THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

Comparison of Algorithms for Cut and Fade Transition Detection for Video

Upesh Patel
Assistant Professor, Department of Electronics & Communication Engineering,
CHARUSAT University, Changa, Gujarat, India
Brijesh Kundaliya
Assistant Professor, Department of Electronics & Communication Engineering,
CHARUSAT University, Changa, Gujarat, India
Kaushal Vala
Assistant Professor, Department of Electronics & Communication Engineering,
CHARUSAT University, Changa, Gujarat, India

Abstract:
Mutual information is common information and similarity metric between two consecutive frames. In this work mutual information is also used as similarity metric of video frames. Mutual information is used mainly in cut detection; its gradual transitions detection performance is so well so, dynamic window-based mutual information shot boundary detection method is applied. This work presents a cut and fade transitions detection method based on mutual information and joint entropy. The experiments based on the proposed method show it reduces effectively shot error detection. Also we use mutual information and joint entropy to detect fade in/out in experiments. Cumulative mutual information use for detect dissolve effect from video.

Keywords: *Cut and fade transitions, mutual information, joint entropy, dissolve boundary, recall, precision, short, medium and long range video*

1. Introduction to Algorithm

1.1. Mutual Information

Mutual information as common information between frames the bigger difference in content the smaller mutual information. We can measure mutual information between frames, and then determine which position shot transitions exists. We can detect cut with mutual information, but gradual transitions detection performance is not good. For gradual transition, we consider not only two consecutive frames but take into account all frames within a certain temporal window W. In this case, the mutual information and joint entropy between two frames are calculated separately for each of the RGB Component. For this case C R (i, j), 0 ≤ i, j ≤ N − 1(N being the number of gray levels in the frame), corresponds to the probability that a pixel with gray level i in frame ft has gray level j in frame ft +1. In other words, C t, t +1 (i, j) equals to the number of pixels which change from gray level i in frame ft to gray level j in frame ft +1, divided by the total number of pixels in the video frame. [10]

The mutual information Ik, l of frames fk, fl for the R component is expressed as,

$$I^{R}_{k,l} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} C^{R}_{k,l}(i, j) \log\left(\frac{C^{R}_{k,l}(i, j)}{C^{R}_{k}(i) C^{R}_{l}(j)}\right)$$

The total mutual information (MI) calculated between frames fk, fl is defined as

$$I(f_k, f_l) = I^{R}_{k,l} + I^{G}_{k,l} + I^{B}_{k,l}$$

The joint entropy Ik, lof framesfk, fl for the R component is expressed as

$$H^{R}_{k,l} = -\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} C^{R}_{k,l}(i, j) \log(C^{R}_{k,l}(i, j))$$

The total joint entropy calculated between frames fk,fl is defined as

$$H_{k,l} = H^{R}_{k,l} + H^{G}_{k,l} + H^{B}_{k,l}$$

In order to detect gradual transitions boundary, we set a temporal window W of size Nw, which is centered around frames fi, fi+1, we calculate the mutual information between (N w/ 2)2 pairs of frames shown in Fig. 1.

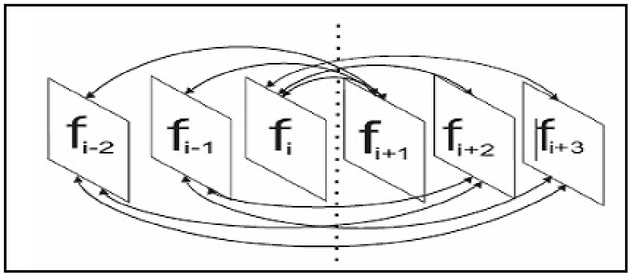

Figure 1: Diagram showing pairs of frames, which contribute to Icumm(i) for a window size of Nw = 6

Then a cumulative measure which combines information from all these frame pairs is calculated as Follows [9]:

$$I_{cumm}(i) = \sum_{k=i-\sigma}^{i} \sum_{l=i+1}^{i+\sigma} I(f_k, f_l)$$

Where σ=NW/2 is half the size of the temporal window.
The procedure is repeated for the whole video sequence by sliding the window over all frames. The terms of the sum provide information on shot gradual transition detection [10].

1.2. Joint Entropy
The average information per pixel of an image is called its entropy.
Entropy is a measure of the distance between two probability distributions. If an image T have N number of pixels, the ith pixel (i=1, 2... n) repeats hi times, then hi/N, thus the entropy H (T) of a source A is defined by [11].

$$H(T) = \sum_{i=1}^{n} P_i \log(\frac{1}{P_i})$$

The gray level histogram of image may count up the number of pixel of each gray level. It's useful in probability of image gray level. Example is, the histogram of two consecutive images A and B is H (i) and H (j) respectively, the probability is PA (i) and PB (j), thus

$$P_A(i) = \frac{H(i)}{\sum_A H(i)}$$

$$P_B(j) = \frac{H(j)}{\sum_B H(j)}$$

The joint histogram of two consecutive images A and B may count the number of pixel-pair of two images in corresponding position. The gray level of image A and image B is i and j respectively so probability is, [11]

$$P_{AB}(i, j) = \frac{H(i, j)}{\sum_{i,j} H(i, j)}$$

The marginal probability distribution is represented by joint probability.

$$P_A(i) = \sum_j P_{AB}(i, j) \quad P_B(j) = \sum_i P_{AB}(i, j)$$

If PA(i) and PB(j) is probability of image A and B respectively. PAB (i,j) is joint probability of image A and B, then the joint entropy is

$$H(A, B) = \sum_{i,j} P_{AB}(i, j) \log(\frac{1}{P_{AB}(i, j)})$$

Thus, the mutual information I (A, B) between frame A and B base on entropy is [J-5],
$I(A, B) = H(A) + H(B) - H(A, B)$
In the RGB image have three components red, green and blue so find mutual information for red plane, green plane and blue plane separately is represented as,

$$I_{t,t+1} = I^R_{t,t+1} + I^G_{t,t+1} + I^B_{t,t+1}$$

Similarly, the total joint entropy is [11]

$$H_{t,t+1} = H^R_{t,t+1} + H^G_{t,t+1} + H^B_{t,t+1}$$

This is equation for joint entropy, based on that we find transition between two consecutive shot.

