

# Entropy Supported Video Indexing for Content based Video Retrieval

P. M. Kamde  
Department of Computer  
Engineering,  
Sinhagad College of  
Engineering, Pune

Sankirti Shiravale  
Department of Computer  
Engineering,  
Marathwada college of  
Engineering, Pune

S. P. Algur, PhD.  
Professor & Chairman,  
Computer Department, Rani  
Channamma University,  
Belagavi- 591156

## ABSTRACT

The increased in availability and usage of on-line digital video has created a need of automated video content analysis techniques, including indexing and retrieving. Automation of indexing significantly reduces the processing cost while by minimizing tedious work. Traditional video retrieval methods based on video metadata, fail to meet technical challenges due to large and rapid growth of multimedia data, demanding effective retrieval systems. One of the most popular solutions for indexing is extracting the features of video key frames for developing a Content Based Video Retrieval (CBVR) system. CBVR works more effectively as these deals with content of video rather than video metadata. Various features like color, texture, shape can be integrated and used for video indexing and retrieval.

Implemented CBVR system is experimented based on integration of texture, color and edge features for video retrieval. Entropy is a texture descriptor used for key frame extraction and video indexing. However entropy, color (RGB) and edge detection algorithms are used for video retrieval. These features are combined in various ways like entropy-edge, entropy- color for result refinement. Dataset is created with the videos from different domains like e-learning, nature, construction etc. By the combination of these features in different ways, we achieved comparative results. Obtained result shows that combining of two or many features gives better retrieval.

## Keywords

CBVR, CBVI, Video indexing, Video retrieval.

## 1. INTRODUCTION

Explosive growth of digital content including image, audio and video on web and as well as on desktop applications has demanded development of new technologies and methods for representation, storage and retrieval of multimedia data. Rapid development of digital libraries and repositories are attempting to achieve efficient techniques for the same. Despite of many initial successes problems persist in the area of effective video retrieval system since decades. Many of the video retrieval systems are presently based on the metadata attributes like name, date of creation, tagged words, annotation, etc. This however leads to unsatisfactory results to users. Content Based Video Retrieval (CBVR) system works more effectively as these deals with content of video rather than video metadata. The increased availability and usage of on-line digital video has created a need for automated video content analysis techniques including indexing and retrieving.

The most unique characteristic of a video document is its ability to convey a rich semantic presentation through the synchronized audio, visual and text presentations over a period of time. In the early days of CBVR research, most efforts have simply borrowed/extended systems and algorithms from image, text and sound retrieval as these types of media have been commonly used much earlier than that of video. Although these strategies has gained some degree of success, but it is not always sufficient for all applications because each video genre has its own syntax, semantic, rules, and formats. Therefore, a CBVR system needs to support diverse requirements from users and applications. For ex: color-texture-shape based indexing generally can be used.

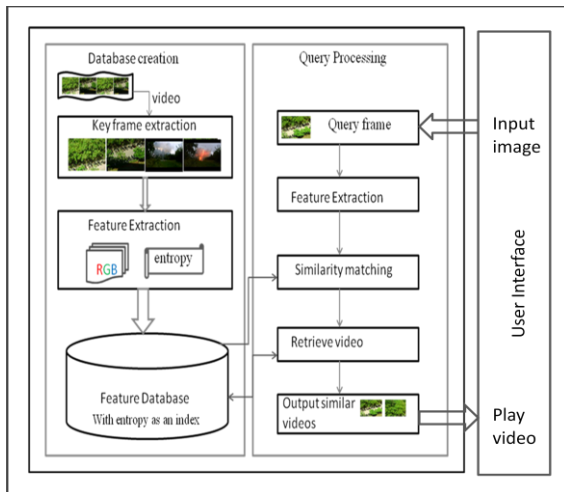
## 2. RELATED WORK

Video indexing is a process of tagging videos and organizing them in an effective manner for fast access and retrieval. Automation of indexing can significantly reduce processing cost while eliminating tedious work. Video processing always is performed on frames which are basic block of video. Group of frames captured together is called shot. Few minutes shot may contain hundreds of frames, which makes video large in size [4]. Storing and processing theses individual frames are memory and computational expensive. Also there is a very minute change of content information between the two consecutive frames of same shot. Selection of frames from single shot may be done to identify the key frame or frames which represent complete shot. These key frames are then used for indexing. Corresponding key frames extracted on shots boundaries are detected. A variety of video indexing techniques have been investigated. Those techniques can be generally classified into four categories based on different cues on that algorithms uses in video indexing and retrieval, including visual features, auditory features, texts, and task models [3]. Various features are extracted from key frames and used to build indices with metadata. In content based video retrieval by giving input query as a text or image is given and required videos are retrieved. According to the input query features are computed and similar videos are retrieved with the help of indices based on features.

