Combined Key Frame Extraction and Object Based Segmentation in Video Processing

[1]K.Ragavan,[2]S.Deepika, [3]M.Manimozhi, [4]S.Venkadadevi
[1]Assistant professor, Department of ECE, Ramco Institute of Technology
[2,3,4]IV year, Department of ECE, Ramco Institute of Technology

Abstract: - Video is the most effective media for capturing the world around us. Application such as multimedia information systems, distance learning uses huge amount of video data. This has lead to an increasing demand of efficient techniques to store, retrieve, index and summarize the video content. Videos are often constructed in the hierarchical fashion: [Frame]->[Shot]->[Scene]->[Video]. As a preprocessing step, frames have to be extracted for object recognition, detection and then for tracking process. To do this, choosing an efficient technique to extract the key Frames from the video is essential.Segmentation of an image is a very important step for image processing. In this paper, three key frame extraction algorithms have been used such as Entropy difference, Histogram, and color histogram difference. And compare those algorithms based on time behaviour.

Keywords: Key frame extraction, Entropy difference, Histogram difference, Color Histogram difference

I. INTRODUCTION:

The video is most important multimedia. It is a combination of frames that is called groups of pictures (GOP). The GOP is sequential arrangement of frames or pictures. Video summarization technique provides tools for selecting the most representative sequences of still or moving pictures that will help users to take quick look through the whole video clip in a limited amount of time. On the web, there are more number of websites allowing users to broadcast videos themselves. As an example, YouTube is a video sharing website where users can upload, browse, and share video clips which may be of different format and different lengths. We know video structurally contains many scenes and each scene composed of many shots as shown in Fig.1. A scene can be defined as a subdivision of a film in which it presents continuous action in one place, or the setting is fixed. Whereas shot is the Sequence of frames captured by single continuous operation of camera. Several shots are combined to make a scene.

Key frame is one type of video abstraction technique for visual indexing. Detection and Tracking of moving objects in a video sequence can be complex due to: Noise in images, Complex object motion, Non-rigid or articulated nature of objects, Partial and full object occlusions, Complex object shapes, Scene illumination changes, and Real-time processing requirements.

To overcome these problems, it is very essential to use an efficient KFE technique on video processing system the technology of the key frame extraction is a basis for video retrieval. The key frame which is also known as the representation frame represents the main content of the video. Using key frames, to browse and query the video data greatly reduces the amount of processing data.

Videos are collections of frames. Frame rate depends on the quality of the individual video. Every bit of frame contains the unique information for every scene change. Extracting the each moment from the video without losing the single information is a difficult task in the video pre-processing. This is done by efficient key frame extraction process. There are many techniques were used to generate key frames [12].

In an object detection process the preprocessing is done with extracting the Key Frame from the video sequence. Technologies for video segmentation and key-frame extraction have become crucial for the development of advanced digital video systems.
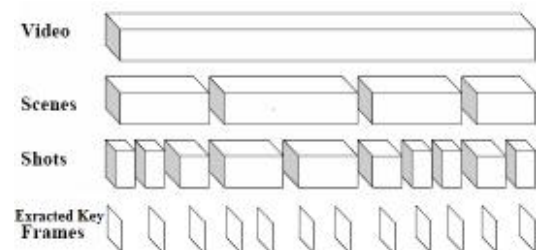

Fig. 1: Structural hierarchy of a video

A.V.Kumthekar et al. proposed that the color histograms have been commonly used for key frame extraction in frame difference based techniques. This is because the color is one of the most important visual features to describe an image. Color histograms are easy to compute and are robust in case of small camera motions. The idea behind histogram based approaches is that two frames with unchanging background and unchanging (although moving) objects will have little difference in their histograms [11].

