

Effective Algorithm for Detection of Dissolve in Presence of Motion and Illumination

Salim Chavan
Research Scholar,
Assistant Professor,
Department of Electronics and
Telecommunication
Engg,S.B.J.I.T.M.R.(M.S.),India

Sudhir Akojwar
Senior Member IEEE,
Professor, Department of
Electronics Engineering,
Rajiv Gandhi College of
Engineering and Technology,
Chandrapur (M.S.) India,

ABSTRACT

The detection of dissolve transition is more difficult than detecting fade in and fade out. In this transition, the last frames in the previous shot fade out and the beginning frames in the next shot fade in i.e. the overlapping of fade out and fade in occurs. The dissolves may be of three or more number of frames. In some videos very short dissolve of even three frames also occurs. So, there are lots of challenges in detection of dissolves. Also illumination and camera / object motion gives rise to false positives thereby degrading the algorithm performance. Many researchers addressed this issue but could not achieve the robustness in the presence of illumination and object / camera motion. Therefore this issue needs to be resolved. An algorithm has been proposed for dissolve detection. In this algorithm, color histogram difference between consecutive frames is calculated and average value of this difference for all consecutive frames is used as a metric for dissolve detection.

Keywords

Dissolve, shot boundaries detection, color histogram difference, recall, precision.

1. INTRODUCTION

Dissolves are the basic artificial effects of gradual transitions in video sequences, which are widely used in TV programs and movies. Through dissolve detection, interesting and important video temporal segments can be easily located and indexed for various applications. Among the various types of gradual transitions, the dissolve-type gradual transition is considered the most common one, but it is also the most difficult one to detect. In most of the existing dissolve detection algorithms, the false/miss detection problem caused by motion is very serious.