## 3. SYSTEM DESIGN AND PROPOSED METHOD

Implemented CBVR system is based on entropy, color, and edge features. Features are extracted and used to build indices with the metadata. System is divided into two modules. First module extracts key frames and database is created with features extracted from key frames. Second module extracts

features of query frame and retrieves similar video from database.



**Fig 1: Architecture of implemented CBVR.**

### 3.1 Module I

This module contains two parts one is key frame extraction and another is feature extraction and database creation. For the key frame extraction, entropy based algorithm is used which is discussed here later. After getting key frames, entropy of that frame, key frame with video metadata like video name is stored in to the database.

### 3.1.1 Key Frame Extraction

Video processing is always performed on frames which are basic block of video. Group of frames captured together is called shot. Few minutes of shot may contain hundreds of frames, which makes video large in size. Storing and processing these individual frames are memory and computational expensive [4]. Also there is a very minute change of content information between the two consecutive frames of same shot. Selection of frames from single shot may be done to identify the key frame or frames which represent complete shot. These key frames are then used for indexing. Corresponding key frames can be extracted after shots boundaries are detected. Simple approach is just to extract first frame of the shot. Better results are obtained by extracting last frame along with the first frame to represent the shot. Key frame extraction is the foundation of video analysis and content based video retrieval. Key-frames provide a suitable abstraction and framework for video browsing.

### Shot Detection Method-

Pixel based, block based, transform based, feature based and histogram based techniques are mainly used for automatic shot boundary detection from raw video stream. Pixel-based methods approach uses pixel difference as key parameter for detecting shot. These techniques are highly sensitive to noise. Block Based techniques work on fundamentals of pixel processing but they operate on image at a time due to which these methods are faster compared to pixel processing techniques. Color histogram based approaches are the most popular, but it does not specify any position of pixel [12]. In this system we are using entropy measure which is one of the texture descriptor. A matrix called as grey level co-occurrence

which represents adjacent pixel pattern is used. From GLC, entropy is computed which measures frequency of adjacent pixel pattern. Following function is used to compute entropy.

Entropy =  $-\sum C(i, j) \log C(i, j)$ , where  $C(i, j)$  is GLC matrix

## Entropy Algorithm

Entropy is a good way of representing the impurity or unpredictability of a set of data since it is dependent on the context in which the measurement is taken. It has been deemed necessary to quantize the image down to 256 colors. Entropies of consecutive frames are computed and if difference is greater than threshold then consecutive frame is considered as new key frame [6].

Key frame extraction algorithm is explained below.

1. Select video clip
2. Select first frame as a key frame
3. Apply color quantization
4. Compute gray level entropies and  
Set value as global entropy
5. Select next consecutive frame
6. Compute entropies of consecutive key frame.
7. compute frame entropies difference
8. if  $\text{diff} > 70\%$  of global entropy  
    then  
        store selected key frame as a real key frame and  
        put the current key frame in buffer  
else  
    if end of clip  
        then go to step 1  
    else  
        go to step 5.
9. Sstop

### 3.1.2 Feature Extraction

The features used for implementing this system are entropy and RGB color space. Entropy is the one of the texture descriptor. Entropy measures frequency of adjacent pixel patterns. RGB color space is widely used for video systems.

### Algorithm

1. Get the key frame
2. Compute Entropy
3. Store entropy, RGB value, key frame and video metadata in database.
4. Use entropy as an index and sort database order by entropy for fast retrieval.

### 3.2. Module II

Second module receives query video frame input from the user. Entropy is calculated with the help of GCL matrix. Some video frames are retrieved for processing from the database whose entropy values are matching with threshold specified in entropy algorithm. After retrieving key frames either of two methods are used to find the similarity. In entropy-Edge method edges are detected and black and white pixels [8] are used in entropy-color method color (RGB value) is used to find similar videos. These methods are explained below.

### 3.2.1 Entropy-Edge Based Method

For query frame and selected key frame, edge is detected using Prewitt edge detection (for various edge detection operators refer Appendix A). Black and white points on the edge are computed. Query frame and key frame are compared based on white points present in both the frames. Similarity measure is applied to compute result and user is presented with videos based on the metadata information of best matched video frames [11]. Detailed implemented algorithms are explained as follow:

#### Algorithm

1. Read the query input video frame name
2. Read User query Frame A
3. calculate entropy of A
4. Search the respective frames from the database those are matching with entropy values.
5. Store those frames in temporary database.
6. for each frame in temporary database  
Compare each frame one by one with query frame and compute percentage\_Matched using similarity matching algorithm.

