETECTING

For input image I', firstly extract feature points with the improved SIFT algorithm, adjust the number and distribution of feature points using the presented method in part II. Get a set of feature points $Q = \{q_i, i = 1, \cdots, M\}$. Center feature points $q_i, i = 1, \cdots, M$, in a round region with radium $4R$ searching three feature points, and construct a quadrilateral with point q_i. Choose the protruding quadrilaterals as watermark detecting regions. For every detecting region, searching watermark inserting points with the same method as watermark inserting, the cross ratios set is that used in watermark inserting process. For pixel $I'(x,y)$ extract watermark bit using following formula.

$$W'(i) = \begin{cases} 1 & if\ \lfloor I_i'(x,y)/D \rfloor \bmod 2 = 1 \\ 0 & if\ \lfloor I_i'(x,y)/D \rfloor \bmod 2 = 0 \end{cases} \quad (15)$$

Where $\lfloor \bullet \rfloor$ means round down, D is the same quantization step used in watermark inserting.

In every chosen local feature quadrilateral the two diagonal lines divide it into four triangles, and watermark is inserted into every triangle repeatedly. In any one triangle of any two local feature quadrilateral detect watermark sequence is $W'(i)$. If the number of bits matched $W'(i)$ and original watermark sequence $W(i)$ is larger than the threshold T, we would judge image I' exist watermark.

Determining whether there are watermarks in image I' is concern to threshold T, which decides the error rate of watermark detector directly.

Two kinds of errors are possible for detector: the false-alarm probability and the miss probability. The false-alarm probability means no watermark inserted but detected while the miss probability means watermark inserted but detected having none. There is a tradeoff between these two error probabilities in selecting detector parameters. Typically, reducing one will increase the other.

The miss probability depends on success detection probability. It is difficult to evaluate the success detection probability of a watermarked bit. It depends on the attacks. So this paper chooses the threshold T only depending on the false-alarm probability.

For an unwatermarked image, the extracted (0.1) bits are assumed to be independent random variables, from which we learn the success probability of every bit is 0.5 Based on Bernoulli trials, match the watermark sequence $W'(i)$ detected from triangle region with original watermark sequence, if there are k bits matched successfully, the false-alarm probability to detect watermark from one triangle can be computed with the following formula.

$$P_{tri} = \sum_{k-T}^{N} (0.5)^N \frac{N!}{k!(N-k)!} \tag{16}$$

Where N is the length of watermark sequence, k is the number of bit matched successfully. T is watermark detecting threshold.

Four triangles in a local feature region are all inserted watermark, if detect watermark from one triangle successfully, we judge existing watermark in the local feature region. The false-alarm probability to detect watermark from one local feature region can be computed from the following formula.

$$P_{qua} = \sum_{i=1}^{4} (P_{tri})^i (1-P_{tri})^{4-i} \frac{4!}{i!(4-i)!} \tag{17}$$

If detect watermark from two local feature region successfully, we judge existing watermark in image I'. Suppose inserting watermark in m local feature region, the false-alarm probability to detect watermark from image I' can be computed with the following formula.

$$P_{false} = \sum_{i=2}^{m} (P_{qua})^i (1-P_{qua})^{m-i} \frac{m!}{i!(m-i)!} \tag{18}$$

Suppose the length of watermark sequence N=25, there are 10 local feature quadrilateral inserting watermark, detect threshold T=20. From the above analysis we can learn the false-alarm probability $P_{false} \approx 2.8 \times 10^{-3}$.

VI. Experimentation Result and Analysis

We choose three gray images Lena, Pepper and Baboon to test our scheme, their size is 512×512. These images have different texture. In experimentation, the distance adjusting parameter r=5, the length of watermark sequence N=25, quantization step D=15, detect threshold T=20. There are 7 local feature regions inserting watermark in image Lena, there are 14 local feature regions inserting watermark in image Baboon, and there are 9 local feature regions inserting watermark in image Pepper.

A. The Visual Quality of Watermarked Image

Fig. 5 includes the watermarked Lena, Baboon and Pepper. Fig. 6 is the difference value images between original images and watermarked ones (the difference value is 30 times of original one). The quality of watermarked image is described by PSNR. If only concern the MSE of local feature regions, the PSNRs of watermarked Lena, Baboon and Pepper are 39.285, 39.259, and 39.345 respectively. If concern the MSE of the whole image, the PSNRs of watermarked Lena, Baboon and Pepper are 46.455, 46.455 and 46.455 respectively. For our scheme, the fist PSNR computing way is proper. The PSNR in this paper is computed with the fist method.

(a)watermarked Lena (b) watermarked Baboon (c) watermarked Pepper

Figure 5. Watermarked images

(a) Lena (b) Baboon (c) Pepper

Figure 6. Difference value images between original images and watermarked ones

PSNR of watermarked image is related to quantization step D and watermark inserting ratio ρ. The lager value of D and ρ will make watermark more robust, but the quality of watermarked image will fall off, and get smaller PSNR.

In section IV part C we have analyzed the relation of quantization step D, watermark inserting ratio ρ and PSNR, and deduce the relation formula (14). Fig. 7 is the result value computed from theory and the experience, and giving their relation curve. The theory value is computed from formula (14) and the experience value is come from three test image.

As can be seen in Fig. 7, with the watermark inserting ratio ρ increasing, PSNR decrease; and with the quantization step D increasing, PSNR decrease. The experience value and theory value fit better. But ρ=1 and ρ=0.5, the experience value and theory value have little

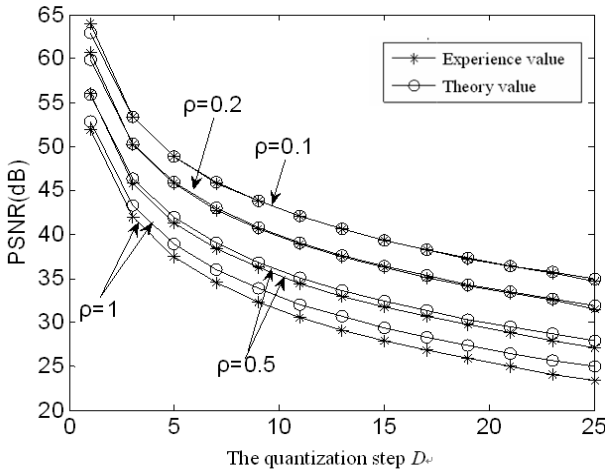


Figure 7. The relation of PSNR, quantization step D and watermark inserting ratio ρ

difference. The main reason is the watermark insert location overlap with the inserting ratio ρ increasing. These experiences validate the formula (12) is correct.

When inserting watermark, set inserting rate based on the area of chosen regions, and the quantization step D is decided by PSNR of watermarked image. Set PSNR=40dB to guarantee the quality of watermarked image.

B. Test and Analyze the Robustness of Watermark

In order to test the robustness of the proposed scheme, we carried three kind experiments for watermarked image. We manipulated the watermarked image with no attack, general signal attacks and geometric attacks respectively, and then detected watermark from the attached image. Detecting results listed in Table 1 and Table 2, and compared with Tang's method [8].

In table 1 and 2, "a" in "a/b" stands for the number of regions that detect watermark successfully, and "b" stands for the number of regions inserted watermark. "×" means failing to detect watermark.

As can be seen in table 1, when there no attack to the watermarked images, the proposed scheme can detect all local feature regions inserted watermark, and the detected watermark bits matched the inserted watermark bits entirely.

Watermark detected results from attacked images by general signal process are also shown in table 1. From the experiments results in table 1 we can see the proposed

TABLE I.
THE EXPERIMENT RESULT UNSER NO ATTACK AND GENERAL SIGNAL PROCESS

Attacks	Watermarked Lena		Watermarked Baboon		Watermarked Pepper	
	The proposed method	Tang'method	The proposed method	Tang'method	The proposed method	Tang'method
No attack	7/7	7/8	14/14	10/11	8/9	4/4
Jpeg 80	6/7	7/8	10/14	9/11	8/9	3/4
Jpeg 70	4/7	7/8	3/14	11/11	4/9	3/4
Jpeg 50	2/7	7/8	×	7/11	2/9	3/4
Jpeg 40	×	5/8	×	5/11	×	1/4
Median filter (3×3)	6/7	×	12/14	2/11	8/9	×
Gaussian filter (3×3)	2/7	5/8	10/14	8/11	9/9	1/4
Additive noise (d=0.1)	×	5/8	×	6/11	2/9	4/4
Gaussian filter (3×3) +Jpeg90	2/7	5/8	10/14	8/11	9/9	2/4

TABLE II.
THE EXPERIMENT RESULT UNSER GEOMETRIC ATTACKS

Attacks	Watermarked Lena		Watermarked Baboon		Watermarked Pepper	
	The proposed method	Tang' method	The proposed method	Tang' method	The proposed method	Tang' method
Remove 5 rows and 17 columns	5/7	×	11/14	3/11	7/9	×
Remove 8 rows and 20 columns	4/7	×	11/14	×	6/7	×
Cropping 10%	2/7	2/8	10/14	2/11	6/7	2/4
Cropping 20%	2/7	×	8/14	×	4/7	×
Shearing-x-5%-y-5%	5/7	×	11/14	2/11	6/7	×
Shearing-x-10%-y-10%	3/7	×	9/14	×	4/7	×
Rotating 1°+ Cropping	6/7	3/8	12/14	3/11	5/7	2/4
Rotating 30°+ Cropping	5/7	×	8/14	×	5/7	×
Rotating 90°	7/7	×	14/14	×	7/7	×
Scaling 0.5	5/7	×	11/14	×	6/7	×
Scaling 1.2	6/7	×	11/14	×	7/7	×
Linear geometric transformation (1.01,0.013,0.009,1.011)	7/7	4/8	12/14	5/11	5/7	×
Project transformation [0.05 0.02; 1.2 0; 1 1; 0 1.05]	6/7	×	12/14	×	5/7	×
Project transformation [0.2 0.1; 1.01 0; 1 1; 0 0.9]	5/7	×	10/14	×	4/7	×
Rotating 30°+Scaling 0.8+Jpeg 90	5/7	×	7/14	×	5/7	×

method is robust to Jpeg compression, median filter and Gaussian filter, but failed to additive noise and Jpeg compression with lager compression rate. For 3×3 median filter attack, the proposed method shows more robust than Tang's method. The main reason is that SIFT algorithm is more robust to median filter than Mexican hat wavelet scale interaction algorithm. Watermark bit is inserted in 3×3 region around the chosen point, which also strengthens the stability to median filter.

For the additive noise with intension 0.1, expect the watermarked Pepper (detect watermark in two local feature regions in watermarked Pepper), the other two images failed to detect watermark. The main reason is that SIFT feature points are sensitive to noise. We also seen the proposed method fails to Jpeg compression with quality factor less than 70. The main reason is that the proposed method inserts watermark into space domain and Tang's method into DFT domain. The low frequency coefficient in frequency domain is more stable to Jpeg compression than pixel in space domain.

The geometric attack include rotation, scaling, translate, cropping, remove column and row, shearing, linear geometric transformation, projective transformation and their combination. Part experiment results are shown in table 2. Compared with Tang's method the proposed method is robust to wider range of geometric attacks. For rotation, Tang's method only resists the rotation under 5°, while the proposed method can resist rotation with arbitrary angle. For projective transformation, Tang's method is no useful, but the proposed method can resist a certain degree projective transformation.

From table 2 we can see the proposed method has better performance than Tang's method in resisting geometric attacks. There are two main reasons: (1) SIFT feature is more stable to geometric transformation than Mexican hat wavelet scale interaction feature. (2) the proposed method choose watermark inserting location based on cross ratio, which is invariance to projective transformation, making watermark robust to projective transform.

VII. Conclusion

Utilizing two kind of invariance, namely feature points extracted by improved SIFT algorithm and cross ratio of collinear points, this paper presents a watermarking method to resisting geometric attack. The method holds the merit of feature-point-based watermarking, and compared with Tang's method, a classical feature-point-based watermarking method, it has stronger ability to resist geometric attack. The main contribution of this paper include: (1) combine the watermarking system, improved SIFT algorithm and get faster computing speed and more stable feature points. (2) Based on the cross ratio invariance to projective transformation, it choose the location of watermark bit, which guarantees watermark inserting and inserting location synchronization. (3) For quantization modulation embedding method, analyze the relation between quantization step and PSNR, and deduce their relation formula, which help decide the quantization step directly according to the given PSNR.

Acknowledgment

This work was supported in part by a grant from Science Technology Project of Henan, China (No. 0624260019) and National Natural Science Foundation of China (No.092102310163 and 082400410210)

References

[1] D. Zhang, J. Zhao, and A. Saddik. "RST-invariant digital image watermarking based on log-polar mapping and phase correlation. *IEEE Trans Circuits Syst Video Technol*, 2003, 13 (8): 753-765.

[2] CY. Lin, M. Wu, and J A. Bloom. "Rotation, scale, and translation resilient watermarking for images". *IEEE Trans Image Process*, 2001,10 (5): 767–782.

[3] S. Pereira, PUNT. "Robust Template Matching for Affine Resistant Image Watermarks". *IEEE Trans on IP*, 2000, 9(6): 1123-1129.

[4] XG. Kang, JW. Huang, and Y. Lin. "Spread-Spectrum Watermarking Resists to Affine Transformation". *Acta Electronic Sinica*, 2004, 32(1): 8-12. (in Chinese)

[5] P. Dong, JG. Brankoy, and NP. Galatsanos. "Digital watermarking robust to geometric distortions". *IEEE Trans on Image Process*, 2005, 14(12): 2140–2150.

[6] MD. Sang, Y. zhao. "Moment Based Multibit Digital Image Watermarking Resisting to RST Attacks". Journal of Electronic and Information Technology, 2007, 29(1): 1251-1256. (in Chinese)

[7] P. Bas, JM. Chassery, and B. Macq. "Geometrically Invariant Watermarking Using Feature Points". *IEEE Trans. Image Processing*, 2002, 11(9): 1014-1028.

[8] CW. Tang, HM. Hang. "A feature-based robust digital image watermarking scheme". *IEEE Trans. Signal. Process*, 2003, 51 (4): 950–959.

[9] XY. Wang, LM. Hou, and J. Wu. "A feature-based robust digital image watermarking against geometric attacks". *Image and Vision Computing*, 2008, 26(7): 980-989

[10] HY. Lee, HS. Kim, and HK. Lee. "Robust image watermarking using local invariant features". *Optical Engineering*, 2006, 45(3): 037002 (1-10)

[11] K. Miklajczyk, C. Schmid. "A performance evaluation of local descriptors". *IEEE Trans Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615-1630.

[12] DG. Lowe. "Distinctive image features from scale-invariant keypoints". International Journal of Computer Vision, 2004, 60(2): 91-110.

Li Jing, received her PH.D. degree in computer software theory from Information Engineering University, Zhengzhou, China in 2009.

Now she is an ASSOCIATE PROFESSOR of School of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou, Henan Province, China. His research fields include artificial neural network,, data mining and image process, etc.

Dr. Jing is also a member of China Computer Federation.

Xiaowen Zhang, is a postgraduate in the department of computer science, University College Cork, Ireland.

A Novel Inpainting Model for Partial Differential Equation Based on Curvature Function

Jiansheng Liu
College of Science, Jiangxi University of Science and Technology, Ganzhou, P. R. China
Email: jxgzjscn@126.com

Mingming Li
Graduate School, Jiangxi University of Science and Technology, Ganzhou, P. R. China
Email: leeag201@126.com

Fangfang He
Graduate School, Zhongshan University,Guangzhou, P.R.China
Email:814187937@qq.com

***Abstract*—For the curvature function of Curvature-Driven Diffusions model (CDD), this paper studies the impact of the CDD on the inpainting, and a novel inpainting model based on partial differential equation (PDE) is proposed which contains five sub-models. For simplification, a discrete model based on this PDE is given. Experimental results show that we can obtain better inpainting effects compared with each other from those sub-models.**

***Index Terms*—Digital Image Inpainting, Partial Differential Equations (PDE), Curvature**

I. INTRODUCTION

Image inpainting can fill up the defect field of image in accordance with certain rules of procedure using information of image, which improve the visualization performance of the original image. In mathematics, two models can be used to describe an inpainting image, for example, Data model(Read existing information) and Prior model(Gets the type original image). In this way, we can take an inpainting problem as optimization problem for an energy function which using mathematics language to express the progressing of image inpainting. Then, we can obtain a data model and a prior model of the image after inpainting[1]. This kind of algorithm always called image inpainting algorithm based on PDE. Another kind of image completion algorithm is mainly used to large objects which may loss of information[2]. Its main algorithm based on texture synthesis, the main idea is to select an element from the boundary of inpainting field firstly and on the element, which can choose the appropriate size of texture objects based on its features[3.4]. Finally we can find the most relatively texture around the inpainting field for it and instead of it secondly.

Bertalmio-Sapiro-Caselles-Ballester (BSCB) model has been raised in literature[5], which is a third-order PDE, its main idea is inpainting along the direction of illumination (means the grayscale image in the same level consisting of a series of points on a line) iterations which considering the information around inpainting field to its inside, then, calculate the discrete gradient vector (means spatial maximum change of direction) of each point stood on edge line of inpainting field with rotating $90^{0\neg}$. This model can be better inpaint image and avoid edge blur, but it leak on stability and poor outcomes represents which is not good for a rich texture image.

Literature [6] proposed a total variation algorithm (TV model) and is a second-order PDE, during the progress of image inpainting, it will make anisotropic diffusion. An advantage of this model is that it can keep edge and guarantee value simply. We know that it comes from the p-Laplace[7] inpainting model, and when the iteration P had been set as a stable value, we can name it as the TV model. Although it works better than the BSCB model about stability and robustness. It may break the theory of visual connection and causes ladder-effect.

Literature [8, 9, 10] improved TV inpainting model in the domestic, these methods make the inpainting effect better than the ordinary TV model. As we know that the most difficult problem of TV model is that it may make staircase effect, which influence our inpainting performance in some extends.

A Curvature-Driven Diffusion (CDD) model has been raised in Literature [11], which is a third-order PDE simulation of diffusion and spreading process. Although it can keep a well inpainting performance, its higher computational complexity makes this model poor because of its relation to the anisotropic diffusion.

Literature [12] locates about fast algorithm of CDD model, it aims to accelerate inpainting time of the traditional CDD model.

Literature [13] puts forward a harmonic inpainting model, an isotropic diffusion with second-order PDE, so it can exclude noise while obscures edge.

Literature [14] submits a p-harmonic model based on literatures [6, 12], they combine the advantages of the TV model and the harmonic model which shorten the inpainting computing complexity and better the inpainting performance.

doi:10.4304/jmm.7.3.239-246

Literature[15] proposes a Criminisi inpainting algorithm, it belongs to the second type of inpainting technology, which is a sample-based texture synthesis algorithm not only image with large bulk of loss information can be repaired, but also details, both in time and visual effects better than other algorithms.

II. BRIEF INTRODUCE FOR TV MODEL AND CDD MODELS

A . TV model

The total variation(TV) model was shown by Tony Chan in 2002, the main idea of the TV model is that it can keep the edge while anisotropic diffusion to obtain a result which can restore all the images. From the equation defined from P-laplace equation

$$\begin{cases}\Delta_p I-\lambda(I-I^0)=0, 1<p<\infty, \\ \mathrm{I}=0\, above \quad \Omega\end{cases},$$

where $I^0=I+N$ is a model contains an additive noise and the relatively energy functional equation is:

$$J[I]=\frac{1}{p}\int_{\Omega}|\nabla I|^p\, d\sigma+\frac{\lambda}{2}\int_{\Omega\setminus D}(I-I^0)^2 d\sigma,$$

Especially when $P=1$, we can regard the image as a smooth piecewise function which is good for establishing models for images on bounded space, then the total variation image inpainting model is:

$$\min J[I]=\int_{\Omega}|\nabla I|\, d\sigma+\frac{\lambda}{2}\int_{\Omega\setminus D}(I-I^0)^2 d\sigma \qquad (1)$$

where $I(x,y)$ is the image grayscale function, reduced it to I, I^0 is the defect image grayscale function, D is the inpainting field, Ω is the total image field. $\Omega\setminus D$ is a known image field, λ is the multiplier of Lagrange, the above formula is a double integration, $d\sigma=dxdy$.

Based on the variation theory[16], we can determine the Euler equations as follows:

$$-\nabla[\frac{\nabla I}{|\nabla I|}]+\lambda_D(x,y)\cdot(I-I^0)=0 \qquad (2)$$

where

$$\lambda_D(x,y)=\lambda\cdot I_{\Omega\setminus D}(x,y)=\begin{cases}\lambda,(x,y)\in\Omega\setminus D\\ 0,(x,y)\in D\end{cases},$$

and

$$I_{\Omega\setminus D}(x,y)=\begin{cases}1,(x,y)\in\Omega\setminus D\\ 0,(x,y)\in D\end{cases}.$$

We should consider that $|\nabla I|$ may be near to 0, so we make the $v=\frac{1}{|\nabla I|_{\varepsilon}}=\frac{1}{\sqrt{|\nabla I|^2+\varepsilon^2}}$ instead of $v=\frac{1}{|\nabla I|}$, while the $v=\frac{1}{|\nabla I|}$ is a small positive number, which makes the strength of diffusion more huge while grads be small and be small while grads be huge.

Euler showed a power definition of curve firstly in research under the action of external forces, we have already know two end point of curve A and B, the strength is d ,the equation is $E(d)=\int_{AB}(\alpha+\beta k^2)ds$, Chan and his team elastic energy into the TV model, then TV-Euler had been showed:

$$E[I\,|\,I^0,D]=\int_{\Omega}\varphi(k)|\nabla I|\, d\sigma+\frac{\lambda}{2}\int_{\Omega\setminus D}(I-I^0)^2 d\sigma,$$

where $\varphi(k)=\alpha+\beta k^2$, $k=div(\frac{\nabla I}{|\nabla I|})$, when $\alpha=1,\beta=0$, this is a TV model.

B. CDD model

In order to make TV model keeps a better visual connectivity, we modified it and considered a CDD model. Modifying the diffusion coefficients of TV model $v=\frac{1}{|\nabla I|}$ to $v=\frac{g(|k|)}{|\nabla I|}$, where

$$g(|k|)=\begin{cases}0,k=0\\ \infty,|k|=\infty\\ >0,0<|k|<\infty\end{cases},$$

In generally, we suppose $g(s)=s^a, s>0, a\geq 1$, while we set $a=1$ here, because it makes diffusion be strong in large curvature while weaker in small curvature, then the corresponding CDD model is:

$$\begin{cases}\frac{\partial I}{\partial t}=\nabla\bullet\left[\frac{g(|k|)\nabla I}{|\nabla I|}\right]+\lambda_D(I^0-I)\\ g(\bullet)\, is \quad increa\sin g \quad function, k=\nabla\bullet\left[\frac{\nabla I}{|\nabla I|}\right]\end{cases} \qquad (3)$$

This is better than TV model of the inpainting performance.