Fig. 1 Dissolve frames in the movie Krish 3

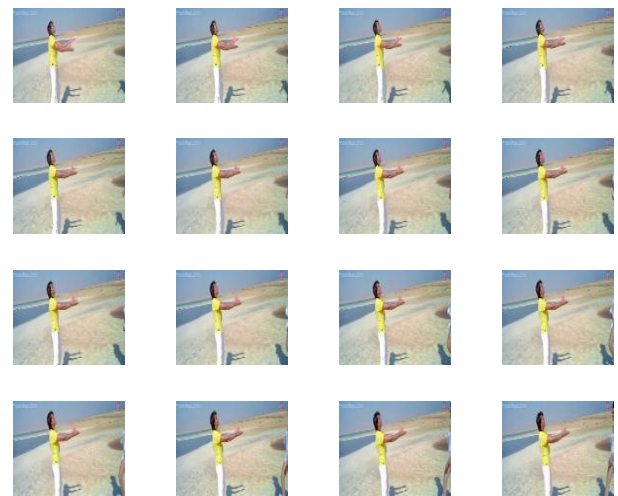


Fig. 2 Dissolve frames with camera motion in the movie Krish 3

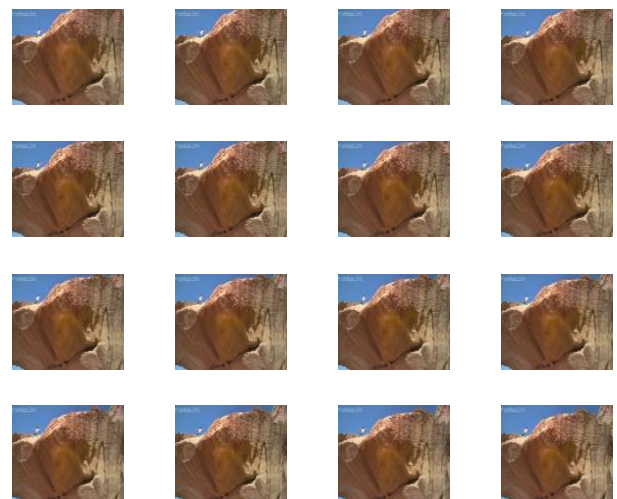


Fig 3. Dissolve frames with object motion in the movie Krish 3

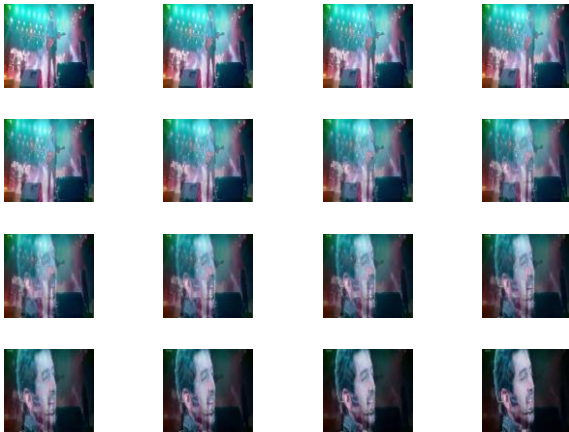


Fig. 4 Dissolve frames with illumination in the movie Ashiqui

The related work for dissolve detection in presence of illumination and motion is given below.

Nagasaka and Tanka [3] proposed the method in which the transitions are detected by using a global inter-frame dissimilarity measure. In addition, the term ‘template matching’ has come to be used to refer to the second method used by Nagasaka to detect shot boundary using pixel pair difference between each pair of consecutive frames. This method takes account of a color pixel as well. Firstly, the difference intensities of two pixels having similar spatial positions between successive frames are computed. Secondly, the accumulated total of differences is compared with a given threshold, to decide if the shot boundary has been detected. To detect dissolve transitions using an edge tracking technique, a method was proposed by Lienhart[4]. Lienhart applied a canny algorithm in order to extract weak and strong edges in the image. After that, two thresholds were compared with the obtained values. Lowest and highest thresholds are used to detect weak (w) and strong (s) edges which are denoted by T_w and T_s . A potential problem of edge detection method that it does not handle rapid changes in bright and dark scene which gives false measures. Hanjalic [6] identify and analyze the major issues related to shot boundary detection problem in detail. Then, on the basis of minimization of the average detection error probability, conceptual solution to the detection of shot boundary is presented in the form of statistical detector. The required statistical function is modeled using a robust metric for visual content discontinuities, which includes the knowledge of the short length distribution, visual discontinuity pattern at shot boundaries and characteristics temporal change of visual features around a boundary. They found that the influence of motion and lighting changes on the detection performance cannot be reduced easily.

Cali et al. [21] have proposed a scheme for identifying the cut and dissolve by using linear prediction for predicting a frame and then prediction error are used to determine existence of short boundaries. Su et al. [9] have proposed an efficient method for detecting dissolve type gradual transitions. They modeled a dissolve based on its nature and filtered out possible confusion caused by the effect of motion via statistics. They used binomial distribution model to distinguish real dissolves from motion. Ling et al. [20] proposed a method which used variance projection function to find distance between the video frames and then a statistical

learning method on SVM is used to determine whether the changes of the distance are caused by gradual transition or not. In [8] the authors presented an effective dissolve detection algorithm, which takes into account the dissolve properties in both temporal and spatial domains. In particular, in the temporal domain, they used frame difference to capture the dissolve characteristics. In the spatial domain, the central area is given bigger weight than the four sides. A fuzzy logic approach [9] to integrate hybrid features for detecting shot boundaries is explained in this paper. In this paper, author proposes a fuzzy logic approach to integrate hybrid features for detecting shot boundaries inside general videos. The fuzzy logic approach contains two processing modes where one is dedicated to detection of abrupt shot cuts including those short dissolve shots, and the other for detection of gradual shot cuts. These two modes are unified by a mode-selector to decide which mode the scheme should work on in order to achieve the best possible detection. The hybrid features used in this paper are color histogram, texture change and edge variance. The fuzzy logic approach performs well as compared to other because generally in other method a threshold is defined and this makes the decision binary but in case of fuzzy logic approach, it will provide the reliability against this by defining a membership function for various features such as color histogram, texture change and edge variance.

