IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

An Enhanced Approach in Novel Content Based Video Retrieval Using Vector Quantization

Prof. Rahul Gaikwad[1], Jitesh R. Neve[2]

Assistant Professor, Department of Computer Engineering., G.F.'s Godavari College of Engineering, Jalgaon, India[1]

Master of Engg, Department of Computer Engineering, G.F.'s Godavari College of Engineering, Jalgaon, MH, India[2]

Abstract: In the recent and modern world, almost everything is in the race to be digitized rapidly. Today, end-user sometimes thinks that video retrieval system based on simple text query is degrading the performance these days. So it's a good time to move on to the content based retrieval approach based systems for searching videos. Revision of the content based video retrieval approach can lead to the effective implementation of this system. Content Based Video Retrieval (CBVR) is the best evolving system for any video retrieval application. The Block Truncation Coding (BTC) is one of the techniques in the CBVR used for color feature extraction. With improvements to the BTC, Thepade's Sorted Ternary Block Truncation Coding (TSTBTC) is also the recent color feature extraction technique. Transform feature extraction is one more technique about extraction of the video. None of the method is right now extended for both of the features like color feature and transform feature of a video. In other hand Vector Quantization (VQ) is the lossy data compression technique. If VQ is used with TSTBTC it can deal with both of the features like color and transform feature of a video. Because the Vector Quantization (VQ) supports hybrid features (i.e. color and transform). VQ is not used before in CBVR, this paper explains the implemented stuff of VQ with CBVR.

Keywords: Content Based Video Retrieval (CBVR); Block Truncation Coding (BTC); Thepade's Sorted Turnary Block Truncation Coding (TSTBTC); Vector Quantization (VQ); Linde-Buzo-Gray (LBG).

I. INTRODUCTION

When we try to explain about a video, we don't need to tell in the brief. Video is just a part of everyone's life. We know that a video is the need of today's life. For sure a strange rise in the storage of the videos over web spaces has been seen recently. Everyone takes video, watches video, uploads video, find video because these are the common operations on the videos. There are many places for searching a video, also effective video retrieval is the timely need. It has many techniques to retrieve a video. Till date content-based video retrieval system has many applications for example – faster searching of video, analysis of visual electronic commerce, isolated instruction execution, digital museums, news event, intelligent web videos management, video surveillance, etc. [1, 2, 3].

A. Typical structure of a video

Video can said as an ordered list of frames presented at a particular frame-rate (i.e. number of frames per second). A frame is nothing but an image extracted at tiny fraction of second. A set of consecutive frames which are arranged in order is called a shot. A Shot consists of meta-data, total number of objects in the image with foreground and background visual information. A typical video structure is shown in fig. 1 in detail [1, 7].

In the recent era, video retrieval is mostly concerned with the text queries. Everyone concentrates on the precise query to get effective and optimal results of the retrieval, this is non-perceptive to the instinct. So the content based video retrieval (CBVR) comes into picture. Content based video retrieval (CBVR) focuses on the actual contents of the video. It focuses on the insides like color, texture, shape, etc. [1, 6].

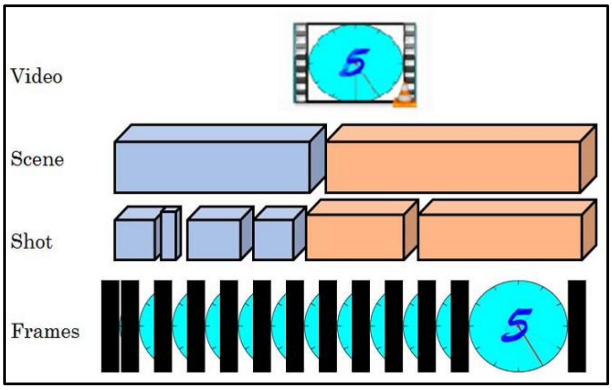

Fig. 1 structure of a video

B. History of video retrieval

Recently video retrieval is accomplished with help of textual features involved with video. The throughput of text based video retrieval system deeply relies upon annotations attached with video. This entire process is manual as well as this entire job is overwhelming in means of time as the annotations are subjective. Videos are typically consisting of audio and images (visual data) represented with a definite frame-rate. Visual contents (i.e. images) of a video involve various colors, outlines, may be

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

gesture or texture. As well as the malformed video data presents video content based on energy dissemination. This energy content of any video in transformed domain is independent of illumination and rotation variations. After that this energy contents can be considered in fractional form to decrease the size of video content for feature extraction, which leads to faster retrieval [5].

