INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

Shot Classification in News Video Based on Color Histogram

Sindhu J.
PG Student, Department of ECE, BGSIT, BG Nagar, Mandya, Karnataka, India
Punith Kumar M. B.
Assistant Professor, Department of ECE, BGSIT, BG Nagar, Mandya, Karnataka, India
M. B. Anadraju
Professor & HOD, Department of ECE, BGSIT, BG Nagar, Mandya, Karnataka, India

Abstract:
This paper presents an approach for classification of shots in the tv news video. Two types of shots are classified here namely anchor shot and report shot. The color histogram is one of the techniques that have been adapted to classify shots. This paper also present an another approach for face detection in news video. The goal of face detection is to locate all regions that contain a face. This project has adopted an algorithm for detecting human faces in videos. Based on RGB function detects the face with a template image. After determines the face, classifies that face as first anchor shot. The experimental results of both approaches will be implemented using MATLAB programming language.

Keywords*: face detection; color histogram; template matching*

1. Introduction

Face detection plays an important role in many applications such as face recognition, video surveillance, facial expression recognition, face tracking, facial feature extraction, gender classification, identification system, document control and access control, clustering, biometric science, human computer interaction (HCI) system and many more. Recently many researches detect face by various methods. First we read video and extract the frames in video and audio particular frame. News video data is a valuable and rich source of information especially on politic and local news. Each video contains one or more stories. One story is represented by one or more shots. The shot is in turn represented by a series of similar frames. This is shown in below figure1.

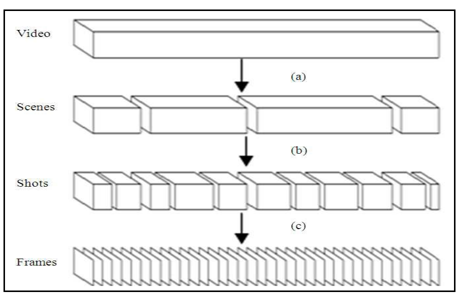

Figure 1: (a) Video into scenes (b) Scenes into shots (c) Shots into Frames

The basic objective of this paper is to use the color histogram features to detect anchor shots in news video. The color histogram for the frames will be computed in RGB color space. This project provides the RGB histogram mean algorithm which used to group similar frames, to build shots. Computing the mean of the RGB histogram for the frames is one of the simplest ways of applying the color histogram features.

In this paper, we propose a color histogram method for segmenting the news videos. This method first uses to read video and extract frames (video and particular frame) at last determines the face and classifies that face as a first anchor shot. After finding face this method segments the news videos into initial stories. Then, these stories are refined by further detection of special types of news stories and the merging of similar stories. The experimental result of implementation the RGB histogram mean has reported a result with the style of news where the anchor person and the background of the anchor shots are motionless. This paper will be implemented using MATLAB programming language.

The rest of the paper is organized as follows: Section 2 describes the proposed work in detail; Section 3 demonstrates our system results; and, Section 4 presents the conclusion and the discussions.

2. Proposed Work

2.1. Face Detection

Face detection is the essential front end of any face recognition system, which locates face regions from video or still image. Various methods have been proposed for face detection. The algorithm uses skin color segmentation and human face features (knowledge-based approach) to segment skin regions from an image to find out a faces. In this paper, uses only two color spaces namely, RGB, and HSV color models are used for skin color segmentation. These color models with thresholds, help to remove non skin like pixels from an image. Using the RGB function we find out the faces in news video.

The RGB color space consists of the three color components: red, green and blue. HSV color model has one major problem associated with RGB (Red, Green, and Blue) color space is that, it does not consider the luminance effect on skin color, which may cause some incorrect information. HSV provides color information as Hue (or color-depth), Saturation (or color-purity) and intensity of the Value (or color-brightness).The skin color pixel's H and S components should satisfy the following conditions for detecting face.

$0 <= H <= 0.25; 0.15 <= S <= 0.9$

Our face detection system consists of three steps. The first step is to classify each pixel in the given image as a skin pixel or a non-skin pixel. The second step is to identify different skin regions in the skin detected image by using connectivity analysis. The last step is to decide whether each of the skin regions identified is a face or not. This is done using two parameters. They are the height to width ratio of the skin region and the percentage of skin in the rectangle defined by the height and width. Figure below describes the block diagram of the face detection.

