ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 2, Issue 1, January 2013

Gesture Interpretation for Video Shot-Boundary Detection

P. Swati Sowjanya, Ravi Mishra

***Abstract* — Unlike text, videos are difficult to analyze. For the analysis of videos these videos should be divided into different shots, by determining the shot boundaries[12]. This division of the videos is termed as Video shot-boundary detection (VSBD), which is useful for various different purposes like editing, indexing, mixing, etc[5]. This detection of boundaries between shots carried out using various methods. In this paper the gesture features along with the color features of a video are used for shot boundary detection.**

***Index Terms*— Color Histogram, Gesture Interpretation, Shot Boundaries, Shots.**

I. INTRODUCTION

Video shot boundary detection means to define the discontinuity between different scenes or shots. These shots can be defined as a number of frames recorded from a source i.e camera continuously. These shots are the basic unit of a video and are very useful for task like video indexing, video surveillance, video editing, video mixing, etc[13] which is possible only by extracting the shots from the videos. These boundaries can be noted by examine of the transitions between different the shots. The transitions can be of various different types. Basically it can be classified as hard cut and soft cut also called as gradual transitions. Again the gradual transitions can be split into wipe, dissolve and fade (fade in and fade out) it can be seen in the flowchart shown in figure (1).

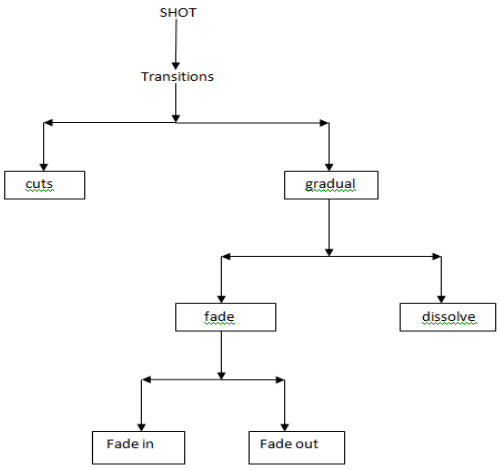
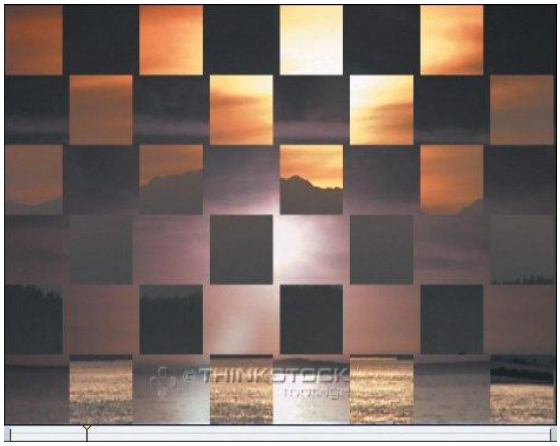


Fig (1):- Classification of shots.

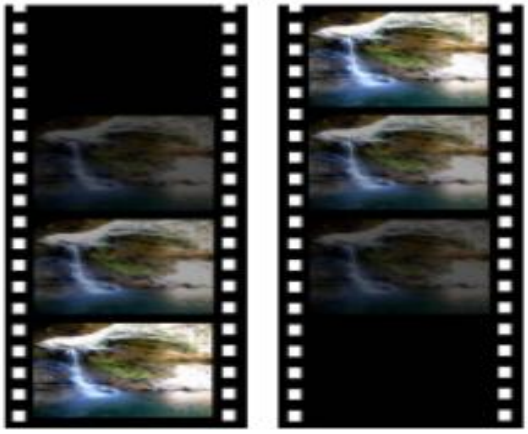
The hard cut is the sudden changes taking place within a frame while the gradual transitions take multiple frames to show the changes. Wipe shows a slow variation in its contents and can move in different shapes like circular wipes, diagonal wipes, horizontal wipes, block wipes, etc. Fade shows the change in the brightness of the frames which starts or ends with multiple solid black frames. In case of fade in the shot starts with black frames and continue till the current frame while a fade out moves from current frame to black frames. Dissolve shows both fade in and fade out in which one set of frames undergoes fade in while another fade outs showing an overlapping between the two shots. The examples of these transitions can be seen in the figure (2) shown below:-

