ISSN No. 0976-5697
Volume 2, No. 5, Sept-Oct 2011
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Applying Wavelet Transform for High Speed Shot Boundary Detection & Key Frame Extraction

Prof.Shruti Bhargava
Department of Electronics and Communication
TIT, Bhopal, Bhopal, India.
bhargava.shruti1987@gmail.com

Shilpa.R.Jadhav*
Department of Electronics and Communication
TIT, Bhopal, Bhopal, India.
jadhav.shilpa05@gmail.com

Anup.V.Kalaskar
Department of Electronics and Communication
TIT, Bhopal, Bhopal, India.
Anupkalaskar5@gmail.com

Abstract: This paper present novel algorithm by applying wavelet transform for high speed boundary detection and key frame extraction. The algorithm differs from conventional methods mainly in the use of image segmentation and image compression.. Matching difference between two consecutive frames is computed with different weight. Shot boundaries are detected with automatic threshold. Key frame is extracted by using reference frame-based approach. Experimentalresults show high speed performance of shot boundary detection by using the proposed algorithms, and key frames represent shot content.

Keywords: shot boundary detection; x^2 histogram; automatic threshold; key frame; wavelet transform.

I. INTRODUCTION

The accurate segmentation of shots in a video sequence is fundamental and an essential functionality for numerous video retrieval and management tasks [1]. Many researchers have proposed algorithms to perform shot boundary detection based on certain features extracted from video frames, such as pixel differences, edge differences, color histograms, etc. Moreover, comparative surveyshave also from a learning theory perspective, it is a natural approach to combine such promising features in order to decide whether a boundary exists or not within a given video sequence. Researchers have actively developed different approaches for intelligent video management, including shot transition detection, key frame extraction, video retrieval, etc.

Many approaches used different kinds of featuresto detect shot boundary, including histogram, shapeInformation, and motionactivity [2-5]. Among these approaches, histogram is the popular approach. However, in these histogram-based approaches, pixels' space distribution was neglected. Different frames may have the same histogram [3-5]. In view of this, Change at aldivided[6] each frame into r blocks, and the difference of the corresponding block of consecutive frame was computed by color histogram; the difference*(i,i+1)*of the twoframes was obtained by adding up all the blocks' difference; in the meanwhile, the difference *V(i.i+1)*between two frames i and *i+1* was measured again without using blocks. Based on *D(i.i+1)* and *V(i,i+1)* short boundary was determined Getting over the drawback of the paper, we propose more efficient algorithms for shot boundary detection and key frame extraction with automatic threshold.

II. IDEA OF SHOT BOUNDARY DETECTION

A. Image Segmentation:

Each frame is divided into blocks with mrow and n Colum. Then the difference of the Corresponding blocks between two consecutive frames is computed finally, the final difference of two frames is calculated by adding up all the difference through different weights

B. Wavelet Transform:

The linear one-dimensional wavelet transform w (a, τ) of a signal function s (t) L_2 (R) is given by

$$W (a. \tau) = \langle s, \psi \rangle = \frac{1}{\sqrt{a}} \int s\ (t)\psi^*\left[\frac{t-\tau}{a}\right] dt$$

Where the symbol L_2 (R) is the symbol of square enterable functions and ψ is fixed function called "The Mother Wavelet", well localized both in time and frequency. ψ* stands for the complex conjugate of ψ ,the variable a is the scale parameter and τ is the shift parameter of wavelet function. Analysis by wavelet transformation can be viewed as a decomposition of the given frame into sub frame. Wavelet based image compression technique is one of the transform coding method in which the given image is first transformed to a different domain and then quantized. The Embedded zero-free wavelet coding algorithm exploits these properties to give excellent compression result.

C. Matching Difference:

There are six kinds of histogram match [8]. Color histogram was used in computing g the matching difference in most literatures. However, through comparing several kinds of histogram matching methods, Nagasaki reached a conclusion that x^2 histogramoutperformed others in shot boundary recognition [8]. Hence, x^2Histograms method is proposed in this method.

III. ALGORITHMS DESCRIPTION

Let F(k) be the k th frame in the video sequence k=1,2…..Fv (Fv denotes the total number of videos) The algorithm of shot boundary detection isdescribed as follows.

