



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 7, Issue 1, January 2019

Efficient Method for Video Summarization Using Support Vector Machine

Mahanthesha Naik H, Dr.Girijamma.H.A

M. Tech, Dept. of CSE, RNS Institute of Technology, Bangalore, Karnataka, India

Professor, Dept. of CSE, RNS Institute of Technology, Bangalore, Karnataka, India

ABSTRACT: There is lot of scope for video summarization in large video collections with clusters of typical categories, according to categories if we perform video summarization it can produce high quality video summaries than the conventional unsupervised methods. In recent years it is very complex for browsing, transformation and retrieval of large amount of videos and also process becomes very much slow, to overcome these problems video summarization has been proposed to make sure the large amount of video content becomes faster transmission and more efficient content indexing and access. In this paper we introduced an efficient method for video summarization by using GLDM and edge histogram features for extracting features from the videos and for optimization of considered videos will use Lazy Greedy optimization technique. SVM (Support Vector Machine) trainer is used to train the knowledge base to create database of the summarized videos and SVM classifier is used for classification of video summaries from different categories of videos.

KEYWORDS: Lazy Greedy optimization technique, SVM classifier and GLDM and edge histogram features.

I. INTRODUCTION

Most videos from YouTube or Daily Motion consist of fast running, clapping and unedited content. As the multimedia content over the internet is increasing day by day, efficient methods for retrieval of these huge amount of data is required. The World Wide Web is the most important source of information now a day. The knowledge and information which we obtain from internet is in the form of textual data, images, videos etc. There are many resources on the internet which people can use to create process and store videos. This has created the need for a means to manage and search these videos. Users would like to browse, i.e., to skim through the video to quickly get a hint on the semantic content. Video summarization addresses this problem by providing a short video summary of a full-length video. An ideal video summary would include all the important video segments and remain short in length. The problem is extremely challenging in general and has been subject of recent research.

With the advent of digital multimedia, a lot of digital content such as movies, news, television shows and sports is widely available. Also, due to the advances in digital content distribution (direct-to-home satellite reception) and digital video recorders, this digital content can be easily recorded. However, the user may NOT have sufficient time to watch the entire video (Ex. User may want to watch just the highlights of a game) or the whole of video content may not be of interest to the user(Ex. Golf game video). In such cases, the user may just want to view the summary of the video instead of watching the whole video.

Thus, the summary should be such that it should convey as much information about the occurrence of various incidents in the video. Also, the method should be very general so that it can work with the videos of a variety of genre. In this paper we introduced an efficient method for video summarization by using GLDM and edge histogram features for extracting features from the videos, which calculates the grey level difference method probability density functions for the pre-processed gray image. This method is used for extracting statistical texture features of a digital image. From each density functions five texture aspects are outlined: contrast, Angular second moment, Entropy, mean and Inverse difference moment. Contrast is defined as the change in intensity between highest and lowest intensity stages in an image for that reason measures the local variations in the grey level. Angular second moment is a measure of homogeneity. If



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

the difference between grey levels over an area is low then these areas are stated to be having better Angular second moment (ASM) values. Mean it offers the average intensity value. and for optimization of considered videos will use Lazy Greedy optimization technique. SVM (Support Vector Machine) trainer is used to train the knowledge base to create database of the summarized videos and SVM classifier is used for classification of video summaries from different categories of videos. It is a non-linear classifier. The idea behind the method is to nonlinearly map the input data to some high dimensional space, where the data can be linearly separated, thus providing great classification performance. Support Vector Machine is a machine learning tool and has emerged as a powerful technique for learning from data and in particular for solving binary classification problems.

