Automatic Video Shot Change Classification using Optic Flow and Robust Pixel Difference Based Method

Amith Kamath B[1], Mahesh P S[2], Meghana Madhusudhan[3], Rakshith N S[4]

[1]Asst. Prof., Dept of CSE, SJBIT, [2,3,4]Student, Dept. of CSE., SJBIT

Abstract: *Shot change detection is an essential step in video content analysis. The field of Video Shot Detection (VSD) is a well exploited area. In the past, there have been numerous approaches designed to successfully detect shot boundaries for temporal segmentation. In this paper we present a procedure to detect and automatically classify Video shots. We present a method to detect shots using optic flow, and a method to classify the shot change into Abrupt/Hard cut or Gradual Transition using Robust Pixel Based Method.*

Keywords: *Video Shot Detection, Shot Boundary Detection, Optic Flow, Corner Detection,Automatic classification.*

I. Introduction

The rapid development of storage and multimedia technologies has made the retrieval and processing of videos relatively easy. Temporal segmentation is a fundamental step in video processing, and shot change detection is the most basic way to achieve it. However, while hard cuts (abrupt transitions) can be easily detected by finding changes in a color histogram, gradual transitions such as dissolves, fades, and wipes are hard to locate.

In practice, however, 99% of all edits fall into one of the following four categories hard cuts, fades, wipes and dissolves. Many shot change detection studies focus on finding low-level visual features, e.g., color histograms and edges, and then locate the spots of changes in those features. We focus on using Optic Flow for Video Shot Detection and after successful detection, Robust Pixel Method for automatic classification.

In the following sections, we present different stages of the proposed algorithm. Feature points extraction and themethodology for video shot detection using optic flow are explained in section II. A methodology is proposed to automatically classify shot change as gradual or abrupt in section III using Robust Pixel. Experimental Results includes test data considered in the current work, Performance on Each Type of Transition and the algorithms accuracy in classification section IV.

II. Feature Point Extraction And Shot Detection Using Optic Flow

A. Feature point extraction using Kanade-Lucas-Tomasi Corner Detector

KLT searches for points where variations in two orthogonal directions are large Based on local structure matrix

$$C_{str} = w_G(r;\sigma) * \begin{bmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{bmatrix} ..$$

Derivatives of the intensity function f(x, y) are calculated in each point. Then, the entries of the matrix C_{str} are obtained for D x D neighborhood of R. Each of the entries is smoothed by Gaussian filter. The diagonal entries will be the two Eigen values

$$C_{str} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Let's now consider the two eigenvalues λ_1 and λ_2 of the matrix C. As C is symmetric and positive semi-definite, both λ_1 and λ_2 are non-negative, and at the location of a corner we have $\lambda_1 >= \lambda_2 > 0$ where both eigenvalues are large. The KLT algorithm compares the smaller eigenvalue λ_2 to a threshold value λ_{min} and if greater saves (x,y) in a potential corner list L. Then it sorts L in decreasing order of λ_2, and scan the sorted list from top to bottom, selecting points in the list in sequence and removing points that fall inside the neighbourhood R of any selected points (in order to have neighbourhood which do not overlap, because those which overlap are probably due to the same corner), until having the required count of features.
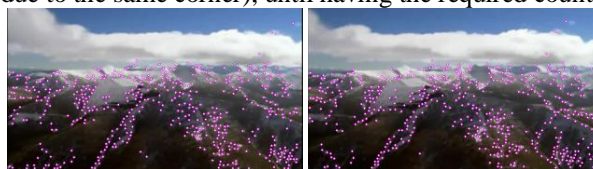

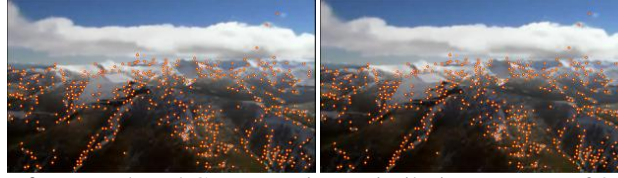**Figure 1.** Corners Detected using KLT & limit of 500 Corners.

Figure 2.   Reduced Corners using a similarity measure of 0.9.