A. Algorithm 1: Shot Boundary Detection:

Step 1:

Partitioning a frame into blocks with mrows and n columns, and *B(i,j,k)* stands for the blockat *(i,j)* in *k th* frame

Step 2:

Computing x2 histogram matching difference between the corresponding blocks between consecutive frames in video sequence.*H(i,j,k)* and *H(i,j,k+1)* stands for histogram of block at (i,j) in the k th and *k+1*th frame respectively. Block'sdifference is measured by the following equation:

$$D_B(k,k+1,i,j)=\sum_{i=0}^{L-1}\frac{\left[H(i,j,k)-H(i,j,k+1)\right]^2}{H(i,j,k)} \qquad (1)$$

where L is the number of gray in an image;

Step 3:

Computing x^2 Histogram difference between two consecutive frames

$$D\ (k,k+1)=\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} D_B(k,k+1,i,j) \qquad (2)$$

Where *Wij* stands for the weight of block at *(i,j);*

Step 4:

Computing threshold automatically:

Computing the mean and standard variance of x^2 Histograms difference over the whole video sequence [7]. Mean and standard variances are defined as follows.

$$MD=\sum_{k=1}^{F_V-1} D\ (k,k+1) \Big/ F_V-1 \qquad (3)$$

$$STD=\sqrt{\sum_{k=1}^{F_V-1}\left(D\ (k,k+1)-MD\right)^2 \Big/ F_V-1} \qquad (4)$$

Step 5:

Shot boundary detection:

Let threshold $T=M\ D\ +a\ \times\ STD$.Shot candidate detection if $D\ (i,i+1) \geq T$,the ithframe is the end frame of previous shot, and the *(i+1)*th frame is the end frame of next shot.

a. Final Shot Detection:

Shots may be very long but not much short, because those shots with only several frames cannot be captured by people and they cannot convey a whole message. Usually, a shortest shot should last for 1 to 2.5 s. For the reason of fluency, frame rate is at least 25 fps, (it is 30 fps in most cases), or flash will appear. So, a shot contains at least a minimum number of 30 to 45 frames. In our experiment, video sequences are downs amp led at 10 fps to improve simulation speed. On this condition, the shortest shot should contain 10 to 20 frames. 20 is selected for our experiment. We formulate a "shots merging principle": if a detected shot contain fewer frames than 20 frames, it will be merged into previous shot, or it will be thought as an independent one.

Definition 1:

Reference Frame: it is the first frame of each shot; General Frames: all the frames except for reference frame; "Shot Dynamic Factor" max(i):the maximum x^2 Histogram within shot *i* ;

b. Dynamic Shot and Static Shot:

A shot will be declared as dynamic shot, if its max(i) is bigger than *MD* ; otherwise it is shot; *Fc(k)* : The k the frame within the current shot, k=1,2,3…………F_{CN} *(k)* (F_{CN} *(k)* is the total number of current shot.)The algorithm of key frame extraction is described as follows.

B. Algorithm 2: Key Frame Extraction:

Step 1:

Computing the difference between all the general frames and reference frame with the above algorithm:

$$D_C\ (1,k)=\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij} D_{CB}(1,k,i,j), k=2,3,4,......F_{CN}$$

Step 2:

Searching for the maximum difference within a shot:

$$\max(i)=\{D_C(1,k)\}_{\max}, k=2,3,4,......F_{CN} \qquad (5)$$

Step 3:

Determining"ShotType" according to therelationship between max (i) and MD: StaticShot(0) or DynamicShot:

$$ShotType_c=\begin{cases}1 & \text{if } \max(i)\geq MD \\ 0 & others\end{cases} \qquad (6)$$

Step 4:

Determining the position of key frame: if ShotTypeC=0, with respect to the odd number of a shot's frames, the frame in the middle of shot is chose as key frame; in the case of the even number, any one frame between the two frames in the middle of shot can be chose as key frame. If ShotTypeC=1, the frame with the maximum difference is declared as key frame.

III. EXPERIMENT RESULTS AND ANALYSIS

These algorithms are tested with shots of different Videos and found result follows:

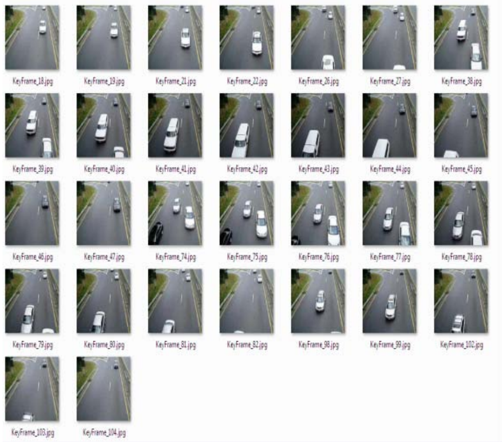
Table I: Shot Boundary Output

Sr.no.	***Frame no***	***Block difference value***
1	Frame no.1	2.252734e+005
2	Frame no.2	3.029871e+005
3	Frame no.3	1.740808e+005
4	Frame no.4	:4.918616e+005
5	Frame no.5	2.299742e+005
6	Frame no.6	9.655472e+005
7	Frame no.7	2.420272e+005
8	Frame no.8	3.452967e+005
9	Frame no.9	2.035338e+005
10	Frame no.10	2.138207e+005
11	Frame no.11	2.306950e+005