Compute the result by using similarity measure

function

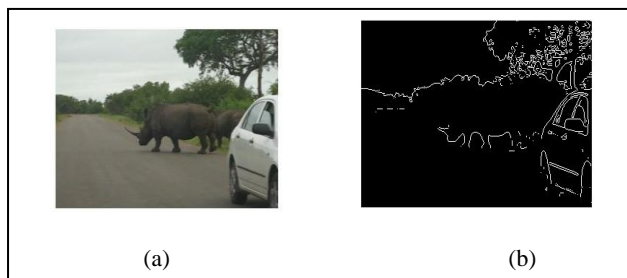
$$\text{Result} = (\text{entropy} * \text{Percentage\_Matched}) / 2;$$

7. Sort the result by order of result
8. Display result

#### Similarity matching algorithm

1. Read frame1, frame2
2. Convert frames into gray level.
3. Apply 'prewitt' edge detection algorithm.
4. Find number of white points present in frame1.
5. Find matched data between two frames by comparing pixel by pixel.  
If pixel (i,j) from both images = 1  
Then match\_data = match\_data+1
6. Compute persenage\_matched =( matched data / total white points.)\*100
7. Return persantage\_matched

Following Fig explains working of similarity matching algorithm. Frame image is converted into gray scale and edge detection algorithm is applied to get black and white points.

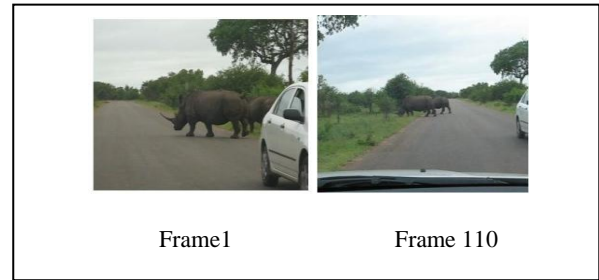


**Fig 2: (a) sample frame image (b) image after applying prewitt edge detection algorithm.**

If we compare image given in Fig 2.a with itself then we get following result.

Similarity matching (fig 2.a,fig 2.a)  
White points=2083  
Matched data=2083

%matched=100



**Fig 3: Similarity matching between two images**

As shown in Fig 3 two different frames from video. If we compare these two images we got following results.

Similarity matching (frame1,frame110)  
White points=2533  
Matched data=99  
%matched=4

Limitations of Entropy-Edge Based method:

1. Entropy is a scalar value which doesn't works on color value of frame. Hence sometimes system can retrieve video images which are highly contrasted in color.
2. As system is using edge detection algorithm which can be applied on fixed size (320\*240) frame hence there is a restriction on video frame size.
3. While using edge detection algorithm for computing result, system is performing pixel by pixel matching hence system is time consuming.

### 3.2.2 Entropy -Color Based Method

In this method average RGB value is used to find nearest matched video key frame. Detailed implemented algorithm is explained as follows:

#### Algorithm

1. Read the query input video frame name
2. Read User query Frame A
3. calculate entropy and average RGB value of A
4. Search the respective frames from the database which match with entropy and RGB values.
5. Compute similarity matching by computing  
Result=Average of difference of entropy \* difference of RGB values.
6. Sort the result by order of result
7. Display result

Advantages of Entropy-Color Based Method

1. Post processing is fast as it is working with numeric values than pixel by pixel.
2. Retrieve videos which are most similar in color.
3. Speed of retrieval is increased and complexity of algorithm is reduced.
4. Suitable in applications which demands similar color videos e.g. nature, water, e-learning videos.

Limitations of Entropy-Color Based Methods:

While retrieving output videos priority goes to color rather than object present in video hence not suitable for object based videos like sports, news videos.

Statistical results obtained by these methods are discussed in following results section.

## 4. RESULTS

Implemented system is using entropy based key frame extraction for video indexing and entropy, color and edge features are used for video retrieval. Results are obtained by using videos from different domains and verified by unit as well as integration testing. Main graphical user interface for implemented CBVR is divided into three parts like play video for playing selected video, display result part showing key frames when extract key frame query is fired and output videos when search for video query is fired. In fire query part either key frame can be extracted from video or videos can be searched with given query input image.