B. Video Shot Detection using Optic Flow
        In the our approach, we consider Video stream / Sequence of images sampled at a frame rate of 5. For each pair of Images(I_1 , I_2) in the video sequence, we detect the corner points using KLT.
Usually every corner detector gives a large number of corner points, many of which are outliers. In order to retain strong and reliable corner points leading to higher matching accuracy, we use two measures viz., cornerness and similarity.
Cornerness is the characteristic property of any interest point (feature point) p and is defined as

$$Cp=|\ \lambda_1^{\ 2} + \lambda_2^{\ 2}\ |$$

where λ_1 and λ_2 are the 2 Eigen values of the feature point. The correspondence between 2 points are measured using the similarity measure S(p,q) is defined as

$$S(p,q)=min\ (Cp,Cq)/max(Cp,Cq).$$

        where p and q are in two images , A point is considered as a good feature if S(p,q)>T. Where T is a variable threshold.If the value of the corner strength exceeds the predefined threshold, then the corner is considered as a prominent feature and is retained for subsequent computations.
Three important calculations are made at this stage. They are number of match points, number of large vectors and orientation of the vectors.Decision is then made using three conditions resulting in classifying the image pair into one of the following cases:
Case 1: No. of Match Points
If there are only few numbers of match points, it implies that a shot change is detected, as there are no match points between the image pair under consideration. If Case 1 is not satisfied, we evaluate Case 2.
Case 2: Number of large Vectors
For large number of match points, we consider the number of large vectors. The possibility of a change in shot is discarded if the number of large vectors are comparatively less than the number of small vectors, The situation is similar to occlusions occurring in a video. If Case 2 is not satisfied, we evaluate Case 3.
Case 3: Orientation of large vectors
For large number of match points and large vectors, we consider orientation of the large vectors. The possibility of a shot change is discarded if the orientation is uniform. This situation is similar to vibration of camera, zooming or panning. If Case 3 is not satisfied, we evaluate Case 4.
Case 4: Large number of match points with large vectors with non uniform orientation, implies the detection of a change in shot.

III.   Automatic Classification using Robust Pixel Method
C. Robust Pixel Method(RPM)
In RPM, we consider a Metric M, computed for each image. Defined as

$$M(I^K|I^{K-1}) = \frac{1}{HW}\sum_{i,j}\rho_{i,j}^K$$

where$I^k$and I^{k-1} are each pair of consecutive images to be compared.H and Ware the image height and width.(i; j) are the coordinates of each one of the pixels in the image. And ρ is

$$\rho_{i,j}^K = \begin{cases} 1 & if\ |I_{i,j}^K - \mu^K| > T_n\ \&\ sign(I_{i,j}^K - \mu^K) = sign(I_{i,j}^{K-1} - \mu^{K-1}) \\ -1 & f\ |I_{i,j}^K - \mu^K| > T_n\ \&\ sign(I_{i,j}^K - \mu^K) \neq sign(I_{i,j}^{K-1} - \mu^{K-1}) \\ 0 & otherwise \end{cases}$$

where$\mu^k$is the mean value of the image I^k, and Tnis a noise threshold (typically Tn= 2).

D. Automatic Classification of Shot Change
The proposed methodology works on two levels, First the Optic Flow method is used to detect the shot change in a video sequence.
For each shot change detected successfully at a frame I_n we use RPM, to calculate the Metric M for a ten frame window i.e., I_{n-5} to I_{n+5}.

Comparing the metric values between the images I_{n-5} to I_{n+5}. The two local minimum values of M are identified as M_{min1} and M_{min2}. The shot is classified as a hard cut only if:

$$| M_{\min 1} - M_{\min 2} | > T_{classification}$$

It is classified as a Gradual transition for all other cases. Where $T_{classification}$ is heuristically set to 0.3.
Example of effectiveness of RPM:

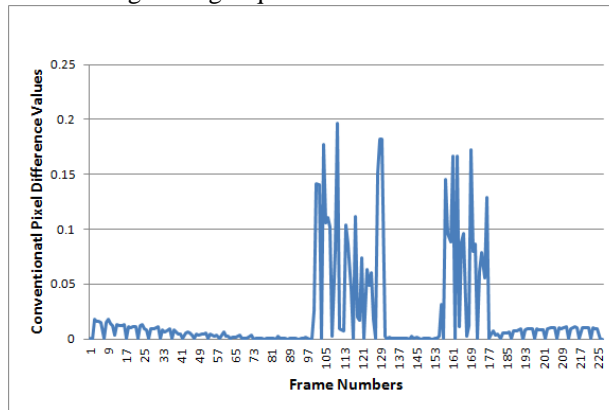**Figure 3.**   Lightening sequence used to demonstrate RPM

