

Performance Evaluation of Algorithm for Detection of Fades in Video Sequences in Presence of Motion and Illumination

S. Chavan¹ and S. Akojwar²

¹Department of Electronics and Telecommunication Engg, S.B.J.I.T.M.R Nagpur, M.S, India.

² Department of Instrumentation Engineering, Government College of Engineering, Chandrapur-442403, M.S, India.
{salimsahil97, sudhirakojwar}[at]rediffmail.com

Abstract. The detection of a fade in and a fade out in the video segmentation is very difficult as compared to the detection of abrupt transition. Mostly, the disturbances are caused due to the object/camera motion and the illumination arising from flashlight, fire, flicker and explosion. This mainly results into false positives. An algorithm has been proposed for the fade in and fades out detection. In the proposed algorithm, frames are decomposed to 5 levels using db-6 discrete wavelet transform. The mean of approximate wavelet coefficients are calculated for each frame and normalized. The variation in normalized mean of wavelet coefficient against frame index gives the presence of fade in and fades out. In our algorithm, no preprocessing or post processing is required and so the computational complexity is reduced to a great extent. The processing time is very small as compared to the previous techniques. The proposed algorithm detects and identifies almost all the cases of fade in and fade out with a negligible miss and the false detection causing due to illumination and motion. Performance comparison of the proposed algorithm with the other existing techniques clearly demonstrated its effectiveness in terms of performance metrics such as Recall, Precision, F1 measure, RSI and detection rate. Another advantage of our algorithm over the previous ones is that the detection rate is almost 100%.

Keywords: *Fade in, Fade out, DWT, recall, precision, F1 measure, RSI.*

1 Introduction and related work

In a video V consisting of n shots, finding the beginning and the end of the each shot is known as the shot boundary detection. The Shot boundary detection is also called as the transition detection. The Shot boundary detection is very important for any video analysis and video application as it permits the segmentation of video into its basic components called as shots. A Shot is a video segment in which one continuous action is captured. For video analysis boundary detection between two video shots is a prior step. In video analysis task, segmentation of a video content into a shot is an essential job. For indexing and video analysis shot boundary detection is a very important task. These shot boundaries/transitions can be categorized into two types: Abrupt/Sharp shot transition and Gradual shot transition. Abrupt/Sharp shot transition is a transition from one shot to another shot. It is also called as cuts or hard cuts. The detection of sharp shot transition is trouble free because the two frames of consecutive shots are not related with to each other. In gradual shot transition the intensity of consecutive frames is gradually increased (fade-in) or decreased (fade-out) or one shot is replaced by another shot in a gradual manner. These are also often known as soft transitions. The gradual transitions take account of camera movements such as panning, tilting, zooming etc. Compared to an abrupt a shot transition, a soft transition is a complex task for detection as the difference between two consecutive shots is reduced. Gradual transitions can be of different types such as fade in, fade out, dissolve and wipe.

The elimination of disturbances due to illumination and motion are the major challenges to the shot boundary detection algorithms. Many researchers have discussed on the methods for avoiding false alarms due to illumination variation during detection of shot boundary such as [3], [8], [27], [9], [10], [28], [29], [2] and also due to the impudence of camera and object motion in the video segmentation [14], [15], [16], [17], [18].

B. Iyer, S. Nalbalwar and R.Pawade(Eds.)
ICCASP/ICMMD-2016. Advances in Intelligent Systems Research.
Vol. 137, Pp. 358-366.

© 2017- The authors. Published by Atlantis Press

This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).



It is a challenging task to develop a single approach which will solve both the issues and hence we propose a method for the detection of fade in and fade out in the presence of abrupt illumination, rapid camera and object motion. The proposed method works in three steps. In the first step, the features are extracted from the frame using the discrete wavelet transform. Wherein we used the db- 6 DWT. The Second stage involves finding potential shot boundaries from the extracted features. In the last stage, a threshold is used to declare the shot boundaries.



Fig.1 (a) Fade in frames in the movie once upon a time in Mumbai

Fig.1 (b) Fade out frames in the movie Ashiqui 2

Related work for a shot boundary detection in the presence of illumination, camera and object motion is given below.