Lawrence et al. [8] have presented an algorithm for shot boundary detection based on the video’s first-order partial derivatives. They tried to make algorithm insensitive to object and camera motion by considering areas with low apparent motion. Su et al. [9] have proposed an efficient method for detecting dissolve type gradual transitions. They modeled a dissolve based on its nature and filtered out possible confusion caused by the effect of motion via statistics. They used binomial distribution model to distinguish real dissolves from motion. Xu et al. [25] have proposed a motion suppression technique for shot boundary detection based on 3D wavelet transform. Motion is characterized in terms of energy and variance, and then motion suppression value is extracted from intensity of motion energy. This motion suppression value is used to suppress motion influence suffered by traditional shot boundary detection algorithm. Jang et al. [11] have proposed a shot transition detection method that compensates motion in a video. They first extracted motion vectors from two consecutive frames by using the size-variable block matching and then adaptive robust estimation method is used to estimate the global motions and eliminate local motions. Park et al. [12] have proposed two-motion based features for shot boundary detection of fast motion sequences. The first feature which is robust to fast motions is developed by considering the motion magnitude. The next feature, block wise motion similarity, is defined by comparing motion directions in consecutive frames. Truong et al. [7] tried to improve the accuracy of cut detection with the help of an adaptive threshold obtained from a local window in the color histogram difference curve. In addition, the existences of these effects were examined based on the mathematical models for producing ideal fades and dissolves. In that procedure, constraints on the characteristics of mean and variance curves of frame luminance were derived to eliminate false positives caused due to disturbances by camera motion and object motion during gradual transitions

Methods for detecting or differentiating flashlight from abrupt transition can be broadly classified as histogram based methods, edge detection based methods and local window based methods. Li D. and Lu H. [3] has proposed a method to

differentiate between shot and flashlights based on difference in area of background in edge image. Weixin K. et al. [23] have used color ratio histogram as illumination invariant metric and local adaptive thresholds have been used to find shot change. Histogram difference is used as a metric to differentiate between abrupt transition and gradual transition [1], and then average intensity difference is used to detect flashlights. Model for flashlight detection is formed by length, strength, brightness, velocity, and impact of the flashlight [33]. Depending on these parameters thresholds are used to detect flashlight. In [22] edge direction, edge position, and edge matching is used to discriminate abrupt scene change from flashlight scenes. Yuliang G. and De X. [23] used twin comparison algorithm to find potential shot change. By using edge direction histogram and inter frame similarities, light change and shot boundary can be detected. Accumulated histogram difference and energy variation is used to identify flashlights in [23]. In [2] frame difference from consecutive frames is detected using subdivided local color histogram comparison. This difference is compressed by using logarithmic transform for efficient detection of flashlights.

An effective algorithm for dissolve detection with camera and object motion is proposed in this paper [9]. This paper provides an approach to overcome the problem of misdetection and false detection by modeling a dissolve transition and selecting a proper threshold. Although there are problems in illumination, this algorithm works with high efficiency. In this paper, dissolve transition is modeled into three categories of pixel as Proponents, Fence sitters and Opponents. Proponents intensity monotonically increases or decreases. Fence sitters are those pixels whose intensity remains unchanged and Opponents are those pixels whose intensity neither monotonically increases or decreases nor remain unchanged. Out of above only Proponents are the part of dissolve transition. In this paper, author has stated one drawback, when dissolve is for long duration misdetection may occur.

2. PROPOSED METHOD FOR DETECTION OF DISSOLVE TRANSITION

The proposed methodology for the detection of dissolve works in three steps. In first step, the color histogram of each frame is calculated. Then the normalized mean of color histogram is obtained. In the second stage, the average value of normalized mean of histogram difference is calculated for detecting gradual transitions. In the third stage fades and dissolve are distinguished.