II. LITERATURE SURVEY

Jeff Donahue et.al [1] has proposed a system on A Deep Convolution Activation Feature for Generic Visual Recognition. In this they compare the efficacy of relying on various video summarizations to define a fixed feature, and report object results that significantly outperform the state-of-the-art on several important vision challenges. We are releasing DeCAF, an open-source implementation of these deep convolution activation features, along with all associated network parameters to enable vision researchers to be able to conduct experimentation with deep representations across a range of visual concept learning paradigms. Zaynab El khattab, et.al [2] has proposed a architecture on Video Summarization: Techniques and Applications. This system involved Video summarization has been proposed to improve faster browsing of large video collections and more efficient content indexing and access. In this paper, they focus on approaches to video summarization. The video summaries can be generated in many different forms. However, two fundamentals ways to generate summaries are static and dynamic. We present different techniques for each mode in the literature and describe some features used for generating video summaries. They finalized with perspective for further research. D. Chen, J [3] proposed a segmentation method based on Markov random field to extract more accurate text characters. This methodology allows handling background gray-scale multimodality and unknown text gray-scale values. Support vector machine (SVM) is used for text verification followed by traditional OCR algorithm. Shih-Wei Sun, Yu-Chiang Frank Wang in proposed a robust moving foreground object detection method followed by the integration of features collected from heterogeneous domains. More focus is on annotating rigid moving objects & considers videos with only one foreground object present. Danila Potapov Et.al [4] has proposed a Category-specific video summarization. This paper presents a novel method for effectiveness performance of a temporal segmentation into semantically-consistent segments, delimited not only by shot boundaries but also general change points. Then, equipped with an SVM classifier. The resulting video assembles the sequence of segments with the highest scores. The obtained video summary is therefore both short and highly informative. Experimental results on videos from the multimedia event detection (MED) dataset of TRECVID'11 show that our approach produces video summaries with higher relevance than the state of the art.

III. METHODOLOGY

Architecture of proposed system is shown in the figure 1 below, we have 2 phases in our system. i.e. training and testing phase. In training phase we extract the frames from video then convert the frame from RGB to gray scale. Extract the features such as GLDM and edge histogram features from the pre-processed images. Apply lazy greedy algorithm for optimization followed by learning the summary and store the features in a database. In the testing phase, take video as input and extract the frames and repeat the steps which we followed in the in the training such as pre-processing, optimization and feature extraction. Once we extract the features from the each frames of query video, compare the features of query frames with stored features in the database and classify the result using RBF SVM classifier.