Alattar [5] has proposed an algorithm based on the variance of pixel intensities. Presence of fades is given by negative spikes of the pixel intensity variance. This is confirmed after ensuring that the first order difference of the mean of the video sequence is relatively constant next to the negative spike. Fernando et al. [21] proposed the theoretical fade creation model based on luminosity to detect fade and dissolve. Modification on the above approach is suggested by Truong et al. [22]. Dissolves, fade-in, and fade-out are also discussed in [23], [4], [6]. Volkmer et al. [24] have presented a gradual transition detection approach based on the average frame similarity and the adaptive thresholds. Albanese et al. [7] have proposed an algorithm which use similarity metric based on the animate vision theory. Bec'os et al. [19] have analyzed the existing methods and proposed a unified detection model for the gradual transition detection. Han et al. [20] proposed a method which uses techniques of wavelet, FCM clustering, Gaussian weighted Hausdorff distance, similarity of color distribution and motion vector based on 3-D wavelet transform to detect fade, dissolve and wipe. Pedro et al. [25] proposed a fade detection using a non-common feature entropy, a scalar representation of the information contained in each video frame. Bezerra and Leite et al. [26] have transformed the video slice into one-dimensional signals obtaining a highly simplified representation of the video content using longest common subsequence between two strings. They developed a chain of mathematical morphology operations over these signals to detect cut, fade and wipe.

Lawrence et al. [10] have proposed an algorithm for shot boundary detection based on the first-order partial derivatives of video.. They developed algorithm insensitive to object and camera motion by considering areas with low apparent motion. Xu et al. [11] have proposed a motion suppression technique for shot boundary detection based on 3D wavelet transform. Jang et al. [12] have proposed a shot transition detection method in which they extracted motion vectors from two consecutive frames by using the size-variable block matching and then adaptive robust estimation method was used to estimate the global motions and eliminate local motions. Park et al. [13] have proposed two-motion based features for shot boundary detection of fast motion sequences. Li D. and Lu H. [3] has proposed a method to differentiate between shot and flashlights based on the difference in the area of background in edge image. Weixin K. et al. [8] have used color ratio histogram. Histogram difference is considered to be a measure to distinguish abrupt transition from gradual transition [1] and then average intensity difference is used to detect flashlights. Limitations of the existing methods are summarized below:

- Difficult to distinguish fade effects from flashlight, fire, flickers and explosion.
- Difficult to detect fades in presence of fast camera motion and object motion.

- Difficult to detect the fade frames with good detection rate.
- Computational complexity and low processing speed.

The rest of the chapter is organized as follows. Section.2 gives a detailed description of the proposed method for fade detection. The effectiveness of the method is validated by experiments on various videos in Section 3. The results of our algorithm as compared to the results of existing algorithms are presented in Section 3.1. Results of algorithm for fade transitions under various disturbances are demonstrated in Section 3.2. whereas Section 4 gives conclusions.

2 The Proposed Method for detection of fade in and fade out

From the extensive literature survey, it has been noted that the wavelet based approaches are more invariant to illumination and motion. Hence we have focused on developing wavelet based approach in which an innovative technique is used.

The proposed method works in three stages. In the first stage, the features are extracted from the frame using discrete wavelet transform wherein we used db-6 DWT. The second stage involves finding potential shot boundaries from the extracted features. In the last a stage, a threshold is used to declare the shot boundaries.

2.1 Extraction of features of an image using DWT

The possibility of using debauches 6 discrete wavelet transform to find out the features in the frames which are insensitive to disturbances caused by illumination and motion in the shot boundary detection is explored.

The process of extracting features is explained stepwise as follows:

- 1) In the proposed algorithm each frame is converted from RGB to Gray level scale.
- 2) Each frame is decomposed to 5 level using db6 discrete wavelet transform.
- 3) This decomposition leads to the extraction of the frame features.

2.2 Detection of potential shot boundaries

To detect the potential shot boundaries, the discrete wavelet transform is used with an innovative technique which was not applied by any researcher previously. The process of detecting potential shot boundaries is explained stepwise as follows:

- 1) The mean of approximate wavelet coefficient for each frame are tabulated in the form of array 'I' and is obtained by the formula

$$I = \frac{\sum_{row} \sum_{col} average\ wavelet\ coefficients}{M \times N}$$

Here,

I = mean of average wavelet coefficients

M = number of rows

N =number of columns

- 2) The array is then normalized forming the normalized vector 'I_N' as shown in fig.4 and normalized mean is obtained by the formula,

$$I_N = \frac{I}{I_{max}}$$

Here,

I_N = Normalized mean of approximate wavelet coefficients

I_{max} = Maximum value of mean of approximate wavelet coefficients

- 3) The normalized vector 'I_N' is plotted against frame index. For test video the nature of the graph is obtained as shown in Figure 2