C. P-harmonic model

We can establish a relative energy function model according to the noise included in the image when $1<p<2$ as follows:

(1): Model I(no noise)

$$J[I]=\int_{\Omega}\frac{1}{p}|\nabla I|^p\, d\sigma$$

(2): Model II(noise)

$$\begin{cases}J_2[I]=\int_{\Omega}\frac{1}{p}|\nabla I|^p\, d\sigma+\frac{\lambda_D}{2}\int_{\Omega}(I-I^0)^2 d\sigma\\ \lambda_D(x,y)=\lambda\cdot I_{\Omega\setminus D}(x,y)=\begin{cases}\lambda & (x,y)\in\Omega\setminus D\\ 0 & (x,y)\in D\end{cases}\end{cases}$$

In general, we discuss the noise model, the energy functional is :

$$J[I]=\int_{\Omega}F(x,y,I,I_x,I_y)d\sigma$$
$$=\iint_{\Omega}F(x,y,I,\frac{\partial I}{\partial x},\frac{\partial I}{\partial y})dxdy$$

After inpainting the image, we can get the value from the following functions:

$$\begin{cases}\dfrac{\partial I}{\partial t}=\nabla\cdot\left(|\nabla I|^{p}\nabla I\right)+\lambda_D(I^0-I) & innerD\\ I=I^0 & inner\partial D\end{cases}$$

III. REALIZATION OF THE PARTIAL DIFFERENTIAL INPAINTING MODEL BASED ON CDD MODEL

A. Model proposed

Based on the foundation of the TV model and the CDD model, we can study the influence of the curvature function $g(\cdot)$ on inpainting performance, here, we contruct the curvature function as follows:

$$g(s)=as^p+b(e^s-1)+cIn(s+1)+d,\ p\geq 1, s>0,$$

where $a,b,c,d\geq 0$.

The purpose of such a construction, firstly, it reaches to discuss the influence for the result of inpainting of these different curvature functions and which contains power functions, exponential functions, logarithmic functions and those functional images are shown in figure 1; Secondly, it summarizes many models which have been introduced before, such as $d=1$,while others are 0, back into TV model; $a=1, p=1$, while others coefficient are 0, obtains a general CDD model[17]. And to some extent, it can supply a gap for TV model during inpainting a large square of defects field which may destroy visual connectivites. So the new established inpainting model is

$$\begin{cases}\dfrac{\partial I}{\partial t}=\nabla\bullet\left[\dfrac{g(|k|)\nabla I}{|\nabla I|}\right]+\lambda_D(I^0-I).k=\nabla\bullet[\dfrac{\nabla I}{|\nabla I|}]\\ g(s)=as^p+b(e^s-1)+cIn(s+1)+d,\ p\geq 1.s>0\end{cases}\qquad(4)$$

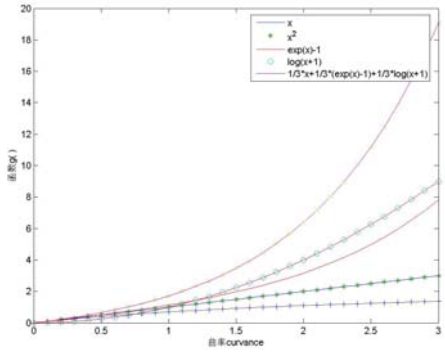


Figure 1. differential curvature function image

when a,p,b,c,d have been set different values, we can get models in special case as its sub-model:

(1) when $d=1$, others are 0,back into TV model, it is a constant model;

(2) when $a=1, p=1$,others are 0, it is a general CDD model and called as Power curvature model;

(3) when $b=1$, others are 0, called as exponential curvature model.;

(4) when $c=1$, called logarithmic model;

(5) when $a=1/3$, $p=1$, $b=1/3$, $c=1/3$, $d=0$, called comprehensive model.

B. Discretization of Model

For calculate conveniently, we need to discrete the above models and put forward discrete model. After discretizing a serials of PDE, the corresponding digital filter has been produced through a digital iteration method, we consult from literature[6,12] to use half-point difference scheme method that is considering for the gradient solutions algorithm of eight neighborhoods of inpainting pixel point.

As shown in figure 2, O is an inpainting pixel point, $\wedge_O=\{N,S,W,E\}$ is O 's four neighborhood collections. $\wedge=\{n,s,w,e\}$ is been used for four half-pixel neighborhoods.

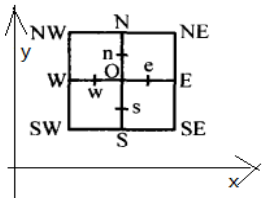


Figure 2. inpainting pixel point and neighborhood

Ordering $v=(v^1,v^2)=\dfrac{g(|k|)\nabla I}{|\nabla I|}$, then give a central difference numerical value for divergence:

$$\nabla\cdot V=\frac{\partial v^1}{\partial x}+\frac{\partial v^2}{\partial y}\approx\frac{v_e^1-v_w^1}{h}+\frac{v_n^2-v_s^2}{h},$$

here, we can set step $h=1$.

In order to avoid $|\nabla I|$ to be 0, we can set parameters as following:

$$|\nabla I|=\sqrt{\varepsilon^2+|\nabla I|^2},$$

or

$$|\nabla I|_{\varepsilon}=\varepsilon+|\nabla I|.$$

Then we calculate the similarity of v_e^1, v_w^1, v_n^1, v_s^1:

$$\begin{cases} v_e^1 = \dfrac{g(|k_e|)}{|\nabla I_e|_\varepsilon}[\dfrac{\partial I}{\partial x}]_e \approx \dfrac{g(|k_e|)}{|\nabla I_e|_\varepsilon}\dfrac{I_E - I_O}{h} \\ |\nabla I_e|_\varepsilon \approx \dfrac{1}{h}\sqrt{\varepsilon^2 + (I_E - I_O)^2 + (\dfrac{I_{NE} - I_{SE} + I_N - I_S}{4})^2} \end{cases}$$

$$\begin{cases} v_w^1 = \dfrac{g(|k_w|)}{|\nabla I_w|_\varepsilon}[\dfrac{\partial I}{\partial x}]_w \approx \dfrac{g(|k_w|)}{|\nabla I_w|_\varepsilon}\dfrac{I_W - I_O}{h} \\ |\nabla I_w|_\varepsilon \approx \dfrac{1}{h}\sqrt{\varepsilon^2 + (I_W - I_O)^2 + (\dfrac{I_{NW} - I_{SW} + I_N - I_S}{4})^2} \end{cases}$$

$$\begin{cases} v_n^2 = \dfrac{g(|k_n|)}{|\nabla I_n|_\varepsilon}[\dfrac{\partial I}{\partial x}]_n \approx \dfrac{g(|k_n|)}{|\nabla I_n|_\varepsilon}\dfrac{I_N - I_O}{h} \\ |\nabla I_n|_\varepsilon \approx \dfrac{1}{h}\sqrt{\varepsilon^2 + (I_N - I_O)^2 + (\dfrac{I_{NE} - I_{NW} + I_E - I_W}{4})^2} \end{cases}$$

$$\begin{cases} v_s^2 = \dfrac{g(|k_s|)}{|\nabla I_s|_\varepsilon}[\dfrac{\partial I}{\partial x}]_s \approx \dfrac{g(|k_s|)}{|\nabla I_s|_\varepsilon}\dfrac{I_S - I_O}{h} \\ |\nabla I_s|_\varepsilon \approx \dfrac{1}{h}\sqrt{\varepsilon^2 + (I_S - I_O)^2 + (\dfrac{I_{SE} - I_{SW} + I_E - I_W}{4})^2} \end{cases} \tag{5}$$

others are similarly.

In order to get the results of the above expression, we still need to calculate $g(|k_\varepsilon|), g(|k_w|), g(|k_n|), g(|k_s|)$. When $g(|k|)=1$, we can get the computational formula $k = \nabla \cdot (\dfrac{\nabla I}{|\nabla I|})$. If the value of $g(\cdot)$ has given, we can obtain the value of $g(|k|)$. Apparently, the inpainting value of the target pixel point O just relates to 3×3 pixel points of its neighborhood (which did not mark out the half-pixels in figure), which has been shown in table 1:

TABLE 1. PIXEL POINT O'S 3×3 NEIGHBORHOOD

NW	N	NE
W	O	E
SW	S	SE

In this way, we need to calculate the value of k_q, where q is the half-pixel point of $\wedge$, which we also choose difference method to calculate it. Assuming that:

$$\frac{\nabla I_q}{|\nabla I_q|} = (\frac{\nabla I_q^1}{|\nabla I_q|}, \frac{\nabla I_q^2}{|\nabla I_q|}) = (v_q^1, v_q^2),$$

$$k_q = \nabla \cdot (\frac{\nabla I_q}{|\nabla I_q|}) = \frac{\partial v_q^1}{\partial x} + \frac{\partial v_q^2}{\partial y}.$$

The calculation formula of k_e is given in the following formulas. We expand it as follows:

$$k = \frac{\partial v_e^1}{\partial x} + \frac{\partial v_e^2}{\partial y} = \frac{v_{eE}^1 - v_{eO}^1}{h} + \frac{v_{eNO}^1 - v_{eSE}^2 + v_{eN}^2 - v_{eS}^2}{4h}$$

But if we can't calculate v_{eE}^1 based on the formula in the 3×3 neighborhood, then we may turn to consider the 5×5 neighborhood, which are shown in table 2 and figure 3.

TABLE 2. PIXEL POINT O'S 5×5 NEIGHBORHOOD

•	•	O_5	O_4	O_3
•	NW	N	NE	O_2
•	W	O	E	O_1
•	SW	S	SE	O_{16}
•	•	•	•	•

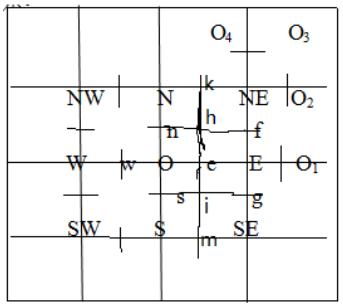


Figure 3. 5×5 Neighborhood and its half-pixel point

Then

$$\begin{cases} v_{eE}^1 = \dfrac{1}{|\nabla I_E|}[\dfrac{\partial I}{\partial x}]_E = \dfrac{1}{|\nabla I_E|}\dfrac{I_{O_1} - I_O}{2h} \\ |\nabla I_E| = \sqrt{(\dfrac{I_{O_1} - I_O}{2h})^2 + (\dfrac{I_{NE} - I_{SE}}{2h})^2} \end{cases},$$

$$\begin{cases} v_{eO}^1 = \dfrac{1}{|\nabla I_O|}\left[\dfrac{\partial I}{\partial x}\right]_O = \dfrac{1}{|\nabla I_O|} = \dfrac{1}{|\nabla I|_O}\dfrac{I_E - I_W}{2h} \\ |\nabla I_O| = \sqrt{(\dfrac{I_E - I_W}{2h})^2 + (\dfrac{I_N - I_S}{2h})^2} \end{cases},$$

$$\begin{cases} v_{eNE}^2 = \dfrac{1}{|\nabla I_{NE}|}[\dfrac{\partial I}{\partial y}]_{NE} = \dfrac{1}{|\nabla I_{NE}|}\dfrac{I_{O_4} - I_E}{2h} \\ |\nabla I_{NE}| = \sqrt{(\dfrac{I_{O_2} - I_N}{2h})^2 + (\dfrac{I - I}{2h})^2} \end{cases}$$

In the same way, we can obtain the values of $v_{eO}^1, v_{eSE}^2, v_{eN}^2, v_{eS}^2$. They are as follows:

$$\begin{aligned} \nabla \cdot [\frac{g(|k|)\nabla I}{|\nabla I|}] &= \nabla \cdot v = \frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} \\ &= \frac{v_e^1 - v_w^1}{h} + \frac{v_n^2 - v_s^2}{h} = \sum_{Q \in \wedge_O, q \in \wedge} \frac{g(|k_q|)}{|\nabla I_q|_\varepsilon}[I_Q - I_O] \end{aligned} \tag{7}$$

where $Q \in \wedge_O, q \in \wedge$ is to say that Q is the pixel point in $\wedge_o$, q is the corresponding half-pixel point of $\wedge$, if Q equal to E, then q equal to e. After setting the above formula to formula (4), we can obtain:

$$I(O)=\frac{\sum_{Q\in\wedge_O,q\in\wedge}\frac{g(|k_q|)}{|\nabla I_q|_\varepsilon}}{\lambda_D(O)+\sum\frac{g(|k_q|)}{|\nabla I_q|_\varepsilon}}I(Q) + \frac{\lambda_D(O)}{\lambda_D(O)+\sum_{Q\in\wedge_O,q\in\wedge}\frac{g(|k_q|)}{|\nabla I_q|_\varepsilon}}I^0(Q) \quad (8)$$

Then, we reformulated it as follows:

$$\omega_q=\frac{g(|k_q|)}{|\nabla I_q|_\varepsilon}, h_Q=\frac{\omega_q}{\lambda_D(O)+\sum_{Q\in\wedge_O,q\in\wedge}\omega_q}, \quad h_O=\frac{\lambda_D(O)}{\lambda_D(O)+\sum_{Q\in\wedge_O,q\in\wedge}\omega_q} \quad (9)$$

Finally, we can obtain:

$$I(O)=\sum_{Q\in\wedge_O,q\in\wedge} h_Q I(Q) + h_O I^0(o) \quad (10)$$

The above expression can be understood as inpainting target pixel values through doing weighted h_Q for its neighborhood $I(Q)$, so it can be considered as weighted inpainting model[18].

Using Gauss-Jacobin iterative method, where n represents its iterations, the above formula can be conveyed as:

$$I(O)=\sum_{Q\in\wedge_O,q\in\wedge} h_Q^{n-1}I^{n-1}(Q)+h_O^{n-1}I^0(O) \quad (11)$$

While in the equation $\sum_{Q\in\wedge_O,q\in\wedge} h_O + h_O = 1$.

C. Model simulation

When progresses with numerical simulation, ordinary we take mask to convince the inpainting field D, then use inpainting algorithm automatically to restore those missing information according to the around information of inpainting field. Depend on the formula (III.10), the steps of improved model for differential image inpainting algorithm are:

(1) read image I and mask information I_D into computer, give the value of λ, ε;

(2) performing (3),(4),(5) on each pixel of mask, until reaching the iterations that in rules;

(3) if pixel locates beyond inpainting field, $\lambda_D = \lambda$, otherwise $\lambda = 0$;

(4) depend on formula (9) to calculate ω_q, h_Q, h_O;

(5) get a new pixel according to the (10) and store it to new image[19].

In the followings simulation test, $\lambda = 0$, $\varepsilon = 0.0001$. We simulate using MATLAB 7.0, some basically procession of image have been related to literature [12].

Test I The comparison chart of visual effect on each sub-model with stunt making and data evaluation.

Figure 4. Original images

Through making stunt for image we can wipe out steel wire in figure 4. Firstly labeling that means using mask to label the steel wire, as has shown in A, this paper have labeled out it by pure white color(255).The inpainting performance will be better if use impure white color and the larger the iterations, the better effect of inpainting.

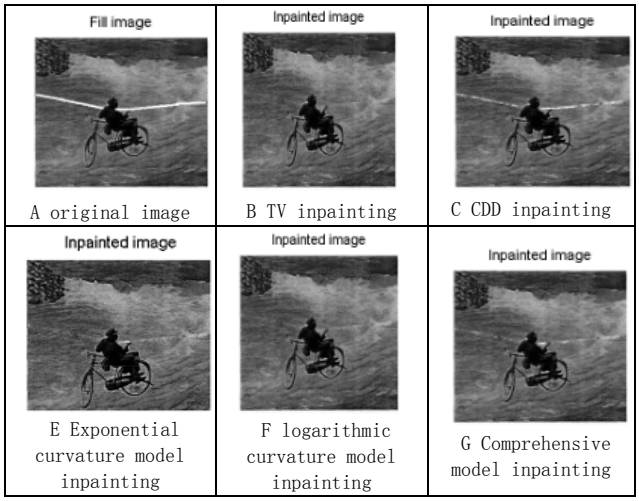


Figure 5. The comparison diagram of each sub-model inpainting performance

Each inpainting image of every sub-model is the inpainting performance during the iteration stays at N=60. So we know that the inpainting performance is very well for the TV model and logarithm curvature model when the inpainting square keeps very small.

TABLE 3. PSNR THE QUALITY EVALUATION METRIC OF EACH SUB-MODEL AND ITERATION DATA COUNT TABLE

	N=10	N=20	N=40
Constant CM(TV)	63.412	74.178	75.735
Power CM(CDD)	53.627	56.932	64.951
Exponential CM	53.06	54.053	57.142
Logarithm CM	54.214	59.132	69.692
Comprehensive CM	53.312	55.343	61.986
	N=60	N=80	N=100
Constant CM(TV)	75.799	75.832	75.819
Power CM(CDD)	71.606	74.189	74.35
Exponential CM	61.686	66.566	70.706
Logarithm CM	74.739	74.732	74.812
Comprehensive CM	68.824	73.339	74.114

From the table 3 we know that inpainting performance is the best for the TV model when deal with these scores in small square fields, then logarithm curvature model, power curvature model, comprehensive curvature model and exponential curvature model. We can conclude that rather the indigestion of $g(\cdot)$ have not represent its advantage for us than the faster it increases the worst the inpainting effect when dealing with smooth and small square fields in the image inpainting.

Figure 6 is the image of table 3, from top to bottom are the images relate to the TV model, logarithm curvature model, CDD model, comprehensive curvature model and exponential curvature model. While the PSNR is a peak value signal noise ratio, which is used for taking measure for the inpainting images, the function is:

$$PSNR = 10 \cdot l\mathrm{g}\left\{\frac{255^2}{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}[\hat{I}(x,y) \_ I(x,y)]^2}\right\}$$

This is better than the usually measures such as SNR:

$$SNR = 10 \cdot l\mathrm{g}\left\{\frac{\sum_{x=1}^{M}\sum_{y=1}^{N}[\hat{I}(x,y)]^2}{\sum_{x=1}^{M}\sum_{y=1}^{N}[\hat{I}(x,y) - I(x,y)]^2}\right\}$$

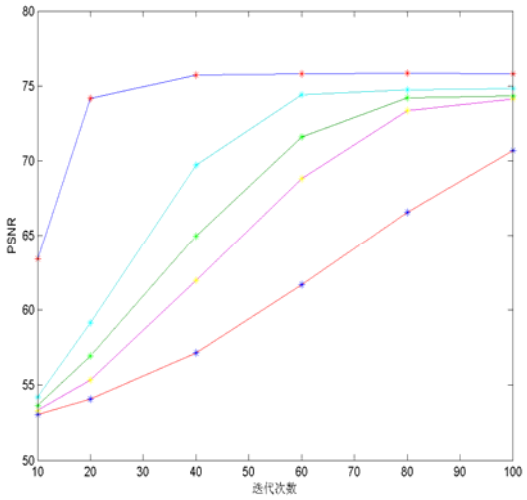


Figure 6. PSNR related to each sub-model and iteration count graphs

The PSNR is a normal measures to use in inpainting images, especially used for some images need to be wiped noise while the bigger the value of it, the better of the effect[20].

Test II. Comparison chart of multicolor image inpainting effect on each sub-model

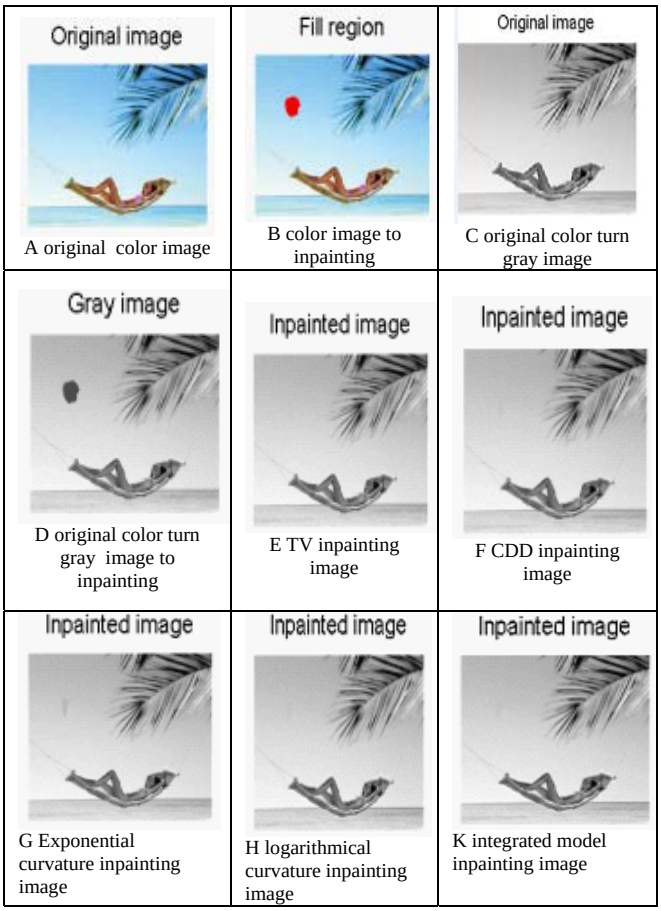


Figure 7. t comparison chart of inpainting effect of multicolor turn to grayscale

The inpainting field of the test which is a red field in a little large square in the multicolor image B, color is [255, 0, 0], from the inpainting effect we know that each model represents its effect slowly. Exponential curvature model still has a part of black field not repaired and TV model, logarithm curvature model, CDD model and comprehensive model have the better inpainting performance.

Test3 Image denoising test for each sub-model

Figure 8. original image

The noise plus in the test denoising image is impulsive noise, and the impulsive probability is 10%, iteration is 2. We know from the result of the test that TV model makes work very functional especially on denoising but it makes a soft edge for image. Yet, logarithm curvature model, exponential curvature model and comprehensive model are not better than the TV model on denoising effect, but they still keep the image clear[21].