Markos Mentzelopoulos and Alexandra Psarrou proposed that key frame extraction using entropy difference. For this each frame we first sort the entropies between all quantization levels and then add them, starting from the highest towards the lowest entropy until it exceed the threshold of the 70% of the total image entropy. With this way of isolate the objects that carrying the most information in the image. For each of the gray level entropies that are used in order to reach the 70 % of the

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCIECC - 2017 Conference Proceedings

total image entropy for the first image of the sequence, then take the absolute difference with the relevant gray-level entropy from the next processed frame. If the sum of the normalized differences is more than 77% then they had to change the content of the frame-sequence and therefore a new key-frame is needed [13]

Histogram differences [6] are most widely used in shot detection. This technique can be used to find the frame whose histogram varies significantly from another frame histogram. Thus we can detect shot boundaries. This approach is less sensitive to motion. A shot boundary is found if more than the threshold block has changed. One major problem of histogram differences is that two images can have exactly the same histograms while the shown content differs extremely. In addition, Region based histogram differences are adapted. Each frame is divided into several blocks. The gray-scale histogram is computed for each block.

II. RELATED WORK:

In this paper is proposed key frame extraction from different scenes of a video clip. Each key frame can represent each related scene and also entirely contains all important information of the scene. After the key frame extraction, the key frames are intended to use in video summarization, feature extraction and other processing so key frame extraction algorithm should not be very complex and time consuming. In a video stream, each video frame is a slightly variation with previous one. However, whenever scenes are changed, visual contents and objects are obviously different between current frame and next one. Hence, different techniques are investigated for KFE.

Importance Of Key Frames In Video Processing:

KFE plays an important role in many video processing applications such as video compression, retrieval, skimming, editing, etc. Each video sequence is a combination of shots. A shot is defined as an unbroken sequence of frames recorded from a single camera, which forms the building block of a video. KFE generally involves selecting one frame from each shot segment called cluster", which represents that video segment. A key frame should follow two main rules to adequately represent its cluster: 1) it should be similar enough to the frames in its cluster and 2) it should tolerably differ from frames in other clusters. In a typical KFE process, first, a set of features are extracted from each frame to form a feature vector. Next, a suitable distance measure is applied to the feature vectors to examine similarity/dissimilarity between frames. Finally, based on the distance measurement in the selected feature space, shot and cluster boundaries are detected and the key frames are extracted.

Based On Entropy Difference:

The first approach we rely on, to extract the key frame, is to estimate the entropy difference between two consecutive frames. First we define entropy. Based on the *information theory* the information in an image can be expressed by a compressed value called *entropy*. Its fundamental premise is that the generation of information can be modeled as a probabilistic function that is measured in the manner that agrees with intuition.

If an image is considered to be the output of an imaginary zero memory "intensity source," we can use the intensity distribution of the observed image to estimate the symbol probabilities. Entropy is the measure of amount of information. For an image entropy is given by [5],

$$H = -\sum_{k=0}^{L-1} p(r_k) \log_2 p(r_k)$$

(eqn.1)

Where,

‘k’ is the luminance of a pixel, $0<k<L-1$,
‘L’ is the number of luminance levels,
‘r_k’ is the luminance value of kth pixel,
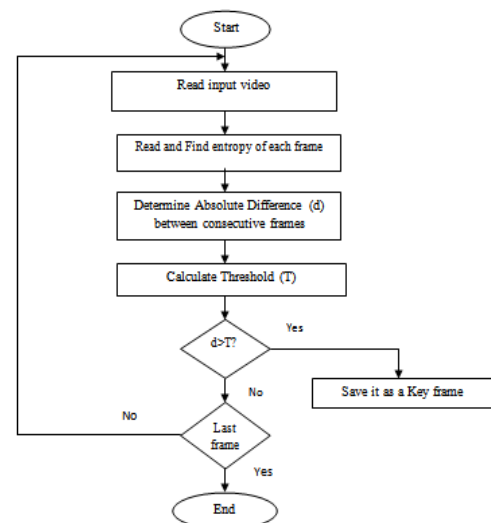‘$p(r_k)$’is the probability of occurrence of r_k.

