An application using Wavelet coefficients for Video-content analysis and retrieval

Lakshmi Rupa G.*
Student, M.Tech IT,
VIT University, Vellore
Tamil Nadu, India
e-mail: lakshmirupa.g@gmail.com

Gitanjali J.
Asst. Prof., SITE
VIT University, Vellore
Tamil Nadu, India
e-mail: gitanjalijayaraman@gmail.com

Abstract: The scheme of video-content analysis and indexing which involving structure analysis, abstraction, feature extraction and indexing poses many challenging research issues. This provides a unique opportunity for using intelligent media analysis techniques for creating advanced searching and browsing techniques to find relevant information quickly and inexpensively. This research mainly is to survey available and potential technologies for video monitoring and mining, the general methods of fast and efficient content-based analysis of video streams and to identify promising directions for research in this challenging area. This involves automatic detection of boundaries between the shots in a video and then those are indexed to form a library, saving the proper features of each shot/frame. This helps in the easy retrieval based on the shot according to the user requirements.

We present an automation technique for video indexing and creation of a digital library, which is composed of stream shots and the wavelet coefficients for these shots. The wavelet coefficients are computed on the image and all the video frames for a full search function in all the frames of the indexed video and can be used for any number of shots/frames.

Keywords: digital library, wavelet coefficient, indexing, image retrieval, shot cut, video streams/ shots.

I. INTRODUCTION

Videos, which are one of the most important media, have been widely used in many applications for information recording. For the increasing requirement of information mining, video analysis has drawn more and more attentions with the purpose to realize effective video content access.

There are a number of methods used for video stream indexing and retrieving. Some of the methods are content based analysis (content-based video data indexing and retrieval systems utilizing the syntactic contents of the video sequences such as color, texture, and shape), motion trajectory indexing and retrieval techniques, Adaptive video indexing technique.

Video database applications include the following,

- Verifying identity at ATM machines;
- Verifying identity for the automated issue of driver's licenses; and
- Searching photo ID records for fraud detection (multiple driver's licenses, multiple welfare claims, etc).

These applications require suitable indexing technique to capture the time-varying nature if video data, and a high performance retrieval strategy. These techniques are not feasible for video indexing, so we use wavelet coefficient [1], [4] for video indexing and retrieval.

The researches on video content mainly involve automatic detection of the boundaries between camera shots. A shot is an unbroken sequence of frames from a camera. A scene is a collection of one or more adjoining shots that focus on an object or objects of interest. There are a number of different types of transitions or boundaries between shots [3].

The wavelet wavelength coefficients are obtained from the video image frames and stored in the database for each frame/shot. Now, for retrieving any shot within streams database, which were built for frames ranged from key frames till the end frame in the same shot and applicable to all shots, we apply a mathematical wavelength to get coefficients. These are the features that are saved. Thus obtained similar images can be used in the application of search engines of videos, which can be done by clustering of the image frames.

Image retrieval systems [6], [7] that compare the query image exhaustively with each individual image in the database are not scalable to large databases. A scalable search system should ensure that the search time does not increase linearly with the number of images in the database.

Many of the image retrieval systems extract specific features from a query image and compare these features with the corresponding pre-computed features of all the images in the database. Here, this process is done by wavelet decomposition.

Searching large databases of images is a challenging task especially for retrieval by content. Most search engines calculate the similarity between the query image and all the images in the database and rank the images by sorting their similarities.

One problem with this exhaustive search approach is that it does not scale up for large databases. So, we use the mathematical way of wavelet coefficients for each image and then compare the values for easy image retrieval.

II. VIDEO INDEXING TECHNIQUE

There are number of methods that are used for video stream indexing and retrieving. Some of the methods are content based analysis (content-based video data indexing and retrieval systems have utilized the syntactic contents of the video sequences such as color, texture, and shape), motion trajectory indexing and retrieval techniques, Adaptive video indexing technique. Video database applications require a suitable indexing technique to capture the time-varying nature of video data, and a high performance retrieval strategy. These techniques are not feasible for video indexing. So we use wavelet coefficient for video indexing and retrieving.