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

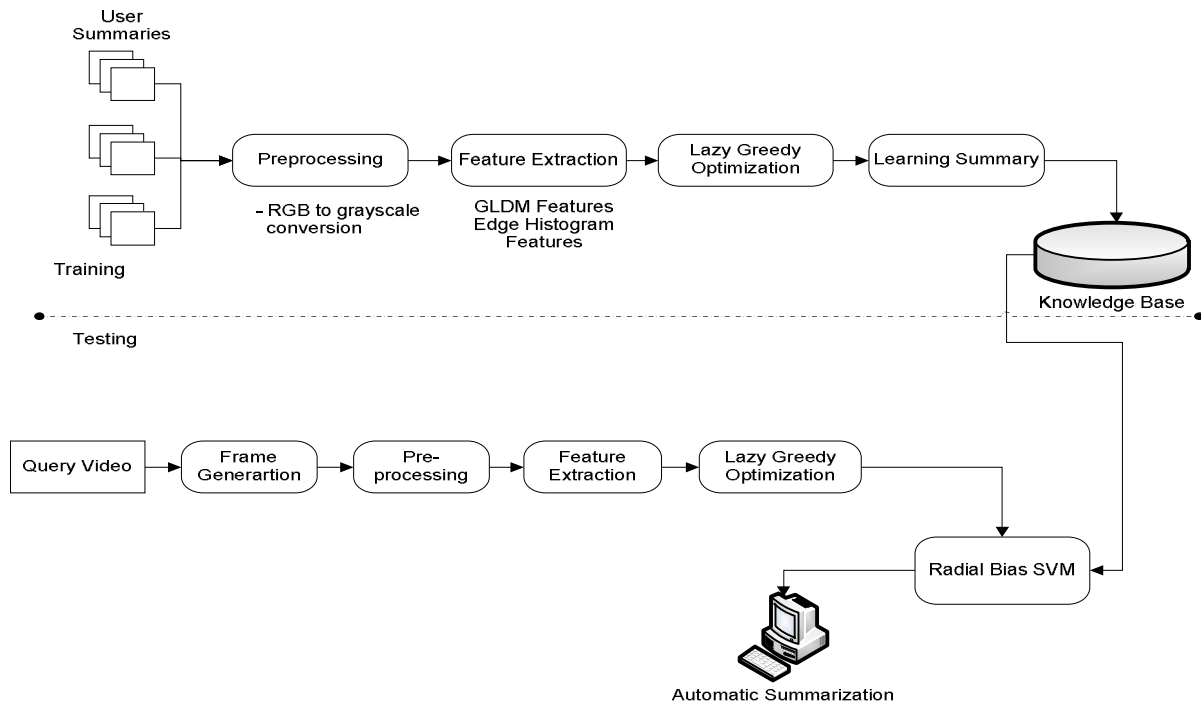


Figure 1:Block Diagram of Proposed System.

A. PRE-PROCESSING

Pre-processing is mainly used to adjust the size of the image, removal of noise, color conversion and isolating objects of interest in the image. Pre-processing is any form of signal processing for which the output is an image or video, the output can be either an image or a set of characteristics or parameters related to image or videos to improve or change some quality of the input. This process will help to improve the video or image such that it increases the chance for success of other processes. In this paper we considered sampled videos as input and those videos are subjected to pre-processing this will resulting in color conversion into gray scale conversion.

B. FEATURE EXTRACTION

First stage of training phase is to read the sampled summarized videos in the database. Then videos are pre-processed. When the image is pre-processed feature extraction is achieved. For feature extraction we use GLDM and Edge histogram features.

I. GLDM (GRAY LEVEL DIFFERENCE METHOD)

The GLDM process calculates the gray level difference method probability density functions for the pre-processed gray image. This method is used for extracting statistical texture features of a digital image. From each density functions five texture aspects are outlined: contrast, Angular second moment, Entropy, mean and Inverse difference moment. Contrast is defined as the change in intensity between highest and lowest intensity stages in an image for that reason measures the local variations in the gray level. Angular second moment is a measure of homogeneity. If the difference between grey levels over an area is low then these areas are stated to be having better Angular second moment (ASM) values. Mean it offers the average intensity value. Entropy is the average understanding per intensity source output. This parameter measures the disease of an image. When the image is just not texturally uniform, entropy could be very large. Entropy is strongly, however inversely, correlated to energy. Inverse difference moment IDM measures the closeness of the distribution of elements in the gray stage Co-occurrence Matrix (GLCM) to the GLCM diagonal. To describe the gray level difference process, let $g(n, m)$ be the digital picture function. For any given displacement $\delta = (\Delta n, \Delta m)$,

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

where Δn and Δm are integers, let $g\delta(n, m) = |g(n, m) - g(n + \Delta n, m + \Delta m)|$. Let $f(|\delta)$ be the estimated probability density function associated with the possible values of $g\delta$ i.e. $f(i|\delta) = P(g\delta(n, m) = i)$ herein our possible forms of vectors δ will be considered. $(0, d), (-d, d), (d, 0), (-d, d)$, where d is inter sample distance. We refer $f(|\delta)$ as gray level difference density functions.

II. EDGE HISTOGRAM FEATURES

Feature extraction is used to extract relevant features for recognition of plant leaves. Figure 2 represents the steps how we are processing the feature extraction using edge histogram features.

