Progress Report

Chun F.Hsu R94922082

November 23, 2006

Abstract

NA now.

Contents

Chapter 1

Progress in individual weeks

SECTION 1.1

7-13-2006

1.1.1 Progress done

- No Paper survey.
- Found a H.264 reference software, but can't encode and decode most of the files, still trying.
- Finish reading the book, MPEG-4 and H.264, and understood some basic ideas of these two codecs.

1.1.2 Progress this week

- Study "Significance Test" which helps decide which video data features are useful.
- Study Digital Image Processing, and focus on Color theory. I will try to know the correct implementation methods of Color Histogram as soon as possible.
- Implement Color Histogram on C++.

1.1.3 Ideas and Criticism

- There is a basic clustering method in Data mining called: "K-means" method, it can do data clustering, or unsupervised learning. This is simple and easy, maybe if some idea of applications come out, this method will help.

SECTION 1.2

7-20-2006

1.2.1 Progress done

- Examine some past read papers and figure out how to represent Color Histogram in the form of Vector.
- Study Color Theory of "Digital Image Processing"and gets basic knowledge in this area.

1.2.2 Progress this week

- Continue study Digital Image Processing, while begins writing color histogram software implementation.
- Read more paper on detail of turning Color Histogram into single vector.

1.2.3 Ideas and Criticism

- I am confused these weeks since I found that Color Histogram isn't a vector, it is simply the probability distribution of color pixels. So I traced the paper I present before, and keep backtracking to its reference paper, and find a method that meet the paper's requirement at last. Their method it to extracted the Color Histogram's DCT's DC value, then doing wavelet transform there are still some questions in this implementation, and I am trying to figure out soon.

}

SECTION 1.3

8-03-2006

1.3.1 Progress done

- Finish YUV file reading tool.
- Study Wavelet transform and get basic knowledge.
- Study Digital Signal Processing, now finished reading the first two chapters.

1.3.2 Progress this week

- Finish writing Color Histogram program, try to merge it with YUV reader.
- Read more paper on detail of turning Color Histogram into single vector.
- Read more paper on detail of implementing Color Histogram interchapter.

1.3.3 Ideas and Criticism

- My research direction will follow these two papers basically[3] [4], and then do demo according to these two papers. Maybe the whole architecture will come out during this implementation.

SECTION 1.4

8-24-2006

1.4.1 Progress done

- Working at new arrived video SK demo, plan to demo their two demo examples, one is **"MPEG-4 Parsing Decoder"**and another is **"RGB Camera"**, I expect to finish these demos this Friday.
- Learn software "Endnote", which is a friendly to user software to manage papers and documents, a snapshot of this software is taken.
- Continue design on the demo of shot detection framework, I've implemented on **Matlab** platform and it's quite easy to implement this kind of system on Matlab, which is more easier to extract images matrix from YUV file, now I am at the stage of implementing Wavelet multi-resolution transform.

1.4.2 Progress this week

- Finish wavelet transform part and collect datas to throw into SVM.
- Organize the papers I read by **Endnote**.
- Read paper, *Temporal Video Segmentation, a survey* and write a report.

SECTION 1.5

8-31-2006

1.5.1 Progress done

- I've tried two demos on Video SK, the first one is MPEG-4 player which work well, another demo needs an A-D converter according to Prof.Yang's lab's experience.
- Continue organize my read papers on Endnote, now it has a list of what I've survey this year.
- Continue design the demo of shot detection framework, now I've tried on C++ version, I hope to end it before this month.

1.5.2 Progress this week

- Collect datas to throw into SVM.
- Organize the papers I read by **Endnote**.
- Read paper, *Temporal Video Segmentation, a survey* and write a report.

SECTION 1.6

9-06-2006

1.6.1 Progress done

- Working on shot detection on video temporal segmentation, this week I tried to transform **.YUV file** into several distinct frames by turning them into **.BMP file**.
- Survey paper, "*Video Shot Detection and Condensed Representation, a Review*"[5]. This paper has made a collection of recent six years topics about video shot detection and it is really worthful reading. I've attached along with this week's report.
- Continuing design the demo of shot detection framework, now I've tried on C++ version, I hope to end it before this month.