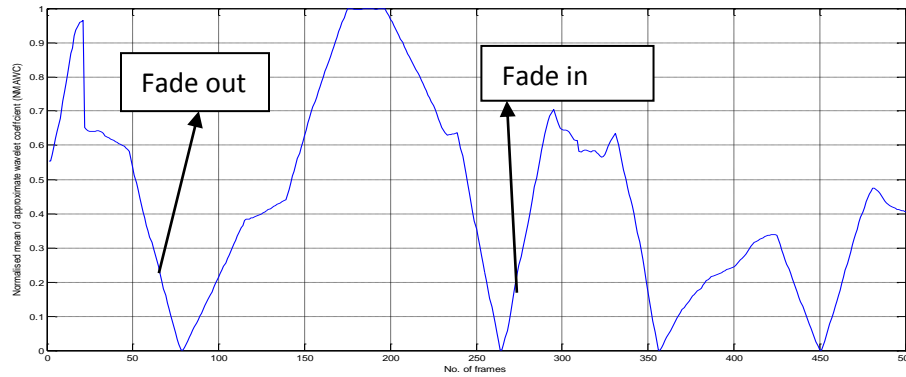


Fig.2 Result of proposed algorithm showing fade in and fade out in test video

- 4) From the graph it is clear that we are getting some valleys. If this valley in the graph goes below 0.1 then the fade in and fade out are present on either side of the valley. Find the index of valley point i.e. I_{\min} .
- 5) The index of '1' or maximum value i.e. I_{\max} is checked from the beginning to minimum point index i.e. I_{\min} .

2.3 Declaration of shot boundaries

After the detection of potential shot boundaries, the declaration of shot boundary as fade in or fade is done as explained below.

- 1) If the index for value '1' or I_{\max} is found then check for next peak point '1' or I_{\max} . ($I_{m-1} > I_m > I_{m+1}$).
- 2) If the peak point index array is empty then no fade out is achieving and if the peak point index array is not empty then choose the last peak point index as a starting point for the fadeout.
- 3) For checking the validity, it is assumed that fade out generally occurs for more than three frames.
- 4) Similarly, the fade in is also detected by finding the peak point index followed by the valley point i.e. I_{\min} index.
- 5) To detect all fade in and fade out present in the video all frames are divided into 'n' overlay in which each overlay consist of 30 frames

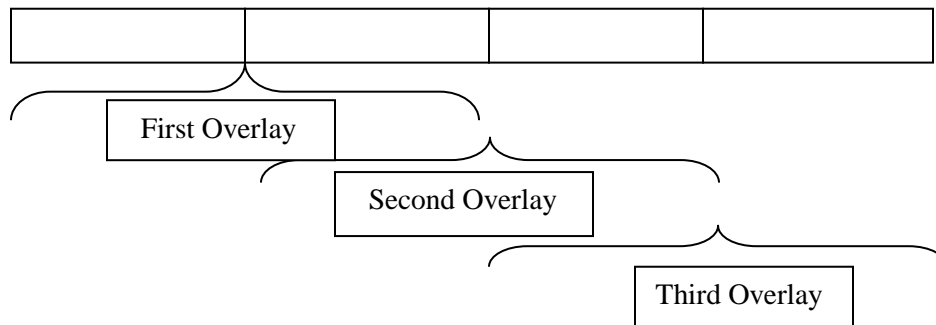


Fig. 3 Overlay structure for boosting processing speed and avoiding miss detection

- 6) First, we check the presence of fade in and fade out in the first overlay i.e. from frame number 1 to 30.
- 7) Then the presence of fade in and fade out is checked for next overlay i.e. from frame number 16 to 45. In this overlay 50% of the frames are being shared. 50% of the overlapping of frames have been taken in each overlay to avoid misdetection of fade in and fade out.

3 Experimental results and discussion

3.1 Test video sequence

To evaluate the effectiveness of the proposed algorithm, experiments are conducted on the six videos mentioned below. We have selected these movies containing lot of camera motion, object motion and illumination, so as to prove robustness of our algorithm in presence of motion and illumination. The proposed algorithm is tested on 1, 44,940 frames. All these frames are observed manually frame to frame with Virtual Dub to note the fade in and fade out transition frames. The six 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 4: “I am in love” Song from “Once upon a time in Mumbai movie”. It contains lighting effect.
- Video 5: “Jumper”. It contains large amount of lighting effect and motion.
- Video 6: “Amplifier”. It contains heavy lighting effect and rapid camera motion.