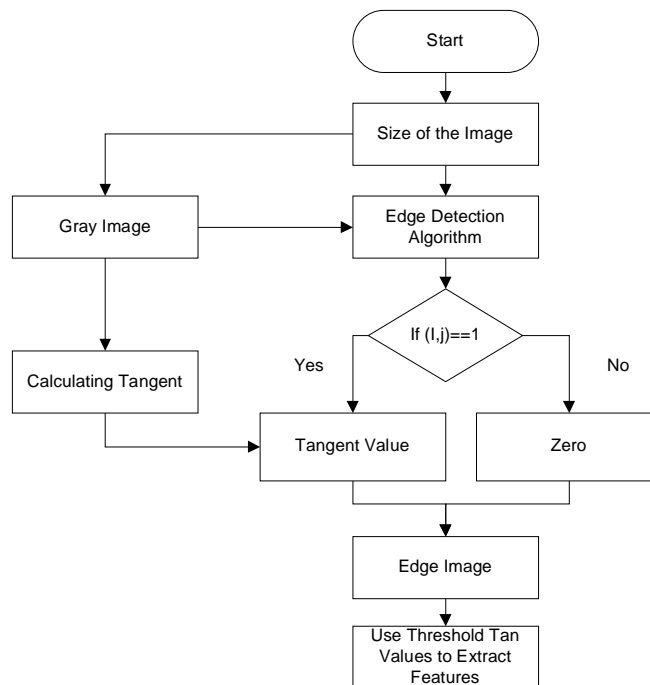


Figure 2: Steps for extracting features.

A histogram is a graphical representation showing a visual impression of the distribution of data. It is an estimate of the probability distribution of a continuous variable and was first introduced by Karl Pearson. A histogram consists of tabular frequencies, shown as adjacent rectangles, erected over discrete intervals (bins), with an area equal to the frequency of the observations in the interval. Once the size of image is know then will go for edge detection in that will find out fallowing attributes Smoothing, finding gradients, Non-maximum suppression, double thresholding and edge tracking by hysteresis. By using these attributes will extract the features from the pre-processed summarized videos.

C. LAZY GREEDY ALGORITHM

A lazy greedy algorithm is an algorithm, with the hope of finding a global optimum this algorithm follows the problem solving heuristic of making the locally optimal choice at each stage. In many problems, a lazy greedy strategy does not in general produce an optimal solution, but nonetheless a greedy heuristic may yield locally optimal solutions that approximate a global optimal solution in a reasonable time. In our approach this algorithm is used for optimization of extracted features from the considered videos for videos summarization. Lazy greedy algorithm is as follows

Start



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 7, Issue 1, January 2019

Step 1: Function (V, Xv, c, w, f, B)
 Step 2: $y_{uc} \leftarrow \text{LAZYGREEDY}(V, Xv, c, w, f, B, \text{uniform cost})$
 Step 3: $y_{cb} \leftarrow \text{LAZYGREEDY}(V, Xv, c, w, f, B, \text{cost benifit})$
 Step 4: return $\arg \max (w^T f(X_v, y_{uc}), w^T f(X_v, y_{cb}))$
 Step 5: End function
 Step 6: Function LAZY GREEDY $(V, Xv, c, w, f, B, \text{type})$
 Step 7: $y \leftarrow \emptyset$ start from an empty solution
 Step 8: $\delta_s \leftarrow \infty, \forall_s \in V$ initialize marginal gain
 Step 9: while $\exists_s \in V \setminus y: c(y \cup \{s\}) \leq B$ do
 Step 10: $cur_s \leftarrow \text{false}, \forall_s \in V \setminus y$ set gain to outdated
 Step 11: while true do
 Step 12: if type=uniform cost then
 Step 13: $s^* \in \text{argmax} \delta_s$ max again
 Step 14: else if type= cost benefit then
 Step 15: $s^* \in \text{argmax} \frac{\delta_s}{c(s)}$ max again by cost
 Step 16: end if
 Step 17: if cur_{s^*} then
 Step 18: $y \leftarrow y \cup \{s^*\};$
 Step 19: break
 Step 20: else
 Step 21: $\delta_s \leftarrow w^T f(X_v, y \cup \{s^*\}) - w^T f(X_v, y)$
 Step 22: end if
 Step 23: end while
 Step 24: end while
 Step 25: return y
 Step 26: end function

Summarizing, the problem of subset selection is difficult to optimize. But if the optimization can be posed as sub modular maximization, we have seen that there exist efficient algorithms, which yield good approximations.