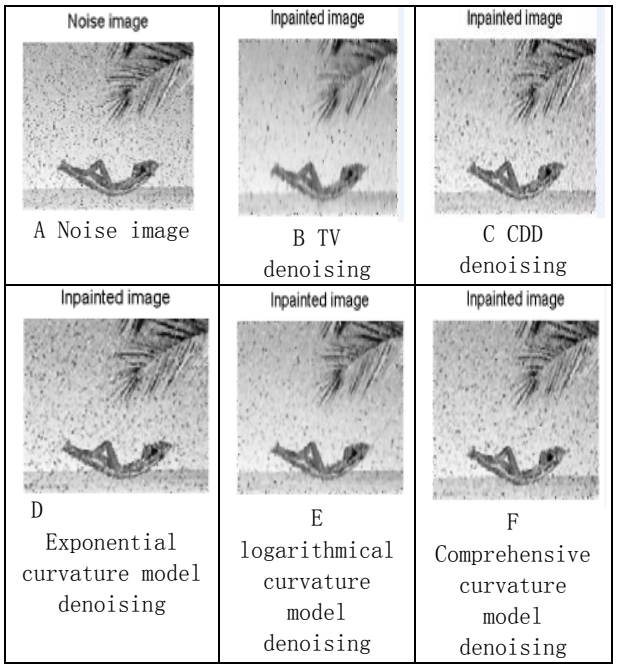


Figure 9. Comparison chart of image denoising

Test 4 Inpainting image and denoising in the same time for logarithm curvature of sub-model

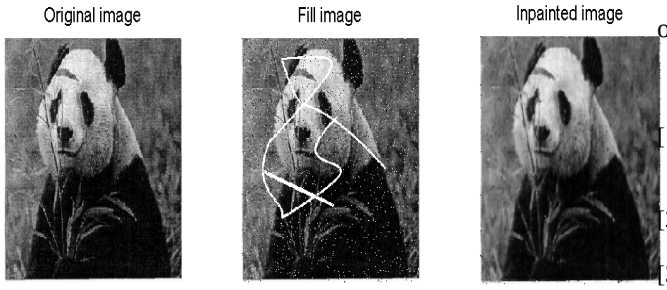


Figure 10. Panda as well as noise and removed scratches

In the figure 10 the iteration is N=40, the impulsive noise in the middle image is impulsive probability is 3%. The last image is the image after denoising and wiping off scores and which looks clearer than the original one.

Test 6 Test of removing text by logarithm curvature model of sub-model

Figure 11. removing text for panda chart

The image which had been pulsed words to the gray image for the middle image of the above images, removing its words on the surface and the color is [255,255,0], if we wipe out some white dot for the image, the new image will looks like removed back to the original image from visual terms. So we know that the logarithm curvature model has a good effect of removing text[22].

D. Concluusion

Based on the TV model and the CDD model in the II, we had modified it according to the advantage and disadvantage of the TV model, we gived out a continual partial diffusion equation model,when all iterations of logarithm function in model have been set differently, we get some other sub-models, which contains those models: TV model, CDD model, exponential curvature model, logarithm curvature model and comprehensive model. For simplisity to study, we put forward discretization model in the whole paper, we considered the area for discreting it in a 5×5 area. So we can obtain the related discretization model for each sub-model, and the new TV model and the CDD model after discreting are different while used with the means of difference equation and difference area are differenr. At last we have simulated it by MATLAB, we test a set of images, then we could convince that discretization model has a well inpainting effect on inpainting images and denoising for small area, and these sub-models are well except exponential curvature model.

E. Acknowledge

We get the support from doctor Liu Songhua in college of science of Jiangxi university of science and technology.

REFERENCES

[1] Bertalmio M,Saprio G,Caselles V,et al.Image inpainting[J].Proceedings of ACM SIGGPAPH,New York,ACM Press, 2000:417-424.
[2] YaDong Wu,Hongying Zhang,Bin Wu,Digital image inpainting technology[M].Beijing: Science press, 2010.3.
[3] Chan T F,Shen J H.Mathematical models for local mom-texture inpainting[J].Society for Industrial and Applied Mathematics, 2001 , 62(3):564~592.
[4] Tang F,Ying Y T,Wang J,Peng Q S.A Novel Texture Synthesis Baed Algorithm for Object Removal in Photographs[J].Ninth Asian Computing Science Conference.China Mai,Thailan,2004:248-258.
[5] A.K.Jain.Partial Differential Equations and Finite-Difference Methods in Image Processing[J].Journal of Optimization Theory and Application.1977,91(23):65-91.
[6] Rudin L,Osher S,Faterni E.Nonlinear total variation based noise removed algorithms.Physical D, 1992, 60(1-4):259-268.
[7] Luminita A.Vese and Stanley J.Osher. Mumerical Methods for p-Harmonic Flows and Applications to Image Processing. AIAM Journal on Numerical Analysis. Vol.40,No.6(2003).PP:2085-2104.
[8] Xiaowei Shaoi,Zhengkai liu,Bi Song.A TV model based adaptive image restoration method[J].Journal of circuits and systems.2004.4(9):113-117.
[9] Yuli Lin, Junhong Zhao, Xuefeng Zhu, Yongjian Hu. Improved TV model of image restroration algorithms[J]. computer engineering and design. 2010,31(4).
[10] Yunyun Xu,Xiaolin Zhu,shubin Huang,kun Zhu.The Application of improved algorithm based on TV model in image inpainting[J].Journal of Hefei University of Technology.2010.12(33):1919-1920.
[11] Tony F.Chan and Jianhong Shen. Mathematical Models for Local Nontexture Inpaintings. Society for Inderstrial and

Applied Mathematics. Vol.62. No.3(Dec.2001-Feb.2002). PP:1019-1043.
[12] Xianglin Li,Fast repair algorithm based on CDD model[J].Computer simulation.2008.10(28):223-227;
[13] Andreas Hofhauser, Carsten Steger and Massir Navab. Harmonic Deformation Model For Edge Based Template Matching. VISAPP 2008.International Conference on Computer Vision Theory and Applications.Pages:75-82.
[14] Jooyoung Hahn,Xue-Cheng Tai,Sofia Borok and Alfred Marcel Bruckstein. Orientation-Matching Minimization for Image Denoising and Inpainting. International Journal of Computer vision. Vol 92. Date of Publication:May.2010. Pages:308-324.
[15] Criminisi A, Perez P and Toyama K. Object removal by exemplar-based inpainting. Computer Vision and Pattern Recognition. Date of Publication:18-20 June 2003. Pages:II-727-II-728.
[16] Chan T F,Kang S HmShen J H.Euler's elastical and curvature based inpainting[J].SIAM Journal of Applied Mathematics,2002,63(2):564~592.
[17] Roesser R. A discrete state-space model for linear image processing. Automatic Control of Journal&Magazines.Date of Publication:Feb.1975.On Pages:1-10.
[18] Zhaoli Zhang, Chunhui Zhao, Xiaodan Mei. Modern image technology and implementation in Matlab[M]. Beijing:People's posts and telecommunications press. 2001.11.
[19] Guy Gilboa and Stanley Osher. Nolocal operators with applications to image processing. AMS subject classifications. 2008.
[20] Saha S and Vemuri R. An analysis on the effect of image features on lossy coding performance. Signal Processing Letters of Journals&magazines.Date of Publication:May 2000. On Pages:104-107.
[21] X Lian, Z Xu, L Feng. Novel Image Inpainting Based on Partial Differential Equation[J]. Computer Engineering . 2009-06.
[22] Chang S.G, Bin Yu, Vetterli M. Adaptive wavelet thresholding for image denoising and compression[J]. Journals&Magazines.Sep 2000.Pages:1532-1546.

Jiansheng Liu, mail, was born Jiangxi of China in 1959.He is an Associate Professor in Jiangxi University of Science and Technology, China. His interests include Information Security, Artificial Intelligence and Numerical Analysis.

Mingming Li, postgraduate student of college of science in Jiangxi University of Science and Technology, the research direction is Intelligence computation and Intelligent recognition.

Fangfang He, postgraduate student of Zhongshan University, Guangzhou,China. the research direction is Numerical Analysis .

A Novel De-noising Model Based on Independent Component Analysis and Beamlet Transform

Guangming Zhang
The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou, China, 215006
Email: gmwell@gmail.com

Zhiming Cui[1,2], Pengpeng Zhao[1] and Jian Wu[1,2]
[1] The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou, China, 215006
[2] JiangSu Province Support Software Engineering R&D Center for Modern Information Technology Application in Enterprise, Suzhou, China, 215104
Email: szzmcui@suda.edu.cn

***Abstract*—Vehicle video key frame processing as an important part of intelligent transportation systems plays a significant role. Traditional vehicle video key frame extraction often has lots of noises, it can't meet the requirements of the recognition and tracking. In this paper, a novel method which is combined independent component analysis with beamlet transform is proposed. Firstly, a random matrix was produce to separate the key frame into a separated image for estimate. Then beamlet transform was applied to optimize the coefficients. At last, the coefficients were selected for image reconstruction by inverse of the beamlet transform. By contrast, this approach could remove more noises and reserve more details, and the efficiency of our approach is better than other traditional de-noising approaches.**

***Index Terms*— beamlet transform; independent component analysis; de-noising; video key frame**

I. INTRODUCTION

In recent years, intelligent transportation system (ITS) has developed quickly. The video vehicle detection system is gradually becoming one of the popular methods for the detection of vehicles to its convenient installation and rich information content provided [1]. Meanwhile the independent component analysis (ICA) [2] as a kind of new signal processing method developed quickly and widely used ,especially in image processing. David L. Donoho and Xiaoming Huo [3] developed a new theory of multiresolution analysis called the beamlet transform. This mathematical transform differs from wavelet and related other mathematical transform. While wavelets offer localized scale/location representation near fixed regions of space with specified scale and location, beamlets have localized scale/location/orientation based on dyadically-organized line segments. This new mathematic transform is based on the wavelet transform and beyond it. It has overcome some limitations of wavelet transform in vehicle video key frame processing. Because of the character of these two image processing method, we combine independent component analysis with beamlet transform to optimize the coefficients for vehicle video key frame de-noising.

Corresponding author：Zhiming Cui, Guangming Zhang
Email: szzmcui@suda.edu.cn, gmwell@gmail.com

As a signal analysis technique, ICA is a useful method for separating the independent signals from overlapping signals [3]. It was greatly developed as a potential statistical technique for blind source separation (BSS). It aims at finding the hidden components inside the original signals, and the components capture the essential structures of the signals. ICA is often used by the image processing. Through making full use of the high-order statistical characteristics of the source, i.e., the fourth-order central moment, ICA can effectively resolve the independent components (ICs) from the measured mixed signals without any additional information about the source signals. It had been widely applied in the signal processing fields, such as biomedical signals, image processing and financial analysis.

Beamlet transform as a newly developed mathematical transform is often used as time-frequency and multiresolution analysis tool in the signal and image processing domain. Beamlet transforms [4] are insensitive to noise, computationally efficient, and able to detect features with high accuracy than the general signal detection algorithms, which are based on pixel-level filtering and can have very poor SNR ratios. So in the view of the combination ICA technique and beamlet transform, this research is initial. For this reason, a vehicle video key frame de-noising model based on ICA and beamlet transform are proposed in the paper.

Vehicle video key frame processing is an important part of ITS construction. We could think of the information of vehicle video as some key frame to contain. By analyzing the characters of vehicle video key frame, we find that when we denoised the vehicle video key frame we should decompose the image by ICA. Because a great deal of vehicle video key frames are taken by the same instrument at different time, so the image decomposition signal could be optimized correctly by beamlet transform.

In this paper the vehicle video key frame which extracted at fixed interval (per second) is shown in Figure 1.

doi:10.4304/jmm.7.3.247-253

Figure 1. vehicle video key frame

The vehicle video key frame sequence extracted is in Figure 2.

These vehicle videos is captured by a video camera with 2,000,000 pixels and 30 frames per second.

There is always some noise produced by the video capture equipments and the processing of transmission. For this reason, this paper proposes a new model which could denoise the vehicle video key frame by ICA and beamlet transform. By experiment, the vehicle video key frame after this approach's processing can display the details much more clearly. It could meet the basic requirement for future processing like image segmentation, object recognition and image annotation.

The rest of this paper is organized as follows. The independent component analysis is given in Sections II. beamlet transform analysis is given in Sections III. Section IV describes the experiment which is using ICA and beamlet transform, then discusses the results. Conclusions are presented in Section V.

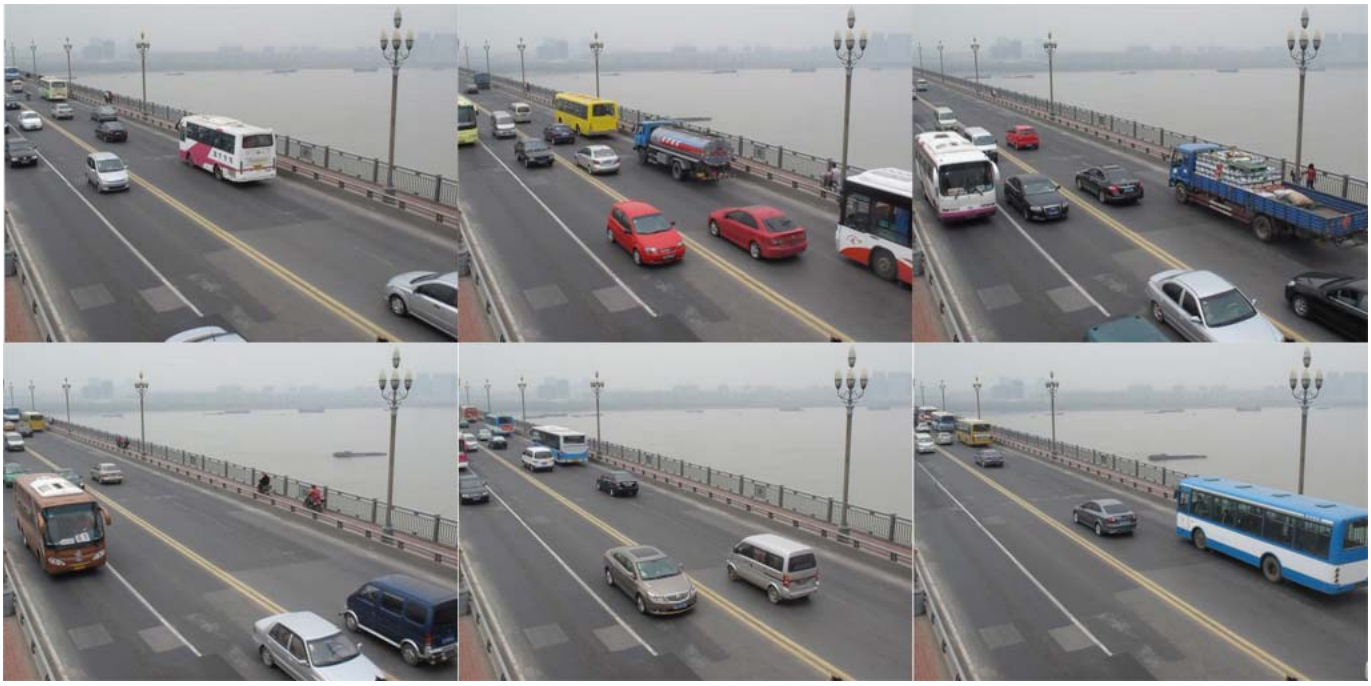


Figure 2. vehicle video key frame sequence

II. Independent Component Analysis

As opposed to the principal component analysis (PCA) that researches the pertinent relation among signals based on the second order statistics, Independent component analysis (ICA) [2]is a modern factor analysis tool developed in the last two decades, which researches the signal's independence relations according to higher order statistics. ICA can decompose the random to many mutual independent components which are the most possible independent. It is widely used for an image processing that can strengthen the signal non-Gauss. In order to get the de-noising independent components after separating, we can exchange it to the clean image. Fast ICA is a fast way of ICA which has good effect to the various noises.

The concept of ICA firstly put forward by Herault in 1988 [6]. The standard ICA model can be defined as follows:

$$\mathbf{x}=\mathbf{As} \tag{1}$$

This model describes that the observation variables are mixed from source variable. The source variable is unknown, could not be observed, and matrix is not known yet. Only random variables can be observed. Consequently, we have to estimate matrix and the independent components according to a assumption: The source variable is statistic independent and non-gauss distributing. Through estimating matrix, we could get the contradictory of A, it also be called abruption matrix. So W estimate could be obtained from s .

$$\mathbf{u=Wx=WAs} \tag{2}$$

In order to separate a set of estimate $W=(W_1,W_2,\cdots,W_n)$ of independent statistic signal source $s=(s_1,s_2,\cdots,s_n)$ by independent component analysis from a set of observation signal $x=(x_1,x_2,\cdots,x_m)$ .The W is independent statistic. Suppose every signal is mixed linearly by independent components:

$$\begin{aligned} x_i &= a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{in}s_n \\ i &= 1,2,\cdots,m. \end{aligned} \tag{3}$$

We supposed each observation variable x_i and each source variable s_j are all random variables and those mean are zero. Then Let the vector x as observation variable, and s as source variable, $A(m\times n)$ as matrix a_{ij}. Separate Mixing Image by ICA model was illustrated as follows:

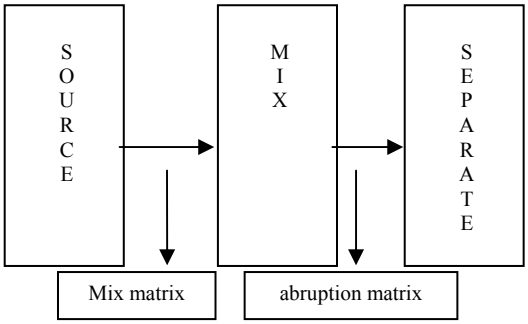


Figure 3. Separate Mixing Image by ICA

Hyvarinen [7-8] put forward the Fast ICA algorithm. By system learning this algorithm find a way to let the projection w^Tx of the cell vector w is the most Gauss. Before running Fast ICA, assume it has pretreatment of ICA, such as get rid of mean and whiten processing. Fast ICA is to find the biggest of non-gauss w^Tx based on fixed-point theory. The Fast ICA separate only one independent component once from observation signal, so it is a fast way of ICA. It is useful to ICA processing.

ICA is used to restore the original sources by estimating the separating matrix. This can be achieved by optimizing contrast functions. Currently, there are a number of contrast functions in use including information maximization, maximum likelihood, high-order cumulants and negentropy. In the Fast ICA algorithm, the initial step is a preliminary whitening of the observations. By a linear transformation, the observations are made uncorrelated and unit-variance. The whitening facilitates the separation of the underlying independent signals, and it can be accomplished by classical PCA processing.

III. BEAMLET TRANSFORM ANALYSIS

The beamlet transform [3] was first put forward by David L. Donoho and Xiaoming Huo. Like the wavelet transform, it is a multiscale transform, with frame elements indexed by scale and location parameters. Unlike the wavelet transform, Beamlets are a simple dyadically organized collection of all line segments at different locations, orientations, and scales. The Beamlet transform is the collection of line integrals along the set of all beamlets. This method allows for the extraction of linear features in noisy image, where traditional methods couldn't process.

We consider an image as a function residing on a [0,1]×[0,1] unit square with "pixels" of size 1/n by 1/n squares arranged in a grid in [0,1]×[0,1].

We set a dyadic square S as the collection of points $\{(x_1,x_2):[k_1/2^j,k_1+1/2^j]\times[k_2/2^j,k_2+1/2^j]\}$ where $0 \le k_1, k_2 < 2^j$ for an integer $j \ge 0$. For clarity, we will could write S(k_1, k_2, j), so that for example S(0, 0, 0) is the unit square [0,1]×[0,1], and so that if we have an n-by-n grid with $n = 2^j$ dyadic, then the individual pixels are the n^2 cells S(k_1, k_2, J), $0 \leqslant k_1, k_2 <$ n.

A. Beamlet Dictionary

To suppose vertices v_1 , $v_2 \in$ [0,1]×[0,1] ,then consider the line segment b = $\overline{v_1v_2}$. Such a segment we call a beam. If we consider only beams connecting vertices (k_1 /n, k_2 /n) at pixel corners, there are order O(n^4) such beams.

Set N = n^2 pixels, and that typical fast image processing algorithms take order O(n^2) = O(N) flops or at most O($n^2\log(n^2)$) = O($N\log(N)$); we are seeking algorithms of order O(n^2) or as near to that as we can get. Relying on collections of cardinality O(n^4) beams as an organizational device will lead to unworkable algorithms, and so we seek a reduced-cardinality substitute.

TO take the collection of all dyadic squares at scales $0 \leqslant j \leqslant J$, fix a quantum of resolution $\delta = 2^{-J-K}$ for $K \geqslant 0$. On each dyadic square, traverse the boundary in clockwise fashion starting at the upper right corner and mark off equispaced vertices a distance δ apart. As δ is dyadic and it divides the perimeter length of every dyadic square with sidelength $\geqslant$ 1/n, there are precisely $M_j = 4\cdot 2^k \cdot 2^{J-j}$ vertices marked out in this fashion on a dyadic square S with side $2^{-j} \geqslant$ 1/n. Call this collection of vertices V (S); label the vertices according to the order they are encountered in the clockwise boundary traverse, so that $V(S) = \{Vi,s : 0 \leqslant i < M_j\}$. If we consider any two dyadic squares which have intersecting boundaries, along the intersection of the boundaries the two squares have the same vertices in common; under our labeling system we might have $Vi,s = Vi',s'$ even though $i \neq i'$. For later use, we let $\upsilon(n,\delta)$ denote the collection of all vertices in all $V(S)$ where S is dyadic of side length $\geqslant$ 1/n, and

we let $L(n)$ denote the latticework of all horizontal and vertical lines in the square at spacing 1/n.

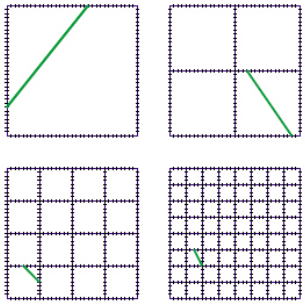


Figure 4. Beamlets at different scales

As for each dyadic square S, consider the collection of all beams connecting vertices on the boundary of S:

$$B_\delta(S) = \{b = \overline{v_{i,s}v_{i',s}} : 0 \le i_1, i_2 < M_j\} \quad (4)$$

There are $\binom{M_j}{2}$ such beams in total.

Definition 1. For given dyadic n and δ, the set of beamlets is the collection

$B_{n,\delta}$ of all beams belonging to some $B_\delta(S)$ for some dyadic square S of sidelength 2^{-j}, 0 ⩽ j ⩽ J.

Some examples of beamlets[3] is shown in Figure 4.