Figure 4.   Graph of Conventional Pixel Difference values for Lightening sequence 1.
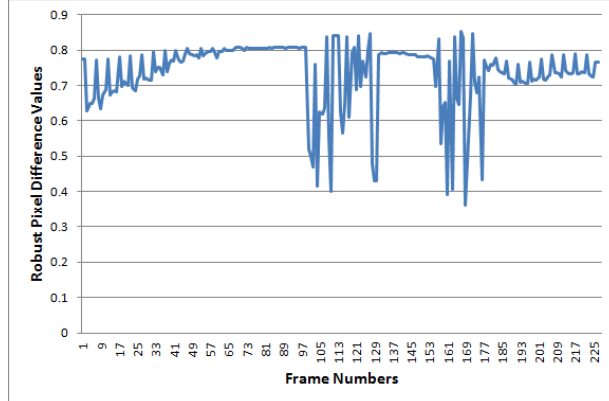
Figure 5.   Graph of Robust Pixel Difference values for Lightening sequence 1.

Results for the shown sequence Using Conventional Pixel Difference method falsely detects video shot changes, Robust Pixel Difference Method does not falsely detect video shot changes, hence proving to be more robust.
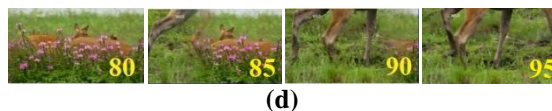
IV. Experimental Results.

(a)

(b)

(c)

(d)
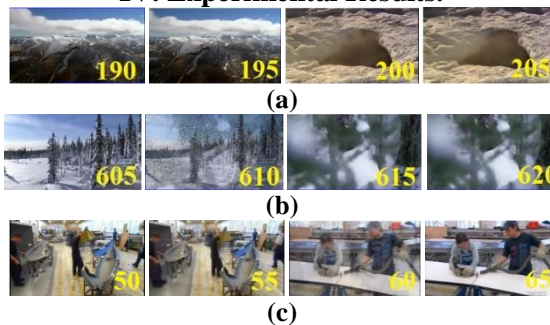
Figure 6. Sample frames in our test set. (a) Hard Cut, (b) Dissolve.,
(c) Fade , (d)Wipe .

In order to evaluate the performance of the shot cutdetection methods, following performance measureswere used.

Recall: the Recall performancemeasure, known as thetrue positive function or sensitivity that corresponds to the ratioof correct experimental detections over the number of all truedetections:

Recall Rate = Correctly Detected/ (Correctly Detected + Falsely Detected)

Precision: the Precision performancemeasure definedas the ratio of correct experimental detections over the number ofall experimental detections:

Precision Rate = Correctly Detected/(Correctly Detected+

Missed Detected)

Figures 7,9,11 and 13 show the corner detection, figures 8,10,12 and 14 show the corresponding flow of corners for the video frames from our test data. Figure 15-18 show the graph of Metric M calculated for a ten frame window at the location of shot change.

Table 1 shows the descriptions of the videos used for performance analysis. It provides information about the video names, the type of transitions in the video, total frames and the temporal width of the transitions in the video. Table 2 shows the Precision, Recall values and accuracy in classification. A precision value of 1 implies that all the shot changes in the video were detected. A recall value of 1 implies that there were no false detections. The table shows the values of each performance parameter for every test video.

V. Conclusion

Video Shot Detection using corners and optical flow provides satisfactory detection of shot boundaries. KLT corner detector provides a computationally faster approach, due to the use of threshold to limit the corner points, cornerness and similarity measure.

The detection of gradual transitions is dependent on the temporal width of the transition, i.e. the methods perform as expected for a temporal transition width of 20 frames.

The proposed method for classification of shot change into Hard cut/ Gradual transition works as expected and has high classification accuracy in our test data for correctly identified shot changes.

References

[1] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classifying scene breaks," in Proc. ACM Multimedia, 1995, pp. 189–200

[2] Comparison of Automatic Shot Boundary Detection Algorithms , Rainer Lienhart1 Microcomputer Research Labs, Intel Corporation, Santa Clara. 2004

[3] Video Shot Boundary Detection Based On Color Histogram.,Jordi Mas, Gabriel Fernandez .Digital Television Center,Spain. 2005

[4] Video Shot Meta-segmentation Based On Multiple Criteria For Gradual Transition Detection,efthymia Tsamoura, Vasileios Mezaris, Ioannis Kompatsiaris,Centre for Research and Technology Greece. 2008

[5] A Svm-based Soccer Video Shot Classification Department of Computer Science, Beijing Institute of Technology, Beijing , 2005,Yi-hua Zhou, Yuan-da Cao, Long-fei Zhang, Hong-xin Zhang

[6] Shot View Classification for Playfield-based Sports Video, Alfian Abdul Halin, Mandava Rajeswari, Dhanesh Ramachandram Computer Vision Lab , Penang, Malaysia , 2009

[7] A Shot Boundary Detection Technique Based On Local Color Moments In Ycbcr Color Space S.A.Angadi and Vilas Naik Natarajan Meghanathan, et al. (Eds): SIPM, FCST, ITCA, WSE, ACSIT, CS & IT , 2012.