3.2 Evaluation criteria

We used Recall, Precision, F1Measure and Retrieval Success Index as an evaluation metrics. These performance metrics are explained below.

$$\text{Recall (R)} = \frac{C}{C + M} \quad \dots\dots\dots (1)$$

$$\text{Precision (P)} = \frac{C}{C + FP} \quad \dots\dots\dots (2)$$

Where, $D = C + M$ is the total number of actual shot boundaries (Ground truth) in the test video sequence, C is the number of shot boundaries correctly detected by the algorithm, M is the number shot boundaries missed by the algorithm, and FP is the false positives detected by the algorithm.

Also to rate the performance of the different algorithms, we used the F1 measure [11]. F1 measure is a harmonic mean of recall and precision and is given below

$$\text{F1 measure} = 2 \times \frac{R \times P}{R + P} \quad \dots\dots\dots (3)$$

$$\text{Retrieval Success Index (RSI)} = \frac{C}{C + M + FP} \quad \dots\dots\dots (4)$$

The proposed system has been implemented and applied on variety of video sequences. The detected shot 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 fade in and fade out transitions in the videos under test. Fig.7 and 8 shows the trade-off between Recall and Precision. It shows that in case of videos 1, 4 and 5 precision drops due to rapid object, camera motion and illumination but our algorithm is capable of distinguishing these effects from fade in and fades out. We have taken the videos containing camera motion and illumination and tested our algorithm on such videos. The results obtained for the recall, precision, f1 measure and RSI are presented in Table 2. 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 a camera motion and an illumination. Some frames containing a motion and an illumination obtained from algorithm are also demonstrated in the figures 9, 10,11 and 12. The results of the algorithm are also shown in figures 13, 14 and 15.

Table 1 The manual data set for fade in and fade out

Video Clips	No. of frames	Fade in	Fade out
Video 1	7425	477	637
Video 2	1905	204	229
Video 3	4475	170	200
Video 4	4825	608	638
Video 5	1,22,110	276	278
Video 6	4200	257	253
Total	1,44,940	1992	2235

Table 2 Results in terms of Performance Metrics

Video Collections	Fade in				Fade out			
	Recall	Precision	F1 measure	RSI	Recall	Precision	F1 measure	RSI
Video 1	100	98.55	99.26	98.55	83	98.14	89.93	81.91
Video 2	100	100	100	100	100	100	100	100
Video 3	100	100	100	100	100	100	100	100
Video 4	100	94.26	97.04	94.26	95.92	100	97.91	95.92
Video 5	100	92.30	95.99	92.30	100	98.23	99.10	98.23
Video 6	100	100	100	100	100	100	100	100

3.3 Comparison with different detection methods

We have implemented some existing algorithms as discussed in chapter 4 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 are given in table 3. For showing the detection rate of our algorithm, we have presented the results in the terms of algorithm detected frames and manually observed frames for one sample movie Ashiqui Remix which contains large amount of illumination in the table 4. Table 5 shows the detection rate in the sample clip “Dil tuhi bata” which contains a rapid camera motion.

Table 3 Comparisons of proposed algorithm with existing algorithms

Algorithms	Videos	Fade in			Fade Out		
		R	P	F1	R	P	F1
Proposed	Video 1	100	100	100	100	100	100
	Video 2	100	100	100	100	100	100
	Video 3	100	100	100	100	100	100
Twin Comparison plus Monochrome frame	Video 1	100	92.81	96.27	100	88.01	93.62
	Video 2	100	74.48	85.37	98.98	63.93	77.68
	Video 3	97.14	94.55	84.99	94.44	80	86.62
Local Key Frame Approach plus Twin comparison	Video 1	87.36	88.29	87.92	82.50	90	86.08
	Video 2	78.44	100	87.91	80.50	98.20	88.47
	Video 3	92.96	92.28	92.61	88.20	85.40	86.77

Table 4 Detection rate for sample clip containing light effect for frames 27-277

<i>Effects</i>	<i>Manually detected frames</i>	<i>Algorithm detected frames</i>	<i>Detection rate in percentage</i>
<i>Fade in</i>	676-696	676-696	100
	317-338	317-338	100
	379-402	379-402	100
	507-534	507-534	100
<i>Fade out</i>	286-317	286-317	100
	363-379	363-379	100
	439-507	439-507	100
	648-674	648-674	100