We suppose that beamlets only connect vertices on the boundary of a dyadic square, so that although the family of beamlets is built from O(n^2) vertices, it contains many fewer than O(n^4) beams. In fact, as $\binom{M_j}{2} \approx M_j^2/2$, we have

$$\#B_{n,\delta} \approx (\mathrm{J}+1) \cdot 8 \cdot 4^k \cdot n^2 \quad (5)$$

For instance, suppose that δ = 1/n. Then there are just 4 vertices Vi,s associated to any dyadic square with sidelength 1/n; these are of course the corners of the squares, and we have

$$\#B_{n,1/n} \approx 8 \cdot (\log_2(n) + 1) \cdot n^2 \quad (6)$$

So as we know, although there are order O(n^2)pixels, and although order O(n^4) beams can be defined based on pixel corners, the collection of beamlets has a cardinality only logarithmically larger than O(n^2). It follows that exhaustive searches through the collection of beamlets can run much faster than exhaustive searches through the collection of beams.

In despite of reduced cardinality, the dictionary of beamlets is expressive. It consists of beams at a variety of scales, locations, and orientations. A relatively small number of beamlets can be used as a substitute for any single beam. In [9] we prove

Lemma 1. Any beam with endpoints anywhere in [0,1]×[0,1] can be approximated within Hausdorff distance $\delta/2+1/n$ by a continuous chain ($b_1, b_2, \ldots, b_m$) of beamlets $b_i \in B_{n,\delta}$ where the number m of beamlets required is bounded by 8 $\log_2(n)$ for n > 2.

The rationale is much like the logic of modern airline traffic. There are very many long connections between a few hubs and relatively fewer short connections between provincial centers . The idea [3] is shown in Figure 5.

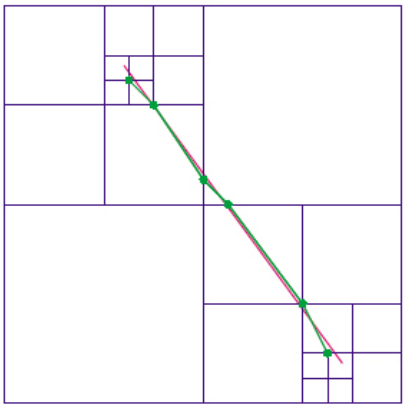


Figure 5. A line segment approximated by a chain of beamlets

B. Beamlet Transform

Set $f(x_1, x_2)$ is a continuous function on [0,1]×[0,1]. The beamlet transform of f is the collection of all line integrals

$$T_f(b) = \int_b f(x(l))dl, \quad b \in B_{n,\delta} \quad (7)$$

Along line segments $b \in B_{n,\delta}$, the integrals has being taken; here $x(l)$ traces out the beamlet b along a unit speed path. The digital beamlet transform of an n×n array (f_{i_1,i_2}) is understood to be the beamlet transform of the function f defined on the continuum by interpolation of the values (f_{i_1,i_2}):

$$f(x_1, x_2) = \sum_{i_1,i_2} f_{i_1,i_2} \phi_{i_1,i_2}(x_1, x_2) \quad (8)$$

where (ϕ_{i_1,i_2}) is a specified family of continuous interpolation functions. There are several ways the functions ϕ_{i_1,i_2} may be chosen. In this paper we use only Average-Interpolating Functions. Let *Pixel*(i_1, i_2) denote the square [i_1/n, (i_1 + 1)/n) × [i_2/n, (i_2 + 1)/n). The functions ϕ_{i_1,i_2} may be chosen to obey the conditions

$$n^2 \int_{pixel(i_{1'},i_{2'})} \phi_{i_1,i_2}(x_1,x_2)dx_1dx_2 = \delta_{i_1,i_{1'}} \cdot \delta_{i_2,i_{2'}} \tag{9}$$

where $\delta_{i,i'}$ is the Kronecker symbol. Then f is a function that obeys

$$f_{i_1,i_2} = Ave\{f \mid Pixel(i_1.i_2)\} \ ;$$

Therefore, the values of f_{i_1,i_2} are viewed as pixel-level averages of the continuous function f. The ensuing definition [3] is illustrated in Figure 6.

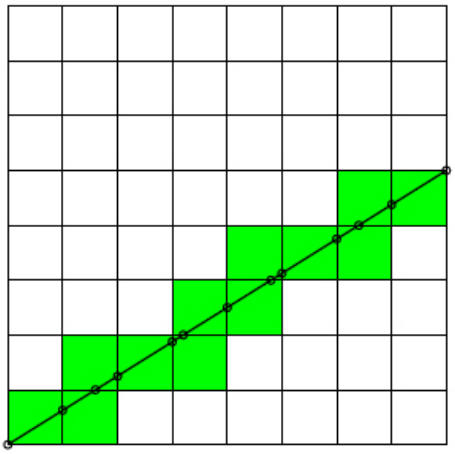


Figure 6. An illustration of piecewise constant interpolation of a digital image, and its associated beamlet transform

Beamlets are connected with various image processing problems ranging from detection to segmentation. There are several levels to 2-D beamlet analysis [10]:

- Beamlet dictionary is a special collection of line segments, deployed acrossb orientations, locations, and scales in 2-D, to sample these in an efficient and complete manner.
- Beamlet transform is the result of obtaining line integrals of the image along all the beamlets.
- Beamlet graph is a graphical structure underlying the 2-D beamlet dictionary, expressing notions of adjacency and allowing to pose algorithms for exploring the space of curves in images in terms of connected multiscale chains of 2-D beamlets, which can be expressed naturally as connected paths in the beamlet graphs.
- Beamlet algorithms are algorithms that for image processing which exploit the beamlet transform and perhaps also the beamlet graph.

IV. Experimental Results and Discussion

In order to apply independent component analysis for image de-noising, a set of estimate of independent statistic signal source is separated from a set of observation signal.

To produce an appropriate matrix is the key part of ICA. This is a random matrix (5×5):

$$A = \begin{bmatrix} 0.2 & 0.1 & 0.4 & 0.1 & 0.2 \\ 0.3 & 0.1 & 0.2 & 0.1 & 0.3 \\ 0.1 & 0.1 & 0.6 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.2 & 0.1 & 0.3 \\ 0.2 & 0.1 & 0.4 & 0.1 & 0.2 \end{bmatrix} \tag{10}$$

By making use of this matrix, we could effectively resolve the independent components from the measured mixed signals without any additional information about the source signals.

For doing the experiment, we get a lot of vehicle video from the Traffic Police Station of Suzhou City. We extracted some key frames for processing. In our experiment, we convert the color key frame into gray one first, then add Gaussian white noise of zero mean noise with 0.006 variance and salt & pepper noise with 0.015 density to the key frame image. Then we de-noise these images separately. The steps of this processing method are as follows:

- The vehicle video key frame image is separated by a random matrix into independent component image.
- Then beamlet transform is using to optimize the coefficients.
- At last, the inverse of the beamlet transform is applied for image reconstruct a new denoised vehicle video key frame image which could be processed for recognition and classification in next research process.

The steps we presented are illustrated as follows:

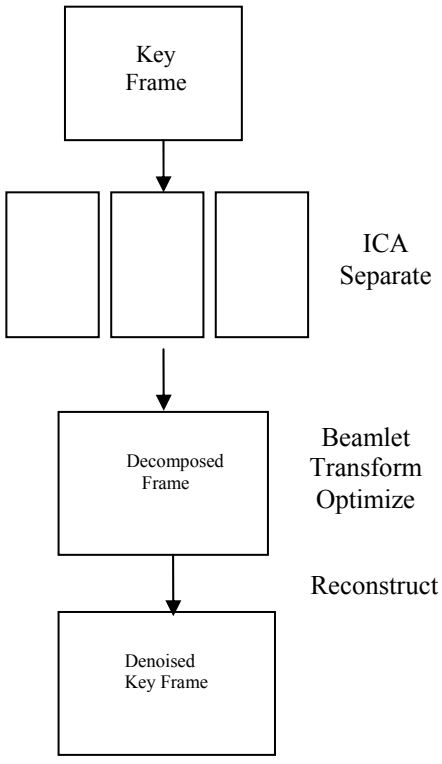


Figure 7. vehicle video key frame de-noising process

In order to prove our method is superior, we use other filter to process the source noise image which is shown in Figure 8(a). By contrast, the efficiency of our method which is shown in Figure 8(d) is better than median and wiener approaches which are shown in Figure 8(b) and Figure 8(c) respectively. It removed the noise effectively, and many details were reserved. The results of different vehicle video key frame image de-noising methods are shown as follows:

(a)

(b)

(c)

(d)

Figure 8. Contrast Several Method of De-noising key frame Image

In order to evaluate the image quality scientifically objectively, reconstruction error was used as a standard objective measure of key frame quality. There are two of the error metrics used to compare the different image de-noising techniques, one is the mean square error (MSE) which is the cumulative squared error between the de-noising and the original image, and another is signal to noise ratio (SNR) which is a measure of the image error. The mathematical formulae can be expressed as follows:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I_{ij} - I_{ij}')^2 \tag{11}$$

$$\mathrm{SNR} = 10\lg[\sum_{i=1}^{M}\sum_{j=1}^{N}(I - I^{''})^2 / \sum_{i=1}^{M}\sum_{j=1}^{N}(I - I^{'})^2] \quad (12)$$

For the sake of contrasting the effect of various de-noising methods, I(x, y) is define as the original key frame image, I'(x, y) is the de-noising image, I''(x, y) is the mean value and M, N are the dimensions of the images. If a value for MSE is lower, that is to say the de-noising image has lesser error, and there is the inverse relation between the MSE and SNR, it create a high value of SNR. It means effect of de-noising is good because it show that the signal is more to noise in the image.

TABLE I. EVALUATION OF VARIOUS DE-NOISING METHODS

Processing Method	SNR	MSE
original	19.24	775
wiener	24.76	217
median	26.03	162
our method	27.24	122

From the table above we can see that our method have a higher value of SNR and a lower value of MSE. So it is superior to median and wiener filtering methods for vehicle video key frame image de-noising.

V. CONCLUSIONS

This paper proposes a new method for vehicle video key frame image de-noising, which is using independent component analysis and beamlet transform. Firstly, a random matrix was produce to separate the key frame into a separated image for estimate. Then beamlet transform was applied to optimize the coefficients. At last, the inverse of the beamlet transform was applied for image reconstruction. The de-noising image has a higher value of SNR and a lower value of MSE. This approach could remove more noises. Further investigations on the use of independent component analysis and beamlet transform for pattern recognition processing are left for future work.

ACKNOWLEDGMENT

This research was partially supported by the Natural Science Foundation of China (60970015) and (61003054), the 2009 Special Guiding Fund Project of Jiangsu Modern Service Industry (Software Industry) ([2009]332-64), the Applied Basic Research Project (Industry) of Suzhou City (SYJG0927) and (SYG201032), the Program for Postgraduates Research Innovattion in University of Jiangsu Province in 2010 (CX10B_041Z) and the Beforehand Research Foundation of Soochow University.

REFERENCES

[1] Yiyan Wang, Yuexian Zou, Hang Shi, He Zhao, "Video Image Vehicle Detection System for Signaled Traffic Intersection," International Conference on Hybrid Intelligent Systems, vol. 1, pp. 222-227, 2009.

[2] P. Comon, "Independent component analysis," Proceedings of the International Signal Processing Workshop on Higher Order Statistics, Chamrousse, France, 1992, pp 29.

[3] D. L. Donoho and Xiaoming Huo, "Beamlets and Multiscale Image Analysis," http://www-stat.stanford.edu/~donoho/Reports/2001/BeamletMSIP051101.pdf, 2001.

[4] L. Ying and E. Salari, "Beamlet Transform Based Technique for Pavement Image Processing and Classification," Proceedings of 2009 IEEE International Conference on Electro/Information Technology, EIT 2009, pp141-145, 2009.

[5] Hyvarinen A, "Independent Component Analysis," John Wiley and Sons, 2001(5):223 - 225.

[6] C. Jutten, J. Herault, "Independent component analysis verus principal component analysis," Europ. Signal Processing Conf. (EUSIPC088), Grenoble, France, 1988.

[7] Hyvarinen. A, Cristescu. R, Oja. E, "Fast algorithm for estimating overcomplete ICA bases for image windows," International Joint Conference on Neural Networks (IJCNN'99), USA, 1999. vol. 2, pp894-899.

[8] Hyvarinen. A, "Fast ICA for noisy data using Gaussian moments," Proceedings of the 1999 IEEE International Symposium on Circuits and Systems, USA, 1999, vol.5, pp57-61.

[9] D.L. Donoho, "Wedgelets: Nearly minimax estimation of edges," Annals of Statistics, vol.27(3), pp859–897, 1999.

[10] D. L. Donoho and O Levi, "Fast X-Ray and Beamlet Transforms for Three-D Data," http://www-stat.stanford.edu/~donoho/Reports/2002/. Three-D-Beamlets.pdf, 2002

Guangming Zhang was born in Suzhou, China, in Feb. 1981. He is now a Ph.D. candidate in the School of Computer Science and Technology, Soochow University, China. His research interests include video processing, machine vision, and pattern recognition.

Zhiming Cui is a professor and doctoral supervisor at the Institute of Computer Science and technology, Soochow University and a CCF senior member. His researches areas are intelligent information processing, image processing, distributing computing, DeepWeb data mining etc.

Pengpeng Zhao received his Ph.D degree in computer science from Soochow University, Suzhou, China, in 2008. His research interests include deep web, data minning and machine learning.

Jian Wu was born in Nantong, 1979, and got master degree from Soochow university, Suzhou, China, in 2004. The main research direction is computer vision, image processing and pattern recognition. Now he is pursuing the doctoral degree.

The Translation Invariant Wavelet-based Contourlet Transform for Image Denoising

Gang Liu[1]
[1]School of Computer Science and Technology, Xidian University, Xi'an, China
Email: gliu@xidian.edu.cn

Jing Liu[2],Quan Wang[1],Wenjuan He[2]
[2]Faculty of Computer Science and Engineering, Xi'an University of Technology, Xi'an, China
Email: liujing@xaut.edu.cn

***Abstract*—A new method of image denoising using wavelet-based contourlet transform (WBCT) is proposed. Due to the lack of translation invariance of WBCT, image denoising by means of WBCT would lead to Gibbs-like phenomena. In the paper,cycle spinning-based technique is applied to develop translation invariant WBCT denoising scheme. Many simulation experiments with images contaminated by additive white Gaussian noise demonstrate that the performance of the proposed approach substantially surpasses that of previously wavelets methods using the cycle spinning both visually and in terms of the PSNR values, especially for the images that include mostly fine textures and contours.**

***Index Terms*—wavelet-based contourlet transform(WBCT), cycle spinning, image denoising, translation invariance**

I. INTRODUCTION

During the acquisition and transmission of image, there always exists noise. We must remove these noise from image to improve the image quality. Traditional denoising methods are classified into two groups: the spatial method such as wiener filter which eliminates the noise in image at the cost of blurring the edge and texture; the other one is using image transform method, for example, the popular wavelet threshold denoise technology.

Over the past decade, wavelet transforms have been paid a lot of attention to many image processing areas such as compression, noise removal, feature extraction, edge enhancement and detection. In denoising, the wavelet with a single orthogonal wavelet function has played an important role because of its ability to capture the energy of an image with few energy transform values[1,2,3]. But wavelet has limited ability in capturing the directional information of the natural images. Aimed at improving the representation sparsity of an image over the Wavelet, some new transforms have been introduced for image denoising. Do and Vetterli developed the contourlet transform(CT)[4,5] based on anisotropy scaling law and directionality which can sparsely represent natural images. It is implemented by Laplacian pyramid and a double filter banks structure that can simultaneously hold multiresolution, localisation, nearly critical sampling, flexible directionality and anisotropy. As we turn to use the transform tool, more advanced denoising methods [6,7,8,9,10,11,12] were proposed. But, due to the redundancy of the Laplacian pyramid, the contourlet transform has a redundancy factor of 4/3 and hence, some approaches have been attempted to introduce non-redundant image transforms. Eslami and Radha proposed a new non-redundant image transform [13], the Wavelet-Based Contourlet Transform (WBCT), with a construction similar to the contourlet transform. The proposed WBCT achieves both radial and angular decomposition to an arbitrary extent and obeys the anisotropy scaling law. Compared to the aforementioned DFB-based non-redundant transforms, the WBCT can easily be realized by applying DFB on the wavelet coefficients of an image. But, the DFB stage of WBCT involves downsampling and therefore, it is shift variant. Translation invariance is a required feature in denoising, which can significantly improve the performance [14,15].

In this paper, we propose a new approach for image denoising based on the Wavelet-Based Contourlet Transform (WBCT). To compensate for the lack of translation invariance property of the WBCT, we apply the principle of cycle spinning to the WBCT which can improve the denoising performance of WBCT. Experimental results show that the significant improvements have feasibility and superiority, and demonstrate that the proposed scheme can achieve better PSNR values when compared with the wavelet transform using cycle spinning (WTCS),and visually,the method is capable of better retaining edges and textures in the denoised images.

This paper is organized as follows: In Section II ,Wavelet-based Contourlet Transform(WBCT) and basic denoising and cycle spinning are discussed. Details of the proposed denoising method is given in Section III . Experimental results are discussed in Section IV. Concluding remarks are given in Section V.

II. THEROY

A. The Wavelet-based Contourlet Transform

The contourlet transform based on a multiscale and multidirectional filter bank developed by Do and Vetterli,

doi:10.4304/jmm.7.3.254-261

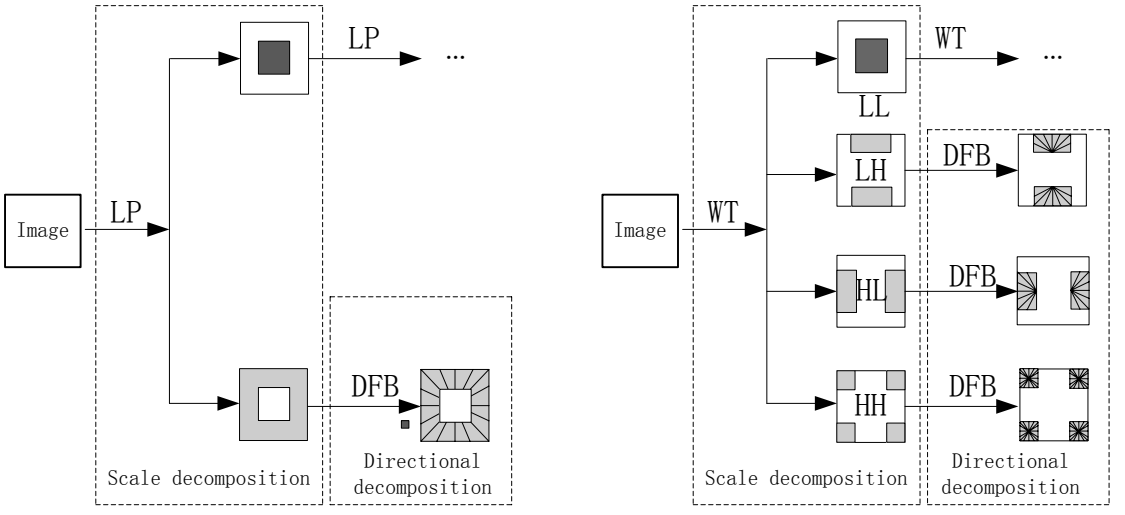
is one of the new geometrical image transforms, which can capture nearly arbitrarily directional information of the natural images. This transform consists of two major stages: the subband decomposition and the directional transform. At the first stage, Laplacian pyramid (LP) is employed, while directional filter banks (DFB) are used for the second stage. The main feature of this transform is the potential to efficiently deal with 2-D singularities such as edges, unlike wavelet which manage point singularities exclusively. We can attribute the success to the two main properties of contourlet transform :(1) the directionality property, i.e. having basis functions at many directions, whereas wavelet has only 3 directions; (2) the anisotropy property, meaning that the basis functions appear at various aspect ratios (depending on the scale), as opposed to equal aspect ratio of wavelet. The main advantage of the contourlet transform over other geometrically-driven representations, e.g. curvelet, is its relatively simple and efficient and wavelet-like implementation using iterative filter banks. Due to its structural resemblance with the wavelet transform, many image processing tasks applied on wavelet can be seamlessly adapted to contourlet. But,the contourlet transform is a redundant image transform based on LP.The Wavelet-Based Contourlet Transform(WBCT) developed by Eslami and Radha , is a new non-redundant image transform with a construction similar to the contourlet .It also consists of two filter bank stages:the first stage provides subband decomposition using wavelet transform rather than the Laplacian pyramid; The second stage of the WBCT is a directional filter bank (DFB), which provides angular decomposition. The graph of the detail decomposition process of CT and WBCT is shown in Fig.1. In the first stage of WBCT, the image is decomposed into the low frequency subband corresponding to the LL and the traditional three highpass bands corresponding to the LH, HL, and HH. In the second stage, perform DFB with the same number of directions on each high frequency subband in a given level,and start from the desired maximum number of directions on the finest level, and decrease the number of directions at every other dyadic scale when proceeded through the coarser levels.

By means of this way, the anisotropy scaling law is achieved. Moreover, the wavelet filters are not perfect in splitting the frequency space to the lowpass and highpass components, the scheme using fully DFB decomposition on each band can compensate for the drawbacks of the wavelet filters. Fig.2(a) shows an example of the WBCT using 3 wavelet levels and applying 8,4,2 number of directions to each level from the finest level to the coarse level.Fig.2(b) shows the example of WBCT representation on Peppers image with size of 512*512. For clear visualizing, the image is only decomposed into three levels, which are then transformed into 16,4,4 directional subbands, respectively. Here, small coefficients are shown in black while large coefficients are shown in white. We see that WBCT can produce the significant coefficients in both location and direction of image contours,and most of the coefficients in the LH subbands are in the horizontal directional subbands (the lower half of the subbands) while those in the HL subbands are in the vertical directional subbands (the upper half of the subbands) .

Because wavelet transform has more advantages than LP decomposition in representing sparsly image, so WBCT can capture image structure features more efficiently. It has been shown to be a better alternative choice than the contourlet transform for image denoising.