[8] Keyframe Based Video Summarization Using Automatic Threshold & Edge Matching Rate. Mr. Sandip T. Dhagdi, Dr. P.R. Deshmukh ,International Journal of Scientific and Research Publications, July 2012.

[9] Real-time shot detection based on motion analysis and multiple low-level techniques.Carlos Cuevas and Narciso Garcia,Grupo de Tratamiento de Imagenes - E.T.S. Ing. Telecomunicacion, 2010

[10] Video Shot Detection using Saliency Measure.Amudha J, Radha D, Naresh Kumar P. International Journal of Computer Applications (0975 – 8887) Volume 45– No.2, May 2012.
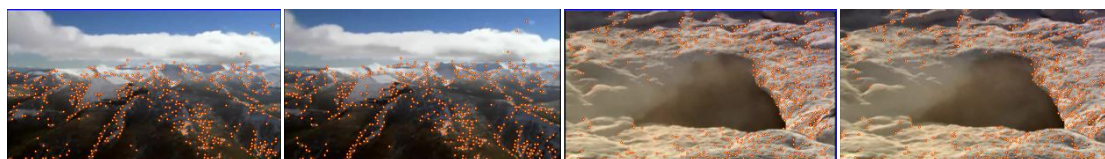
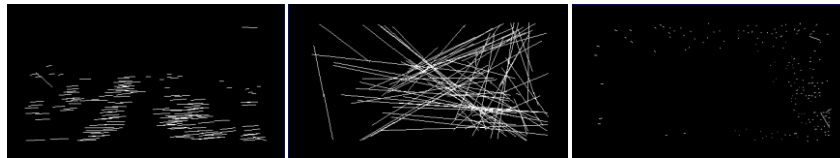**Figure 7.** KLT corner detector, Demonstration of 'Hard Cut' from Test data (a)

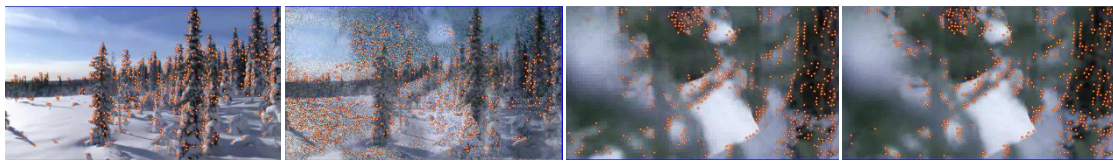**Figure 8.**   Flow of corner points from Fig.7

Figure 9.   KLT corner detector, Demonstration of 'Dissolve'from Test data (b).

Figure 10.   Flow of corner points from Fig.9

Figure 11.   KLT corner detector, Demonstration of 'Fade'from Test data (c).
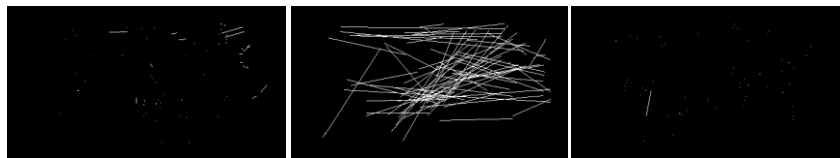

Figure 12.   Flow of corner points from Fig.11

Figure 13.   KLT corner detector, Demonstration of 'Wipe'from Test data (d).
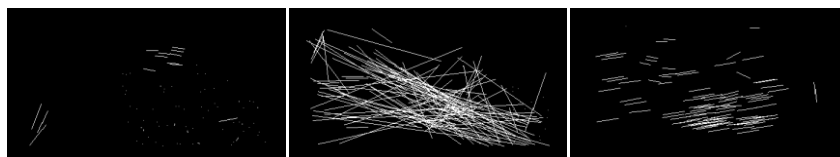