2. Simulation Results

2.1. Test Data and Evaluation Criteria: -

The performances of the implemented algorithms are evaluated based on the recall and precision criteria. Recall is defined as the percentage of desired items that are retrieved. Precision is defined as the percentage of retrieved items that are desired items.

$$\text{Re}\ call = \frac{Correct}{Correct + Missed}$$

$$\text{Pr}ecision = \frac{Correct}{Correct + FalsePositives}$$

In order compare the overall performance of the algorithms, F measure, which combines recall and precision results with equal weight, is adopted:

$$F_1(recall, precision) = \frac{2 \times recall \times precision}{recall + precision}$$

In the Simulation Results sections, the results are presented in an accumulated manner so that we will be able to compare the results of the SBD algorithm under test with the results of the previously tested algorithms

2.2. Simulation Results for Mutual Information and Joint Entropy

<u>2.2.1. Small Ranges Video Clip</u>

2.2.1.1. Cut Boundary

Simulation result for mutual information. In this figure mutual information is small if cut is detected. In this small video five cut transitions are detected so at five places the mutual information is small.

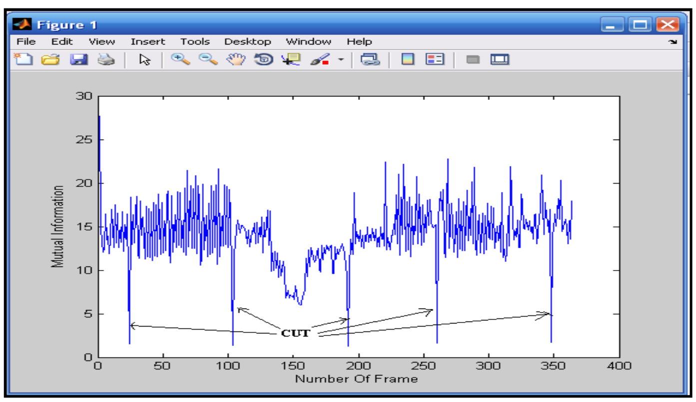

Figure 2: Mutual Information of cartoon clip for cut transition

Simulation result for joint entropy. In this figure joint entropy is high if cut transition is detected.

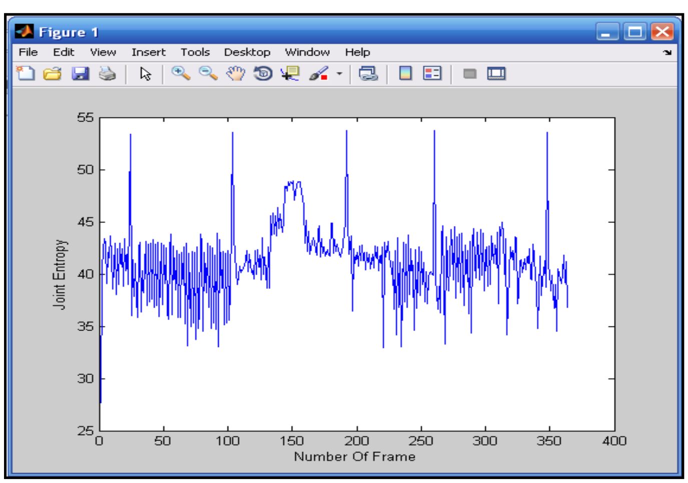

Figure 3: Joint Entropy for cartoon clip

2.2.2.2. Fade Boundary

For fade transition take same cartoon clip and check out the difference in result. if fade transition detected the value is down. See mutual information is slowly down during fade transition.

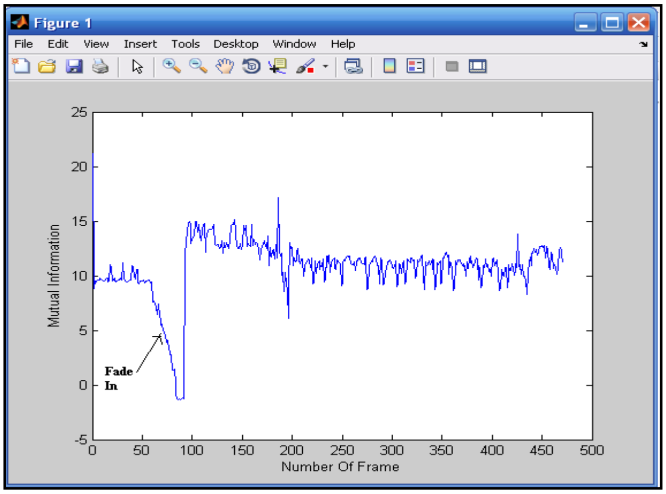

Figure 4: Mutual Information for fade effect

if fade transition detected the value is down. See joint entropy is slowly down during fade transition.