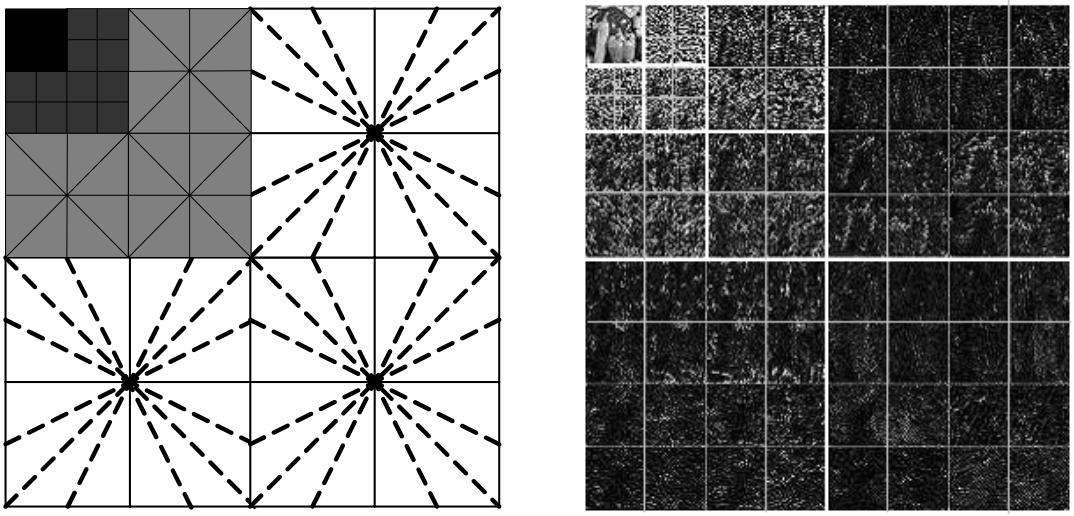
B. Basic Denoising

Image denoising by wavelet thresholding was introduced by Donoho and Johnstone in [14]. Let $\boldsymbol{x}$ be noiseless image and $\boldsymbol{y}$ be the image corrupted image with independent Gaussian noise $\boldsymbol{n}$, the noise-corrupted image



(a) The first level decomposition of the CT (b) The first level decomposition of the WBCT

Figure 1.Comparison of decomposition process of CT and WBCT.



(a) The plot of WBCT using 3 wavelet levels and 8 directions at the finest level.

(b)WBCT representation of Peppers image

Figure 2. A flow graph of WBCT scheme

is defined as follows:

$$y = x + n \tag{1}$$

Perform the orthogonal wavelet transform on the noise-corrupted image $\boldsymbol{y}$ and then zero out the detail coefficients that fall below a certain threshold ,which are likely to contain mostly noise. An inverse wavelet transform is applied to the thresholded image to yield an estimate for the noise-free image $\boldsymbol{x}$.Threshold $\boldsymbol{T}$ can be calculated as below:

$$T = \hat{\sigma}\sqrt{2\log^{N}} \tag{2}$$

$$\hat{\sigma} = \frac{median(|d_i|)}{0.6745} \tag{3}$$

where $\boldsymbol{N}$ is length of the corresponding data, $\boldsymbol{d_i}$ represents all of the coefficients in the finest level.The wavelet threshold denoise includes hard threshold denoise and soft threshold denoise. Hard-threshold consists of replacing each wavelet coefficient $\boldsymbol{d_i}$ by the value $\boldsymbol{d'_i}$,where

$$d_i' = \begin{cases} d_i & if \quad d_i > T \\ 0 & else \end{cases} \tag{4}$$

Each wavelet coefficient $\boldsymbol{d_i}$ of Soft-threshold is replaced by the value $\boldsymbol{d'_i}$, where

$$d_i' = \begin{cases} \mathrm{sgn}(d_i)(|d_i| - T) & if \quad d_i \geq T \\ 0 & else \end{cases} \tag{5}$$

This method was remarkable for its simplicity, it performs well under a number of applications of image denoising. But since the denoise principle of threshold is to set finescale wavelet coefficients to zero, and the feature also lies in the high frequency bands, it brings about visual artifacts.

Since the wavelet-based contourlet transform with a construction similar to the wavelet transform, has the compaction property that there is only a small number of large coefficients,and all the rest coefficients are very small. We can apply threshold-based image denoising to the domain of wavelet-based contourlet transform, however,it can not escape to eliminate visual artifacts or blur the feature. what should we do to resolve it?

C. Cycle Spinning

The wavelet-based contourlet transform and wavelet transform are not time invariant.Consequently, if the noisy image is shifted in time, denoised,and then shifted back, the result will, in general, be different from the estimate obtained from denoising without shifting. One way of improving upon basic wavelet transform denoising is through cycle spinning, where the image to be denoised is translated by various time shifts. For a shift variant wavelet transform $\boldsymbol{P}$, if the following procedure is applied to the noisy image $\boldsymbol{y}$:

$$\hat{x}_{i,j} = \frac{1}{k_1 k_2}\sum_{i=0,j=0}^{k_1 k_2} X_{-i,-j}(P^{-1}(T[P(X_{i,j}(y))])) \tag{6}$$

$$\hat{x} = \frac{1}{k_1 k_2}\sum_{i=0,j=0}^{k_1 k_2} \hat{x}_{i,j} \tag{7}$$

Where $\boldsymbol{X_{i,j}}$ and $\boldsymbol{T}$ respectively represent the 2-D circulant shift and its threshold operators , $\boldsymbol{k_1,k_2}$ are the maximum number of shifts on the row and col direction, $\boldsymbol{P^{-1}}$ expresses the inverse transform. The cycle spinning estimate $\hat{x}$ is obtained by simply linearly averaging the $\boldsymbol{k_1 \times k_2}$ estimates(As in (7)). The errors in the estimates are not completely dependent. Consequently averaging these estimates yields a reduction in noise. Similar to wavelet, if an image of size $\boldsymbol{N \times N}$ is decomposed using WBCT, the maximum number of decomposition levels in the scale stage will be $\boldsymbol{k}(\boldsymbol{N=2^k})$, and therefore, the maximum number of shifts are $\boldsymbol{k \times k}$ in the row and column directions. We applied this procedure to WBCT and

achieved superior performance in our denoising experiments as briefly demonstrated in Section IV.

III. ALGORITHM DESCRIPTION

Since the WBCT can decompose image in multi-scale and multi-direction, and using the idea of cycle spinning ,we know that the edge can be kept best when the anisotropic filter's long axes is in accord with the edge,and with the angle between the edge and the anisotropic filter's long axes becomes larger the edge becomes blurrier. The DFB decomposition can be seen as the DFB filter's long axes is placed at different orientations and which resulting in the many directional subbands. Every directional subbands has a decomposition direction and the edge in this direction has a largest gray scale comparing with which in any other direction .

On the base of the idea that the coefficients of image WBCT lying in the edges is larger than the noise WBCT coefficients, we can compare the WBCT coefficients with a threshold and kill the smaller coefficients to remove the noise from image.

The whole denoising algorithm consists of 6 steps as follows:

1) Shift the image in each direction, and the maximum numbers of shifts will be $k(N=2^k)$;

2) Perform WBCT on the shifted image .

3) Apply hard thresholding to the resulting coefficients.

4) Perform inverse WBCT on the thresholded coefficients;

5)Implement inverse shift the processed image,and get the noise-free image $\hat{x}_{i,j}$.

6)Average over the all results using (7) to get the denoised image $\hat{x}$.

IV. EXPERIMENTAL RESULTS

To evaluate the proposed scheme, we applied four approaches to four images with size of 512×512: Barbara, Lena, Cameraman and Peppers. The four approaches are: the wavelet transform (WT), the translation invariant wavelet transform (WTCS) and the wavelet-based contourlet transform(WBCT), in addition to the proposed method based on WBCT using cycle spinning (WBCTCS). We used biorthogonal Daubechies 9/7 wavelet transform. For the first stage of WBCT, we also used the same biorthogonal filters and applied 4 levels and 8 directions decomposition at the finest level. The images are corrupted by a zero-mean Gaussian noise with a standard deviation of σ , ranging from 20 to 100. Since hard-thresholding usually yields better results than soft-thresholding, we used hard-thresholding with a fixed

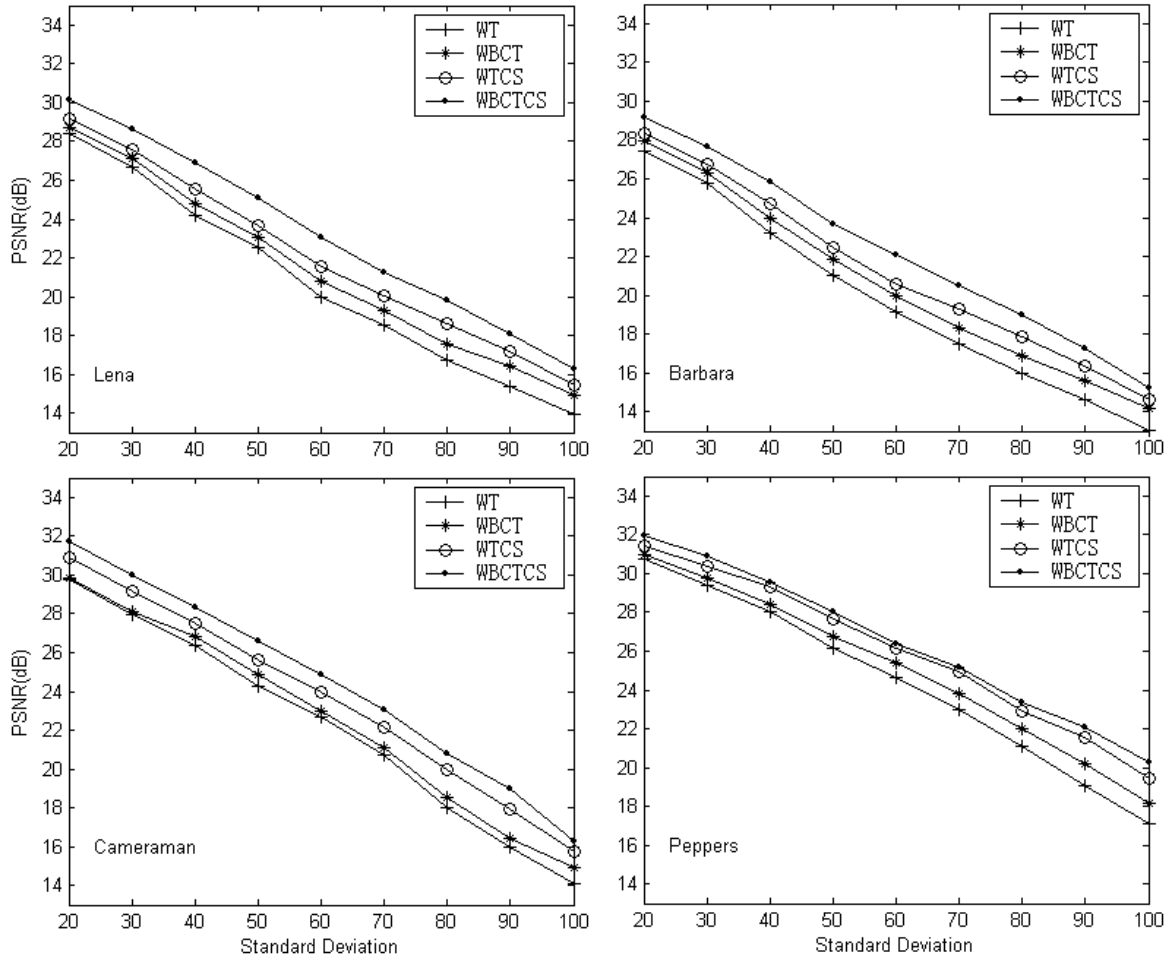


Figure 3.PSNR curves of other denoising methods relative to our method

threshold value equal to $T=3\sigma$ [Ref.14], here. Table I shows the PSNR value achieved by the four methods, respectively. It is clear that the use of the new method improves the quality of denoising image. For the Cameraman and Peppers image, which are more smooth images, and hence they are "wavelet-friendly" images, the WTCS performs almost the same as the WBCTCS at a range of the input noise power. However, in case of the images containing mostly textures and contours such as the Barbara and Lena images, the WBCTCS yields significant improvements up to 0.7 dB over the WTCS. The PSNR vs. standard deviation curves for the images are provided in Fig.3. Our method is seen to clearly

TABLE Ⅰ.PSNR VALUES FOR THE FOUR IMAGES APPLIED TO THE VARIOUS METHODS(σ=20)

	Lena	Barbara	Cameraman	Peppers
WT	28.4322	27.4251	29.7387	30.7387
WBCT	28.7156	27.9530	29.8135	30.9357
WTCS	29.1364	28.3121	30.9251	31.4459
WBCTCS	30.1378	29.1024	31.7154	31.9269

(a) Original image

(b) Noisy image (σ= 20, PSNR=20.2dB)

(c) Denoised image using WT(PSNR=27.4dB)

(d) Denoised image using WTCS(PSNR=28.3dB)

(e) Denoised image using WBCT(PSNR=27.9dB)

(f) Denoised image using WBCTCS(PSNR=29.1dB)

Figure 4. Comparison of denoising results of Barbara image contaminated with a Gaussian noise of σ=20

outperform the other three over the entire range of noise levels, it achieved better results at all cases. To visually compare the estimated images, we show the denoised images of Barbara image when σ=20 in Fig.4. We can see that there are many visual artifacts in the experimental result of the WBCT due to the Gibbs-like phenomena; however ,we could reduce these artifacts to a large extent and achieve the superior PSNR values by using our proposed method. To compare visibility of the artifacts of the various denoising methods, Fig.5 and Fig.6 show two small different part of the Barbara image. As can be seen, Our method provides fewer artifacts as well as better preservation of edges and other details. significantly more levels of detail and texture are retrieved by the proposed scheme, most of the visual artifacts due to the Gibbs-like phenomena in the WBCT denoising are reduced by using cycle spinning. From the comparison of the outcomes of both sets of experiments, we may conclude that our method contributes significantly to the performance advantage.

V. CONCLUSION

We propose an efficient method of image denoising. We utilize the cycle spinning algorithm in developing a translation invariant WBCT-based denoising. Our experimental results clearly demonstrate that the proposed scheme can provide smoothness and better edge preservation, especially for those images possessing detailed textures. we can eliminate most of the visual artifacts resulting from the wavelet transform denoising function that uses hard thresholding .The experimental evaluation showed that the proposed methods have far better performance than the translation invariant wavelet denoising.

ACKNOWLEDGEMENG

This work is supported in part by the National Nature Science Foundation of China (Grant No. 61070045) and the Scientific Research Program of Shaanxi Provincial Education Department(Grant No.12JK0724).

REFERENCES

[1] D.Gleich and M.Datcu, "Wavelet-based SAR image despeckling and information extraction, using particle filter," IEEE Transactions On Image Processing, 2009,vol.18,no.10, 2167-2184.

[2] D.Gleich ,M. Kseneman and M.Datcu, "Despeckling of TerraSAR-X data using second-generation wavelets," IEEE Geoscience and Remote Sensing Letters, 2010,vol.7,no.1,68-72.

(a) Denoised image using WT(PSNR=27.4dB)

(b) Denoised image using WTCS(PSNR=28.3dB)

(c) Denoised image using WBCT(PSNR=27.9dB)

(d) Denoised image using WBCTCS(PSNR=29.1dB)

Figure 5. Comparision of denoising results of Barbara image(cropped to 200×200 for visibility of the artifacts).

(a) Denoised image using WT(PSNR=27.4dB)

(b) Denoised image using WTCS(PSNR=28.3dB)

(c) Denoised image using WBCT(PSNR=27.9dB)

(d) Denoised image using WBCTCS(PSNR=29.1dB)

Figure 6. Comparision of denoising results of Barbara image(cropped to 200×200 for visibility of the artifacts).

[3] Q.Sun, L. C. Jiao and B. A.Hou, "Synthetic aperture Radar image despeckling via spatially adaptive shrinkage in the nonsubsampled contourlet transform domain," Journal Of Electronic Imaging, ,2008,vol.17,no.1,2392-2404.

[4] M. N. Do and M. Vetterli, "Contourlets: a directional multiresolution image representation," Proceedings of International Conference on Image Processing, Rochester, September, 2002 vol,no.1, 357–360.

[5] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," IEEE Transaction on Image Process, 2005 vol.14,no.120, 2091–2016.

[6] Z.J.Luo and Y.Q.Wu, "A method of target detection in infrared image sequence based on contourlet transform," Sigmal Processing, Vol24,No. 4 ,2008,pp. 676-679.

[7] D.Liang, "Image enhancement based on the Nonsubsampled Contourlet Transform and adaptive threshold," Acta Elsectonica Sinica, 2008 vol.36,no. 3,.527-530.

[8] F.n Wang,"Image denoising using nonsubsampled contourlet transform",ComputerApplications.2007,vol.27, no.10,2516-519.

[9] Z.j Luo, y.q Wu, "A method of target detection ininfrared image sequence based on contourlet transform",Sigmal Processing,2008.8, vol.24,no.4,676-679.

[10] X.j Guo, Z.l Wang , "Nonsubsampled Contourlet image denoising based on Inter-scale correlations", Journal of Opto Electronics ·Laser, 2007,vol 18,no.9,1116-1119.

[11] W.y.Wu and Y.Q.Wu, "Method of infrared in targets detection based on Contourlet transform," Infrared and Laser Engineering, 2008, vol.37,no.1,136-138.

[12] R. Eslami,H. Radha, "Wavelet based contourlet transform and its application to image coding," Proceedings of of the International Conference on Image Processing, 2004, vol.5,3189-3192.

[13] R.R.Coifman and D. L.Donoho, "Translation invariant denoising,wavelets and statistics,"Springer Lecture Notes in Statistics 103,New York:Springer-Vedag,1995,125-150.

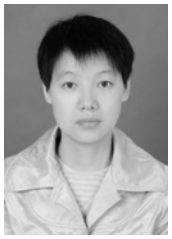
[14] M. S. Mallat, "A wavelet tour of signal processing," Academic Press, 3rd Edition, 2008,535-610.

[15] Q Pan, G.p Yan Y.k Zhang, "Mechanism of simultaneous contrast in biological vision for high-pass filtering",Signal Processing, 2008,vol.24,no.2,281-285.

Gang Liu born in Hunan China, July 1969. He received the B.S degree from Wuhan university, Wuhan,China, in 1992 and ph.D degree from Xian Jiaotong University ,Xi'an, China, in 2005. His major is multimedia & intelligent information processing.

He is currently professor in School of Computer Science and Technology, Xidian University. He has nineteen years experience of teaching undergraduate and postgraduate students and guided number of post-graduate theses. Presently he is

guiding five M.S students in the area of digital image processing, network.



Jing Liu born in Qiqihaer Heilongjiang province China, in 1973/03. Received the B.S. degree and the M. S. degree in electronic information, both from Xi'an University of Technology in 1994 and 2000 respectively. In July 2003, she joined Xi'an University of Technology . Her research interests include digital image processing and signal processing.

Quan Wang born in Weinan Shannxi province China, in 1969/06. Received the B.S. degree in computer sciences, from Xidian University ,Xi'an, China, in 1992 .In June 1992,he joined Xidian University.His research interests include digital image processing.

Wenjuan He born in Sanyuan Shannxi province China, in 1973/09. Received the B.S. degree in computer sciences, from Xi'an University of Technology in 1997.In August 2001,she joined Xi'an University of Technology .Her research interests include signal processing.

An Improved Method of Detecting Edge Direction for Spatial Error Concealment

Yan Zhao, Hexin Chen, Shigang Wang
School of Communication Engineering, Jilin University, Changchun, China
Email: {zhao_y, chx@jlu.edu.cn};wangshigang@vip.sina.com

Moncef Gabbouj
Institute of Signal Processing, Tampere University of Technology, Tampere, Finland
Email: Moncef.Gabbouj@tut.fi

***Abstract*—Sobel operator is widely used to detect the edge direction in most spatial error concealment methods. However, Sobel operator may not work when detecting edge directions in the case of existing line in the image. This paper designs an advanced method of detecting edge directions, which can estimate the direction of both step edge and line correctly. The benefit of the advanced edge detection method is that all edge directions including step edges and lines can be detected properly. The proposed method is applied in a spatial error concealment method. Experimental results show that the advanced method yields better objective and subjective quality than that of using Sobel operator for determining edge direction only.**

***Index Terms*—edge direction, Sobel, line, error concealment**

I. INTRODUCTION

Spatial error concealment uses correlation between adjacent pixels in an image to recover the lost pixels. There have been many studies on spatial error concealment methods. The edge feature in neighboring blocks of the corrupted block is often required to estimate in the spatial domain in order to recover edge information inside the corrupted blocks. The Sobel operator is widely used to detect the edge direction in many spatial error concealment methods. A content-adaptive error concealment algorithm was proposed in [1], which calculated local edge gradient magnitude and angular direction by Sobel operator and concealed the edge block by directional interpolation. The local edge gradient magnitude and angular direction at the pixel were calculated by convolving pixel values with the 3×3 Sobel operators in [2] to implement the hybrid error concealment based on block content. A spatial error concealment algorithm using directional extrapolation used Sobel masks to determine the direction of edge traversing the to-be-recovered pixel, which was recovered by corresponding directional extrapolation [3]. The algorithms in [4] and [5] applied a horizontal and vertical Sobel mask on the luminance value of each pixel at the missing area boundaries to detect the edges of the missing block, which was further concealed by directional interpolation and main direction smoothing. The Sobel operator with constant 1/4 was adopted in [6] to calculate the amplitude and angle of the edge gradient to restore edge components of the lost macroblocks for I frame concealment. Fine directional interpolation for spatial error concealment was proposed in [7], which used Sobel masks to roughly detect edge orientations for further refinement step. The algorithm of concealment of damaged block using projection onto convex sets in [8] used 3×3 Sobel mask operators as edge orientation detector to determine if a missing bock is with a particular orientation for further directional filtering. The early work in [9] selected Sobel operator for gradient estimation for multidirectional interpolation for spatial error concealment. The algorithm in [10] used also 3×3 Sobel operators to determine the potential edge direction of the 4×4 block, which is concealed by Intra_4×4 prediction in H.264/AVC corresponding to the potential edge direction.

Edges play an important role in the subjective image quality because the human visual system is sensitive to the structural information revealed by edges. Therefore, restoring edge features of the lost macroblocks becomes a necessary job for spatial error concealment. The accuracy of edge detection determines the correct estimation of the edge features and determines further the performance of spatial error concealment algorithm. However, using the Sobel operator to detect edge direction does not always work. It may give wrong edge direction if the edge is a line since it is a first-order operator. Therefore, the detection of edges and their directions is not accurate for corrupted blocks with lines.

In this paper, a new method is designed to find out wrong directional estimation detected using Sobel operator and an advanced method is also proposed to estimate the direction of lines correctly. Thus, the edge direction accuracy can be improved significantly. The proposed method is applied in a spatial error concealment algorithm to conceal the corrupted blocks. The paper is organized as follows. Section 2 and 3 describe the proposed method of detecting edge direction and concealment. In Section 4, experimental results are given. Conclusions are drawn in Section 5.