12	Frame no.12	1.043614e+006
13	Frame no.13	1.545241e+006
14	Frame no.14	9.570712e+005
15	Frame no.15	2.300548e+006
16	Frame no.16	2.232479e+006
17	Frame no.17	1.643245e+006
18	Frame no.18	3.115060e+007
19	Frame no.19	2.238738e+006
20	Frame no.20	4.988129e+006
21	Frame no.21	3.637578e+007
22	Frame no.22	9.083900e+006
Sr.no.	***Frame no***	***Block difference value***
23	Frame no.23	4.059998e+006
24	Frame no.24	6.239042e+006
25	Frame no.25	3.184043e+006
26	Frame no.26	2.128465e+007
27	Frame no.27	4.234818e+006
28	Frame no.28	1.517516e+007
29	Frame no.29	9.038847e+005
30	Frame no.30	9.291212e+005
31	Frame no.31	1.816369e+006
32	Frame no.32	1.377436e+006
33	Frame no.33	2.366793e+006
34	Frame no.34	2.499197e+006
35	Frame no.35	4.265971e+006
36	Frame no.36	4.664410e+006
37	Frame no.37	1.645068e+007
38	Frame no.38	:2.491203e+007
39	Frame no.39	2.655367e+007
40	Frame no.40	3.462600e+007
41	Frame no.41	6.263576e+006
42	Frame no.42	3.812128e+007
43	Frame no.43	3.823555e+007
44	Frame no.44	3.300014e+007
45	Frame no.45	3.445289e+007
46	Frame no.46	3.606135e+007
47	Frame no.47	:1.312662e+006
48	Frame no.48	2.046394e+006
49	Frame no.49	2.479815e+006
50	Frame no.50	3.438300e+006
Sr.no.	***Frame no***	***Block difference value***
51	Frame no.51	:1.741411e+006
52	Frame no.52	1.660074e+006
53	Frame no.53	1.407927e+006
54	Frame no.54	1.806781e+006
55	Frame no.55	1.556794e+006
56	Frame no.56	6.257250e+005
57	Frame no.57	4.741886e+005
58	Frame no.58	8.067111e+005
59	Frame no.59	8.242188e+005
60	Frame no.60	1.019812e+006
61	Frame no.61	1.279052e+006
62	Frame no.62	1.089395e+006
63	Frame no.63	:1.080301e+006
64	Frame no.64	1.494956e+006
65	Frame no.65	1.529653e+007
66	Frame no.66	2.561012e+006
67	Frame no.67	3.227772e+006
68	Frame no.68	3.426323e+006
69	Frame no.69	4.467072e+006
70	Frame no.70	5.090183e+006
71	Frame no.71	:6.058270e+006
72	Frame no.72	:4.459274e+006
73	Frame no.73	1.645334e+007
74	Frame no.74	2.056309e+007
75	Frame no.75	6.206481e+006
76	Frame no.76	2.683221e+007
77	Frame no.77	3.083760e+007
78	Frame no.78	2.422882e+007
Sr.no.	***Frame no***	***Block difference value***
79	Frame no.79	3.405632e+006
80	Frame no.80	2.789725e+007
81	Frame no.81	:3.033992e+007
82	Frame no.82	9.515336e+005