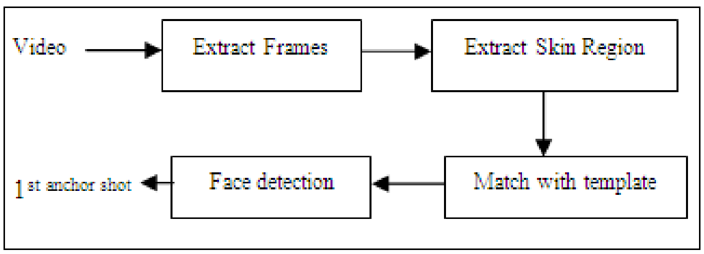

Figure 2: The Face detection system

2.2. Shot Boundary Detection

Histograms are the most common method used to detect shot boundaries. Color histogram is used to detect anchor shots in news videos. As we discussed in the introduction, face information is effective for anchor shot classification. Shot boundary detection is often the first step in algorithms that achieve other video analysis tasks, for example condensed video representation. This is because it is the most basic temporal video segmentation task, and fundamentally relates to the video production. Therefore, it remains an obvious choice for segmenting the video into digestible parts. Moreover, the idea of shot boundary detection has been present since the first days of motion pictures as those who worked in the industry, perceptually segmented their work into a hierarchy of partitions. Designing an appropriate shot boundary technique is complicated by the fact that there are different categories of shot changes. Most models are based on abrupt shot changes, however there are also cases of dissolve fade and wipe.

Shot boundary detection method based on the differences of color histograms between frames has been presented. Usually, in the digital video, frames are represented in RGB color space, 24 bits/pixel images (8 bits for Red, Green, Blue colors). The shot boundary detection methods used in are impressive, however in order to overcome the problems of discard false alarms and missed shots, improved methods are necessary.

The algorithm for above proposed work

- Step 1: Read video and extract frames (particular video and audio frames).
- Step 2: Detect and extract the skin region.
- Step 3: For getting the accurate result detect mouth and eye region.
- Step 4: Determines the face and classifies that as 1st anchor shot.
- Step 5: Finding RGB mean for selecting frame (shot) using below equation

$$RGB\ mean = sum\ R + sum\ G + sum\ B / m * n * 3$$

- Step 6: Finding RGB mean for rest of the frames in video.
- Step 7: Differentiate and compare with threshold and find anchor shot array.
- Step 8: Filter out actual anchor frame for anchor shot array based on anchor shot frame rate.

Each frame is stored as an m-by-n-by-3 data array that defines red, green, and blue color components for each individual pixel where *m* is the width of the frame and *n* is the height represented in above equation.

3. Implementation and Results

There are many ways to implement such approach which calculate the RGB histogram mean and classify the anchor shots. Several programming languages could be used to implement the RGB histogram mean algorithm. The RGB histogram mean was implemented using the MATLAB programming language which is easier and fast in process comparing with JAVA. The RGB histogram mean implementation process was divided to three stage using three different functions. In the first stage, the video data was extracted to frames and then RGB histogram mean was calculated for one frame per second. In this algorithm it is not important to detect the boundaries of the shots. Hence, selecting frame each second and calculate the RGB histogram mean for that frame is reduced the processing time. The outputs from this function are the frame number and the RGB histogram mean value for the frame which is a double value in the range [0,255]. The output from the first function will be used in the second one to group similar frame to build the shots.

4. Results

The RGB histogram mean has been implemented and tested using 5 news videos in two languages and from four different television networks. There are outputs for each of the three function of the code. The final output is an array which contains information about the anchor shots. It has been experimented and compared with the anchor shots detected manually for each video source. Figure 3 illustrates the RGB mean differences of selected frames related to two different shots. The first four frames from the left related to an anchor shot. They have similar RGB mean in the range. The following four frames show the beginning of a report shot. They have different RGB mean in the range [90, 93]. It provides a clear change in the RGB mean level from the anchor shot and the report shot.