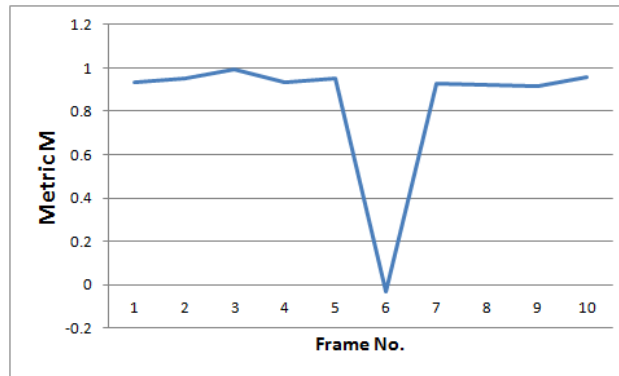**Figure 14.**   Flow of corner points from Fig.13

Figure 15.   Graph of Metric M computed for a ten frame window around location of detected Shot change shown in Hard cut, figure 7.
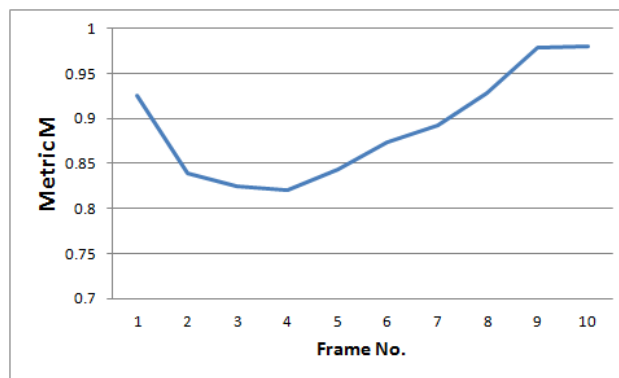

Figure 16.   Graph of Metric M computed for a ten frame window around location of detected Shot change shown in Fade, figure 11.
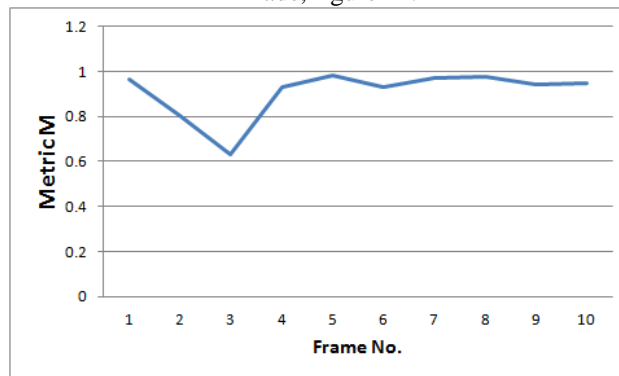

Figure 17.   Graph of Metric M computed for a ten frame window around location of detected Shot change shown in Dissolve, figure 9.
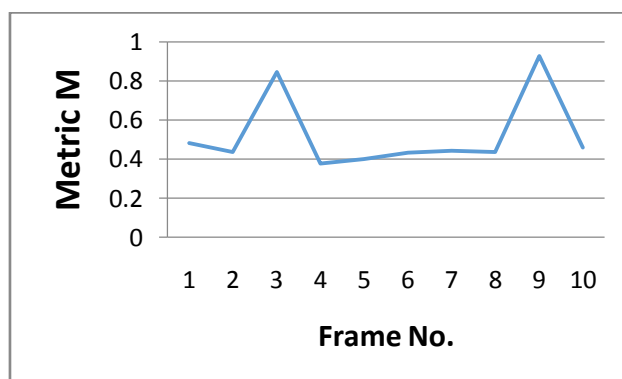

Figure 18.   Graph of Metric M computed for a ten frame window around location of detected Shot change shown in Wipe, figure 13.

Results & Analysis:

Table 1. Test videos used for analysis

Video Name	Type of Transitions	Temporal width of Transition	No. of Transitions	Total Number of Frames
CUT 1.avi	Abrupt/Hard cut	2frames	4	390
Dissolve 1.avi	Dissolve/Gradual Transition	20frames	4	940
Fade 1.avi	Fade/Gradual Transition	20frames	4	1149
Wipe 1.avi	Wipe/Gradual Transition	20frames	4	1953

Table 2. Precision and Recall of the Methods discussed above.

Video Name	VSD using Optic Flow		Classification Results	Classification Accuracy of correctly identified shot change
	Precision	Recall		
CUT 1.avi				
	1	1	4 Hard cuts	100%
Dissolve 1.avi	1	1	4 Gradual Transitions	100%
Fade 1.avi	1	1	4 Gradual Transitions	100%
Wipe 1.avi	1	1	4 Gradual Transition	100%