Corresponding author: Yan Zhao;

doi:10.4304/jmm.7.3.262-268

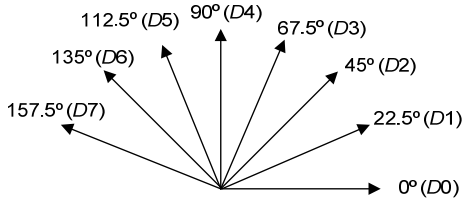


Figure 1. Eight edge directions.

II. METHOD OF DETECTING EDGE DIRECTION

In general, there are two types of edges in an image: step edge and line. It is worth noting that the line is with only one pixel width here. The angle of the edge direction in most error concealment is rounded to the nearest 22.5°, which means that there are 8 different directions between 0° and 180° as shown in Fig.1

The operational flow of estimating the edge direction is sketched in Fig. 2. We firstly detect the edge direction by using traditional Sobel operator. Then we check if the direction is correct by using our proposed detection method. If the direction detected by Sobel operator is correct, the direction is output as the step edge direction, otherwise, we may check if it is a line and if so we will determine the direction of the line by using our proposed method. Otherwise, we may determine there is not edge in the image.

A. Detecting Edge Direction of Step Edge

Sobel operator is an effective tool which is widely used for determining direction of the step edge. In our proposed method, the 3×3 Sobel operator is used for detecting edge direction firstly. In order to estimate the edge direction existed in a subimage or a mask with size 3×3 as shown in Fig.3, the simplified calculation below is developed.

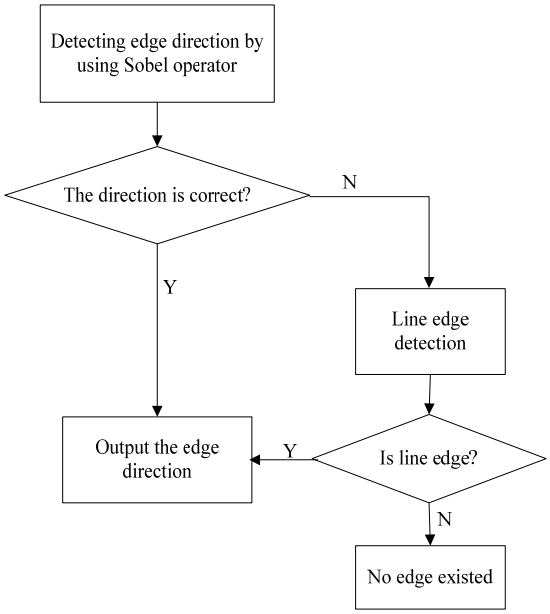


Figure 2. Flow chart of estimating edge direction

TABLE I.
RELATIONSHIP BETWEEN THE RANGES OF $\tan\theta$ AND EIGHT DIRECTIONS.

Range of $\tan\theta$	Direction
$-1/4 \le \tan\theta < 1/4$	D0
$1/4 \le \tan\theta < 3/4$	D1
$3/4 \le \tan\theta < 3/2$	D2
$3/2 \le \tan\theta < 9/2$	D3
$\tan\theta \ge 9/2$ or $\tan\theta < -9/2$	D4
$-9/2 \le \tan\theta < -3/2$	D5
$-3/2 \le \tan\theta < -3/4$	D6
$-3/4 \le \tan\theta < -1/4$	D7

$$\tan\theta = \frac{G_y}{G_x} \tag{1}$$

where θ is the angle of the edge direction.

$$\begin{aligned} G_y &= (x_2 + 2x_5 + x_8) - (x_0 + 2x_3 + x_6) \\ G_x &= (x_6 + 2x_7 + x_8) - (x_0 + 2x_1 + x_2) \end{aligned} \tag{2}$$

The relationship between ranges of $\tan\theta$ and eight directions is illustrated in Table 1. Sobel operator is effective for estimating the correct direction of the step edge. However, it might provide wrong result if the edge is a line, e.g., if the mask is like that shown in Fig.4. Obviously, this mask is in a line direction of 0°, but, the result given by Sobel operator is 45°. Therefore, it is necessary to find out the case when Sobel operator cannot work properly and design a method to detect line direction correctly. After the edge direction is detected by Sobel operator, the correction of the direction can be checked by:

$$\Delta x = \left|x_{i+1} - x_i\right| + \left|x_i - x_{i-1}\right| \tag{3}$$

where x_i is the center of the mask, that is x_4 in Fig.3, x_{i+1} and x_{i-1} are two adjacent pixel values of x_i along the detected direction which are given by:

x_0	x_1	x_2
x_3	x_4	x_5
x_6	x_7	x_8

Figure. 3 The 3×3 subimage or mask.

100	100	99
10	11	10
100	99	100

Figure. 4 An example of 3×3 mask

$$\begin{cases} \begin{cases} x_{i+1}=x_5 \\ x_{i-1}=x_3 \end{cases} & \text{if the direction is D0;} \\ \begin{cases} x_{i+1}=(x_2+x_5+0.5)/2 \\ x_{i-1}=(x_3+x_6+0.5)/2 \end{cases} & \text{if the direction is D1;} \\ \begin{cases} x_{i+1}=x_2 \\ x_{i-1}=x_6 \end{cases} & \text{if the direction is D2;} \\ \begin{cases} x_{i+1}=(x_1+x_2+0.5)/2 \\ x_{i-1}=(x_6+x_7+0.5)/2 \end{cases} & \text{if the direction is D3;} \\ \begin{cases} x_{i+1}=x_1 \\ x_{i-1}=x_7 \end{cases} & \text{if the direction is D4;} \\ \begin{cases} x_{i+1}=(x_0+x_1+0.5)/2 \\ x_{i-1}=(x_7+x_8+0.5)/2 \end{cases} & \text{if the direction is D5;} \\ \begin{cases} x_{i+1}=x_0 \\ x_{i-1}=x_8 \end{cases} & \text{if the direction is D6;} \\ \begin{cases} x_{i+1}=(x_0+x_3+0.5)/2 \\ x_{i-1}=(x_5+x_8+0.5)/2 \end{cases} & \text{if the direction is D7.} \end{cases} \quad (4)$$

where x_0, x_1, … x_8 are the pixel values in the mask as shown in Fig.3. If Δx is larger than a certain threshold α, it means the direction detected by Sobel operator is not correct and the line edge direction will be detected further. Otherwise, the direction of the edge is considered correctly and the magnitude of the edge can be calculated by:

$$mag_s = \sqrt{G_x^2 + G_y^2} \quad (5)$$

B. Detecting Edge Direction of Line

If the direction detected by Sobel operator is not correct, the line and its direction in the 3×3 mask should be further detected. In order to detect line direction, firstly the eight line magnitude values corresponding to eight directions should be calculated by:

$$\begin{cases} m_{D0}=2(x_3+x_4+x_5)-x_0-x_1-x_2-x_6-x_7-x_8 \\ m_{D1}=2((x_3+x_6+0.5)/2+x_4+(x_2+x_5+0.5)/2)-x_2-x_3 \\ \quad -x_5-x_6-(x_1+x_4+0.5)/2-(x_4+x_7+0.5)/2 \\ m_{D2}=2(x_2+x_4+x_6)-x_0-x_1-x_3-x_5-x_7-x_8 \\ m_{D3}=2((x_6+x_7+0.5)/2+x_4+(x_1+x_2+0.5)/2)-x_1-x_2 \\ \quad -x_6-x_7-(x_3+x_4+0.5)/2-(x_4+x_5+0.5)/2 \\ m_{D4}=2(x_1+x_4+x_7)-x_0-x_3-x_6-x_2-x_5-x_8 \\ m_{D5}=2((x_0+x_1+0.5)/2+x_4+(x_7+x_8+0.5)/2)-x_0-x_1 \\ \quad -x_7-x_8-(x_3+x_4+0.5)/2-(x_4+x_5+0.5)/2 \\ m_{D6}=2(x_0+x_4+x_8)-x_1-x_2-x_5-x_3-x_6-x_7 \\ m_{D7}=2((x_0+x_3+0.5)/2+x_4+(x_5+x_7+0.5)/2)-x_0-x_3 \\ \quad -x_5-x_8-(x_1+x_4+0.5)/2-(x_4+x_7+0.5)/2 \end{cases} \quad (6)$$

where x_0, x_1, … x_8 are the pixel values in the mask shown in Fig.3. Then, the maximum line magnitude and its direction are obtained by:

$$mag_l = \max(m_{Di}),\ i = 0,1,...,7 \quad (7)$$

$$Dm = \arg\max_{Di}(m_{Di}),\ i = 0,1,...,7 \quad (8)$$

If the maximum line magnitude mag_l is larger than a certain threshold β, a line in the mask and the line direction Dm are considered. Otherwise, the line magnitude of the direction is set zero, i.e.,, mag_l=0 for direction Dm.

III. Concealment By Pixel Extrpolation

The pixels in the missing block are recovered pixel-by-pixel in our algorithm. Assume the pixel to be recovered is *y*, then eight edge directions which have chance to traverse *y* are shown in Fig.5.

We use eight masks with the size of 3×3, each for one direction, to determine if there is an edge that can be covered by one of the eight directions traversing *y*. Pixels in the mask are denoted by x_0 to x_8 as shown in Fig.3. The nearest pixel to *y* along any direction is denoted y_1 and the second nearest pixel is denoted y_2. Example positions of y_1 and y_2 on directions D2 and D7 are shown in Fig.6. The 3×3 masks for directions D0, D1, D2, D4, D6 and D7 are centered at y_2, that is, the pixel x_4 in the 3×3 mask for these directions is y_2. The 3×3 masks for directions D3 and D5 are centered at y_1, that is, the pixel x_4 in the 3×3 mask for directions D3 and D5 is y_1. For each direction from D0 to D7, we detect the edge direction and its magnitude (mag_s or

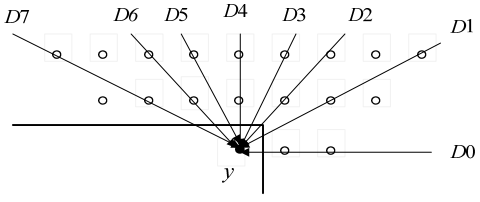


Figure. 5 Eight edge directions which have chance to traverse the to-be-recovered pixel.

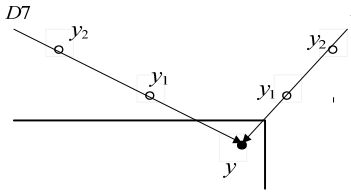


Figure. 6 Example positions of y_1 and y_2 along directions D2 and D7.

mag_l) in its corresponding 3×3 mask by using the above edge detection method. The direction with the maximum magnitude (mag_s or mag_l) is considered as the edge direction which traverses the pixel to be recovered. If the magnitudes of eight directions are all zero, it indicates that there is not any edge traversing the pixel to be recovered. In other words, the pixel *y* is supposed to be in a flat region.

If there is an edge direction traversing the pixel *y*, it will be concealed through extrapolation of two pixels y_1 and y_2 along the determined direction as:

$$y = \frac{2}{3} y_1 + \frac{1}{3} y_2 \tag{9}$$

If pixel *y* is in a flat region, it will be recovered based on weighted pixel averaging [11].

The order of pixels to be recovered in the missing block is from the outer layer to the inner layer [3].

IV. Experimental Results

In order to evaluate the performance of the proposed method of detecting edge direction, the simple Sobel operator detection in [3] is replaced by our proposed method. It is worth clarifying that eight directions as shown in Table 1 are used in our experiment instead of ten directions in [3].

The 512×512 Lena image and the 512×384 StockCurve image are used to evaluate the performance of the proposed algorithm. The StockCurve image is the typical image appeared in the stock market, which is full of lines. Both cases of 8×8 block and 16×16 block are considered in our experiments because the case of 8×8 block was used in most previous spatial error concealment methods and the case of 16×16 block was considered in some of previous studies on error concealment. In the experiment, $\alpha = 10$ and $\beta = 200$ for both cases of 8×8 block and 16×16 block. The traditional PSNR is measured to evaluate the quality of recovered image. PSNR is defined by:

$$\mathrm{PSNR} = 10 \log \frac{255^2}{\frac{1}{\mathrm{MN}} \sum_{x=0}^{\mathrm{M}-1} \sum_{y=0}^{\mathrm{N}-1} [f(x,y) - F(x,y)]^2} (\mathrm{dB}) \tag{10}$$

where M is the width of the image, N is the height of the image, $f(x, y)$ and $F(x, y)$ are the pixel values of the original image and the reconstructed image respectively.

Firstly the case of 8×8 block with loss rate 25% is tested. Table 2 lists the PSNR comparison results between our proposed method and the previous work in [3]. The difference between these two methods is that the advanced method of detecting edge direction in the error concealment is used in our proposed method, while the Sobel operator is only used for detecting edge directions in [3].

TABLE II.
PERFORMANCE COMPARISON FOR 8×8 BLOCK

Image	Previous [3] (dB)	Proposed (dB)
Lena	33.22	33.45
StockCurve	23.81	25.25

Table 2 shows that the PSNR performance for Lena image has 0.23dB gain with our proposed method comparing with that of the previous method. While for StockCurve image, which has more lines in the image, the PSNR performance has about 1.44 dB gain against the previous method. Fig.7 illustrates the comparison and difference of the corresponding reconstructed Lena images recovered by the previous work [3] and the proposed method: (a) is the original image, (b) is the corrupted image, (c) is the reconstructed image using algorithm in [3] and (d) is the reconstructed image given by the proposed technique.

Fig.8 shows the comparison of the corresponding reconstructed StockCurve images recovered by the previous work [3] and the proposed method: (a) is the original image, (b) is the corrupted image, (c) is the reconstructed image using algorithm in [3] and (d) is the reconstructed image given by the proposed method. Based on the comparison shown in Fig.7 and Fig.8, it is obviously that the proposed algorithm can recover the corrupted block with better subjective quality, especially for images with more lines.

For better subjective evaluation, the enlarged portions of the StockCurve image are shown in Fig.9. By comparing (c) and (d), it is clearly shown that the line is recovered much better by using the proposed method.

The second experiment was carried out with the case of 16×16 block with loss rate 10%. Table 3 lists the PSNR comparison results between our proposed method and the previous work in [3]. The results in Table 3 demonstrate that both Lena image and StockCurve image have better objective quality comparing with the previous method. The concealed results of Lena image are shown in Fig.10. They are: (a) the damaged image, (b) the concealed image using algorithm in [3] and (c) the concealed image given by the proposed technique. Fig.11 shows the concealed results of StockCurve image: (a) is the damaged image, (b) is the concealed image using algorithm in [3] and (c) is the concealed image given by the proposed technique. It can be seen that the concealed image by using the proposed method looks much better.

V. Conclusions

In this paper, an advanced edge detection method, which can estimate both step edge and line direction correctly, is proposed. The proposed method of detecting edge direction is applied in a spatial error concealment algorithm in order to evaluate its performance.

Experimental results show that the proposed edge direction detection method can provide more accurate edge direction for further spatial error concealment. Therefore, both objective and subjective of the recovered images can be improved significantly, especially for images with more lines.

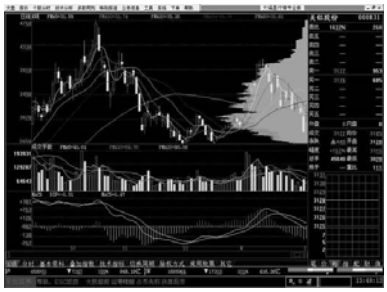
(a) Original Lena image

(b) Corrupted image

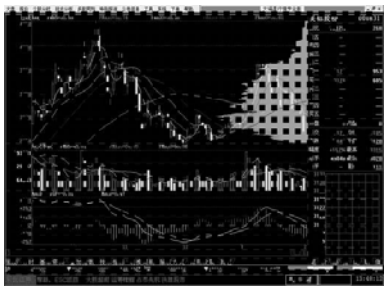
(c) Reconstructed image by [3],
PSNR = 33.22dB

(d) Reconstructed image by proposed method,
PSNR = 33.45dB.

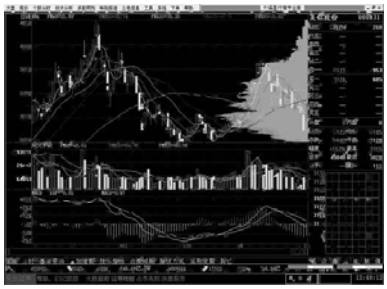
Figure. 7 Comparison results of Lena image for the case of 8×8 block



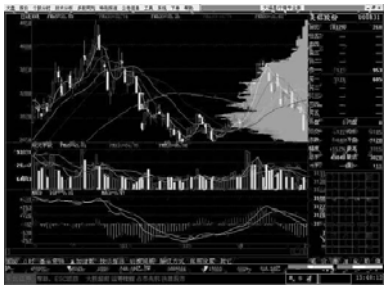
(a) Original StockCurve image



(b) Corrupted image



(c) Reconstructed image by [3],
PSNR = 23.81dB.



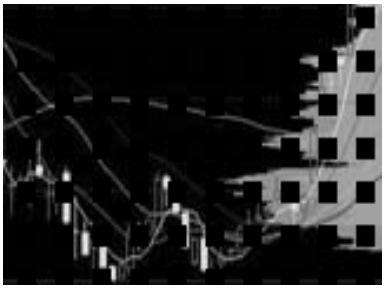
(d) Reconstructed image by proposed method,
PSNR = 25.25dB

Figure. 8 Comparison results of StockCurve image for the case of 8×8 block

ACKNOWLEDGMENT

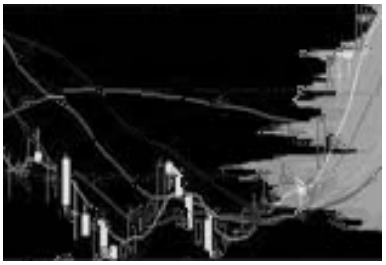
This work was supported by the project of National Natural Science Foundation of China under Grant 60832002, 61171078 and in part by the Research Fund for Doctorial Program of Higher Education of China under Grant 20110061110084 and the Outstanding Youth Foundation of Jilin University under Grant 200905018.

(a)



(b)

(c)

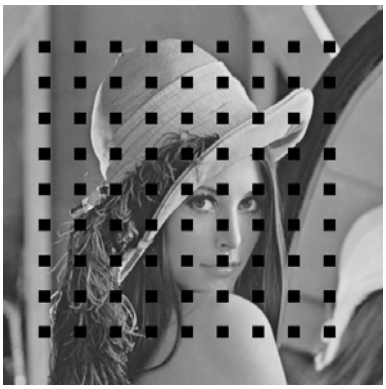


(d)

Figure. 9 Comparison of the zoomed portions in Fig.8.

TABLE III.
PERFORMANCE COMPARISON FOR 16×16 BLOCK

Image	Previous [3] (dB)	Proposed (dB)
Lena	35.91	36.00
StockCurve	29.34	29.41



(a) Corrupted image

(b) Concealed image by [3], PSNR = 35.91dB

(c) Concealed image by proposed method, PSNR = 36.00dB

Figure. 10 Comparison results of Lena image for the case of 16×16 block

REFERENCES

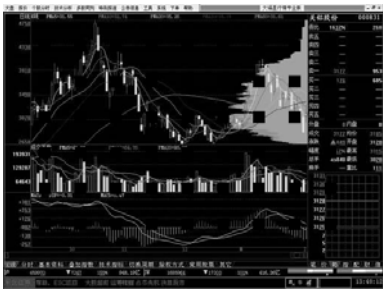
[1] R. Zhang, Y. Zhou. and X. Huang, "Content-adaptive spatial error concealment for video communication", IEEE Trans. Consumer Electronics, vol.50, no.1, pp. 335-341, 2004.

[2] M.-H. Jo, H.-N. Kim and W.-J. Song, "Hybrid error concealments based on block content", IET Image Processing, vol.1, no.2, pp.141-148, 2007.

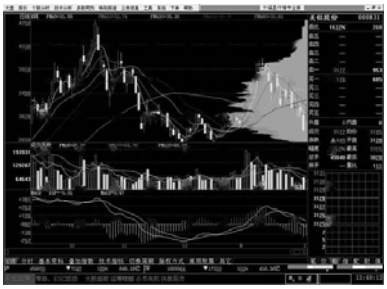
[3] Y. Zhao, H. Chen, X. Chi, J. S. Jin, "Spatial Error Concealment Using Directional Extrapolation", Proceedings of Digital Image Computing: Techniques and Applications, pp. 278-283, 2005.

[4] O. Nemethova, A. Al Moghrabi, M. Rupp, "An Adaptive Error Concealment Mechanism for H.264/AVC Encoded Low-Resolution Video Streaming", Proceedings of 14th European Signal Processing Conference (EUSIPCO), pp. 1-5, 2006.

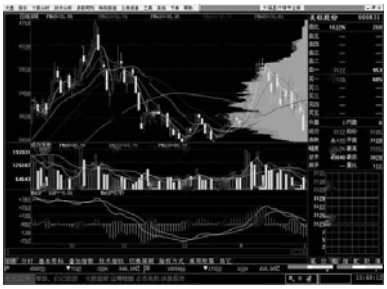
[5] O. Nemethova, A. Al Moghrabi, M. Rupp, "Flexible Error Concealment for H.264 Based on Directional Interpolation", Proceedings of the 2005 International



(a) Corrupted image



(b) Concealed image by [3], PSNR = 29.34dB



(c) Concealed image by proposed method, PSNR = 29.41dB.

Figure. 11 Comparison results of StockCurve image for the case of 16×16 block

Conference on Wireless Networks Communications and Mobile Computing", vol.2, pp. 1255-1260, 2005.

[6] Wei-Ying Kung, Chang-Su Kim and C.-C. Jay Kuo, "Spatial and Temporal Error Concealment Techniques for Video Transmission Over Noisy Channels", IEEE Trans. on Circuits and System for Video Technology, vol.16, no.7, pp.789–802, 2006.

[7] Wonki Kim, Jasung Koo and Jechang Jeong, "Fine Directional Interpolation for Spatial Error Concealment", IEEE Transactions on Consumer Electronics , vol.52, no.3, pp.1050-1056, 2006.

[8] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projections onto convex sets", IEEE Trans. Image Processing, vol.4, pp.470-477, 1995.

[9] Wilson Kwok and Huifang Sun, "Multi-directional interpolation for spatial error concealment", IEEE Trans. Consumer Electronics, vol.39, pp. 455-460, 1993.

[10] Y. Zhao, D. Tian, M.M.Hannukasela, M. Gabbouj, Spatial Error concealment Based on Directional Decision and Intra Prediction, IEEE International Symposium on Circuits and Systems, Volume 3, pp.2899-2902, 2005.