C. Content Based Video Retrieval (CBVR)

From this we got to know that retrieving videos with the efficient way is now became an hour's need. Because video retrieval is the fundamental operation in numerous video oriented applications. Now a days, majority of the video retrieval systems mostly focuses on query-string for the retrieval of results which is not intuitive for user's instinct, so there is the need of content based video retrieval (CBVR). The result of the CBVR heavily relies on the actual contents of the videos like shape, color, texture, etc. [1, 4].

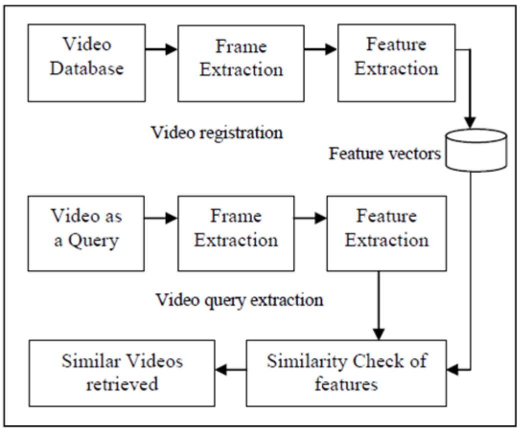

Fig. 2 Flow of typical CBVR process

Content based video retrieval always depends on the verbal depiction of some substratal extracted features such as shape, color, texture, etc. that can be extracted automatically from videos themselves. Queries to a specific CBVR system are very frequently stated as color feature-vector of a video and the similarly evaluation can also be done on the query and the most matching video-results are retrieved [1, 2].

II. VECTOR QUANTIZATION

Vector Quantization (VQ) is one of the well-known data compression techniques in lossy way. VQ algorithm contains the practice of clustering. VQ prepares the code-book for each and every frame just same as a signature. And on top of it, it is evaluated in accordance with a specific VQ algorithm [1]. In simple words, VQ is simply a algebraic mapping function which maps 'k' number of dimensional vector space to a actual finite set CB = {C_1, C_2… C_N}. The given set CB is called as a code-book. Which is usually involving N number of code-vector and each code-vector C_i= {c_{i1}, c_{i2}....c_{ik}} is of 'K' dimension. There are many different Codebooks sizes in the multiples of the power of 2. For example {2, 4, 8, 16, 32, 64, 128, 256, 512, …..}. Basically they all are in the multiples of 2n [1, 4].

Vector Quantization can be intuitively used by many platforms like mobile, tablets, standalone systems, servers, etc. for their use in many applications, such as detection of speech, biometric applications, CBIR, etc. [11].

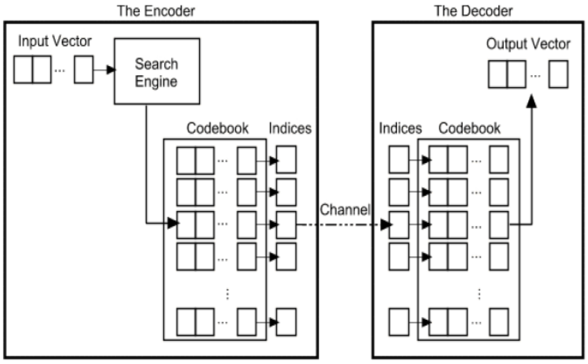

Fig. 3 Encoder and decoder in Vector Quantization

The method of generating codebook defines that how one video-frame is split into numerous 'k' vectors of dimensional training as mention above C= {C_{i1}, C_{i2} ..C_{ik}} and various techniques of codebook generation are applied. Here the process of encoding states that, a typical video frame is unglued into some vectors into 'k' dimensions. After that each vector is encoded by an index of Code-Word by 'look-up table' method. The encoded results of Code-Word are then kept in an 'index-table'. During the process of decoding, the receiver block uses the same 'Code-Book' to transform that index back to its resultant Code-Word for recreating the image again [4].