Table 5 Detection rate for sample clip containing rapid camera motion

<i>Effects</i>	<i>Manually detected frames</i>	<i>Algorithm detected frames</i>	<i>Detection rate in percentage</i>
<i>Fade in</i>	191-218	191-217	100
	319-342	319-342	100
<i>Fade out</i>	166-191	166-191	100
	286-315	286-315	100

From the above tables it has been concluded that our algorithm achieves 100% detection rate even in presence of heavy lighting and camera motion

3.4 Demonstration of fade in and fade out transitions obtained using a proposed algorithm

The frames containing fade in, fade out in the videos under test are shown in fig.4a & 4b. fig.5 shows the results of our algorithm implemented in MATLAB for detecting fade in and fade out.



Fig. 4. (a) Demonstration of Fade in frames in video song I am in love (b) Demonstration of fade out frames in the movie once upon a time in Mumbai

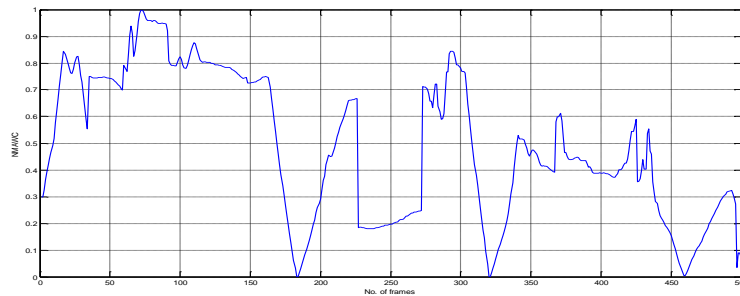


Fig 5 Variation in normalized mean of approximate wavelet co-efficient with frame index (Shot boundaries at frames 161,185,301,321,436 and 460).

4 Conclusion

In any shot boundary detection algorithm, the major challenges faced by all the researchers have been the disturbances caused due to the abrupt changes in illumination or the rapid object and camera motion. These problems are usually producing false alarms. In this paper, we have focused on the elimination of these problems by using a novel approach. In our algorithm, we have used the db-6 discrete wavelet transform in which each frame is decomposed to 5 levels. It is observed that there has been a dramatically removal of the camera motion, the object motion and an illumination to a great extent. Then we applied an innovative approach for detecting the presence of fade in and fade out in the video. It is observed that our algorithm gives the detection rate almost above 90% which is not achieved in the previous approaches. Also the processing speed is high and it is computationally less complex to a great extent. Our proposed algorithm achieved a relatively better trade-off between Recall and Precision with high F1 measure as compared to the other tested algorithms on the same video data. The proposed algorithm is successful in eliminating disturbances due to illumination change and fast motion when a camera follows the object. However our method is sensitive to unusual cases when the background in the consecutive frames changes rapidly in addition to the appearance and the disappearance of the multiple objects in the same scene. Since the main applications of shot boundary detection are in video indexing and retrieval, we have calculated retrieval success index (RSI). Very few researchers have justified by giving attention towards this parameter. It is concluded that the RSI for all the videos under the test is also above 80%. It clearly means that our algorithm is effective for the video retrieval applications.

References

- [1] Zhang D., Qi W., and Zhang H. J., "A new shot boundary detection algorithm", PCM, LNCS Springer-Verlag, vol. 2195, pp. 63–70, 2001.
- [2] Cheol K., Cheon Y., Kim G., and Choi H., "Robust scene change detection algorithm for flashlights", ICCSA, LNCS, Springer-Verlag, vol. 4705, pp. 1003–1013, 2007.
- [3] Li D. and Lu H., "Avoiding false alarms due to illumination variation in shotdetection", IEEE Workshop on Signal Processing Systems, pp. 828–836, 2000.
- [4] Drew M. S., Li Z-N, and Zhang X., "Video dissolve and wipe detection via spatio-temporal images of chromatic histogram differences", In Proc. IEEE Int. Conf. Image Process., vol. 3, pp. 929–932, 2000.
- [5] Alattar A. M., "Detecting fade regions in uncompressed video sequences", in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, pp. 3025–3028, 1997.
- [6] Lienhart R., "Reliable dissolve detection", In Proc. SPIE Storage and Retrievalfor Media Databases, vol. 4315, pp. 219–230, 2001.
- [7] Albanese M., Chianese A., Moscato V., and Sansone L., "A formal model forvideo shot segmentation and its application via animate vision", MultimediaTools and Applications, vol. 24, pp. 253–272, 2004.
- [8] Weixin K., Ding X., Lu H., and Songde M., "Improvement of shot detectionusing illumination invariant metric and dynamic threshold selection", LNCS, Springer-Verlag, vol. 2195, pp. 63–70, 2001.