D. SUPPORT VECTOR MACHINE (SVM)

Support vector machine (SVM) is a non-linear classifier. The idea behind the method is to nonlinearly map the input data to some high dimensional space, where the data can be linearly separated, thus providing great classification performance. Support Vector Machine is a machine learning tool and has emerged as a powerful technique for learning from data and in particular for solving binary classification problems. The main concepts of SVM are to first transform input data into a higher dimensional space by means of a kernel function and then construct an OSH (Optimal Separating Hyper Plane) between the two classes in the transformed space. For plant leaf classification it will transform feature vector extracted from leaf's contour. SVM finds the OSH by maximizing the margin between the classes. Data vectors nearest to the constructed line in the transformed space are called the support vectors. The SVM estimates a function for classifying data into two classes. Using a nonlinear transformation that depends on a regularization parameter, the input vectors are placed into a high-dimensional feature space, where a linear separation is employed. To construct a nonlinear support vector classifier, the inner product (x, y) is replaced by a kernel function $K(x, y)$, as in (1)

$$f(x) \text{sgn} \left(\sum_{i=1}^1 \alpha_i y_i K(x_i, x) + b \right) \quad (1)$$

where $f(x)$ determines the membership of x . We assume normal subjects were labeled as -1 and other subjects as +1. The SVM has two layers. During the learning process, the first layer selects the basis $K(x_i, x)$, $i=1, 2, \dots, N$ from the given set of kernels, while the second layer constructs a linear function in the space. This is equivalent to finding the

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

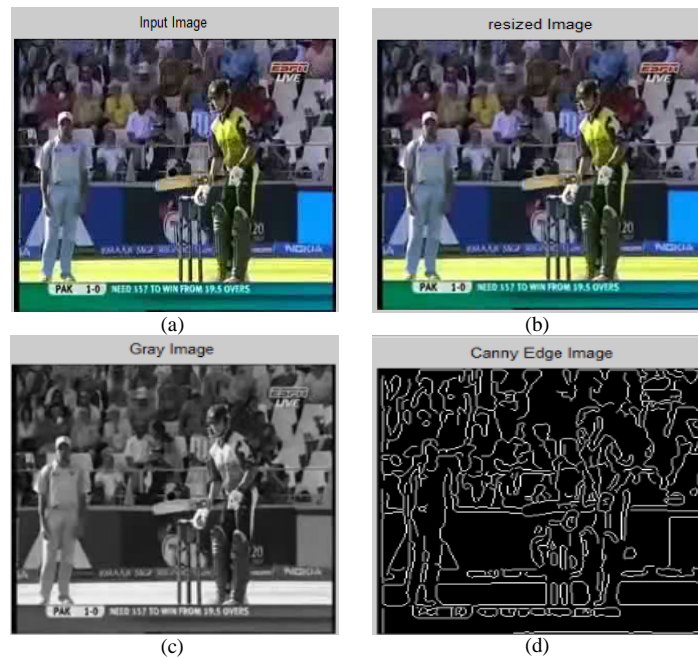
Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

optimal hyper plane in the corresponding feature space. The SVM algorithm can construct a variety of learning machines using different kernel functions. Main advantage of SVM is it has a simple geometric interpretation and gives a sparse solution. Unlike neural networks, the computational complexity of SVMs does not depend on the dimensionality of the input space One of the bottlenecks of the SVM is the large number of support vectors used from the training set to perform classification tasks.

E. EXPERIMENTAL RESULT

Figure 3 represent the overall experimental results of the proposed system. Aswe extract the frames from input video , the input image is as shown in Figure 3(a) and then by applying pre-processing techniques will get resized and RGB to gray scale converted image as shown in Figure 3(b) and 3(c) respectively, next will apply canny edge detection algorithm to get Figure 3(d) as shown below. Next will apply feature extraction algorithm to get key frames and event frames as shown in Figure 3(e) and 3(f) respectively. The output image is as shown in Figure 3(g).