1. In this algorithm color histogram is calculated for each frame and it is tabulated in the form of array D.
2. The mean of color histogram is calculated for each frame and it is normalized to form an array M and it is plotted.
3. Next, the average level of array M is calculated. The values of array M which are above average level are stored in an array N. Fades and dissolve are detected when the values of M are above average level. From observation of plotted graph, it is cleared that if the consecutive values of an array N are more than eight frames then it may belongs to dissolve or fades. The values which are less than eight frames are eliminated and an array corresponding to fade and dissolve is formed i.e. D_c .

4. The array D_c consists of n numbers of rows. The values of each row corresponding to fades and dissolve. To distinguish between fades and dissolve the first and last point of each row are checked, if these values are found to be very close then that row will correspond to dissolve.
5. The first point of the row corresponding to the start of dissolve and last point corresponding to the end of dissolve as shown in graph.

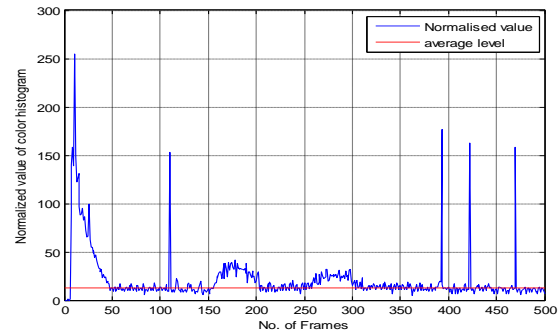


Fig. 5 The variation in normalized value of color histogram with frame index

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Test video sequence

To evaluate the effectiveness of the proposed algorithm, experiments are conducted on the five videos mentioned below. These movies containing lot of camera motion, object motion and illumination are selected so as to prove robustness of our algorithm in presence of motion and illumination. The proposed algorithm is tested on 1, 33,919 frames. All these frames are observed manually frame to frame with Virtual Dub to note the dissolve transition frames. The five videos used for testing our algorithm are given below.

- Video 1: Ashiqui 2 movie with high amount of light effect.
- Video 2: Jeene laga hoon song from Ramaiya Vastavaiya.
- Video 3: Dil Tu Hi Bata song from Krish 3 movie with complete camera and object motion.
- Video 3: I am in love Song from Once upon a time in Mumbai movie. It contains lighting effect.
- Video 4: Jumper movie. It contains large amount of lighting effect and motion.
- Video 6: Amplifier movie. It contains heavy lighting effect and rapid camera motion.

3.2 Evaluation criteria

The proposed system has been implemented and applied on variety of video sequences. The output of segmentation process of a given video package is a sequence of segment boundaries. These boundaries are compared with those detected manually using Virtual Dub from the same video sequences as a basis for evaluation. The table 1 shows the manual data set for dissolve transitions in the videos under test. Fig 6 shows the trade-off between Recall and Precision. We have taken the videos containing camera motion and illumination and tested our algorithm on such videos. Table 2 shows results for detecting shots for each of the four video

collections using our algorithm. The results obtained for the recall, precision, f1 measure and RSI are presented in Table 3. Most of the approaches give false measure in presence of rapid camera and object motion or illumination. But our approach gives improved results even in the presence of camera motion and illumination.

Table 1 Manual data set for dissolve transitions in the videos

Video Collections	Dissolve Frames
Video1	1514
Video 2	203
Video 3	123
Video3	175
Video4	212
Video6	219
Total	2513

Table 2 Results for detecting shots for each of the five video collections

Video Collections	Correct Detections	Missed Detections	False Detections
Video1	1704	133	60
Video 2	190	13	14
Video 3	143	6	03
Video3	172	06	04
Video4	200	12	05
Video6	212	6	12

Table 3 Results obtained for the recall, precision, f1 measure and RSI

Video Collections	Recall	Precision	F1 mMeasure	RSI
Video1	92.8	93.13	93.33	57.20
Video 2	93.13	92.65	92.90	56.74
Video 3	93.56	97.13	94.22	91.61
Video3	96.22	97.17	96.59	93.95
Video4	93.33	96.14	94.23	90.90
Video6	94.53	94.03	94.33	91.8

3.3 Comparison with different detection methods

We have implemented some existing algorithms as discussed in chapter 3 for comparing their results with our proposed algorithm. It has been observed that in the existing algorithms false alarms due to camera motion, object motion and illumination reduces the precision. These challenges due to disturbances caused by motion and illumination are properly addressed in our algorithm. It is also computationally fast and less complex. The comparison results with some basic algorithms for some videos are given in table 4.