III. WAVELET DECOMPOSITION OVERVIEW

Wavelet transform has become a popular tool for multi resolution image analysis [4]. Here, we compute the approximation coefficients matrices CH, CV and CD (horizontal, vertical and diagonal), obtained by wavelet decomposition for the input matrix of the extracted frames from some shots of a stream.

The videos are stored in a compressed MPEG format and these indexing techniques may include transform domain techniques using Fourier transform, cosine transform, Karhunen-Loeve transform, Sub-bands and wavelets and spatial domain transform using vector quantization and fractals.

Wavelets are families of functions generated from one single prototype function by dilation and translation operation [2]. There are some predefined families of wavelets, namely, Haar, Daubechies, Coiflets, and Symlets.

The principle of image wavelet decomposition is considered by taking an image having one column. The decomposition stage includes the processing of the images matrix by columns at first using wavelet (high-pass) and scaling (low-pass) function followed by row down sampling by factor D.
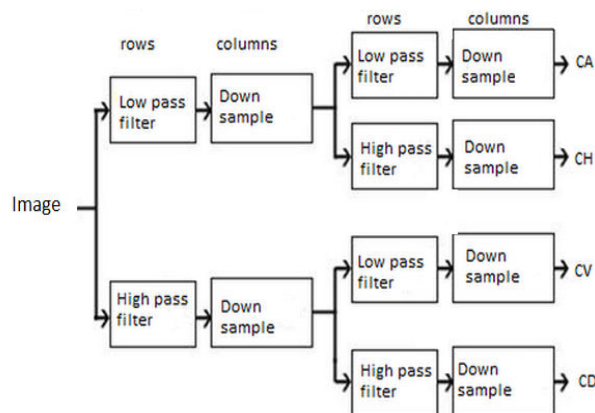

Figure.1 2D wavelet decomposition followed by down sampling

Taking a selected column of the image matrix G (n,m) as signal

$$\left\{ x(n) \right\}_{n-0}^{N-1} = [x(0), x(1),....x(n-1)]^T$$

This signal can be analyzed by a half-band low pass filter represented by the scaling function with its impulse response, and the corresponding high-pass filter represented by the wavelet function based upon impulse response.

IV. PROPOSED DIGITAL LIBRARY SYSTEM

Video mining is extracting information from video data using image and video processing techniques. E.g.: detecting special events, finding the similar video clips. The amount of digital video content has grown extensively during recent years, resulting in a rising need for the development of systems for automatic indexing, summarization, and semantic analysis.

In this project, we implement an algorithm for indexing and retrieving video streams by creating digital library for video frames separating the stream shots and then indexing the shots frames based on wavelet coefficients for each frame in the shot.(Fig 2). First, we detect the shot cut out of the whole video stream. Compute horizontal, vertical, diagonal wavelet coefficient matrix for each frame in a video stream and save them into the database. Using this wavelet coefficient retrieve the corresponding video shot.

In this System, we detect the shot cut out of the whole video stream. Then, we compute horizontal, vertical, diagonal wavelet coefficient matrix for each frame in a video stream and save them into the database. Using this wavelet coefficient retrieve the corresponding video shot.

Retrieval is performed by comparing features of a query image with corresponding features of images stored in a database and presenting the user with the images in the database that have the most similar features.

The goal of this system is to allow for the user to search for given input image in the video.
The components of this work include,
   i.   Indexing the video stream
   ii.   Building the wavelet coefficients database for the video stream based on shots.
   iii. Retrieving video shots based in the input image.
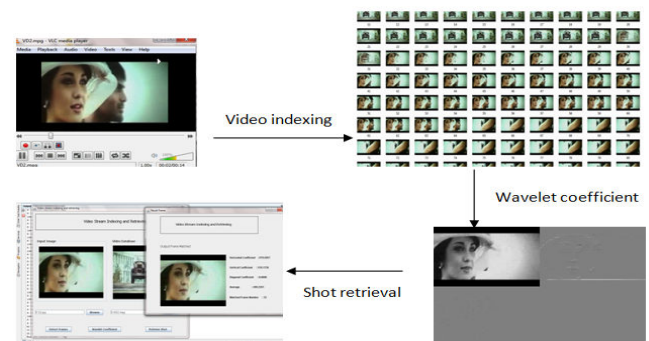

Figure.2 Architectural representation of the system

A. Indexing Video shots/frames

The shot cut out of the whole stream based on any type of cutting, i.e., gradual cut or sharp cut is detected. Then, the features (wavelet coefficients matrices) are saved to help in retrieving the targeted shot. A shot is a series of pictures represented in a continuous action in time and space.