83	Frame no.83	9.183184e+005
84	Frame no.84	6.327894e+005
85	Frame no.85	7.926676e+005
86	Frame no.86	7.424013e+005
87	Frame no.87	7.223795e+005
88	Frame no.88	:7.930851e+005
89	Frame no.89	:1.136819e+006
90	Frame no.90	2.802542e+006
91	Frame no.91	3.241871e+006
92	Frame no.92	3.743190e+006
93	Frame no.93	:2.640398e+006
94	Frame no.94	2.253193e+006
95	Frame no.95	3.129525e+006
96	Frame no.96	1.731097e+007
97	Frame no.97	6.908658e+006
98	Frame no.98	2.756800e+007
99	Frame no.99	6.642552e+006
100	Frame.no.100	1.704278e+007
101	Frame no.101	1.828782e+006
102	Frame no.102	:4.445243e+007
103	Frame no.103	2.719461e+007
104	Frame no.104	2.154940e+006
105	Frame no.105	7.612865e+005
Sr.no.	***Frame no***	***Block difference value***
106	Frame no.106	9.674090e+005
107	Frame no.107	6.521334e+005
108	Frame no.108	1.169921e+006
109	Frame no.109	1.626765e+006
110	Frame no.110	1.856963e+006
111	Frame no.111	1.327165e+006
112	Frame no.112	1.612671e+006
113	Frame no.113	2.237268e+006
114	Fram no.114	3.684151e+006
115	Frame no.115	3.752450e+006
116	Frame no.116	1.805379e+006
117	Frame no.117	:1.434608e+006
118	Frame no.118	1.498452e+006
119	Frame no.119	1.359151e+006

Mean:7711831.886797,
Std Deviation:11187076.248366
Threshold:18898908.135163



<u>After Determining “Shot Type” according to the relationship between max(i) and MD.Extracted Key Frames are follows</u>

Writing key frame to file
Frame 18 has a block difference threshold
Writing key frame to file
Frame 21 has a block difference threshold
Writing key frame to file
Frame 26 has a block difference threshold
Writing key frame to file
Frame 38 has a block difference threshold
Writing key frame to file
Frame 39 has a block difference threshold
Writing key frame to file
Frame 40 has a block difference threshold
Writing key frame to file
Frame 42 has a block difference threshold
Writing key frame to file
Frame 43 has a block difference threshold
Writing key frame to file
Frame 44 has a block difference threshold
Writing key frame to file
Frame 45 has a block difference threshold
Writing key frame to file
Frame 46 has a block difference threshold
Writing key frame to file
Frame 74 has a block difference threshold
Writing key frame to file
Frame 76 has a block difference threshold
Writing key frame to file
Frame 77 has a block difference threshold
Writing key frame to file
Frame 78 has a block difference threshold
Writing key frame to file
Frame 80 has a block difference threshold
Writing key frame to file
Frame 81 has a block difference threshold
Writing key frame to file
Frame 98 has a block difference threshold
This is dynamic shot

IV. CONCLUSION

We proposed new model to improve the high speed shot boundary detectionand key frame extraction.We also compress the image by using wavelet transform. We first segmented each frame, and then employed differentweights to compute the matching difference and threshold. By using the automatic threshold, wedetected boundary, and by using image compression image get compress. Next, reference frame-basedapproach was provided to extract key frame. Our experimental results show that the proposedboundary detection algorithm can provide accurate and reduced boundary result with image compression. The extracted key framescan satisfactorily represent the content of video In order to further improve accuracy and clearity.

V. REFERENCES

[1]. L. Zhu, M .S Guido, and Aggelos K.K., “MIMAX optimal video summarization”, IEEE Transaction on Circuits and System for Video Technology, Vol.15, No.10, October 2005, pp. 1245-1256 .

[2]. A.Hanjalic, “Shot-boundary detection: Unraveledand resolved?”, IEEE Transaction on Circuits and System for Video Technology., Vol.12, No.2, February, 2002, pp. 90-105

[3]. Z.Cernekova, I. Pitas, and C Nikou, “Information theory-based shot cut/fade detection and video summarization”, IEEE Transaction on Circuits and System for Video Technology, Vol.16, No.1 January 2006, pp. 82-91

[4]. N. Babaguchi, Y. Kawai,T. Ogura, and T. Kitahashi“Personalized abstraction of broadcasted Americanfootball video b y highlight selection”, IEEE Transaction On Multimedia, Vol.6,No.4, August 2004, pp. 575-586

[5]. A. Hanjalic, “Multimodal approach tomeasuring excitement in video”, Proceedings of International Conference on Multimedia and Expo ICME 03[C] Vol.2, July 2003, pp. 289-292

[6]. Y. Chen g, X. Yang, and D. Xu, “A method for shot boundary detection with automatic threshold”, TENCON’02.Proceedings. 2002 IEEE Region10conferenceonComputers, Communication, Control and Power Engineering[C], Vol.1, October 2002

[7]. Y. Zhuang, Y. Rui, T.S. Huang, and S. Mehrotra“Adaptive key frame extraction using unsupervised clustering”, Proceeding.ICIP’98[C], Chicago, IL, 1998, Vol.1, pp. 866-870.

[8]. D.J. He, N. Geng, “Digital image processing”, Professor of Indian University. July, 2003