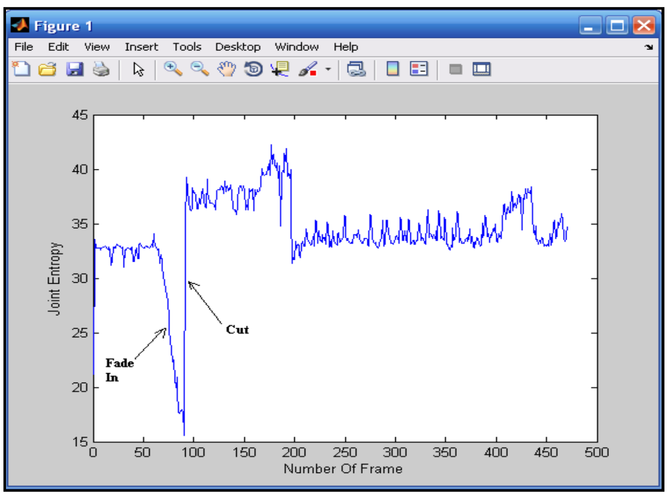

Figure 5: Joint Entropy for fade effect

2.2.2.3. Dissolve Boundary

For dissolve boundary, set a temporal window W of size NW, which is centered around frames fi, fi+1, calculate the mutual information between (NW/2)2 pairs of frames.

Then cumulative measure which combines information from all these frame pairs is calculated as follows:

$$I_{cumm}(i) = \sum_{k=i-\sigma}^{i} \sum_{l=i+1}^{i+\sigma} I(f_k, f_l)$$

Where σ = NW/2 is half the size of the temporal window.

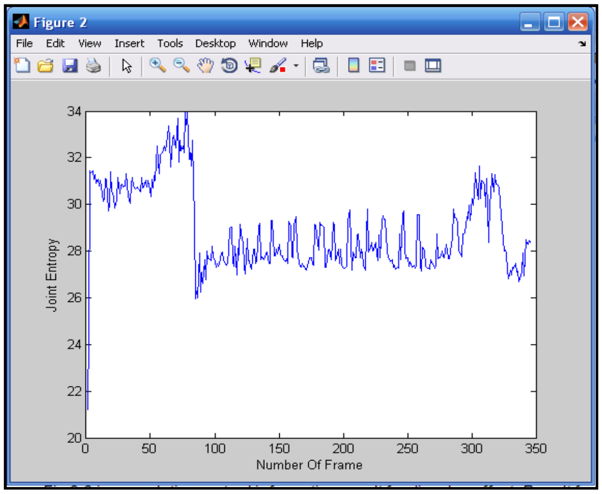

Figure 6: MI for Dissolve

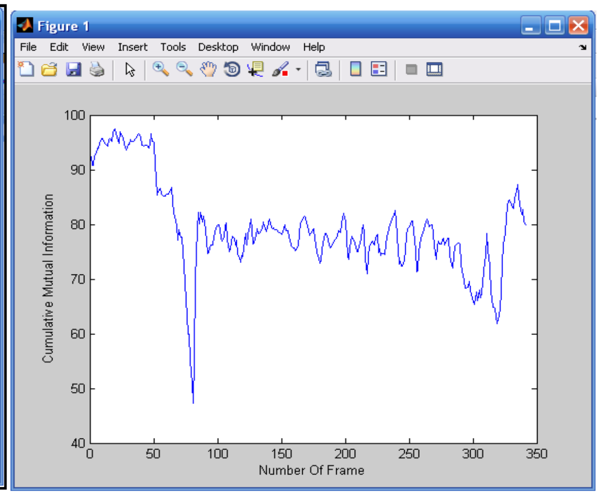

Figure 7: For Icumm with window size Nw =6.

Next results are from medium range of videos with uncompressed (.avi) format.

2.2.2. Medium range video clip (6000 frames)

For Uncompressed (.avi) format

- Movie – Die-Hard 1[1988]
- Frame – 1-6000
- Format - .avi (Uncompress)
- Algorithm – Mutual Information

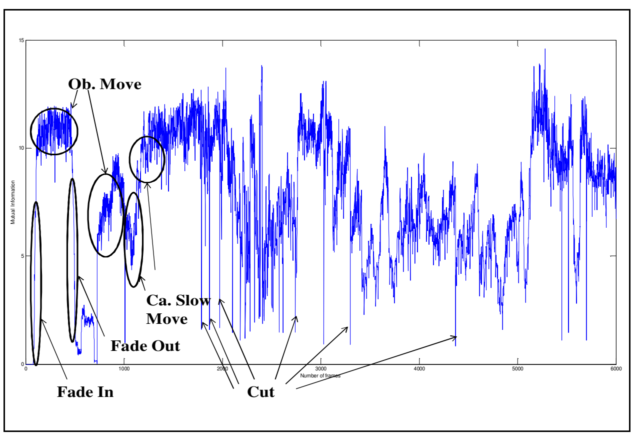

Figure 8: Total 1-6000 Frames]

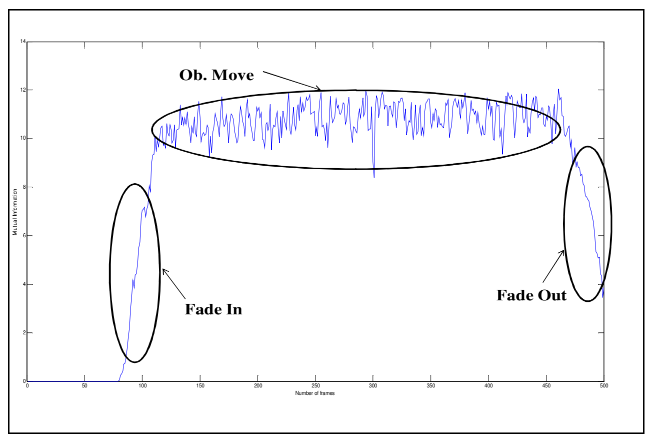

Figure 9: Frame 1-500]

Fig (8) is mutual information result for range 1-6000 from movie Die-Hard 1.