(a) Hard cut

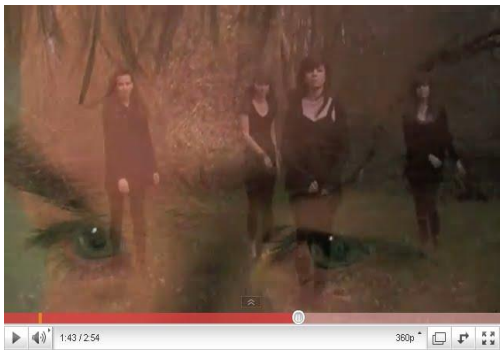
Manuscript received
***P. Swati Sowjanya**, ME Scholar, SSCET, CSVTU, Bhilai, (swatisowjanya@gmail.com). Bhilai, INDIA,*
***Mr. Ravi Mishra**, Sr. Assistant Professor, SSCET, CSVTU, Bhilai, (ravimishra_elex@yahoo.co.in). Bhilai, INDIA*



(b) Wipe transition.



(c) Fade in (left) and Fade out (right)



(d) Dissolve

Fig (2):- Examples of transitions types.

There are various different features that can be used to check for these types of discontinuities between the frames[3]. A few that has already being used are:-

- Pixel differences[9].
- Edge differences[4].
- Statistical differences.
- Compression difference.
- Motion vector[11].

1) Pixel differences:

Two common approaches:

- Calculate pixel-to-pixel difference & Compare the sum with a threshold
- Count the number of pixels that change in value more than some threshold & Compare the total number against a second threshold.

2) Statistical differences:

- Divide image into regions
- Compute statistical measures from these regions (e.g., mean, standard deviation)
- Compare the obtained statistical measures

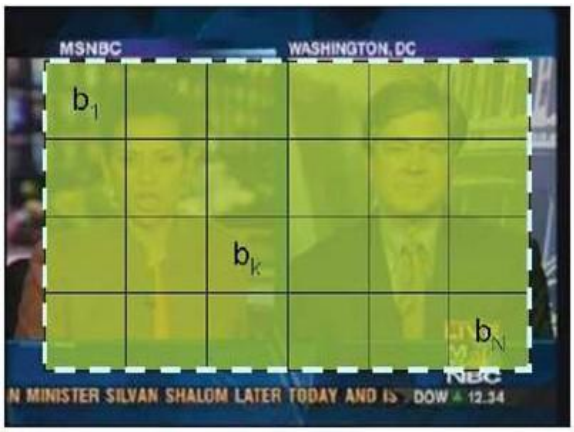


Fig (3)

3) Compression differences:

- Use differences in the discrete cosine transform (DCT) coefficients of JPEG compressed frames as the measure of frame similarity.
- Avoid the need to decompress the frames

4) Edges differences:

- The edges of the objects in the last frame before the hard cut usually cannot be found in the first frame after the hard cut,
- The edges of the objects in the first frame after the hard cut in turn cannot be found in the last frame before the hard cut.
- Use Edge Change Ratio (ECR) to detect hard cuts Edges/contours: Based on edge change ratio (ECR). Let σn be the number of edge pixels in frame n, and Xnin and Xn-1out the number of entering and exiting edge pixels in frames in frames n and n-1, respectively. The edge change ratio ECRn between frames n-1 and n is defined as:

$$ECR_n = \max(X_n^{in} / \delta_n, X_{n-1}^{out} / \delta_{n-1})$$

5) Motion vector:

Use motion vector to determine discontinuity.