[11] Ye-Kui Wang, M. M. Hannuksela, Viktor Varsa, Ari Hourunranta and Moncef Gabbouj. The error concealment feature in the H.26L test model. Proc. of ICIP 2002, vol.II, pp.729-732, 2002.

Yan Zhao was born in Jilin, China, in 1971. She received the B.S. degree in communication engineering in 1993 from Chan-g-chun Institute of Posts and Telecommunications, the M.S. de-gree in communication and electronic in 1999 from Jilin Univ-ersity of Technology, and the Ph.D. degree in communication and information system in 2003 from Jilin University.

She has been a postdoc researcher in the Digital Media Institute of Tampere University of Technology in Finland from Mar.2003 to Dec.2003. From Mar. 2008 to Aug.2008, she was a visiting p-rofessor in the Institute of Communications and Radio-Frequency Engineering in the Vi-enna University of Technology. She currently is an associate professor of communication enginee-ring. Her research interests include image and video coding, multimedia signal processing and error concealment for audio and video transmitted over unreliable networks.

Dr. Zhao is a member of IEEE.

Hexin Chen was born in Jilin, China, in 1949. He received the M.S. and Ph.D. degrees in communication and electronic in 1982 and 1990 from Jilin University of Technology, respectively.

He has been a visiting scholar in the University of Alberta from 1987 to 1988. From Feb.1993 to Aug.1993, he was a visiting professor in Tampere University of Technology in Finlan-d. He currently is a professor of communication engineering. His research interests include image and video coding, multidimensional signal processing, image and video retrieval and audio and video synchronization.

Shigang Wang was born in Jilin, China, in 1962. He received the B.S. degree in 1983 from Northeastern University, the M.S degree in communication and electronic in 1998 from Jilin University of Technology, and the Ph.D. degree in communication and information system in 2001 from Jilin University.

He currently is a professor of communication engineering. His research interests include image and video coding, multidimensional signal processing and stereoscopic and multi-view video coding.

Moncef Gabbouj received his BS degree in electrical engineering in 1985 from Oklahoma State University, Stillwater, and his MS and PhD degrees in electrical engineering from Purdue University, West Lafayette, Indiana, in 1986 and 1989, respectively.

Dr. Gabbouj is currently an Academy Professor and Professor at the Department of Signal Processing at Tampere University of Technology, Tampere, Finland. His research interests include multimedia content-based analysis, indexing and retrieval; nonlinear signal and image processing and analysis; and video processing, coding and communications.

Palmprint Image Processing and Linear Discriminant Analysis Method

Shuang Xu
Dalian Nationalities University/Information & Communication Engineering, Dalian, China
Email: xushuangcong@163.com

Jifeng Ding
Dalian Nationalities University/Information & Communication Engineering, Dalian, China
Email: hitdjf@163.com

***Abstract*—In this paper, the method of processing and linear discriminant analysis of palmprint image is proposed. The palmprint image processing focuses on the location and segmentation which involves rotation and transition. By means of finding the two locate points about the index finger and middle finger, ring finger and little finger, the palmprint image is rotated and corrected a new coordinate system is created, which determines the region of interest (ROI). Linear discriminant analysis is a method of Gabor plus improved two-dimensional linear discriminant (Gabor+I2DLDA) which is improved 2DLDA method by integrating the Gabor wavelet representation of palm images is proposed. Traditional two-dimensional linear discriminant (2DLDA) method eliminates the column relevance of the image, while improved 2DLDA is a direct 2DLDA method which is defined in the traditional 2DLDA basis. The Gabor wavelets are used to extract palmprint features. The proposed Gabor+I2DLDA yields greater palmprint recognition accuracy while reduces the dimension.**

The experiment results show that location accuracy and recognition accuracy. The effectiveness of the proposed method is also verified using the PolyU palmprint databases for palmprint recognition.

***Index Terms*—palmprint image processing; location and segmentation; region of interest; linear discriminant analysis; Gabor wavelets; palmprint recognition**

I. INTRODUCTION

In recent years, biometrics has been used to refer to the emerging field of information technology devoted to automated identification of individuals using biological trains, which include palmprint, fingerprint, hand-geometry, face, DNA, iris, voice, retina, gait, ear and dynamic signature etc [1-2]. While palmprint recognition is an important kind of biometrics aiming to recognize human identity through some unique features in human palms, such as the principle lines, wrinkles, ridges, texture, minutiae and singular points etc. Compared with the other physical characteristics, palmprint has several advantages: stable features, low cost, non-intrusive and high user acceptance.

Palmprint image processing which includes palmprint segmentation, location, and normalization is a key technique for palmprint identification. Palmprint location and segmentation has two methods: the square location and segmentation [3-9] and the inscribed circle location and segmentation [10-13]. In this paper, a square location and segmentation method is proposed which is to use edge tracking algorithm to determine the intersection of the angle sub-line with the palm of the ring finger and little finger, and with the palm of the index finger and middle finger, and then cut out a square region of interest (ROI) according to the intersection. Disadvantage of this approach is the intersection that angle sub-line is not easy to confirm, so the effective area extracted palmprint is not high accuracy. An inscribed circle segmentation method need to search the inscribed circle, thus it needs long preprocessing time, higher complexity.

To solve the above problem in palmprint segmentation, we propose a new method of palmprint image processing. Experimental results show that the method can still precisely position and extract ROI of palmprint even if the same hand images produce translation and rotation. Therefore, this method can speed up the pre-speed, lower computational complexity, and enhance matching accuracy.

Palmprint recognition depends on the choice of features used by the classifier, it is common to start with a given set of features and then attempt to derive an optimal subset of features leading to high classification performances. Feature extraction is a key technique for palmprint recognition. Researchers have recently developed many algorithms [14-17]. Gabor wavelets are extensively used to extract feature for biometric recognition [18-22]. But the dimension of the extracted Gabor wavelets feature matrix is often very high due to convolution of biometric image with a series of Gabor filters. Therefore, there are some subspace projection methods to reduce feature matrix dimension.

A powerful tool used for feature extraction and data reduction is the linear discriminant analysis (LDA), which maximizes the ratio of the trace of the between-class scatter to the trace of the within-class scatter matrix. However, LDA suffers from the singularity problem of the within-class scatter matrix caused by the limited number of training samples. This problem is referred to as the small sample size problem. To address the problem, the two dimensional LDA (2DLDA) has been of wide

doi:10.4304/jmm.7.3.269-276

concern. The 2DLDA method was originally developed by Li and Yan [23]. 2DLDA is mainly focused on constructing the image covariance matrix directly using the original image matrices. In contrast to the covariance matrix of LDA, the size of the image covariance matrix using 2DLDA is much smaller. As a result, 2DLDA is easier to evaluate the covariance matrix accurately and to require less time to determine the corresponding eigenvectors. While traditional 2DLDA which simultaneously diagonalizes the image between covariance matrix and the image within covariance matrix, the between covariance matrix is usually whitened as part of the method.

To solve this problem, the paper proposed Gabor+I2DLDA which avoid this whitening process, because this whitening process can be shown to be redundant and also leads to increased computational complexity. In the proposed method, Gabor wavelets first capture the properties of spatial localization, orientation, spatial frequency and phase relationship. The I2DLDA method is then applied to project palmprints from the high-dimensional palmprint space to a significantly lower- dimensional feature space, in which the palmprints from the different palms can be discriminated much more efficiently.

II. PALMPRINT IMAGE PROCESSING

A. Binary of Palmprint image

In this paper, Hong Kong Polytechnic University providing the standard palm database is used. The standard palm image is shown in Fig. 1, which collected palmprint image is high quality and easy to be processing. Since the original palmprint images have system noise, it is not binary directly. If it's binary directly, the binary image will have a lot of noise contour points which cause by inaccurate locate of the image. Before binarization processing, the original palmprint image makes to do median filtering. The median filter is the key to gray level threshold selection. The experiment makes use of the input image histogram to select thresholds. The results of the histogram are shown in Fig. 2. Based on the results, the local minimum threshold gray is selected which is range form 11 to 255. Because the entire database of samples collects in a stable environment, a fixed threshold of 30 gray is chosen.

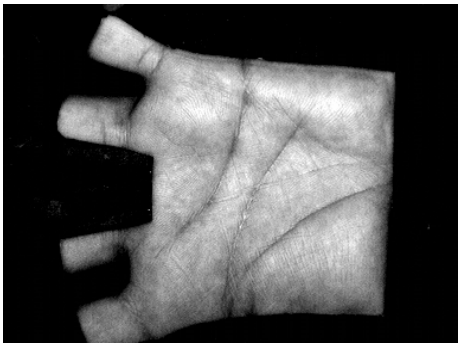


Figure 1. Original palmprint image

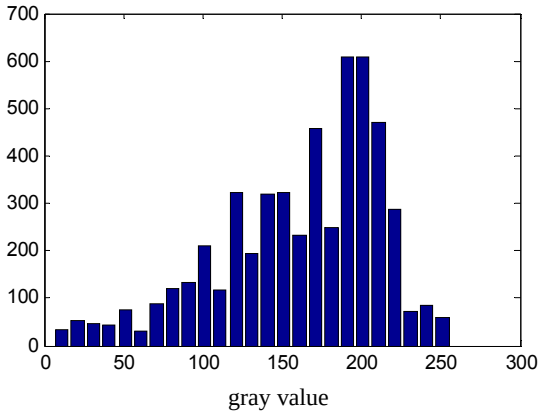


Figure 3. Histogram of the input image

After the filtered binary image is processing, a binary image is get. Then it uses boundary tracking methods to track the binary image borders, converts the border point and re-establish the boundary of the new image which is shown in Fig. 3. After the above steps, a new set of binary images are obtained and binary image processing is completed. In the binary image, the 0 points of the pixel value are black points and the 255 points of pixel value are white points.

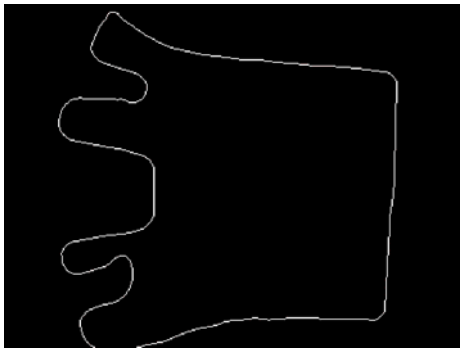


Figure 4. The new images showing the boundary

B. Getting Palmprint Region of Interest

The location and segmentation of palm image are a vital link in the palmprint preprocessing. The quality of location and segmentation results has a direct relationship to the region of interest of accuracy. The location and segmentation of palm image provide a consistent frame of reference to the feature extraction and matching palm prints, so palmprint matching and feature contrasts have the uniform standards. The correct location and segmentation can the lessen image noise and reduce the difficulty of feature matching algorithm, with the result that can ensure the accuracy and effectiveness of identification systems.

The location and segmentation of palm image is a series of different adjustment, which can split out the palm center active area to extract the palm feature, thus it realizes a variety of feature matching identification. Effective area of the center of palm often is called the

region of interest (ROI). For the same palm images, ROI is the same. ROI contains the most important features of palmprint image information and all of the palmprint images are present in this region. According to the location and segmentation to extract the ROI, it brings down the image matching the difficulty and increases robustness of the system.

The ROI extraction methods include the following main steps.

(i) To separate the fingers and palms.

(ii) To find the two valley points of the index finger and middle finger, ring finger and little finger.

(iii) To rotate image based on the two valley points and correct image position.

(iv) To create coordinate system according to valley points and determine ROI.

The new method does not rely on mathematical modeling, but it analyzes the main forms of palmprint, to complete the segmentation.

(i) To separate the fingers and palms. Binarization processing of images that reflect the shape of a hand, including the index finger, middle finger, ring finger, little finger and palm of the upper and lower boundaries of the upper and lower boundaries. As the locate point of the selection mainly depends on four fingers, only the border of four fingers is considered to get rid of the palm of the border in the study. The separation main idea of the fingers and palm is to store each boundary point of binary image as an array column by column vector of the form. Four fingers are total 8 edges each column, which thinner at the edges is just 8 points and wide at the edges is larger than 8 points. Whereas the palm area is 2 edges each column, which thinner at the boundary is 2 points, the wider parts close to the heel of your is more than 8 points. The ROI is only interested in a column vector of 8 points, and they meet as in (1).

$$l = \begin{cases} l+1 & C_n = 8 \\ 0 & C_n \quad \text{other} \end{cases} \tag{1}$$

Here C_n is the number of boundary points for each column and l is the number of column vectors to the number of boundary points for the 8. After several trials the total number of 8 points each column vector is generally greater than or equal to 10, and the boundary points is equal to 8 multiplied by l. which is used to create a new binary image. It can fit the broad contours of four fingers, which is shown in Fig. 4.

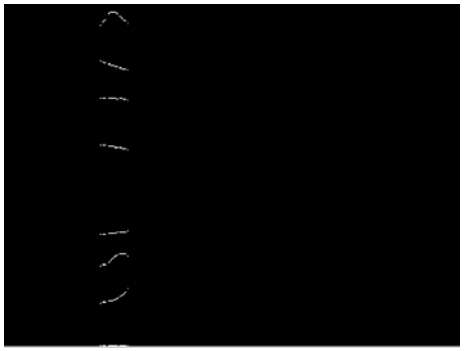


Figure 5. Four fingers outline image

(ii)To determine the location points. The four fingers outline doing the binary image processing extracts every boundary and reflects the form of an array P_i. The array $P_i(1)$ is to show all points of index finger upper border. The array $P_i(2)$ is to express all points of index finger under border. The array $P_i(3)$ is to represent all points of middle finger upper border. The array $P_i(4)$ is to indicate all points of middle finger under border. The array $P_i(5)$ is to show all points of ring finger upper border. The array $P_i(6)$ is to express all points of ring finger under border. The array $P_i(7)$ is to represent all points of little finger upper border. The array $P_i(8)$ is to indicate all points of little finger under border. The locate points are two crossing points. One is the crossing point between points of index finger under border and points of middle finger upper border. The other is the crossing point between points of ring finger under border and points of little finger upper border. so long as the $P_i(2)$, $P_i(3)$, $P_i(6)$ and $P_i(7)$ for the direction of the boundary point closest to the palm is extracted , the locate points can be determined. Finally, row coordinate of the array is stored in four variables y_1, y_2, y_3 and y_4, and they satisfy as in (2), (3), (4) and (5).

$$y_1 = P_i(X_{2,}2), \quad if\ y_1 \geq P_i(X_{2,}2) \quad X_2 = 2:8:N-6 \tag{2}$$

$$y_2 = P_i(X_2+1,2), \quad if\ y_1 \geq P_3(X_2+1,2) \quad X_2 = 2:8:N-6 \tag{3}$$

$$y_3 = P_i(X_2+4,2), \quad if\ y_3 \geq P_i(X_2+4,2) \quad X_2 = 2:8:N-6 \tag{4}$$

$$y_4 = P_i(X_2+5,2), \quad if\ y_4 \geq P_i(X_2+5,2) \quad X_2 = 2:8:N-6 \tag{5}$$

The variable y_1 means first column coordinates points of the second row in four contours which represent the lower boundary of the index finger. The variable y_2 says first column coordinates points of the third row in four contours which express the upper boundary of the middle finger. The variable y_3 expresses first column coordinates points of the sixth row in four contours which stand for the lower boundary of the ring finger. The variable y_4 indicates first column coordinates points of the seventh row in four contours which show the upper boundary of the little finger.

The variable X_2 is to state the second line of four fingers contour outline which represents the lower boundary of the index finger. It is a column for every 8 points for every one until two locate points are found. The locate point in the index finger and middle finger is between the variable y_1 and the variable y_2, which is defined as $M1(x_1, y_1)$. The locate point in ring finger and little is between the variable y3 and the variable y4, which is defined as $M2(x_2, y_2)$. The locate point image is shown in Fig. 5.

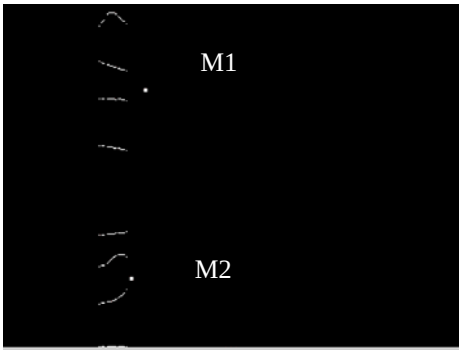


Figure 6. The locate point image

(iii) To rotate and correct the image. As the acquired image can cause translation and rotation of palm, the palm image must be rotated and corrected after determining the locate points. According to the locate points the rotation angle are calculated, as in (6).

$$\theta = \arctan \frac{y_2 - y_1}{x_2 - x_1} \tag{6}$$

The original image makes clockwise rotation when θ is greater than 0. The original image makes counter-clockwise rotation when θ is less than 0. The vertical axis of the two locate points is equal after palm image rotation, so the original image does not rotate when θ is equal to 0.

Let the image coordinates of any point is to be (x, y), and the midpoint of two locate points is to equal (x_0, y_0). The rotational image obtained the new coordinates $(\overline{x}, \overline{y})$ with the midpoint of two anchor points as the center of rotation, which is expressed as (7) and (8).

$$\overline{x} = x_0 + (x - x_0) \times \cos\theta + (y - y_0) \times \sin\theta \tag{7}$$

$$\overline{y} = x_0 \pm (x - x_0) \times \sin\theta + (y - y_0) \times \cos\theta \tag{8}$$

The "+" sign indicates the image clockwise rotation in the "±" of equation (8) and the "-" sign states the image counter-clockwise rotation in the '±' of equation (8) indicates the image. Fig. 6 illustrates the rotational and corrected image.

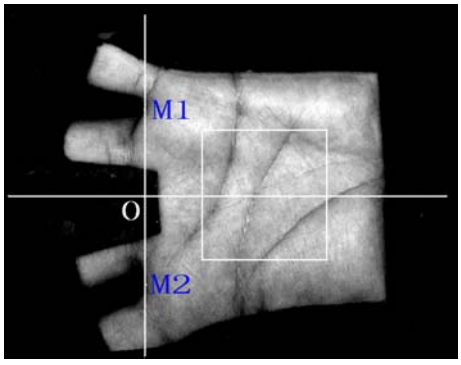


Figure 7. Extraction of the ROI

(iv) To create coordinate system and determine ROI. After the above processing of Palm images, a new coordinate is established, whose X-axis is connection line with M1 and M2, and Y-axis is vertical line through the midpoint of M1M2 line. The intersection of the X-axis and Y-axis is for the origin of the coordinate system which is shown in Fig. 6. According to establish the coordinate ROI is received. In light of follow-up feature extraction and feature matching requirements, the ROI of different palm samples for the same hand palm is acquired whose fixed size is 128 × 128.

C. Experimental results and analysis

In the experiments we adopt test palmprint images of the PolyU-I and PolyU-II palmprint database [24]. the PolyU-I and PolyU-II palmprint database which is a result of palmprint recognition project carried out at Hong Kong Polytechnic University is used for the evaluation of the recognition performance. The PolyU-I palmprint database contains 600 grey-scale images of 100 with six samples for eachpalm. The PolyU-II palmprint database contains 7700 grey-scale images of 385 with twenty samples for eachpalm. The samples in public databases were captured by a CCD based palmprint device. The resolution of all original palmprint images is 384×284 pixels at 75 dpi. Six palm samples of the same hand palm are shown in Fig. 7. The rotational and corrected results are shown in Table I.

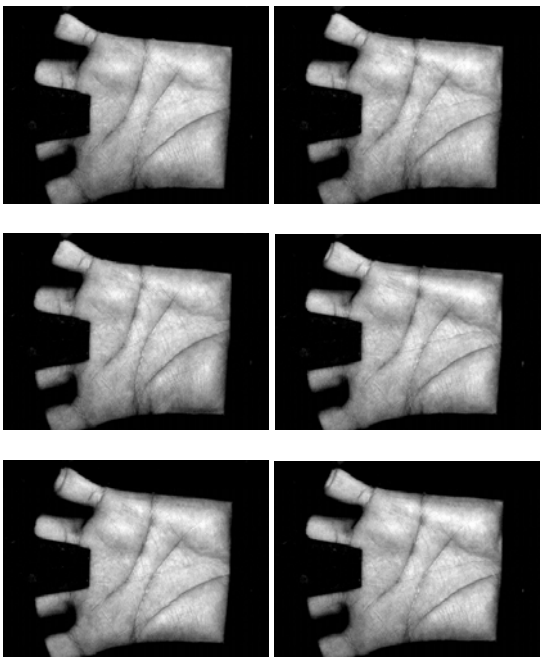


Figure 8. Six palm samples of the same hand palm

Table 1 includes locate point coordinates and rotation angle of six palm samples of the same hand palm. The rotation angle is calculated based on equation (6). To align different palmprint images for matching, the central part of a palmprint, whose size is 128×128, is cropped to represent the whole palmprint by using above method in this paper. ROI is shown in Fig. 8 for six samples for

eachpalm. Table II and Table III show the rate of correct location.

TABLE I. THE SAME HAND PALM ROTATION CORRECTION RESULTS OF SIX PALM SAMPLES

Palmprint image number	M1 coordinate	M2 coordinate	Rotation angle
1	(68,123)	(221,107)	-0.1042
2	(72,121)	(228,109)	-0.0768
3	(70,116)	(231,108)	-0.0496
4	(75,124)	(233,108)	-0.1009
5	(77,131)	(231,106)	-0.1609
6	(77,128)	(234,109)	-0.1204

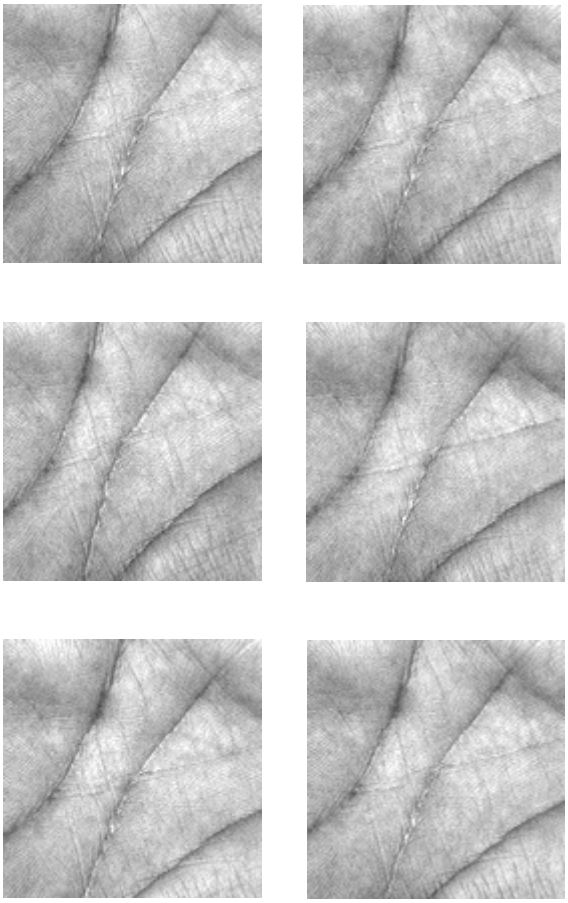


Figure 9. Six palm samples ROI of the same hand palm

Table II demonstrates the location and segmentation results of this method in PolyU-I database which contains 600 palm images, where the number of correct location is 594 images and location accuracy is 99%. While correct location is 590 images and location accuracy is 98.3% for [30] proposed the method in PolyU-I database. Table III illustrate the location and segmentation results in 2000 palm images of, where the number of correct location is 1976 images and location accuracy is 98.8%. While correct location is 1964 images and location accuracy is 98.2% for [30] proposed the method in PolyU-II database. Besides, consuming time of CPU is also obtained which is Pentium dual-core @ 2.80GHz and 2.79GHz, 1.99GB RAM for ROI extraction. They respectively are 4.62s and 11.38s for this method. They respectively are 5.20s and 13.24s for this method. By means of the palm image analysis to the error location, the error is mainly due to not enough extended palm and inaccuracy placement. If the palm is stretched and the palm position is standard, the accuracy can be increased. Experimental results show that the method is high accuracy of location.