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 7, Issue 1, January 2019

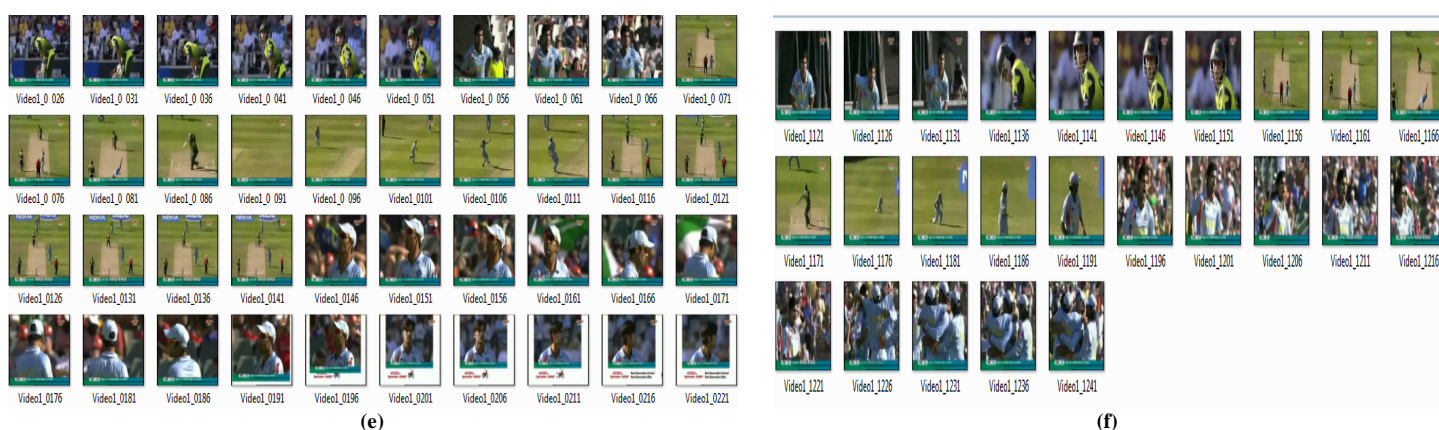


Figure 3: (a) Input Image; (b) Resized Image; (c) Gray Scale Image; (d) Canny Edge Image; (e) key frames; (f) Event Frames; (g)Result



(g)

F. CONCLUSION

In this paper we have proposed a new method for video summarization by using Lazy greedy algorithm for optimization .Using GLDM feature extraction method we achieved the effective results of the proposed architecture. In this paper we introduced an efficient method for video summarization by using GLDM and edge histogram features for extracting features from the videos and for optimization of considered videos will use Lazy Greedy optimization technique. SVM (Support Vector Machine) trainer is used to train the knowledge base to create database of the summarized videos and SVM classifier is used for classification of video summaries from different categories of videos. Our experimental results demonstrate the potential and generality of our method.

REFERENCES

1. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. DeCAF, “ A Deep Convolution Activation Feature for Generic Visual Recognition,” PP.1310.1531, 2013.
2. Zaynab El khattabi, Youness Tabii, Abdelhamid Benkaddour, “Video Summarization: Techniques and Applications”, International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol. 9, No. 4, 2015
3. D. Chen, J. Odobez, H. Bourlard, “Text detection and recognition in images and video frames, Pattern Recognition 37” pp. 595 – 608, 2004.
4. Danila Potapov, Matthijs Douze, Zaid Harchaoui, Cordelia Schmid, “Category specific videoSummarization”, ECCV 2014 - European Conference on Computer Vision, Sep 2014.
5. R. Gomes and A. Krause, “Budgeted Nonparametric Learning from Data Streams.”, ICML, Vol. 4, PP. 1562-1873, 2010.



ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 1, January 2019

6. S. Yeung, A. Fathi, and L. Fei-Fei, "Video Summary Evaluation through Text", Vol. 5, P.5824-5967, 2014.
7. S. Zhu, Z. Liang, and Y. Liu, Automatic Video Abstraction via the Progress of Story, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, vol. 6297, pp. 308-318, 2010.
8. R. Laganière, R. Bacco, A. Hocevar, P. Lambert, G. Païs, and B. E. Ionescu, "Video summarization from spatio-temporal features," in TVS'08 Proceedings of the 2nd ACM TREC'Vid Video Summarization Workshop, pp. 144-148, 2008.
9. Y. Gao, W.-B. Wang, J.-H. Yong and H.-J. Gu, "Dynamic video summarization using two-level redundancy detection," Multimedia Tools and Applications, vol. 42, pp. 233-250, 2009.
10. Ngo, C., H. Zhang, and T. Pong, "Recent Advances in Content-based Video Analysis", international journal of image and graphics, 2001.