Fig(4)

6) Histograms comparison:

The most common method used to detect shot boundaries.

- Provides good trade-off between accuracy and speed
- The simplest histogram method computes gray level or color histograms of the two images. If the bin-wise difference between the two histograms is above a threshold, a shot boundary is assumed.
- Several extensions available: Using regions, region weighting, different distance metrics.

Many approaches have be proposed for the detection of shots[1,7,8], of these the method using color histogram is the simplest and provide significant result with the hard cuts. Recent work presents a method with motion activity, which uses motion vectors[14] which can be used efficiently but is complex and time consuming[10], all these methods can produce efficient results but are sensitive to luminance as well as camera motion.

In this paper we are using gesture interpretation method for the detection of shot boundaries.

II. Gesture Interpretation

The gesture interpretation uses spatio-temporal analysis for checking the changes occurring in the shots. The spatio-temporal analysis can be said as the property that shows both space as well as time variations in it, for example consider a cricket match here if a batsman is going to hit a ball, consider the movement in his gesture that slowly varies with time as well as space as he hits the ball.

Gesture Interpretation characterizes the image very well and is also resistant to luminance and small changes or transitions. It treats the whole image as one object[2]. In a shot due to the appearance and disappearance of the objects, the color histogram works inefficiently and those produces incorrect results and the gesture features works good in these cases.

III. Methodology

For using the gesture interpretation[6], firstly we resize the image into 128*128 pixels, and then we use a filter to divide each frame into 3 scales and 8 orientations, thus forming 24 filters which are again divided into 16 different parts thus forming (16*24) 384 dimensional vectors of which the top 64 are selected of the determination of shot boundaries. The 64 vectors are then used for decision making as shown below:-

Video showing Hard cut:-

Frames:-

These features are taken as (1…N) & time varying signal $X_i(t)$, where, $i \in 1..N$

Now, if

$$dX_i(t)/dt \approx 0 \quad (1)$$

then it shows no transition.

If

$$|dX_i(t-1)/dt| \,\&\, |dX_i(t+1)/dt| \approx 0 \quad (2)$$

then it shows hard cut.

Else if

$$|dX_i(t-1)/dt|, |dX_i(t)/dt| \,\&\, |dX_i(t+1)/dt| \geq 0 \quad (3)$$

Then it shows the gradual transition.

Then we use z-score matrix for the task of outlier detection and also draw the histogram for the color features of the frame.

Where,

$$z = (x - \mu) / \sigma$$

shows the z-score matrix.

The color features shows a significant result in case of hard cut detection.