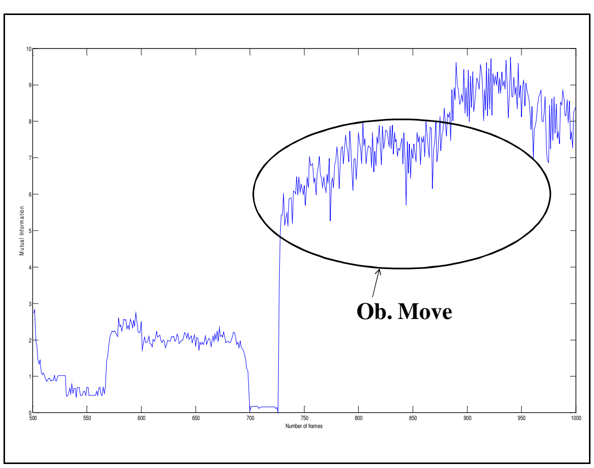

Figure 10: Frame 500-1000]

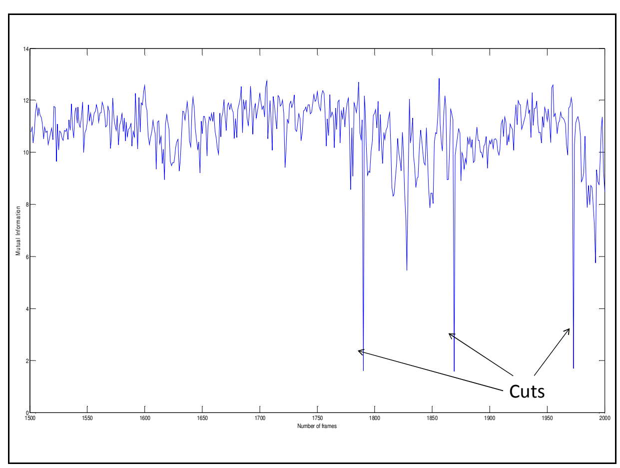

Figure 11: Frame 1500-2000]

Mutual information result for fade-in, fade-out, camera movement, object movement, cut from fig 8. is observed. Mutual information value is zero if frame is completely black.
Mutual information near to zero if fade or black frame detect in video and if cut transition present in video mutual information value is very small at that place, see in fig (10) mutual information is very small.

- Movie – Die-Hard 1[1988]
- Frame – 1-6000
- Format - .avi (Uncompress)
- Algorithm – Joint Entropy

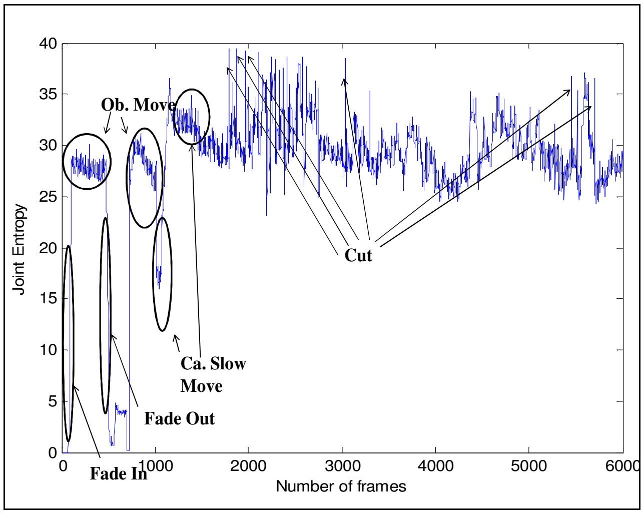

Figure 12: Total Frames]

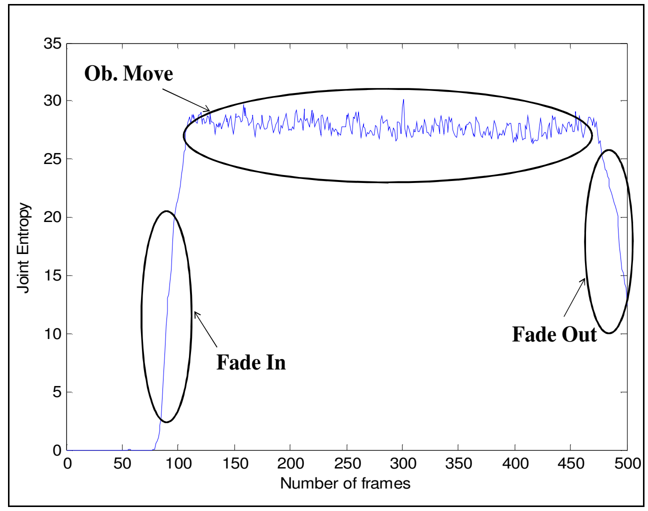

Figure 13: Frame 1-500]