Following Fig 4 shows graphical user interface of the system.



Fig 4: GUI of Implemented System.

### Extract Key Frame

Extract key frames is one of the modules of the system which is responsible for video indexing and database creation. Entropy field is used for creating index. In this module shot is detected and key frames are extracted. Video is processed frame by frame and shot is detected based on entropy difference and key frames are extracted with features entropy, RGB value. Extracted key frames with its frame numbers are displayed on the screen which is shown in Fig 4. After extraction is complete, video extraction complete message is displayed.



Fig 5: Extracted Key Frames for Video Boxcorrect.avi.

### Limitation of Entropy Based Key Frame Extraction Algorithm

Entropy Based key frame extraction algorithm is affected by several factors like editing effects, fade, flash, sunlight. In this case, consecutive frame based algorithms becomes unreliable as they produce redundant key frames for same shot as shot boundaries. Fig 6 shows redundant key frames that are generated due to sunlight variation.

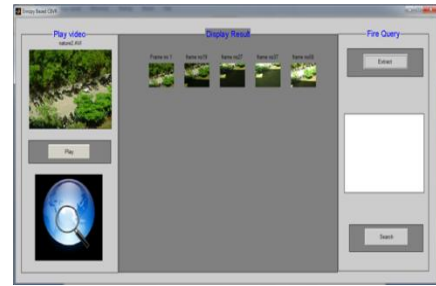


Fig 6: Redundant Key Frames due to Sunlight Effect.

## 4.2 Search Video

Search for video module is used for video retrieval. To search the video, image is given as an input query. With that input image similar frames from the database are retrieved. After post processing top three key frames with its video name, key index and matching percentage (in terms of either difference or similarity) are displayed on the screen in sorted order. Selected video can be played in play video module.

### Results Obtained by Entropy-Edge Based Method

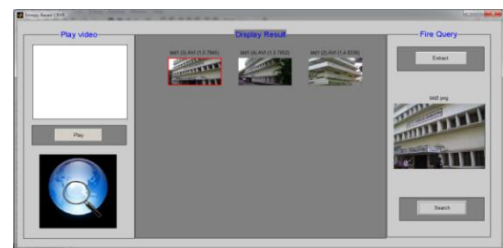


Fig 7: Edge Based Video Retrieval in Construction Domain.

Fig 7 shows resultant videos for given input image which is a non-key frame from construction domain. Fig 8 shows videos in news domain for non key input image where third video frame can have similar human body posture (with very less matching %) which are different in color. In edge based algorithm matching percentage get affected due to small variations in object e.g. body movement.

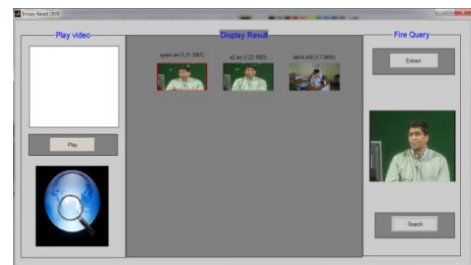
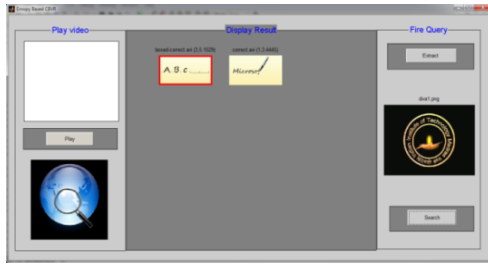


Fig 8: Edge based Video Search in News Domain.

### Limitation of Entropy-Edge Based Method

Entropy counts frequency of adjacent pixel patterns. It doesn't give any color information. Entropy of plane red image and plane yellow image can be the same and if edge detection algorithm is applied on both the images then we get

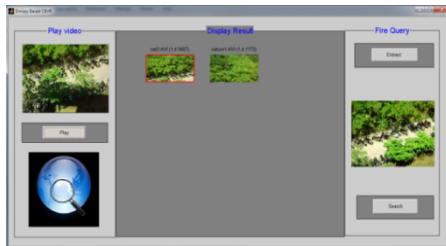
100% matching. Highly contrast videos can also retrieved as shown in Fig 9 though respective videos are not present in database.



**Fig 9: Contrast Color Video Retrieval for Video not Present in Database.**

### Results Obtained by Entropy-Color Based Method

Entropy-color based method retrieves video which are similar in color as shown in Fig 10.