1.6.2 Progress this week

- Collect datas to throw into SVM.
- Organize the papers I read by **Endnote**.
- Read paper, *Temporal Video Segmentation, a survey* and write a report.

SECTION 1.7

9-13-2006

1.7.1 Progress done

- Now we can tried transform **.YUV file** into several distinct frames by turning them into **.BMP file**. This part is almost done except a GUI interface enable us to see the frames individually.
- Survey Xilinx board for CSIE department's board shopping plan.
- Continuing design the demo of shot detection framework, now I've tried on C++ version, I hope to end it before this month.

1.7.2 Progress this week

- Collect datas to throw into SVM.
- Organize more papers I read by **Endnote**.
- Read paper, *Video Shot Detection and Condensed Representation, a Review* and write a report.

SECTION 1.8

9-21-2006

1.8.1 Progress done

- I was at the Color Histogram stage of the shot detection this week, which I finished the transforming of YUV to RGB and then to BMP work, it is a really dirty work and took such a long time to implement it, however, Color Histogram seems working, and I am ready to go on the next stage.
- Survey paper, "*Improving Color Based Video Shot Detection*"[5], which they proposed a new measure to improve color-based method of shot detection. It can be applied to different areas of application, such as Video Browsing, Indexing, and to form the basis of semantic content analysis.

1.8.2 Progress this week

- Collect datas to throw into SVM.
- Organize more papers I read by **Endnote**.

1.8.3 Ideas and Criticism

Shot detection can be an interesting start point to all other applications, since it does the foundamental *Intelligent Video Processing*such as the segmentation of source video and video classification. Many approaches has been proposed, some of them are in compression domain and other's are not. They also don't rely on a single classifier such as SVM, which there might be more efficient and "enough" algorighm to do this woork. Indeed, it is a hard work for me alone. I really expect for having more lab members to work and discuss together on this huge project and gets the best efficancy.

SECTION 1.9

9-28-2006

1.9.1 Progress done

- This week I've finished the simplest color-histogram shot detector and test two video, and consequently work well. Nevertheless, restricted to its foundamental behaviour, some problems emerge during the parameter selection stage. Especially on the " **Threshold**" used to determine the boundary of frames, is fully "manual" work which has to define ourselve. Furthermore, it requires a better interface to handle I/O work and shows the frames which has chosen to be the shot boundary frame. However , we can move on improving software predicting accuracy.
- Survey paper, "*Improving Color Based Video Shot Detection*"[5]. I've made a brief article analyzizing and critizizing it.

1.9.2 Progress this week

- Do enhancement on the propotype software by adding wavelem transform and block color histogram.

SECTION 1.10

10-05-2006

1.10.1 Progress due this day

- Besides working on the GUI interface design and algorithm enhancement on Color Histogram, there isn't significant progress this week to show. However, I've turned the 256 color histogram into the 64 color histogram; this enhancement can save much space and memory. The original color histogram only focus on Y frame, which the present histogram can handle 3 dimensions color space R, G, and B. It also saves memory by only extract the prefix 2 bits, instead of 8. Although there are three features of colors, it's memory requirement for each frame is reduced from 2^{24} bits to 2^6 bits, without lossing much prediction performance as compare to the formal one.
- I've tried to design a window frame to show each picture frames. It is a tough task, and I want to have it done before next Monday.

1.10.2 Progress this week

- Do enhancement on the propotype software by adding wavelet transform, more specific speaking, the Canny wavelet.

1.10.3 Ideas and Criticism

SECTION 1.11

10-12-2006

1.11.1 Progress till this day

- Working on 64histogram implementation, verified the RGB picture correctness, and found a huge size of data sets of metadata.
- Coordinate board and organized the TA job this week.
- Make a plan and a sketch of system I will build.

1.11.2 Progress this week

- Finish the enhancement on the propotype software by adding wavelet transform, more specific speaking, the Canny wavelet.
- Survey paper relevant to semantic sport video analyzing and classification.

1.11.3 Ideas and Criticism

A plan emerged im my mind this week, despite my shot detection implementation, I came up an idea of a interesing topic, semantic sport indexing and browsing, with connection to internet service server. Why I pull internet in is while I collect datas from the real-time processing video, maybe, in a small handheld device, it will connect through the Internet and access the server database to do classification or indexing browsing work; while the browsing work rely on the existed software or mining program, a software-hardware codesign will be made into reality. Therefore, we can divide this work into three parts, shot detection and information extraction–the hardware's duty, communication and error preventing and retrieving system, and the final, server of video database to recognize the feedback data. I believe this framework will be very interesting and contributes much.