Fig. 2: Flow chart for entropy difference based key frame extraction

Based On Histogram Difference:

The second approach is to estimate the histogram difference between the two consecutive frames. Histogram of an image is a graphical representation of the probability of the occurrence of the luminance value of the pixels in that image.

Histograms of an image are the basis for numerous spatial domain processing techniques. Histogram computation of an image helps to succeed in other image processing techniques such as enhancement. For analysis purpose of an image also histogram gives a great deal of support. The image will be dark if its histogram distribution is on the left side of the graph. Similarly the image will be light if the distribution of luminance values are in the right side of the graph. Further, it he intensity distribution is in the central part it will be of low contrast, and if the intensity distribution is even in the entire graph it will be a high contrast image.

Histogram normalization and histogram matching techniques helps to improve the quality of the image. We directly use the histogram distribution of the two frames to identify the difference between them and hence utilize it to identify the key frames of the given video sequence. The steps which are explained to identify the key frames of the

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCIECC - 2017 Conference Proceedings

video sequence based on histogram difference are very similar to that of entropy difference based method. The difference is that entropy will be the compressed value giving the information about the image and histogram is the distribution information about the image.
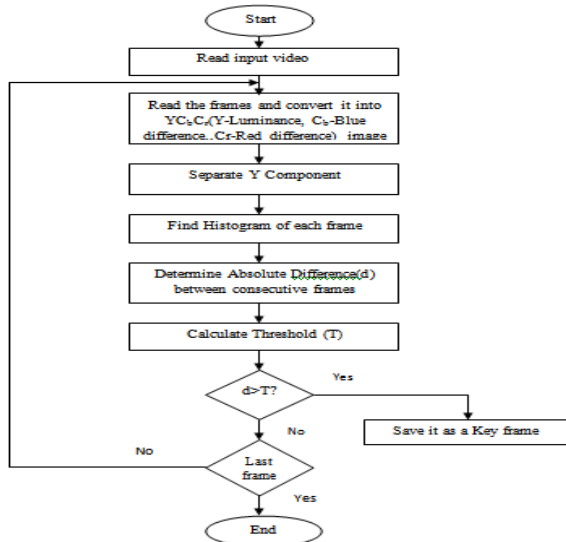

Fig. 3: Flow chart for Histogram difference based key frame extraction

Normalized histogram computation is given by [6],

$$p(r\_k) = n\_k / MN$$

(eqn.2)

Where,

'k' is the luminance value of a pixel,

'rk' is the luminance value of kth pixel,

'p(rk) is the probability of occurrence of a rk,

'MN' is the dimension of the image.

From (eqn.1) and (eqn.2) it is clear that both the methods are based on the probability of occurrence of the luminance levels in the image, i.e. the frame of the video sequence.

Based On Color Histogram Difference:

The color histograms have been commonly used for key frame extraction in frame difference based techniques. This is because the color is one of the most important visual features to describe an image. Color histograms are easy to compute and are robust in case of small camera motions. The idea behind histogram based approaches is that two frames with unchanging background and unchanging (although moving) objects will have little difference in their histograms.
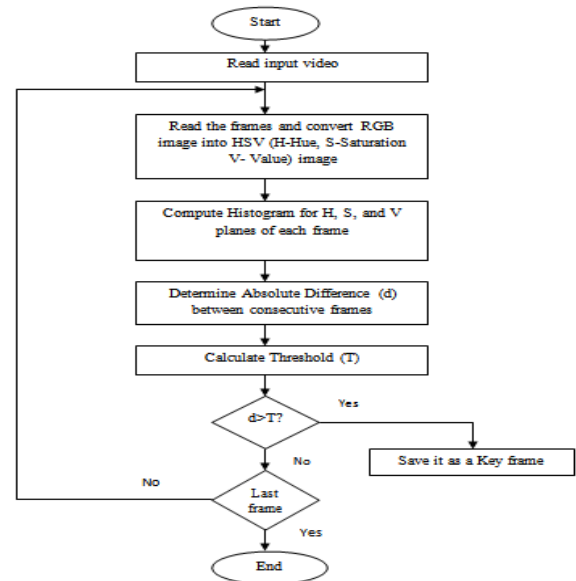