<u>2.2.3. Large Range Video Clip</u>

large range of three different videos has divided in some range of frames like: 1-90000 frames from Die-Hard 1, 1-90000 frames from Shrek-1 and 1-86500 frames from The Last Air-Bender.
This large range first calculates cut transition, fade transition and dissolve transition manually with use of virtual dub software.
Cut transition is detected from video with use of joint entropy because mutual information is down to zero if cut transition in video and joint entropy is higher if cut transition in video.
If use mutual information for cut boundary and set threshold for that, black frame also calculates in cut transition because mutual information less then threshold means cut transition is there.
First find joint entropy of video, after that for threshold calculate mean μ and standard deviation σ of the feature vector distances as follows. [12]

$$\mu = \frac{1}{N-1}\sum_{i=1}^{N-1} H(f_k, f_{k+1})$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1} (H(f_k, f_{k+1}) - \mu)}$$

Where N is number of frames in the video sequence and H is joint entropy between two consecutive frames.

$$T = \mu + A\sigma$$

Where, A is pre-defined constant [c].

$$H(f_k, f_{k+1}) > T$$

If joint entropy value is higher than threshold means cut transition is there. First only use sum of mean and standard deviation means set constant value is one. Then change constant value find cut transition value.

2.2.3.1. Cut Transition

- Movie: – Die-Hard 1

Cut transition value is 744 In Die-Hard 1 calculated with use of virtual dub software. 1-90000 frames from this movie

Constant Value (A)	Threshold Value (T)	Cut transition
1	33.5248	12702
1.5	35.7828	4764
1.75	36.9118	2430
1.85	37.3633	1814
2.04	38.2213	927
2.06	38.3116	861
2.065	38.3342	851
2.07	39.1390	833
2.075	38.3794	811
2.08	38.4020	799
2.085	38.4245	786
2.095	**38.4697**	**745**

Table 1: Cut transition with threshold value for Die-Hard 1

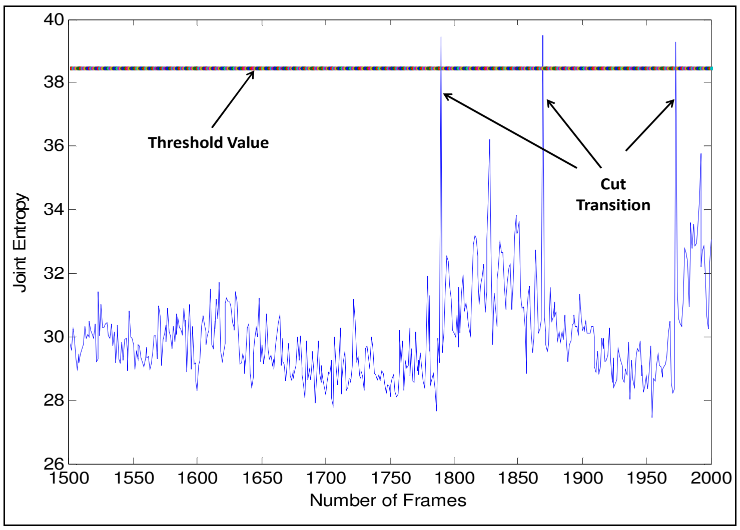

Figure 14: Cut transition with threshold value for Die-Hard 1

It is clear from fig that joint entropy is greater than threshold if cut transition is there.

- Movie: - Shrek-1

Cut transition value is 974 in shrek-1, this is cartoon movie. I have taken 1-9000 frames from this movie.

Constant Value (A)	Threshold Value (T)	Cut transition
1	34.9915	13569
1.5	36.9296	7105
2	38.8677	2682
2.15	39.4491	1807
2.3	40.0306	1226
2.35	40.2244	1048
2.36	40.2631	1019
2.37	40.3019	992
2.377	**40.3290**	**976**

Table 2: Cut transition with threshold value for Shrek-1

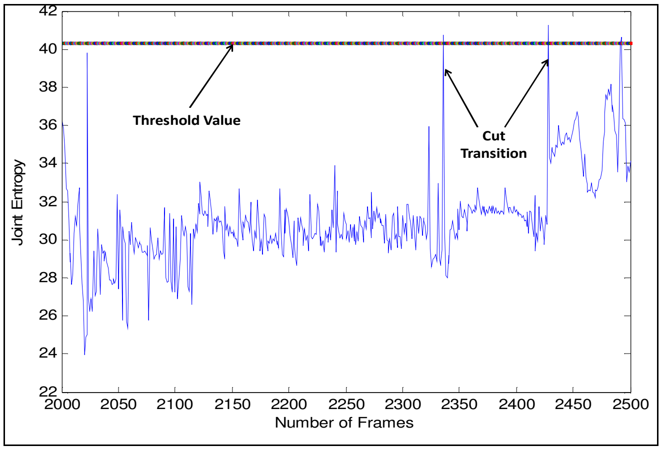

Figure 15: Cut transition with threshold value for Shrek-1

Threshold method detects good cut transition result but some time transition may occur at less than threshold value.

- Movie: - The Last Air-Bender

Cut transition value is 468 in The Last Air-Bender, this is entertainment movie. I have taken 1-9000 frames from this movie.