It is observed that our proposed algorithm used better results than other algorithms used for comparison. We tested our algorithm on the videos containing fast camera, object motion and illumination. Due to these disturbances, the precision is affected adversely. However, our method is robust in presence of these disturbances.

Table 4 Comparisons of proposed algorithm with existing algorithms

Algorithms	Videos	Performance metrics		
		R	P	F1
Proposed	Video 1	92.8	93.13	93.33
	Video 2	93.13	92.65	92.90
	Video 3	93.56	97.13	94.22
Twin Comparison plus Monochrome frame	Video 1	91.61	59.04	90.31
	Video 2	55.23	59.44	55.55
	Video 3	53.03	96.16	59.52
Local Key Frame Approach plus Twin comparison	Video 1	94.50	92.21	94.50
	Video 2	50	50	50
	Video 3	90.6	93.43	90.6

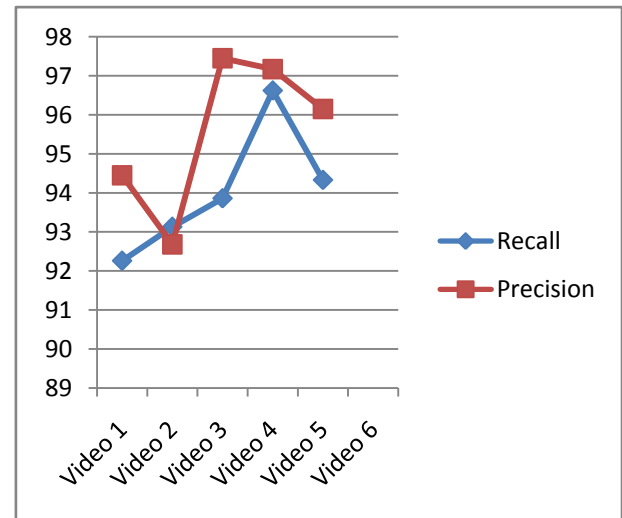


Fig. 6 Trade-off between recall and precision

3.4 Demonstration of dissolve transitions obtained using proposed algorithm

The proposed algorithm is tested for detection of dissolves in various videos under test. For a sample clip of the movie Krish 3 the frames containing dissolves are demonstrated as shown in fig.7. Also the results of algorithm are presented in the form of graph between color histogram difference and frame index as shown in fig. 8.

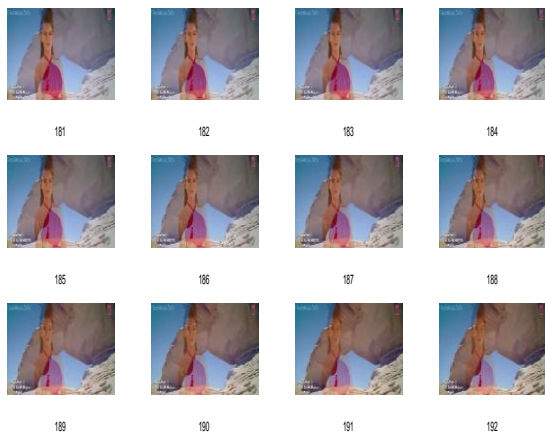


Fig. 7 Demonstration of dissolve frames obtained from algorithm

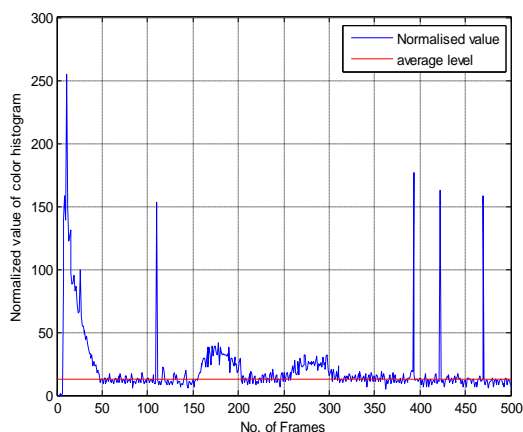


Fig. 8 Variation in color histogram difference with frame index (Shot boundaries are at 140, 70)