Fig. 4: Flow chart for Color Histogram difference based key frame extraction

Experimental Results:

We did our experiment with the 'Wildlife' video, consists of a total of 901 frames. First the average entropy difference between consecutive two frames is calculated, which is 0.01744. Threshold value used is 0.19814 (11 times the average entropy difference). When the entropy difference between two consecutive frames exceeds the threshold, the latter frame is taken as key frame. In this way 8 key frames (Frame index: 2, 110, 189, 196, 339, 504, 607 and 721) are identified. Then the average histogram difference is calculated, which is 3870.8. As used in the case of entropy based procedure, we fixed the threshold 23224.8 (6 times the average histogram difference between two consecutive frames). When the histogram difference between two frames exceeds the threshold, the latter is taken as key frame. In this we way 8 key frames (Frame index: 2, 110, 189, 196, 339, 504, 607 and 721) are identified.

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCIECC - 2017 Conference Proceedings

These threshold values used are identified experimentally. The threshold value less than 10 times the average entropy difference for entropy based method and less than 5 times the average histogram difference for histogram based method had given wrong identification of the key frames. Too much of frames are identified as key frames out of which many frames violates the condition to be the key frame. With 10 times the average entropy difference (0.1744) as threshold for entropy based method had given 10 key frames (2, 110, 189, 190, 193, 196, 339, 504, 607, 721) and with 5 times the average histogram difference (19354) as threshold for histogram based method had given 9 key frames (2, 110, 189, 190, 196, 339, 504, 607, 721). Frames 190 and 193 were similar to 189 and could not be the correctly identified key frames as they may not be said independent of other frames.

Fig. 5 Sample frames from video

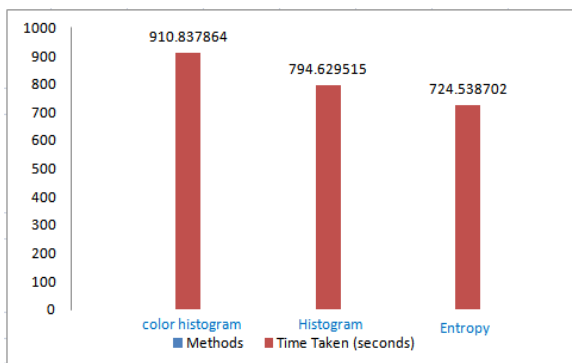Fig.6 Key frames extracted from the Entropy difference method

Fig.7 Key frames extracted from the Histogram difference method

Fig. 8 key frames extracted from the color histogram method

Comparison Table:

Methods	Time Taken (seconds)	Key Frame Numbers
Color Histogram Difference Based KFE	910.837864	2,126,203,204,205,214,354,519,625,738
Histogram Difference Based KFE	794.629515	2,126,203,204,214,354,519,625,738
Entropy Difference Based KFE	724.538702	2,126,192,203,214,354,625,738

Future work:

Now the KFE methods have been implemented, process of image segmentation will be planned.

IV. APPLICATIONS

The goal of the distance learning is to provide quality of learning that is effective and comparable with the traditional classroom environment.

Telemedicine is used to provide the good health care services where direct services are difficult to provide due to some geographical environment. It is combination of audio, video, electronic information which provides diagnosis, consultations, and procedures to the patients at remote site here also key frame will provides a major help.

Key frame extraction is useful for enhancing the functionality of the interactive televisions.

In multimedia key frames is important to extract and analyze the important information required for user such as digital libraries, video retrieval etc.