Constant Value (A)	Threshold Value (T)	Cut transition
1	34.6852	12874
1.5	37.1875	3554
1.6	37.6880	2351
1.7	38.1885	1609
1.8	38.6890	1016
1.85	38.9392	807
1.9	39.1894	632
1.95	39.4397	477
1.952	39.4497	472
1.953	39.4547	470
1.955	**39.4647**	**462**

Table 3: Cut transition with threshold value for The Last Air-Bender

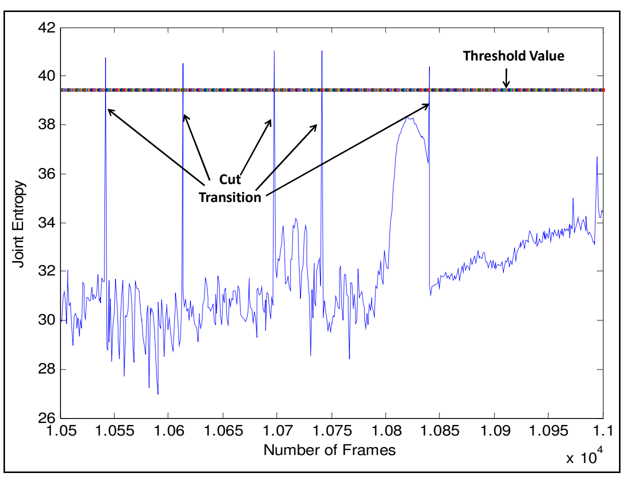

Figure 16: Cut transition with threshold value for the last Air-bender

This is figure for The Last Air-Bender cut transition detects with threshold. Observe from figure cut transition is greater than threshold value.

2.2.3.2. Fade Transition

The key problem of FOI detection is the recognition of monochrome frame, since there is at least one monochrome frame within the FOI transition but monochrome frame seldom appears elsewhere. One dominant characteristic of monochrome frame is its low mutual

information of consecutive frame. Thus mutual information feature is utilized in FOI detection process. FOI detection process is described as follows.
Step 1: - Detect monochrome frame. A monochrome frame is declared once its mutual information feature is below a given threshold, which is heuristically determined.
Step 2: - Judge the type of entering transition, abrupt or gradual. If it is gradual, search the fade out boundary of the previous shot.
Step 3: - Judge the type of exiting transition, abrupt or gradual. If it is gradual, track the fade in boundary of the next shot.
Threshold value used for fade transition is shown below: -

$$T = \mu + \sigma$$

For fade transition we need lower threshold value that's why I used scaling factor for decrease threshold value.

$$T_f = \frac{sf}{100} * T$$

Where sf (scaling factor) = 1, 2, 3 ……
Die-Hard 1, Shrek-1 and Air-Bender movies have 3, 4 and 10 fade transitions respectively. Calculate scaling factor and threshold for every movie. Find how many fade transition is correctly detected from this method; result is so good.

Name of Movie	Threshold (T=μ+σ)	Scaling Factor	T_f (Fade Threshold)	Desired Fade	Detected Fade
Die-Hard 1	10.6915	2	0.2138	3	5
Shrek-1	12.5658	2	0.2513	4	7
Air-Bender	10.9490	3	0.3285	10	15

Table 4: Threshold with Scaling Factor for Fade Transition

Movie	Desired	Detected	Correct	Miss	F-P	R	P	F1
Die-Hard 1	3	5	2	1	2	0.66	0.5	0.57
Shrek-1	4	7	4	0	3	1	0.57	0.72
Air-Bender	10	15	9	1	5	0.99	0.64	0.75

Table 5: Recall, Precision and F1 for fade transition

From the figures and tables data it is clear that Mutual information gives better results against fade transition. But some time completely black or white frame enter in video it also gives smaller mutual information and algorithm calculate this frame in fade transition.
Mutual information gives low value base on that I find fade transition and it gives good result for correct fade transition. Observe from table we miss minimum value of fade transition.