IV. Result and Discussion

The results obtained from the different videos can be seen below:-

Frame 30 Frame 31 Frame 32

Frame 33 Frame 34

Fig 5:- frames for hart cut video

Output of gesture interpretation:-

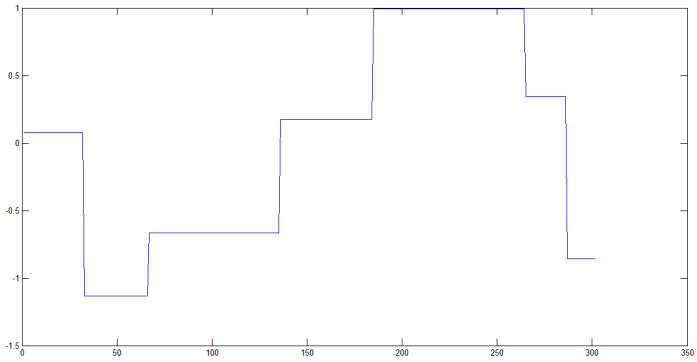


Fig 6:- gesture interpretation output

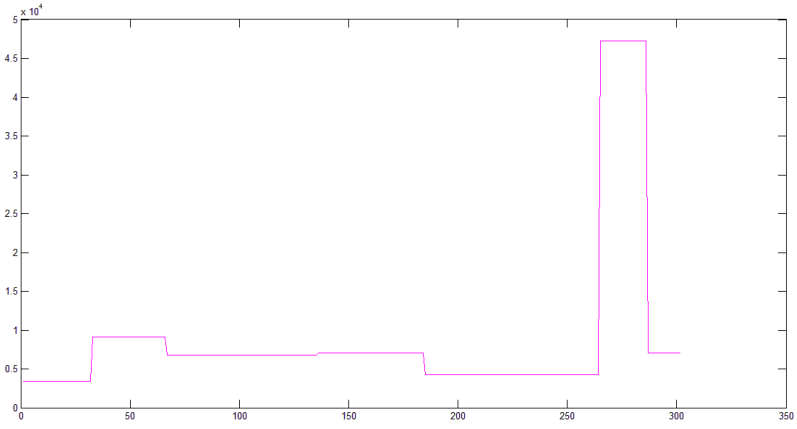


Fig 7:- output for color features

For video with shatter (gradual) transition:-

Frames:-

Frame 2 Frame 3 Frame 4

Frame 5 Frame 6

Fig 8:- Frames

Output for Gesture interpretation:-

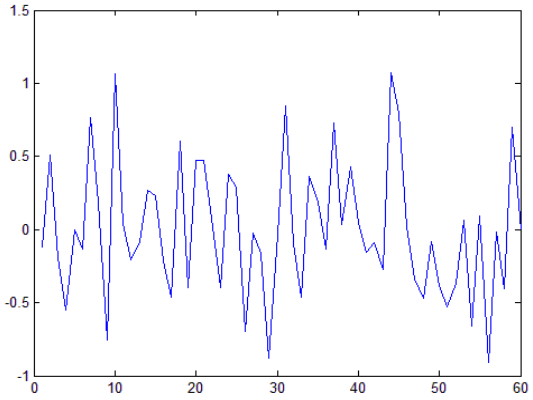


Fig 9:- output for gesture interpretation

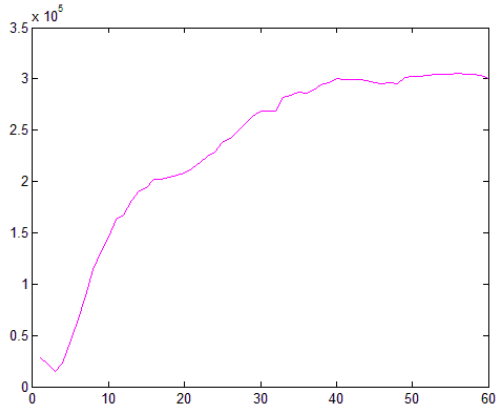


Fig 10:- output for color features

V. CONCLUSION

In this paper we have proposed the Gesture Interpretation features along with the color features for the detection of shot boundaries in videos. The gesture features were presented as a task for outlier detection which is useful for the detection of gradual changes in the videos with the help of z-score matrix. We observe that Gesture Interpretation is better than color

features for detecting hard cuts. However, color performs better for gradual transitions. The best performance for the system is: 84% precision and 91% recall for hard cuts. The best system performance for gradual transitions is 60% precision at 70% recall.

ACKNOWLEDGMENT

Here we would like to acknowledge all the people associated with this project that has helped in one or the other manner towards the completion of the project. I also like to thank to my guide who has helped me in this project. I also acknowledge my friends and family members.

REFERENCES

[1]. Alan F. Smeaton, Paul Over, Aiden R. Doherty, "Video Shot Boundary Detection: Seven Years of TRECVid Activity," Preprint submitted to Elsevier 12 March 2009

[2]. Birinci, M.; Diaz-de-Maria, F.; Abdollahian, G.; Delp, E.J.; Gabbouj, M.; "Neighborhood matching for object recognition algorithms based on local image features", Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE, Jan 2011, pp 157-162

[3]. Murat Birinci, Serkan Kiranyaz, Moncef Gabbouj, "Video Shot Boundary Detection By Structural Analysis Of Local Image Features", 2011

[4]. Don Adjeroh, M. C. Lee, N. Banda, Uma Kandaswamy, "Adaptive Edge-Oriented Shot Boundary Detection", NSF ITR, 2009, pp 1-24.