This step is mainly for automated indexing and image retrieval application providing an efficient access to huge video archives.

Figure 3. Indexing of video streams

B. Wavelet Coefficient database

The wavelet coefficient is applied on each frame in all stream shots and stored in a database. They are computed by the horizontal, vertical and diagonal wavelet's extended matrix scaling. Also compute the standard deviation for the above values. Store all the coefficient values into the database.
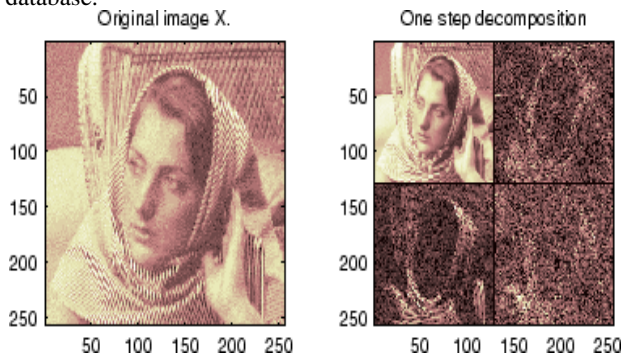
Figure 4. 2D image decomposition input frame, approximation coefficient matix CA, diagonal CD, horizontal CH, and vertical CV

After the decomposition, the image consists of approximation coefficient matrix CA in the upper left, the coefficient matrix for diagonal CD in the upper right, the coefficient matrix for horizontal CH in the lower left, and the last coefficient matrix is vertical CV in lower right.

Then we compute standard deviation for each of the above matrix, thus obtain CH1, CV1, and CD1 for the indexed frames. Save CH1, CV1, and CD1 in the database.

C. Shot Retrieval

A mathematical wavelet transformation is applied here to get coefficients for retrieving a shot within the streams database. These are the features which will be saved. The wavelet transform coefficient is obtained as CH2, CV2, and CD2.

Then we compute the correlation coefficient for ch=(CH1,CH2), v=(CV1,CV2), and cd=(CD1,CD2) to obtain the best correlation between database frames and the input frame. Then, we compute the mean for (ch, cv, cd), to obtain the mean number for input frame and one frame from database frames. The maximum value in this mean matrix gives the output matching frame.

IV. INDEXING AND RETRIEVAL ALGORITHM SUMMARY

Here, we consider a video stream composed of 30 shots and 420 frames as an example of short shots, fast content rate of change which helps to prove the efficiency of the algorithm.

I. *Indexing algorithm:*

- Apply shot detection technique for the video stream.
- Calculate the horizontal wavelet coefficients for each frame ranged from 1 to 12 for testing, and then the same for all frames in the shot within each shot and the results are matrices with same size (the first 11 frames for each of the 30 shots.
- Repeat the last step for vertical and diagonal.
- Compute standard deviation for each of the aboce matrix, then we get CH1, CV1, and CD1 for the indexed frames.
- Save CH1, CV1, and CD1 in the database.

II. *Retrieving algorithm:*

- Compute wavelet transform coefficients as in the above section to the input frame. Thus, we obtain CH2, CV2, and CD2.

- Compute the correlation coefficient for ch=(CH1,CH2), cv=(CV1,CV2), cd=(CD1,CD2), to get the best correlation between database frames (all 420 frames), and the input frame.
- Compute the mean for (ch,cv,cd), then we get mean number for input frame and one frame from database frames, and so on.
- Select the maximum value in this mean matrix; we get the output matching frame, which this value represents in this matrix.

V. RESULTS AND DISCUSSION

As a description for table's structure, the first column represents the test input frame image. The second column represents the test input frame number, and the third column represents the calculated valure from wavelet coefficients normalized by multiplying each value from wavelet coefficients normalizes. The fourth, fifth and sixth columns is calculated wavelet coefficients. The seventh column represents the shot number. The eighth column represents the retrieved of frame we compare the input frame with, within the associated shot. The ninth column represents the retrieved frame.