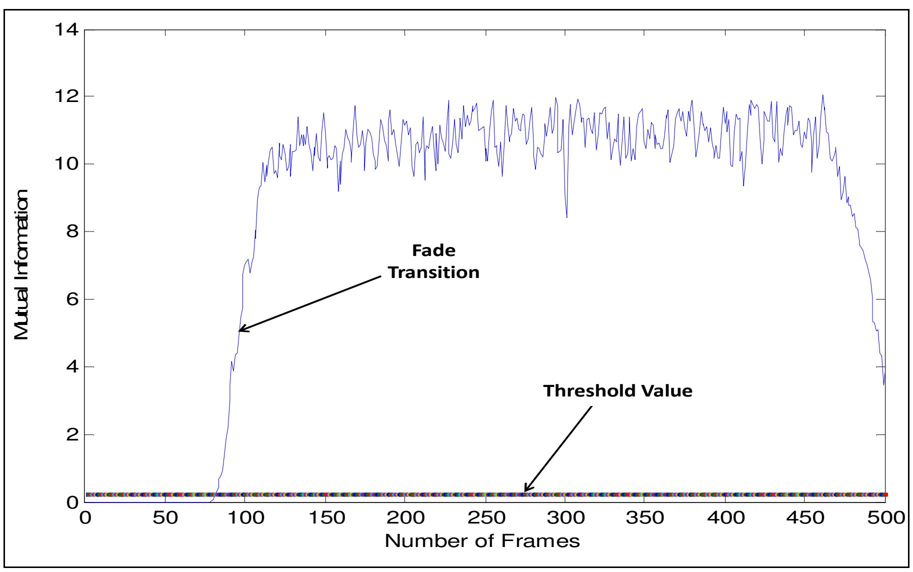

Figure 17: Fade transition with threshold value from Die-Hard 1

2.2.3.3. Dissolve Transition
For Dissolve Transition I used Cumulative Mutual Information with window size (NW) 6.

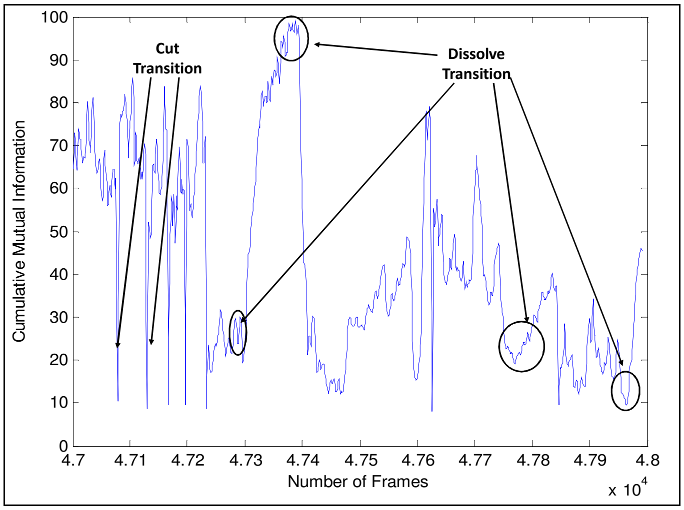

Figure 18: Dissolve transition with window size (NW) = 6

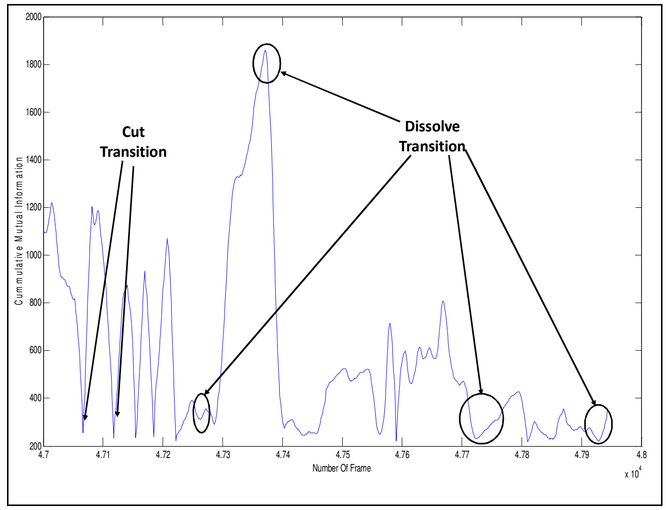

Figure 19: Dissolve transition with window size (NW) = 30

These two figures are cumulative mutual information for dissolve transition with different window size. The last Air-Bender movie used for dissolves effect with Range 47000-48000.

Observe from fig if we increase the window size of cumulative mutual information result is clear for dissolve effect. The difference between figures can be observed easily.

Movie	Frame Range	Algo.	Effect	Desired	Correct	Miss	F. P.	R	P	F1
Shrek	1 To 6000	MI/JE	Cut	57	48	9	5	0.84	0.905	0.87
			Fade	4	3	1	3	0.75	0.5	0.6
		Icumm	Dissolve	1	1	0	2	1	0.333	0.5
	37800 To 45300	MI/JE	Cut	53	47	6	12	0.88	0.796	0.83
			Fade	-	-	-	-	-	-	-
		Icumm	Dissolve	9	7	2	3	0.77	0.7	0.73
The Last Air Bender	31000 To 37000	MI/JE	Cut	28	23	5	6	0.82	0.793	0.80
			Fade	3	2	1	2	0.66	0.5	0.57
		Icumm	Dissolve	2	1	1	1	0.5	0.5	0.5
	46000 To 52000	MI/JE	Cut	25	24	1	9	0.96	0.727	0.82
			Fade	-	-	-	-	-	-	-
		Icumm	Dissolve	6	5	1	2	0.83	0.714	0.76
Die Hard 1	1 To 6000	MI/JE	Cut	22	18	4	26	0.81	0.409	0.54
			Fade	5	4	1	4	0.8	0.5	0.61
		Icumm	Dissolve	-	-	-	-	-	-	-
	96000 To 102000	MI/JE	Cut	69	55	14	14	0.79	0.79	0.79
			Fade	-	-	-	-	-	-	-
		Icumm	Dissolve	-	-	-	-	-	-	-

Table 6: Analysis of Mutual Information Algorithm

Movie	Frame Range	Algo.	Effect	Desired	Correct	Miss	F. P.	R	P	F1
Shrek	1 To 6000	MI/JE	Cut	57	52	5	4	0.9122	0.9285	0.9203
			Dissolve	1	0	1	-	0	0	0
		Icumm	Fade	4	3	1	-	0.75	1	0.8571
	37800 To 45300	MI/JE	Cut	53	45	8	2	0.8490	0.9574	0.9
			Dissolve	9	7	2	-	0.7777	1	0.875
		Icumm	Cut	28	28	0	-	1	1	1
The Last Air Bender	31000 To 37000	MI/JE	Dissolve	2	1	1	-	0.5	1	0.6666
			Fade	3	3	0	-	1	1	1
		Icumm	Cut	25	23	2	4	0.92	0.8512	0.8846
	46000 To 52000	MI/JE	Dissolve	6	4	2	-	0.6666	1	0.8
			Cut	22	20	2	5	0.9090	0.8	0.8510
		Icumm	Fade	5	2	3	-	0.4	1	0.5714
Die Hard 1	1 To 6000	MI/JE	Cut	69	65	4	3	0.9420	0.9558	0.9489
			Cut	57	52	5	4	0.9122	0.9285	0.9203
		Icumm	Dissolve	1	0	1	-	0	0	0
	96000 To 102000	MI/JE	Fade	4	3	1	-	0.75	1	0.8571
			Cut	53	45	8	2	0.8490	0.9574	0.9
		Icumm	Dissolve	9	7	2	-	0.7777	1	0.875