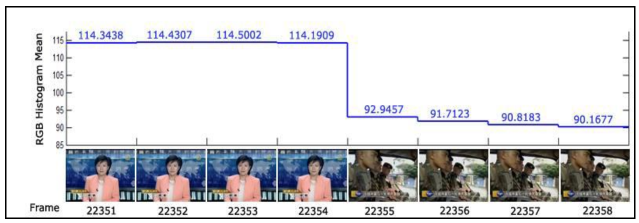

Figure 3: RGB Mean values for sequence frames

The differences in the RGB histogram mean for similar frames are less than the differences of different frames. Assuming that the anchor shot will be no less than three seconds (120 frames per second in the rate 30 fps), the RGB mean is used to group similar frames to produce shots from the video contents. The shot that has duration less than three second (less than 120 frames in rate of 30frame/second) will be ignored. It is important to define variable which will be used to group similar frames. By testing some video data, it is possible to find the values that could be used as a range of the differences in RGB histogram mean. Hence, if the differences between RGB mean is within this range, then the fames will group together to build a shot. The rage values will be defined in the implementation section.

Each extracted shot will have a *key-frame*. The key-frame is a frame that related to the shot and it selected from the middle of the shot content. It will be used to classify similar shots together to find the anchor shots. There are some attributes affect the result of the detection and classification.

The given below table shows classification of shots using the threshold values. The given threshold values are different for same and different channels. The same channels have different anchors so the threshold values are different for different anchors also.

The given table displays the news video data that have been used to test under threshold values

Channel	T1 (Anchor Shot)	T2 (Reporter Shot)
TV9	0	2
NDTV	2	8
TV9	8	20
ETV	0	8
DD1	8	16

Table 1

The given table shows the news video data that have been used to test and evaluate the above algorithm. For different videos given the different frames and match that with the original frames. Using the above algorithm, we getting an equal time length compared with the input video timings.

Channel	Language	Frames	Duration (mm:ss)
TV9	Kannada	15316	10:12
NDTV	English	3000	03:27
TV9	Kannada	2000	04.36
ETV	Kannada	1500	02:42
DD1	Kannada	1500	02:37

Table 2

5. Conclusion

Color histogram features have been used widely related with shot detection. The simplest way of applying the color histogram technique is to calculate the mean value of the RGB histogram for the frames and group similar frames to detect a shot. MATLAB has become the most popular programming environment for the video and image processing. It provides tools and libraries that could be used to extract frame and process the video in less time comparing with other programming languages.

The implementation of the algorithm has been tested using 5 video data with more than 10 hours time in two different languages and from four television networks. Recall and precision are the performance measurement method used to test the performance of the algorithm.

The RGB histogram mean algorithm which has been implemented and test in this paper registered accurate and efficient result with videos where the anchor person and the background of the anchor shot are motionless. However, with the style of the videos where the anchor appearance is always changeable from shot to shot the algorithm registered some limitation.

6. References

1. Ming-Hsuan Yang et.al. "Detecting faces in Images: a survey", IEEE transaction on Pattern analysis and machine intelligene, vol. 24, no.1 2002.
2. Sanjay singh et.al, "A robust skin color based face detection algorithm", Tamkang Journal of Science and Engineering vol.6, no.4, pp227-234, 2003.
3. Ide I, Yamamoto K Tanaka H "Automatic Video Indexing Based on Shot Classification" Springer-Verlag Berlin pp87-102, 1999.
4. J. Mas and G. Fernandez, "Video shot boundary detection using color histogram.", Notebook Papers TRECVID2003, Gaithersburg, Maryland, NIST, October (2003).
5. R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5, 696–706, 2002.
6. R.Gonzalez and R.Woods, "Digital image Processing", second edition, Prentice hall, 2002.
7. Peter Peer, Franc Solina, "An Automatic Human Face Detection Method.
8. H. Zhang, Y. Gong, S. W. Smoliar and S. Y. Tan, Automatic Parsing of News Video, Proc. IEEE Conf. on Multimedia Computing and Systems, 1994