SECTION 1.12

10-19-2006

1.12.1 Research Done This Week

- I've optimized my design in Matlab, which I continued enhancing the color correctness, and it shows pretty well . The four shot boundaries are as follow:
- I've checked the **TRECVID** website trying to find some authorized video clips, and couldn't get these files. The reason is these video's length is over 135 hours, it must be shipped by the TREC organization. However, these data sets are for competitions and unfortunately this year's requiring deadline has expired. I'll find something else, and the worse case might be some sport video which I own, or some news video download from Internet.
- A paper suggested [5]that the most fitable color space may be the YIQ color space, so I've made an alternate functions that can deal with this color space. The paper suggested that we only extract I and Q, and the reason is they think it's the color difference that matters, not the intensity. As many as these heuritic approach, which I considered that all of them work certainly, but some careful examination must made to the video source. So I will included as much as possible while I am looking for suitable video testing data set.
- My system prototype is still rough, however, I switch efficiently in C and Matlab. While producing color histogram and evaluating algorithm at one side, I produced chars and fancy GUI at another by using Matlab.

1.12.2 My Progress next week

- Search news or sports video, and easily embedded video decoder to put it in my design(in Matlab).
- Study Wavelet Transform and implement it.

- Read more paper on Shot detection in *temporal* and *spatial* characteritics. The temporal features will be time and distance between two frames, while the spatial features will be histogram difference and the derivative of difference, in color or intensity.

SECTION 1.13

10-26-2006

1.13.1 Research done this week

- I've read two papers, one is about designing histogram interchapter used at the measurement of spatial distance of two frames using its color histogram, it includes proofs and statistic meaning wrote in detail. Another is a application paper by IBM which is about video summarization in pervasive device. The latter which doesn't mention its algorithm a lot, instead they talk about the framework and relations between blocks and blocks. I've regard this paper as a very intersting one.
- I am still working on the interface designing work. Unfortunately, no big or significant progress this week in this part.

1.13.2 Progress next week

- Construct an user-interface which can play a video, analysize a video, and display the hiearachy of a video.
- I've to work on *Video Signal Processing* final project this week, as the result, the programming progress may delay one or two weeks.

SECTION 1.14

11-2-2006

1.14.1 Research done this week

- I've configure my software so it can accept two informations from color histogram, one of them is histogram interchapter and another is histogram difference. Combining these two we can get a better result than using only one of them. I've used a yuv data to test it, and I've adjusted the threshold to see its effect. The result snapshot:
  http://www.csie.ntu.edu.tw/~r94082/progress/test_1500.png
  The result frames package:
  http://www.csie.ntu.edu.tw/~r94082/progress/test_1500.rar
  There are some problems, one of them is when using these two algorighms, inevitably they will meet the same problem when a dark item appear and cover the white background. This needs to be solved.

subsectionProgress next week

- Tune my system and add another algorithm helping adjusting the result.
- Require the test data from CMLab, as far as I know, they have some good test data.

SECTION 1.15

11-09-2006

1.15.1 Research done this week

- I add an additional test to my design, which is the "Run test", this test is used to detect basic transition shot. It's basic design is as follow:

  1. Set another threshold
  2. If threshold expired, add one to run
  3. If run expired the run threshold(Set at the beginning). Record it as transition shot.
  4. Combine the former histogram comparison and output the shot frames.

- Chin-Wen and I made a short clip in our lab for testing, the file is :
  test file
  There are also shot detection experiment on this, which is:
  Shot collection file

- I had asked CMLab about the test data they own, and CYY ask me to have a short chat Friday afternoon, able to get data or not depends on this meeting.

1.15.2 Progress next week

- Begin to add motion vector and DCT values belongs to H.264 transcoding from the film I made.
- Enhance my code by adding motion vector adjustment.
- Survey surveillance video paper especially on pedestrian identification and extraction.