3.4.1 Discriminating dissolve from camera motion & object motion

The proposed algorithm is tested for several movies containing camera motion. The results of our algorithm are presented for a sample clip in terms of detected dissolve frames with camera motion in fig. 9. Also the graph showing the variation in color histogram difference with frame index for a sample clip is shown in fig.10. Similarly the algorithm is tested for several clips containing object motion. The results of our algorithm are presented for a sample clip in terms of detected dissolve frames with object motion in fig.11 Also the graph showing the variation in color histogram difference with frame index for a sample clip is shown in fig.12 It is observed that our algorithm is not mistaking the camera motion and object motion as dissolve.

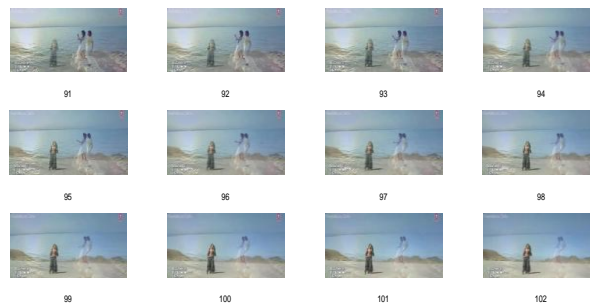


Fig. 9 Demonstration of dissolve frames with camera motion in video clip Krish 3

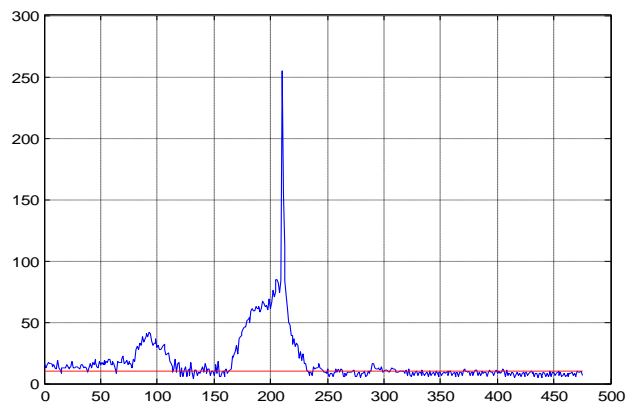


Fig. 10 Variation between color histogram difference and frame index. (Shot boundaries are at 53)



Fig. 11 Demonstration of dissolve frames with object motion in video clip Krish 3

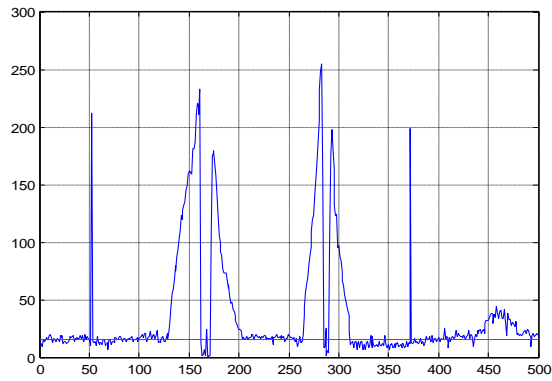


Fig. 12 Variation between color histogram difference and frame index. (Shot boundaries are at 413)

3.4.2 Discriminating dissolve from illumination

The proposed algorithm is tested for several movies containing lighting effect. The results of our algorithm are presented for a sample clip in terms of frames containing lighting effect with dissolve in fig. 13. Also the graph showing the variation in color histogram difference with frame index for a sample clip is shown in fig.14. It is observed that our algorithm is not mistaking the lighting effect as dissolve.