V.CONCLUSION:

The above comparison of various KFE techniques successfully detects the shot boundary key frames. KFE has lead to an increasing demand of efficient techniques to store, retrieve, index & summarize the video. In this paper various KFE techniques for extracting key frames have been implemented. Since many techniques for KFE have been implemented, the advantages, disadvantages and time behavior of each method have been understood.

REFERENCE:

[1] AriffaBegum ,S. &Askarunisa, A (2014) "Performance Analysis of Various Key Frame Extraction Methods for Surveillance Applications", International Journal of Emerging Technology and Advanced Engineering, Vol .4, No.9

[2] Aziz Makandar & Daneshwari Mulimani (2016), "Key frame extraction and Object Detection in the Sports Video", International Conference on Soft Computing Techniques in Engineering and Technology (ASCTET)

[3] Chandra Shekhar Mithlesh & Dolley Shukla (2016) "A Case Study of Key frame Extraction Techniques", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.5, No.3

[4] Chinh Dang & Hayder Radha (2015) "RPCA-KFE: Key Frame Extraction for Video Using Robust Principal Component Analysis", IEEE Transactions On Image Processing, Vol.24, No.11

[5] Ganesh. I. Rathod et al (2013) "An Algorithm for Shot Boundary Detection and Key Frame Extraction Using Histogram Difference", International Journal of Emerging Technology and Advanced Engineering, Vol.3, No.8, Difference,International Journal of Emerging Technology and Advanced Engineering,Vol. 3, No. 8

[6] Guozhu Liu & Junming Zhao (2009), "Key Frame Extraction from MPEG Video Stream", Proceedings of the Second Symposium International Computer Science and Computational Technology (ISCSCT'09) Huangshan, P. R. China, 26-28, pp. 007-011

[7] Kasturi, R.&Jain,R.(1991), "Dynamic vision, in Computer Vision: Principles (R. Kasturi and R. Jain, eds.)", pp. 469-480, IEEE Computer Society Press

[8] Kikukawa, T. & Kawafuchi, S. (1992), "Development of an automatic summary editing system for the audiovisual resources", IEEE Trans. on Electronics and Information, J75-A, pp. 204-212

[9] IjyaChugh et al (2016) "Techniques for Key Frame Extraction: Shot segmentation and feature trajectory computation", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.22, No.12

[10] KhinThandar Tint&KyiSoe (2013),"Key Frame Extraction for Video Summarization Using DWT Wavelet Statistics", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET, Vol. 2, No. 5

[11] Kumthekar, A.V et al (2013) "Key frame extraction using color histogram method", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.2, No.4, pp.207-214

[12] Priyanka, U. et al (2016) "Analysis of Various Keyframe Extraction Methods", International Journal of Electrical and Electronics Research, Vol.4, No.2, pp.35-40

[13] Markos Mentzelopoulos & Alexandra Psarrou,"Key-frame extraction algorithm using entropy difference"

[14] Sanjoy Ghatak (2016) "Key-frame extraction using Threshold technique", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.1, No.8, pp.67-74

[15] Sanjoy Ghatak et al (2013) "Extraction of Key Frames from News Video Using EDF, MDF AND HI Method for News Video Summarization" International Journal of Engineering and Innovative Technology, Vol.2, No.12

[16] Sheena & Narayanana, N. K.(2015) "Key-frame extraction by analysis of histograms of video frames using statistical methods", 4th International Conference on Eco-friendly Computing and Communication Systems, ELSEVIER Procedia Computer Science

[17] Soham Sarkar & Swagatam Das (2013) "Multilevel Image Thresholding Based on 2D Histogram and Maximum Tsallis Entropy A Differential Evolution Approach", IEEE Transactions On Image Processing, Vol. 22, No. 12

[18] Truong,B.T, et al (2000), "New enhancements to cut, fade, and dissolve detection processes in video segmentation", ACM Multimedia 2000, pp. 219-227.