**Fig 10: Video Retrieval in Nature Domain.**

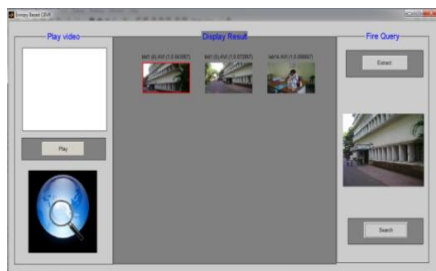
### Limitation of Entropy-Color Based Method

In this method priority goes to color than present object present as shown in Fig 11.



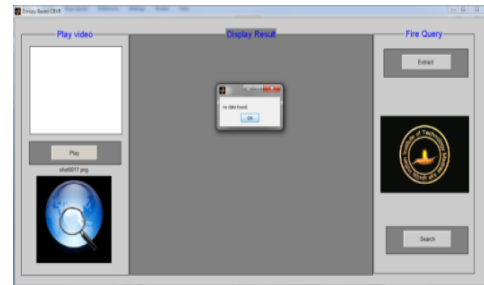
**Fig 11: Color Based Video Retrieval in e-learning Domain.**

Comparing result obtained in Fig 7 and Fig 12 in construction domain conclude that entropy-edge based method is suitable in object specific applications.



**Fig 12: Color Based Video Retrieval in Construction Domain.**

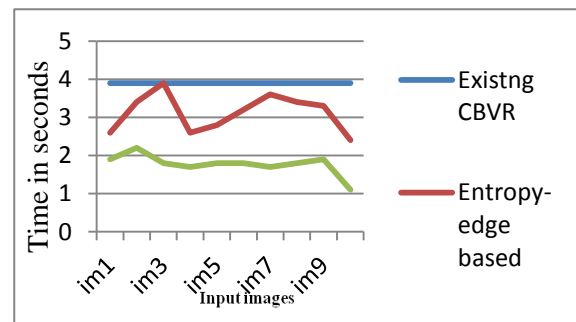
Whereas comparing result obtained in Fig 9 and Fig 13 we can conclude that entropy-color based method restricts highly color contrast videos and suitable in domains like nature scene video.



**Fig 13: Result for Image whose Video not Present in Database.**

### Performance Analysis

Performance of implemented methods is measured by calculating number of videos retrieved for post processing as shown in Fig 14. Where x-axis shows input image and y-axis shows number of videos retrieved for post processing for the input image. In existing CBVR all videos present in database are retrieved. While in implemented system entropy based indexing is used. Videos matching with input entropy are retrieved for post processing in entropy-edge based method. Whereas videos matching with input entropy and RGB color are retrieved for post processing in entropy-color based method.



**Fig 14: Performance Analysis of Implemented Methods.**

### Time Analysis

Time required for retrieving resultant videos for given input image for implemented methods is computed and shown in Fig 15. Where x-axis shows input image and y-axis shows time required to retrieve resultant videos for the input image in seconds. Retrieval time is dependent upon number of videos retrieved for post processing, time increases as number of video increases. Existing CBVR retrieves all videos present in database for post processing hence require constant retrieval time for all images. While in implemented system retrieval time will vary as system works with subset of database. Retrieval time is large in entropy-edge based method as it works on pixel wise comparison. Whereas retrieval time is less in entropy-color based method as it works with scalar values. Retrieval time is average in entropy-color-edge as it combines both methods.



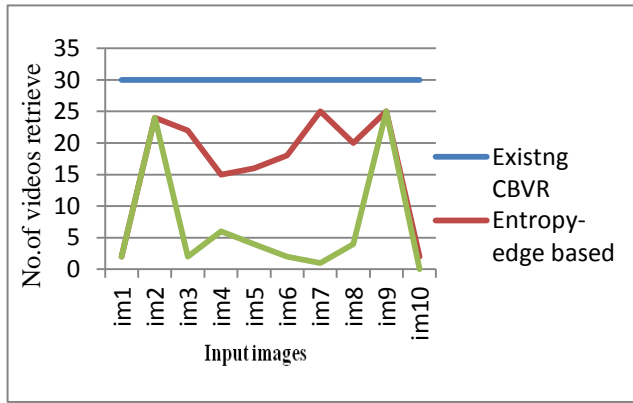


Fig 15: Time Analysis of Implemented Methods.