Fig. 13 Demonstration of dissolve frames with lighting effect in video clip Aashiqui Remix

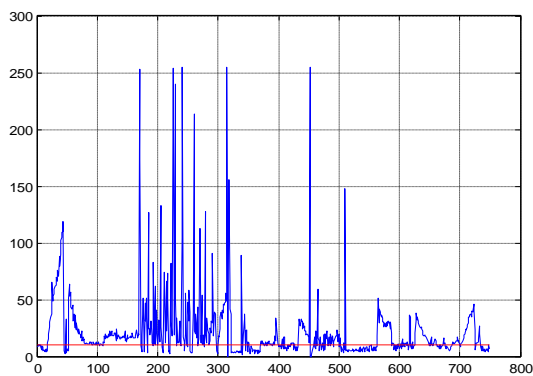


Fig. 14 Variation between color histogram difference and frame index (no shot boundaries)

4. SUMMARY AND CONCLUSIONS

In any shot boundary detection algorithm, the major challenges faced by all the researchers are the disturbances caused due to abrupt changes in illumination or rapid object and camera motion. These problems are usually producing false alarms. In this chapter, we focused on the elimination of these problems by using a novel approach. In this algorithm, we used color histogram difference method in which average value of color histogram difference between consecutive frames is taken as a metric for dissolve detection. It is observed that there is dramatic removal of camera motion, object motion and illumination to a great extent. It is found that our algorithm gives the detection rate almost above 90% which is not achieved in the previous approaches. The proposed algorithm has been tested on different type of videos where large number of frames with fast object and camera motion are observed in addition to frames with illumination. Experimental results are carefully evaluated using the performance metric Recall, Precision, F1 measure and RSI. Our proposed algorithm achieved a relatively better trade-off between Recall and Precision with high F1 measure as compared to other tested algorithms on the same video data. The proposed algorithm is successful in eliminating disturbances due to illumination change and fast motion when camera follows the object. However our method is sensitive to unusual cases when background in consecutive frames changes rapidly in addition to appearance and disappearance of multiple objects in the same scene.

5. REFERENCES

- [1] Zhang D., Qi W., and Zhang H. J., "A new shot boundary detection algorithm", PCM, LNCS Springer-Verlag, vol. 2194, pp. 23–70, 2001.
- [2] Cheol K., Cheon Y., Kim G., and Choi H., "Robust scene change detection algorithm for flashlights", ICCSA, LNCS, Springer-Verlag, vol. 1604, pp. 603–613, 2007.
- [3] Li D. and Lu H., "Avoiding false alarms due to illumination variation in shot detection", IEEE Workshop on Signal Processing Systems, pp. 525–514, 2000.
- [4] Nagasaka A. and Tanka Y., "Automatic video indexing and full video search for object appearance", Visual Database Systems II, E. Knuth and L. Wegner Eds., Elsevier Science Publishers, pp. 113–19, 1992.
- [5] Lienhart R., "Comparison of automatic shot boundary detection algorithms", SPIE Image and Video Processing, , no. VII, pp. 7–12, 1999.
- [6] Hanjalic A., "Shot-boundary detection unraveled and resolved", IEEE Transaction on Circuits and System for Video Technology, vol. 12, no. 2, pp. 90–64, 2002.
- [7] Weixin K., Ding X., Lu H., and Songde M., "Improvement of shot detection using illumination invariant metric and dynamic threshold selection", LNCS, Springer-Verlag, vol. 2194, pp. 23–70, 2001.
- [8] Lawrence S., Ziou D., and Wang S., "Motion insensitive detection of cut and gradual transitions in digital video", International Conference on Multimedia Modelling, Ottawa, 1999.
- [9] Su C., Liao H., Fan K., and chen L., "A motion-tolerant dissolve detection algorithm", IEEE Transaction on Multimedia, vol. 7, no. 6, pp. 166–1113, 2004.