TABLE II. LOCATION RESULTS IN POLYU-I DATABASE

	Number of test images	Number of correct location	Number of error location	Location accuracy	Location time(s)
this method	600	594	6	99%	4.62
proposed method of [16]	600	590	10	98.3%	5.20

TABLE III. LOCATION RESULTS IN POLYU-II DATABASE

	Number of test images	Number of correct location	Number of error location	Location accuracy	Location time(s)
this method	2000	1976	24	98.8%	11.38
proposed method of [16]	2000	1964	36	98.2%	13.24

III. LINEAR DISCRIMINANT ANALYSIS METHOD

A. . Gabor wavelets transformation

Gabor wavelets provide more precision, multi-channel frequency characteristics to show the mechanism of image space and can reflect the feelings of biological neurons in the visual field. The Gabor wavelets have been found to be particularly suitable for image decomposition and representation when the goal is to find local and discriminating features.

The Gabor wavelets (kernel, or filter) functions are Gaussian modulated by complex sinusoids. Gabor wavelets can be defined as follows [20]

$$\psi_{\mu,\nu}(z)=\frac{\left\|k_{\mu,\nu}\right\|^2}{\sigma^2}e^{-\|k_{\mu,\nu}\|^2\|z\|^2/2\sigma^2}[e^{jk_{\mu,\nu}z}-e^{-\sigma^2/2}] \quad (9)$$

where μ and ν define the orientation and scale of the Gabor kernels, $z(x,y)$ is the variable in a spatial domain and $k_{\mu,\nu}$ is the wave vector, which determines the scale and orientation of Gabor wavelets, $k_{\mu,\nu}=k_\nu e^{i\phi_\mu}$, where $k_\nu=k_{\max}/f^\nu$ and $\phi_\mu=\pi\mu/8$. $k_{\max}$ is the maximum frequency, and f is the spacing factor between kernels in the frequency domain [7], σ is the standard deviation of the Gaussian envelop determining the number of oscillations. It is usual to use Gabor wavelets at five different scales, $\nu\in\{0,\ldots,4\}$ and eight orientations,

$\mu \in \{0,\ldots,7\}$ with the following parameters: $\sigma = 2\pi$, $k_{\max} = \pi/2$ and $f = \sqrt{2}$ [25].

The Gabor wavelets representation of an image is the convolution of the image with a family of Gabor kernels as defined by (9). In a given palm image, $I(x, y)$, the convolution is defined as

$$F_{\mu,\nu}(z) = I(z) * \psi_{\mu,\nu}(z) \tag{10}$$

where $F_{\mu,\nu}(z)$ is the convolution result corresponding to the Gabor kernel at orientation μ and scale ν. We then derive an augmented Gabor feature vector X which is defined as follows

$$X = \{F_{0,0}^T, F_{0,1}^T, \ldots, F_{4,7}^T\} \tag{11}$$

where T is the transpose operator.

B. GABOR+I2DLDA palmprint recognition

Since the Gabor outputs are two-dimensional matrices, it is necessary to reduce the dimension in rows or columns. Therefore we apply the I2DLDA method criterion to reduce the dimension.

The improved 2DLDA method will be derived which is more suitable for extracting the discriminant palmprint features. This method is referred to I2DLDA. It applied on the Gabor wavelet outputs which are robust to variations because of illumination and pose.

Let the training sample set $X = \{x_{ij} | i = 1,2,...,L; j = 1,2,...,M_i\}$ is a $N \times M$ matrix, where N is the dimension of the training samples, M is the training samples, and x_{ij} is the jth training sample in classes i. There are L known pattern classes of training sample set, where M_i is the number of training sample set in classes i, μ_i the mean of the training sample set in classes i and μ the global mean of all training sample set.

The traditional 2DLDA method is to find an optimal projection vector to make the maximum divergence ratio of the between-class scatter to the within-class scatter, which is to meet the Fisher criterion function

$$J(\phi) = \frac{\phi^T S_B \phi}{\phi^T S_W \phi} \tag{12}$$

where T is the transpose operator, S_B and S_W is respectively the between-class scatter matrix and the within-class scatter matrix，defined as follows

$$S_B = \frac{1}{M}\sum_{i=1}^{L} M_i (\mu_i - \mu)^T (\mu_i - \mu) \tag{13}$$

$$S_W = \frac{1}{M}\sum_{i=1}^{L}\sum_{j=1}^{M_i} (x_{ij} - \mu_i)^T (x_{ij} - \mu_i) \tag{14}$$

Clearly, the optimal projection direction ϕ is the largest eigenvector corresponding to the $S_W^{-1} S_B$, which are obtained by solving the generalized eigenvalue equation $S_B\phi = \lambda S_W \phi$.

Traditional 2DLDA method eliminates the column relevance of the image. This paper proposed a direct 2DLDA method which can eliminate the row correlation in the traditional 2DLDA basis. The between-class scatter matrix and the within-class scatter matrix are described by

$$S_B' = \frac{1}{M}\sum_{i=1}^{L} M_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{15}$$

$$S_W' = \frac{1}{M}\sum_{i=1}^{L}\sum_{j=1}^{M_i} (x_{ij} - \mu_i)(x_{ij} - \mu_i)^T \tag{16}$$

To meet the Fisher criterion function

$$J(\xi) = \frac{\xi^T S_B' \xi}{\xi^T S_W' \xi} \tag{17}$$

The vector ξ is obtained by applying the eigen-decomposition to the matrix $S_W'^{-1} S_B'$, if S_W' is non-singular. However, our aim is to accurately compute the eigenvectors of S_W' and S_B' containing the most energy of the images. Suppose that the eigenvectors $\{\alpha_i, i = 1,2,\cdots, p\}$ corresponding to the largest eigenvalues of S_W', then the projection direction ξ can be defined as:

$$\xi = Q\alpha_i, \qquad Q = [q_1, q_2, \cdots, q_k] \tag{18}$$

where $[q_1, q_2, \cdots, q_k]$ are the eigenvectors corresponding to the largest eigenvalues of S_B'. Then formula (19) can be expressed as

$$J(\xi) = \frac{\alpha_i^T (Q^T S_B' Q)\alpha_i}{\alpha_i^T (Q^T S_W' Q)\alpha_i} \tag{19}$$

According to following formula，the new criterion function is defined as

$$J(\alpha_i) = \frac{\alpha_i^T S_B'^Q \alpha_i}{\alpha_i^T S_W'^Q \alpha_i} \tag{20}$$

where $S_B'^Q = Q^T S_B' Q$ and $S_W'^Q = Q^T S_W' Q$. It is easy to show that $S_B'^Q$ is a diagonal matrix whereas $S_W'^Q$ is not diagonalised. 而 $\{\alpha_i, i = 1,2,\cdots, p\}$ are the eigenvectors corresponding to the p largest eigenvalues.

Therefore, x_{ij} which is the jth training sample in classes i satisfies the following linear relationship.

$$I_{ij} = x_{ij} Q\Omega \tag{21}$$

where $\Omega = [\alpha_1, \alpha_2, \cdots, \alpha_p]$.

Because the traditional 2DLDA method produces white noise obtaining the optimal projected vector, it leads to high computational complexity and low recognition rate. The I2DLDA method can solve this problem.

C. GABOR+I2DLDA Technique

In palmprint recognition low feature dimension and low computational complexity is a key link. The Gabor transformed image is two-dimensional matrix feature which feature dimension is high, so the transformed image using Gabor need reduce feature dimension and eliminate the relevance of the column. The proposed Gabor+I2DLDA method just is able to meet the

requirement, whose between-class scatter matrix and within-class scatter matrix are defined as

$$S_B'^Q = \frac{1}{NPQ}\sum_{i=1}^{L} N_i PQ(\bar{F}_i - \bar{F})(\bar{F}_i - \bar{F})^T \quad (22)$$

$$S_W'^Q = \frac{1}{NPQ}\sum_{i,j,\mu,\nu} (F_{\mu,\nu}^{ij} - \bar{F}_i)(F_{\mu,\nu}^{ij} - \bar{F}_i)^T \quad (23)$$

where $\bar{F}_i = \frac{1}{N_i PQ}\sum_{j,\mu,\nu} F_{\mu,\nu}^{ij}$ is the mean of the training samples in classes i, $\bar{F} = \frac{1}{NPQ}\sum_{i,j,\mu,\nu} F_{\mu,\nu}^{ij}$ is the total mean to the training samples. Because of $S_B'^Q = Q^T S_B' Q$ and $S_W'^Q = Q^T S_W' Q$, the best projection matrix in the lower dimensional subspace can be defined as

$$Y_{\mu,\nu}^i = F_{\mu,\nu}^{ij} Q\Omega \quad (24)$$

where $F_{\mu,\nu}^i$ is $d \times p$ feature matrixes。

D. Results and discussion

All the tests are finished to palmprint recognition in the PolyU-I database and the PolyU-I database. Before beginning to the tests the above proposed location and segmentation method is used. The palm region of interest shows in Fig.8 for a hand with 6 samples.

The first analysis evaluates the number of eigenvectors Ω required to optimize the recognition. We designed a series of experiments for varying number of Ω eigenvectors. Table IV show that Gabor+I2DLDA, I2DLDA and 2DLDA are the optimal recognition accuracy in the Poly-I database. The highest recognition rate of 2DLDA is 93.33 when number of Ω eigenvectors equal to 11. The highest recognition rate of I2DLDA is 97.00 when number of Ω eigenvectors equal to 7. The highest recognition rate of Gabor+I2DLDA is 99.00 when number of Ω eigenvectors equal to 11. It is clear that number of different eigenvectors corresponding to the recognition rate is different and Gabor+I2DLDA recognition rate is higher than other methods.

TABLE IV. COMPARISON OF THE RECOGNITION RATE (%) FOR VARYING Ω 'S EIGENVECTORS

Method	Number of Eigenvectors for Ω				
	5	***7***	***9***	***11***	***13***
2DLDA	90.33	90.33	92.00	93.33	92.00
I2DLDA	95.00	97.00	96.33	96.33	96.33
GABOR+ I2DLDA	97.00	97.00	98.67	99.00	98.67

Palmprint recognition is a small sample problem. Therefore, we design a series of tests to compare the performance of varying method under conditions where the sample size is varied. Here, we perform five tests with a varying number of training samples. Table V presents the best recognition rate of three different methods in PolyU-I database. Table VI shows the best recognition rate of the different algorithms in the PolyU-II database of 2000 image samples. In the PolyU-I database Gabor+I2DLDA average recognition accuracy is 96.60%, while I2DLDA and 2DLDA average recognition accuracy is 94.27% and 91.13%, respectively. In the PolyU-II database Gabor+I2DLDA average recognition accuracy is 96.52%, while I2DLDA and 2DLDA average recognition accuracy is 94.08% and 90.60%, respectively.

Therefore, it can be concluded that the number of training samples is more, the recognition rate is higher. the proposed method is the highest recognition rate for different number of training samples.

TABLE V. COMPARISON OF THE BEST RECOGNITION RATE (%) TO DIFFERENT TRAINING SAMPLES IN POLYU-I DATABASE

Method	Training samples/class					
	1	***2***	***3***	***4***	***5***	***average***
2DLDA	86.67	88.00	90.33	93.33	97.33	91.13
I2DLDA	90.00	91.67	95.00	96.33	98.33	94.27
GABOR+ I2DLDA	91.67	96.33	97.00	98.67	99.33	96.60

TABLE VI. COMPARISON OF THE BEST RECOGNITION RATE (%) TO DIFFERENT TRAINING SAMPLES IN POLYU-II DATABASE

Method	Training samples/class					
	1	***2***	***3***	***4***	***5***	***average***
2DLDA	85.00	87.25	90.65	93.10	97.00	90.60
I2DLDA	90.00	90.95	95.20	96.00	98.25	94.08
GABOR+ I2DLDA	91.25	96.00	97.35	99.00	99.00	96.52

IV CONCLUSIONS

The Palmprint image processing about segmentation and location is to extract from the palmprint appropriate reference point, establish reference coordinate system, and acquire ROI of palmprint. The result of segmentation and location affects the accuracy of ROI. Only the precise extraction of ROI can ensure the accuracy and effectiveness to identification system, which has very important significance in palmprint identification system. Through a careful analysis, it is found that the main reason of fault location palmprint images is not enough palm stretch, fingers separated and incorrect hand placement. If the location and posture of the hand are further regulated, we can further reduce the location time and increase accuracy of location. Therefore, a simple and effective location and segmentation method is proposed.

To aim at the shortcomings of 2DLDA subspace linear discriminant analysis in palmprint recognition, the paper proposed Gabor+I2DLDA linear discriminant analysis method which can eliminate the image columns correlation to extract the optimal discriminant information. Experimental results have shown the

superiority of the method proposed palmprint processing and recognition in terms of accuracy and speed.

REFERENCES

[1] D.Zhang, A.Kong, J.You, M.Wong, "Online Palmprint identification", IEEE Trans, Pattern Anal.Mach.intell, vol.25, pp. 1041-1050, 2003.

[2] Q. Xiao, "Technology review-biometrics- technology, application, challenge, and computational intelligence solutions", Computational Intelligence Magazine, vol. 2, no.2, pp. 5-25, 2007.

[3] Wu X, Wang K, Zhang D, "Wavelet Energy Feature Extraction and Matching for Palmprint Recognition", Journal of Computer Science and Technology, vol.20, no.5, pp. 411-418, 2005.

[4] SATOSHI I, KOICHI I, TAKAFUMI A., "A practical palmprint recognition algorithm using phase information", Thampa, FL: 19th International Conference on Pattern Recognition, pp.1-4, 2008.

[5] Zhang D, Lu G, Li W, "Palmprint recognition using 3-D information", IEEE Trans on Systems, vol.39, no.5, pp. 505-519, 2009.

[6] Yue F, Zuo W, Zhang D, "Orientation selection using modified FCM for competitive code-based palmprint recognition", Pattern Recognition, vol.42, no.3, pp. 2841-2849, 2009.

[7] Guo Z, Zhang D, Zhang L, "Palmprint verification using binary orientation co-occurrence vector", Pattern Recognition Letters, vol.30, no.5, pp. 1219-1227, 2009.

[8] Zuo W, Zhang H, Zhang D, "Post-processed LDA for face and palmprint recognition-What is the rationale", Signal Processing, vol.90, no.8, pp. 2344-2352, 2010.

[9] Shang P, Li T, "Multifractal characteristics of palmprint and its extracted algorithm", Applied Mathematical Modeling, vol.33, no.3, pp. 4378-4387, 2009.

[10] LIAMBAS C, TSOUROS C, "An algorithm for detecting hand orientation and palmprint location from a highly noisy image", Alcala DeHenares, Spain: Proceedings of IEEE International Symposium on Intelligent Signal Processing, pp.1-6, 2007.

[11] Tian Q, LI Z, Zhu Y, "A Novel Palmprint Segmentation and Recognition Algorithm", Cambridge UK: International Conference on Intelligent Computation Technology and Automation, pp 273-276, 2010.

[12] CHOGE K, OYAMA T, KARUNGARU S, "A circle-based Region-Of-Interest segmentation method for palmprint recognition", Fukuoka, Japan: ICROS-SICE International Joint Conference, pp. 4993 -4997, 2009.

[13] LI Yan, WU Guifang, DAI Gaole, "A New Algorithm to Extract Contour Feature Points of Palmprint", Microelectronics & Computer, vol.27, no.5, pp. 90-94, 2010.

[14] G. Lu, D. Zhang and K.Wang, "Palmprint recognition using eigenpalms features," Pattern Recognit., vol.24, no.10, pp. 1463–1467, 2003.

[15] X. Wu, D. Zhang and K. Wang, "Fisherpalms based palmprint recognition," Pattern Recognit, vol.24, no.15, pp. 2829–2838,2003.

[16] D. Zhang, W. K. Wong, J. You, and M. Wong, "Online palmprint identification," IEEE Trans. Pattern Anal. Mach. Intell, vol.25, no.9, pp. 1041–1050,2003.

[17] X. Wu, D. Zhang, K. Wang, and B. Huang, "Palmprint classification using principal lines," Pattern Recognit., vol.37, no.10, pp. 1987–1998,2004.

[18] C. Liu and H. Wechsler, "Gabor feature based classification using the enhancedfisher linear discriminant model for face recognition," IEEE Transaction on Image Processing, vol.11, no.4. pp. 467–476, 2002.

[19] C. J. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition," IEEE Transaction on Neural Networks, vol.14, no. 4, pp. 658–669, 2003.

[20] C. Liu, "Gabor-based kernel PCA with fractional power polynomial models for face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 5, pp. 572–581,2004.

[21] Y. Pang, Y. Yuan and X. Li, "Gabor-base region covariance matrices for face recognition," IEEE Trans. Circuit Syst. Video Technol., vol. 18, no. 7, pp.989–993, 2008.

[22] X. Yang, Y. Zhou, T. Zhang, and J. Yang, "Gabor phase based gait recognition," Eletron. Lett., vol. 44, no. 10, pp. 620–621, 2008.

[23] MING L, YUAN B. 2D-LDA: a statistical linear discriminant analysis for image matrix[J]. Pattern Recognition Letter, 2005,26(5):527-532.

[24] Zhang D. The PolyU palmprint database[DB/OL]. http://www.comp.polyu.edu.hk/biometrics/, 2009.

[25] D. J. Field, "Relation between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am., pp. 2379–2394, 1987.

Shuang Xu received master from Dalian Maritime University, China in 2004. She is currently a lecturer in college of Information & Communication Engineering, Dalian Nationalities University, China. Her main research interests include image processing and pattern recognition.

Jifeng Ding received master from Harbin Institute of Technology, China in 2006. He is currently a lecturer in college of Information & Communication Engineering, Dalian Nationalities University, China. His main research interests include information processing and intelligent sensor technology.

Call for Papers and Special Issue Proposals

Aims and Scope.

Journal of Multimedia (JMM, ISSN 1796-2048) is a scholarly peer-reviewed international scientific journal published bimonthly, focusing on theories, methods, algorithms, and applications in multimedia. It provides a high profile, leading edge forum for academic researchers, industrial professionals, engineers, consultants, managers, educators and policy makers working in the field to contribute and disseminate innovative new work on multimedia.

The Journal of Multimedia covers the breadth of research in multimedia technology and applications. JMM invites original, previously unpublished, research, survey and tutorial papers, plus case studies and short research notes, on both applied and theoretical aspects of multimedia. These areas include, but are not limited to, the following topics:

- Multimedia Signal Processing
- Multimedia Content Understanding
- Multimedia Interface and Interaction
- Multimedia Databases and File Systems
- Multimedia Communication and Networking
- Multimedia Systems and Devices
- Multimedia Applications

JMM EDICS (Editors Information Classification Scheme) can be found at http://www.academypublisher.com/jmm/jmmedics.html.

Special Issue Guidelines

Special issues feature specifically aimed and targeted topics of interest contributed by authors responding to a particular Call for Papers or by invitation, edited by guest editor(s). We encourage you to submit proposals for creating special issues in areas that are of interest to the Journal. Preference will be given to proposals that cover some unique aspect of the technology and ones that include subjects that are timely and useful to the readers of the Journal. A Special Issue is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

The following information should be included as part of the proposal:

- Proposed title for the Special Issue
- Description of the topic area to be focused upon and justification
- Review process for the selection and rejection of papers.
- Name, contact, position, affiliation, and biography of the Guest Editor(s)
- List of potential reviewers
- Potential authors to the issue
- Tentative time-table for the call for papers and reviews

If a proposal is accepted, the guest editor will be responsible for:

- Preparing the “Call for Papers” to be included on the Journal’s Web site.
- Distribution of the Call for Papers broadly to various mailing lists and sites.
- Getting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Instructions for Authors.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

Special Issue for a Conference/Workshop

A special issue for a Conference/Workshop is usually released in association with the committee members of the Conference/Workshop like general chairs and/or program chairs who are appointed as the Guest Editors of the Special Issue. Special Issue for a Conference/Workshop is typically made of 10 to 15 papers, with each paper 8 to 12 pages of length.

Guest Editors are involved in the following steps in guest-editing a Special Issue based on a Conference/Workshop:

- Selecting a Title for the Special Issue, e.g. “Special Issue: Selected Best Papers of XYZ Conference”.
- Sending us a formal “Letter of Intent” for the Special Issue.
- Creating a “Call for Papers” for the Special Issue, posting it on the conference web site, and publicizing it to the conference attendees. Information about the Journal and Academy Publisher can be included in the Call for Papers.
- Establishing criteria for paper selection/rejections. The papers can be nominated based on multiple criteria, e.g. rank in review process plus the evaluation from the Session Chairs and the feedback from the Conference attendees.
- Selecting and inviting submissions, arranging review process, making decisions, and carrying out all correspondence with the authors. Authors should be informed the Author Instructions. Usually, the Proceedings manuscripts should be expanded and enhanced.
- Providing us the completed and approved final versions of the papers formatted in the Journal’s style, together with all authors’ contact information.
- Writing a one- or two-page introductory editorial to be published in the Special Issue.

More information is available on the web site at http://www.academypublisher.com/jmm/.