- [9] Lienhart R., "Reliable transition detection in videos: a survey and practitioner's guide", *Int. J. Image Graph*, vol. 1, no. 3, pp. 469–486, 2001.
- [10] 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.
- [11] Xu Y., De X., Tengfei G., Aimin W., and Congyan L., "3-DWT based motionsuppression for video shot boundary detection", *Springer-Verlag, KES 2005, LNAI 3682*, R. Khosla et al.(Eds.), pp. 1204–1209, 2005.
- [12] Jang S., Kim G., and Choi H., "Shot transition detection by compensating forglobal and local motions", *Springer-Verlag, FSKD 2005, LNAI 3614*, L. Wary and Y. Jin (Eds.), pp. 1061–1066, 2005.
- [13] Park M., Park R., and Lee S., "Efficient shot boundary detection for action movies using blockwise motion-based features", *Springer-Verlag, ISVS 2005, LNCS 3804*, G. Bebis et al.(Eds.), pp. 478–485, 2005.
- [14] Shahraray B., "Scene change detection and content-based sampling of video sequences", *Digital Video Compression : Algorithms and Technologies, SPIE*, vol. 2419, pp. 2–13, 1995.
- [15] Zabih R., Miller J., and Mai K., "A feature-based algorithm for detecting and classifying scene breaks", *Proc. ACM Multimedia*, San Francisco, CA, pp. 189–200, 1995.
- [16] Yeo B. and Liu B., "Rapid scene analysis on compressed video", *IEEE Trans action on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 533–544, 1995.
- [17] Idris F. and Panchanathan S., "Review of image and video indexing techniques", *Journal of Visual Communication and Image Representation*, vol. 8, no. 2, pp. 146–166, 1997.
- [18] Ford R., Roboson C., Temple D., and Gerlach M., "Metrics for shot boundary detection in digital video sequences", *Multimedia System*, vol. 8, pp. 37–46, 2000.
- [19] 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. 293–307, 2005.
- [20] Han B., Ji H., and Gao X., "A 3D wavelet and motion vector based method for wipe transition detection", *Proc. ICSP*, , no. 2, pp. 1207–1210, 2004.
- [21] Fernando WAC, Canagarajah C. N., and Bull D. R., "Fade and dissolve detection in uncompressed and compressed video sequences", *In Proc. of IEEE, ICIP*, pp. 299–303, 1999.
- [22] Truong B. T., Dorai C., and Venkatesh S., "New enhancements to cut, fade, and dissolve detection processes in video segmentation", *In ACM Multimedia*, pp. 219–227, 2000.
- [23] Lienhart R. and Zaccarin A., "A system for reliable dissolve detection in video", *In Proc. IEEE ICIP, Greece*, pp. 406–409, 2001.
- [24] Volkmer T., Tahaghoghi S. M. M., and Williams H. E., "Gradual transition detection using average frame similarity", *Proc. IEEE Computer Society Conf. on Computer Vision and pattern recognition, CVPRW04*, 2004.
- [25] Pedro J. S., Dom'inguez S., and Denis N., "On the use of entrophy series for fade detection", *CAEPIA2005, LNAI4177*, R. Mar'in et al. (Eds.), Springer-Verlag, pp. 360–369, 2005.
- [26] Bezerra F. N. Leite N. J., "Using string matching to detect video transitions", *Pattern Anal. Application*, vol. 10, pp. 45–54, 2007.
- [27] Ruiloba R., Joly P., Marchand-Maillet S., and Quenot G., "Towards a stan dard proptocol for the evaluation of video-to-shots segmentation algorithms", *European Workshop on[Content Based Multimedia Indexing, Toulouse*, 1999.
- [28] Daubechies I., "Ten lectures on wavelets", *Society for Industrial and Applied Mathematics, CBMS-NSF Regional Conference series in Applied Mathematics*, no. 61, 1992.
- [29] Ford R. M., Temple D., and Gerlach M., "Metrics for shot boundary detection in digital video sequence", *Multimedia Systems*, vol. 8, pp. 37–46, 2000.
- [30] Zuzana ˇcernekov'a and Ioannis Pitas, "Information theory-based shot cut/fade detection and video summarization", *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 16, no. 1, pp. 82–91, 2006.