SECTION 1.16

11-16-2006

1.16.1 Progress self-check this week

- Begin to add motion vector and DCT values belongs to H.264 transcoding from the film I made.
  Didn't finish quite well, one reason is I want to try to operate directly on the data from CMlab, In order to avoid redundant work, I decided to do this part after acquire their data next Monday.
- Enhance my code by adding motion vector adjustment.
  Not finish yet. I need to understand what does it like and how do they present in H.264 file. This part need two weeks to achieve it.

- Survey surveillance video paper especially on pedestrian identification and extraction.
  I had collect several papers. The one I am reading now is:[7]. Below is my comment:

  1. All surveillance video research has involved in background subtraction, that is, to identify or classify objects and background. Some use statistical methods, some use classifier. One paper uses SVR(Support Vector Regression), a closed relative to SVM to adjust regions that doesn't belong to background, and then track that object. The object is not limit to human, cars or motorcycle or any moving objects are included, too. Problem: It needs cameras to be fixed, or the background model(scene model) will be confused. It also needs to be adaptively to light condition, when operating in long hours, lighting will change and it will still leave a hard problem to keep scene model constantly. Problem II: Most researcher agree that in the detection stage, the performance depends on fast and effective updating work.
  2. Detection can be done after model has set up. They think we shouldn't treat detection and tracking sequentially and independently, but simutaneously. In their approach, detection enable tracking, and tracking feedback detection to adaptively optimize the threshold and improve system robustness.
  3. After detection and tracking, to localize the boundrary box of object, they uses simple approach to give high confidence score on object tracking, which they used projection histogram. Problem: Why don't we use color histogram instead of projection histogram, are they different? According to the first color histogram author's research, when object rotates or move, in the very same region the color histogram stays the same. Answer: Their reason is using projection histogram, the information of object's position and shape would take into consider into the threshold optimization stage, which will iteratively enhance tracking and detection on the same object. Then it will suppress the false alarm setup by noise.

- Overall evaluation: Not well. No significant effort due to new topic surveying.

1.16.2 Progress plan next week

- Talk with MK and finalize my master thesis topic first. Then prepare all the information and be able to show a proposal next Friday.
- Enable to extract motion vector and DCT values from H.264.
- Survey single camera event detection methods.

SECTION 1.17

11-23-2006

- Progress done this week

- After some survey paper work this week, I have summarized the procedure as follow:

  The procedure for event detection in surveillance video

  1. A foreground and background separating mechanism: To train a model containing only of background, excludes all the foreground object.
  2. Building individual object model, and use these models to determine and detect when intrusion event occurs. Then, mark the frame number and begins to record.
  3. After the detect object leaves the scene, record all the actions and events it cause in its journey, such as motion trajectory, simple gestures, etc.

- In first item above, to eliminating the foreground at a surveillance video to get a clean background, a technique call median filter can be applied. Below is a demonstration,
  A paper has suggest that if we change the input from spatial to temporal