Table 7: Analysis of Block-Based χ^2 Histogram Algorithms

Movie	Frame Range	Effect	Desired	Correct	Miss.	F. P.	R	P	F1
Shrek	1 To 6000	Cut	57	54	3	2	0.94	0.96	0.95
		Dissolve	1	1	0	0	1	1	1
		Fade	4	3	1	0	0.75	1	0.85
	37800 To 45300	Cut	53	51	2	7	0.96	0.87	0.91
		Dissolve	9	7	2	3	0.77	0.7	0.73
The Last Air Bender	31000 To 37000	Cut	28	27	1	6	0.94	0.81	0.88
		Dissolve	2	2	0	2	1	0.5	0.66
		Fade	3	2	1	0	0.66	1	0.8
	46000 To 52000	Cut	25	25	0	4	1	0.86	0.92
		Dissolve	6	4	2	3	0.66	0.57	0.0.61
Die Hard	1 To 6000	Cut	22	22	0	3	1	0.88	0.93
		Fade	5	4	1	0	0.8	1	0.88
	96000 To 102000	Cut	69	68	1	8	0.98	0.89	0.93

Table 8: Analysis of Color-Histogram Difference Algorithms

Movie	Frames	Effects	Desired	Correct	Missed	F. P.	Recall	Precision	F1
The Last Air bender	(31000-37000)	Cut	28	21	7	-	0.75	1	0.857
		Dissolve	2	1	1	-	0.5	1	0.666
		Fade	3	2	1	-	0.666	1	0.8
	(46000-52000)	Cut	25	25	0	5	1	0.833	0.909
		Dissolve	6	4	2	-	0.666	1	0.8
		Fade	0	0	0	1	-	0	-
The Shrek	(1-6000)	Cut	57	57	0	3	1	0.95	0.974
		Dissolve	1	1	0	1	1	0.5	0.666
		Fade	4	2	2	-	0.5	1	0.666
	(37800-45300)	Cut	40	40	0	5	1	0.8889	0.941
		Dissolve	9	4	5	-	0.444	1	0.615
The Die Hard1	(1-6000)	Cut	21	19	2	-	0.904	1	0.95
		Fade	2	1	1	-	0.5	1	0.666
	(96000-102000)	Cut	37	37	0	2	1	0.948	0.973

Table 9: Analysis of χ^2-Histogram Algorithms

3. Conclusion
Compare with Different algorithms like Block-Based X2 Histogram Algorithm, Color-Histogram Difference Algorithm and X2-Histogram Algorithm. it is concluded color histogram difference and block base X2 histogram method good for cut transition. Mutual Information is good for fade transition. For Dissolve Cumulative mutual information give good result.

4. References

i. CHENG W. G., LIU Y. M. and XU D. " Shot boundary detection based on the knowledge of information theory", IEEE International Conference Neural Networks & Signal Processing, Nanjing, China, 2003: 1237-1241.
ii. Torsten B. and Jean P. T. "Shot boundary detection with mutual information". IEEE International Conference on Image Processing, Greece, 2001(3): 422-425.
iii. XIA D. Y., DENG X. F. and ZENG Q. N, "Shot boundary detection based on difference Sequences of Mutual Information". Fourth International Conference on Image and Graphics, 2007: 389-394.
iv. Zuzana C. and Ioannis P, "Information theory based shot cut/fade detection and video summarization". IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(1): 82-9.
v. T. M. Cover, J. A. Thomas, "Elements of Information Theory", New York: John Wiley & Sons, Inc., 1991.
vi. Torsten Butz, Jean-Philippe Thiran, "Shot Boundary Detection with Mutual Information", In: Proc. 2001 IEEE Int. Conf. on Image Processing (ICIP2001), Thessaloniki, Greece, 2001, no. 3, pp. 422-425.
vii. Dan Jiang, "Information Theory & Coding," Hefei: Press of University of Science & Technology of China, 2001 (In Chinese).
viii. T.M. Cover and J.A. Thomas, "Elements of information Theory", John Wiley & Sons, Inc., 1991.
ix. Cernekov'a Z., Nikolaidis N. and Pitas I. "AIIA shot boundary detection at TRECVID 2006", TREC Video Retrieval Evaluation, 2006.
x. ZHAO Huan, HU Bin, ZHENG Min, LI Xiu-huan," Shot boundary detection based on mutual information and canny edge detector", Oct. 2009, Volume 6, No.10 (Serial No.59),Journal of Communication and Computer, ISSN1548-7709, USA
xi. XIA D. Y., DENG X.F. and ZENG Q.N,"Shot boundary detection based on difference Sequences of Mutual Information". Fourth International Conference on Image and Graphics, 2007:389-394.
xii. Maheshkumar H. Kolekar, Somnath Sengupta,"video shot boundary detection: a comparative study of three popular approaches", Electronics and Electrical Communication Engineering Department, Indian Institute of Technology, Kharagpur, India.