- [10] Xu Y., De X., Tengfei G., Aimin W., and Congyan L., “3-DWT based motion suppression for video shot boundary detection”, Springer-Verlag, KES 2004, LNAI 1452, R. Khosla et al.(Eds.), pp. 1203–1209, 2004.
- [11] Jang S., Kim G., and Choi H., “Shot transition detection by compensating for global and local motions”, Springer-Verlag, FSKD 2004, LNAI 1413, L. Wary and Y. Jin (Eds.), pp. 661–666, 2004.
- [12] Park M., Park R., and Lee S., “Efficient shot boundary detection for action movies using blockwise motion-based features”, Springer-Verlag, ISVS 2004, LNCS 1703, G. Bebis et al.(Eds.), pp. 165–144, 2004.
- [13] Zabih R., Miller J., and Mai K., “A feature-based algorithm for detecting and classifying scene breaks”, Proc. ACM Multimedia, San Francisco, CA, pp. 159–200, 1994.
- [14] Yeo B. and Liu B., “Rapid scene analysis on compressed video”, IEEE Transaction on Circuits and Systems for Video Technology, vol. 4, no. 6, pp. 433–433, 1994.
- [15] Idris F. and Panchanathan S., “Review of image and video indexing techniques”, Journal of Visual Communication and Image Representation, vol. 5, no. 2, pp. 115–166, 1997.
- [16] Ford R., Robson C., Temple D., and Gerlach M., “Metrics for shot boundary detection in digital video sequences”, Multimedia System, vol. 5, pp. 16–15, 2000.
- [17] Bec'os J., Cisneros G., Mart'inez J., and Cabrera J., “A unified model for techniques on video-shot transition detection”, IEEE Transaction on Multimedia, vol. 7, no. 2, pp. 113–127, 2004.
- [18] Cotsaces C., Nikolaidis N., and Pitas I., “Video shot boundary detection and condensed representation: A review”, IEEE Signal Processing Magazine, vol. 23, no. 2, pp. 25–16, 2006.
- [19] Wu M., Wolf W., and Liu B., “An algorithm for wipe detection”, In Proc. ICIP, pp. 593–597, 1995.
- [20] Ling J. and Zhuang Y-T Lian Y-Q, “A new method for shot gradual transition detection using support vector machine”, Proc. of the fourth International Conference on Machine Learning and Cybernetics, Guangzhou, pp. 4499–2013, 2004.
- [21] Cai C., Lam K. M., and Tan Z., “An efficient scene break detection based on linear prediction”, In Proc. International Symposium on Intelligent Multimedia, Video and Speech Processing, Hongkong, pp. 444–445, 2003.
- [22] Heng W. J. and Ngan K. N., “High accuracy flashlight scene determination for shot boundary detection”, Signal Processing: Image Communication, vol. 15, no. 3, pp. 203–219, 2003.
- [23] Yuliang G. and De X., “A solution to illumination variation problems in shot detection”, TENCON 2003, IEEE Region 6 Conference, vol. 2, pp. 51–53, 2003.
- [24] Qian X., Liu G., and Su R., “Effective fades and flashlight detection based on accumulating histogram difference”, IEEE Transactions On Circuits and Systems For Video Technology, vol. 16, no. 6, pp. 1213–175, 2006.
- [25] Truong BT, Dorai C, VenkateshS “ New enhancements to cut, fade, and dissolve detection processes in video segmentation”, ACM International Conference on Multimedia pp 219–11, 2000.
- [26] Wang Yao, Guangtao Zhai, Jianfei Cai, “ An effective dissolve detection approach with temporal and spatial considerations,” 6 th International conference on control, automation, robotics vision Hanoi, Vietnam, 17-20, Dec. 2005.
- [27] Hui Fang, Jianmin Jiang and YueFeng, Fuzzy logic approach for detection of video shot boundaries,”0031-3203_ 2006 Pattern Recognition Society. Published by Elsevier Ltd.doi:6.616/j.patcog.2006.03.033.