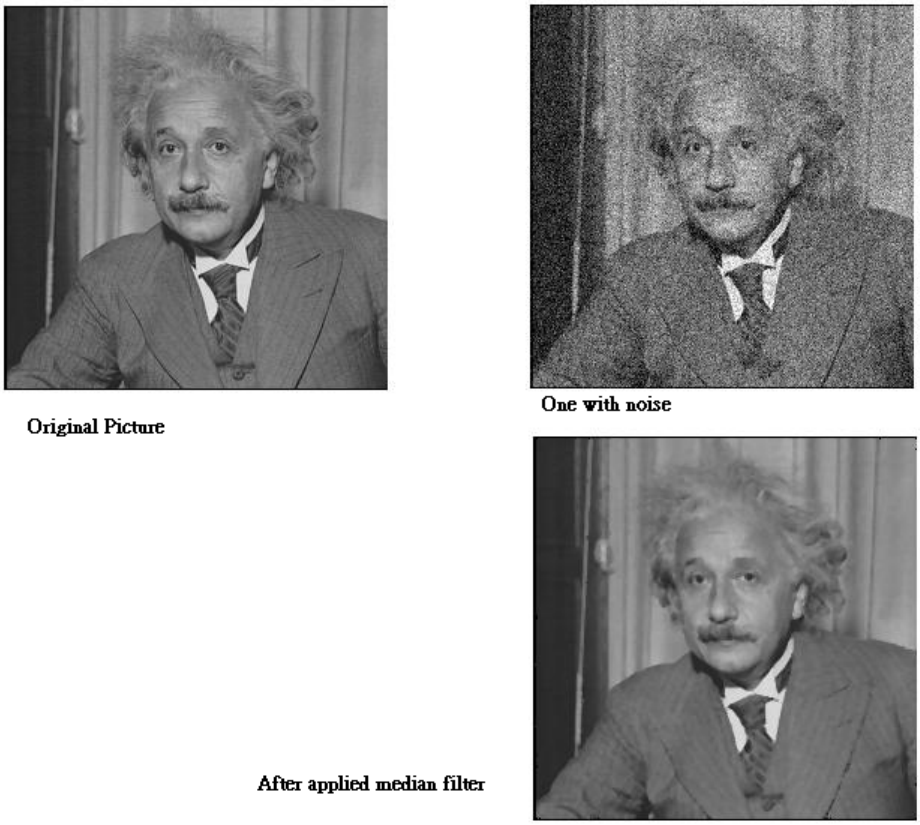

Figure 1.1: Median filter easy demo

timeline, there will emerge a simple but powerful scheme to do background

elimination. The method they proposed is, if we collect a great amount of frames, e.g, 90 frames, that is, three seconds of video. Then apply median filter to the individual pixels among the 90 frames and reconstruct the video background scene withous blurring the shape and edges.

- Because there are no sufficient video for me to do median filtering and see if we can reconstruct the background, so I only uses pictures to do a simple demo. I will try to borrow a DV to record one or two hours to be our test data. If possible, will lab have our own DV so we don't need to borrow from other lab?

1.17.1 Progress to be done next week

- Keep working through the rest part of my procedures above.
- Record some video using a fixed-position video.
- Applying the scene model using SVR and detect object.
- Read three more papers talk about tracking method.

Chapter 2

Progress Schedule

date	Progress Schedule	Remark
5/15–5/31	Learn hot to use libsvm	Finish
6/01–6/30	Implementing system of my chosen paper	In progress
7/01–7/08	Planning summer research and system implementation.	**Done**
7/15–7/30	Implementing Color Histogram on single vector.	Done in Matlab and C
9/01–9/15	Implementing YUV file to BMP file	**Done**
11/16–11/30	Develop an object extractor of H.264	Not yet
11/16–11/23	Survey single camera event detection methods	In progress

Chapter 3

November plan

- Read at least 20 papers in 3d-trajectory
- Survey H.264 and be able to :
  1. Extract DCT DC
  2. Extract motion vector
  3. Extract motion predict vector?
- Implement a 3D-object-tracking with a boundary box.

Bibliography

[1] Jin Huang Charles X.Ling, "Using AUC and Accuracy in Evaluating Learning Algorithms," *IEEE TRANSACTION ON KNOWLEDGE AND DATA ENGINEERING, VOL 17, NO.3,* 2005

[2] Vasileios Mezaris, Ioannis Komparsiaris, and Michael G.Strintzis, "A Knowledge-Based Approach to Domain-Specific Compressed Video Analysis,"*IEEE international Conference on Image-Processing*,2004

[3] HuaMin Feng, A Chandrashekhara, T.S. Chua. An Automatic Temporal Multi-resolution Analysis Framework for Shot Boundary Detection. MMM 2003, the 9th International Conference on Multi-Media Modeling, 2003, 224

[4] Huamin Feng, Wei Fang, Sen Liu, Yong Fang. A New General Framework for Shot Boundary Detection and Key-Frame Extraction,MIR05, November 10-11, 2005, Singapore.

[5] Cotsaces, C., N. Nikolaidis, et al. (2006). "Video shot detection and condensed representation. a review." Signal Processing Magazine, IEEE 23(2): 28-37.

[6] W.Zhao, J. Wang, D. Bhat, K. Sakiewicz, N. Nandhakumar, W. Chang. "Improving Color Based Video Shot Detection", IEEE, 1999.

[7] Junxian, W., G. Bebis, et al. (2006). Robust Video-Based Surveillance by Integrating Target Detection with Tracking. Computer Vision and Pattern Recognition Workshop, 2006 Conference on.