[5]. Liangda Li, Ke Zhou, Gui-Rong Xue, Hongyuan Zha, Yong Yu, "Video Summarization via Transferrable Structured Learning", International World Wide Web Conference Committee (IW3C2), March-April 2011, pp 287-296.

[6]. Baber, J.; Afzulpurkar, N.; Dailey, M.N.; Bakhtyar, M.; "Shot boundary detection from videos using entropy and local descriptor", Digital Signal Processing (DSP), 2011 17th International Conference on Zhao Huan, July 2011, pp 1-6

[7]. Ali Amiri and Mahmood Fathy, "Video Shot Boundary Detection Using QR-Decomposition and Gaussian Transition Detection", EURASIP Journal on Advances in Signal Processing Volume 2009 (2009), Article ID 509438, 12 pages

[8]. Hua-Yong Liu; Tingting He, "Using Multimodal Analysis for Story Segmentation of News Video", Artificial Intelligence, 2009. JCAI '09. International Joint Conference on, April 2011, pp 124-127

[9]. Bo Han; Weiguo Wu; "Video scene segmentation using a novel boundary evaluation criterion and dynamic programming", Multimedia and Expo (ICME), 2011 IEEE International Conference on, July 2011, pp 1-6

[10]. Moshe Guttmann, Lior Wolf, Danny Cohen-Or. "Content aware video manipulation". Computer Vision and Image Understanding xxx (2011) xxx–xxx.

[11]. Wu-Chih Hu, "Real-Time On-Line Video Object Segmentation Based On Motion Detection Without Background Construction", International Journal of Innovative Computing, Information and Control ICIC International Volume 7, Number 4, April 2011, pp. 1845-1860.

[12]. Paul Browne, Alan F Smeaton, Noel Murphy, Noel O'Connor, Seán Marlow, Catherine Berrut, "Evaluating and Combining Digital Video Shot Boundary Detection Algorithms", Conference or Workshop Item, Feb-2009

[13]. Nicola Anselmi, Dongmo Sylvestre Aime, Rossi Maurizio, "Shot Boundary Detection in OpenCV", 2011.

[14]. Abdelati Malek Amel, Ben Abdelali Abdessalem and Mtibaa Abdellatif, "Video shot boundary detection using motion activity descriptor", JOURNAL OF TELECOMMUNICATIONS, VOLUME 2, ISSUE 1, APRIL 2010, pp 54-59.

Ms. P. Swati Sowjanya did her B.E. in Electronics & Telecomm. Engineering in year 2010. She is pursuing her M.E in Communication Engineering from SSCET, Bhilai in final year. She has passed her schooling form St. Xavier's Durg. Her main fields of interest are: Video & Image Processing, Embedded Systems and Microprocessors.

Mr. Ravi Mishra did his B.E. in Electronics & Telecomm. Engineering in year 2002. He acquired his M. Tech. (Computers) degrees from COEP, Pune, India in years 2008. He has worked as a faculty member in Engineering Academy, Bhilai during year 2002-2004. He has worked as a Head R&D in Indus Global Software Solution, Bhilai during year 2006-2008. He has worked as a Senior Lecturer in Electronics & Telecomm. Engineering department at CSIT, Durg, during 2008 – 2009, India. He has worked as a Senior Lecturer in Electronics & Telecomm. Engineering department at DIMAT, Raipur during 2009 - 2010 India. Presently he is working with the Electrical & Electronics Engineering department at SSCET, Bhilai, India. His main fields of interest are: Video & Image Processing, Embedded Systems.