## 5. CONCLUSION

Content based video retrieval system is implemented using entropy based shot detection method. Entropy feature is used for indexing the video with the combination of RGB color values and video metadata. The entropy based key frame extraction algorithm performs very well when the image background is distinguishable from the objects. Entropy, edge and color features are used for video retrieval. Database with the videos from different domains is created and it is observed that combination of various features in different ways gives comparative result. Combination of entropy and edge features can work better in object specific application domains like news video. Videos retrieved using entropy- edge method may not be similar in color and time consuming. Hence combining entropy and color can improve the result by retrieving videos which are more similar in color. Entropy-color method is suitable in applications which demands similar color videos e.g. nature, water, e-learning videos. While retrieving output videos priority goes to color rather than object present in video. Work can be enhanced further on video retrieval method. Performance can be improved by adding more features.

## 6. REFERENCES

- [1]. Ballan, L. Bertini, M. Serra, G. Del Bimbo, "Video Annotation and Retrieval Using Ontologies and Rule Learning", *Multimedia IEEE*, 2010.
- [2]. Destrero, C. De Mol, F. Odone, and A. Verri, "A Sparsity-Enforcing Method for learning Face Features", *IEEE Transactions On Image Processing*, Vol. 18, No. 1, January 2009.
- [3]. Dongsong Zhang, Nunamaker, J.F., "A natural language approach to content-based video indexing and retrieval for interactive e-learning", *Multimedia, IEEE Transactions*, vol: 6, June 2004.
- [4]. Farouk H., Elsalamony H. A, "Digital library creation based on wavelet coefficients for video stream indexing and retrieving", *Signal Processing Systems*, vol: 1, 2010.
- [5]. Guozhu Liu, Junming Zha, "Key Frame Extraction from MPEG Video Stream", *ISCSCT 09*, pp. 007-011, 26-28, Dec. 2009.
- [6]. Gresle P, O, Huang T S, "Gisting of Video Documents: A Key Frames Selection Algorithm Using Relative Activity Measure," In: *The 2nd Int.Conf.on Visual Information Systems*, 1997.
- [7]. Gargi U. and R. Kasturi, "An Evaluation of Color Histogram Based Methods in Video Indexing", *International Workshop on Image Databases and Multimedia Search*, Amsterdam, pp. 75-82, August 1996.
- [8]. Mori, M. Sawaki, M. Yamato, J., "Robust Character Recognition Using Adaptive Feature Extraction" *Image and Vision Computing New Zealand, IVCNZ 2008*, 26-28 Nov. 2008.
- [9]. Markos M., Alexandra P, "KeyFrame Extraction Algorithm using Entropy Difference", *ACM*, 2004.
- [10]. Natarajan P., B. Elmieh, R. Schwartz, and J. Makhoul, "Videotext OCR using Hidden Markov Models," *Proceedings Sixth International Conference on Document Analysis and Recognition*, pp. 947 – 951, 2011.
- [11]. Patel B. V., Deorankar, A. V.; Meshram, "Content based video retrieval using entropy, edge detection, black and white color features", *Computer Engineering and Technology (ICCET)*, Vol: 6, 2010.
- [12]. Rong Zhao, "Video shot detection using color anologram and latent Semantic Indexing: from Content to Semantics", *citeseerx.ist.psu.edu*.
- [13]. Sankirti S., Kamde P., "Video OCR for video indexing", *International Journal of Engineering and Technology*, Vol.3, No. 3, pp. 287-289, 2011.
- [14]. Tsai D., S. Lai, "Independent Component Analysis-Based Background Subtraction for Indoor Surveillance", *IEEE Transactions On Image Processing*, Vol. 18, No. 1, January 2009.
- [15]. Vrochidis, S. Moutzidou, A.; King P., "VERGE: A video interactive retrieval engine", *Content-Based Multimedia Indexing (CBMI)*, 2010.
- [16]. Wolf W, "Key frame selection by motion analysis," *Proc IEEE Int Conf Acoust, Speech and Signal Proc*, 1996.
- [17]. Zhan, W. Yu, "A method of key frame extraction combining global and local information," *Application Research of Computers*, vol.24, no. 11, pp. 1-4, 2007.
- [18]. Zhu Yingying, Zhou Dongru, "An Approach of Key Frame Extraction from MPEG Compressed Video," *COMPUTER ENGINEERING AND APPLICATION*, pp.13-14, 2003.
- [19]. Zhang Jidong, Chen Du, "Video searching technology based on content," *TV ENGINEERING*, pp.17-19, 2002.
- [20]. Zhang Z, Wu J, Zhong D, "An Integrated System for Content based Video Retrieval and Browsing," *Pattern Recognition*, vol.4, pp.643, 1997.