Here, we start by the test input frame in the first column, which have its location number in the stream in column two. By the calculations of ch, cv, and cd we get the maximum number of the average of them in third column. This maximum number tends to the matched keyframe, which exists in last column with the shot number and location in the stream specified.

This novel idea applied may help in many applications like digital library; especially in retrieval which need to extract automatically the shot in an indexed stream by just having any frame of the shot. The different sources of noises effecting the video stream effect the algorithm because of different wavelet coefficients. This is scalable to any number of video streams, any number of shots, for shot indexing and wavelet coefficient calculations.

VI. CONCLUSION

In this paper, we innovate and implement an algorithm for indexing and retrieving video streams by creating digital library of video frames separating the stream shots and then indexing the shots frames based on the wavelet coefficients for each frame in the shot.

Thus, we achieve indexing of the video streams frames, which is a part of digital library creation process and wavelet coefficients for each frame is also calculated. We can even retrieve any shot from the digital library efficiently. Thus created digital library has many applications, specially for retrieval of any frame of the shot automatically in the indexed stream. This is scalable to any number of video streams, any number of shots, for creating the shot indexing and wavelet coefficient calculations. This novel algorithm for video digital library helps in indexing and retrieval based on wavelet transform. This also helps to save a feature representing each stream frames which helps in archiving and retrieving process on stream shots.

Table 1: Summary of Results

Input Image	Input frame number	Maximum value in all stream shots Average of (ch, cv, cd)	Cd	Cv	Ch	Shot Number	Output frame no.	Output Frame Matched
	12	442.2373	409.4530	496.3002	420.9585	1	1	
	58	590.4133	517.8125	654.5256	598.9020	2	49	
	136	959.2345	962.9787	958.2576	873.9633	3	124	
	178	630.2696	762.6789	520.7567	607.3818	4	168	
	210	750.6128	747.6382	753.5437	750.4148	5	202	
	235	638.9499	594.6578	695.6793	628.9867	6	230	
	306	757.8675	650.7632	736.5672	884.8762	8	297	
	344	968.9843	980.5627	965.4723	968.4592	9	342	
	360	766.8746	800.5483	706.5921	705.5927	11	355	

VII. ACKNOWLEDGMENT

VIII. REFERENCES

[1] D. Tsai and S. Lai, "Independent Component Analysis-Based Background Subtraction for Indoor Surveillance", IEEE Transactions On Image Processing, Vol. 18, No. 1, January 2009

[2] A. Smeulders, M.Worring, S. Santini, A. Gupta, and R. Jain, "Content based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349–1380, December 2000..

[3] M. Nixon Alberto S. Aguado, "Feature Extraction and Image Processing", Second edition, book Elsevier Ltd., 2008.

[4] Tao Li, Qi Li, Shenghuo Zhu, Mitsunori Ogihara, "A Survey on Wavelet Applications in Data Mining", SIGKDD Explorations, Volume 4, Issue 2, Pg 49-68.

[5] K. Yoon, D. F. DeMenthon, and D. Doermann, "Event detection from MPEG video in the compressed domain," in Int. Conf. on Pattern Recognition, Barcelona, Spain, 2000.

[6] Y.Alp Aslandogan and Clement T. Yu, "Techniques and Systems for Image and Video Retrieval", IEEE Trans. On Knowledge and Data engineering, Vol.11, No. 1, 1999.

[7] V. Kobla, D.S. Doermann, and K-I. Lin, "Archiving, indexing and retrieval of video in compressed domain," in SPIE Conference on Multimedia Storage and Archiving Systems, 1996, vol. 2916, pp. 78-89

[8] A. W. M. Smeulders et al., "Content-Based Image Retrieval at the End of the Early Years, "IEEE Transactions on Pattern Analysis and Machine Intelligence (2000): 1349-1380.

[9] M. S. Lew et al., "Content-based multimedia information retrieval: State of the art and challenges," ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP) 2, no. 1 (2006): 1-19.