A. Linde-Buzo-Gray (LBG) technique for CodeBook generation

One of the popular scheme which is mostly used to generate codebook is the Linde-Buzo-Gray (LBG) algorithm. In this technique, the centroid is calculated over the first Code-Vector for the training set. Further two vectors C1 & C2 are generated by adding and subtracting a constant error value to from the 'Code-Vector'. Using 'Euclidean Distance' all of the vectors are compared with vectors C1 & C2 and two clusters are created on the basis of the nearest of C1 or C2. Following is the algorithm for Code-Book generation using LBG Code-Book generation technique [4].

Stage 1: Read an image and fragment the R (red), G (green) and B (blue) color components of an image into uncommon chunks and change each of the block to a vector form a training vector set.

Stage 2: Compute (generate) the centroid of a training vector set that contained; It can be also called as Code-Vector.

Stage 3: Code-Vector we can determine value of a constant error vector which can be added and subtracted to form two Code-Vector (i.e. C1 and C2).

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

Stage 4: We can calculate the distance between training vector and Code-Vector C1 and C2 using 'Euclidean Distance' similarity measure. After that we divided the cluster into two on nearest basis.
Stage 5: Re-compute the centroid (Code-Vector) for clusters to obtain in the above stage 4.
Stage 6: Re-iterate from stage 2 to stage 5 till the required codebook size is obtained
Stage 7: halt

III. PROPOSED CONTENT BASED VIDEO RETRIEVAL USING LBG

There are two phases in the retrieval system. First one is registration phase which divides every video into many key-frames and feature-vector is computed over them. Here in registration phase feature vector database of all of videos is constructed. After first phase, the second one (query execution phase) defines that how the queried video should be processed further to form a feature-vector in a same manner as in first phase.
The query (i.e. feature vector in query) is equated with the available data-set of the feature-vector using various similarity measures. Throughput of this system is computed over the precision. Results found most relevant are better for the retrieval system.

There are two major phases in the video retrieval process those phases are as follows –

A. Registration phase
Whenever a new video is included in the collection, it usually evaluated to calculate the color feature vector. However that color feature vector normally represents the ratio of pixels of each of the color within the same video. After calculating color feature vector for each video, it is arranged in the data-source [5].

B. Retrieval Phase
There are numerous approaches of retrieving videos based on color similarity. It is explained there above, but most of those are variations of the same core idea. Here to extract the feature vectors, the BTC applied over video frames [5].

IV. EXPERIMENT ENVIRONMENT

Experiment environment defines the standard sets of entities which we thought useful to identify the best fits or pros/cons of the proposed systems. We tried categorizing them in following way.

A. Video test bed
We know that, there are several types of videos with categorization of size, resolution, frame rate, no of frames, etc. This means that there is no standard (pre-defined) data set for CBVR system. So here we have taken a standard set of 500 videos with 10 different categories (i.e. 50 videos per category) to observe the performance and result analysis of the proposed technique. The test bed used for the experiment over 10 categories of the video as shown in fig. 4

B. Platform details
The software platform required for trialing this technique is MATLAB R2012a. It's done on the systems having CPU core i5 and 8 GB of RAM.
We experimented the CBVR using VQ by firing queries of all 500 videos over the data set. Let's see further analysis in the results and discussions point.

Fig. 4 Video test bed used for experiment

V. RESULTS AND DISCUSSION

Here we have extracted the results after applying the proposed algorithm over the videos present in the test bed. We just have extracted the output of the algorithm i.e. Code-Vector with different code-book sizes. The code-book size differs in the power of 2; this means it will be {2^1, 2^2, 2^3, 2^4, …} so similarly the set of values for code-book is { 2, 4, 8, 16, 32, 64, 128, …}.
We have processed the test bed of 500 videos with code-book size up to 64 (i.e. up to 26). On that data we have calculated the mean precision over the results of the processing for Euclidean Distance similarity measure for matching or searching the video. The performance graph is as described in the fig. 5.

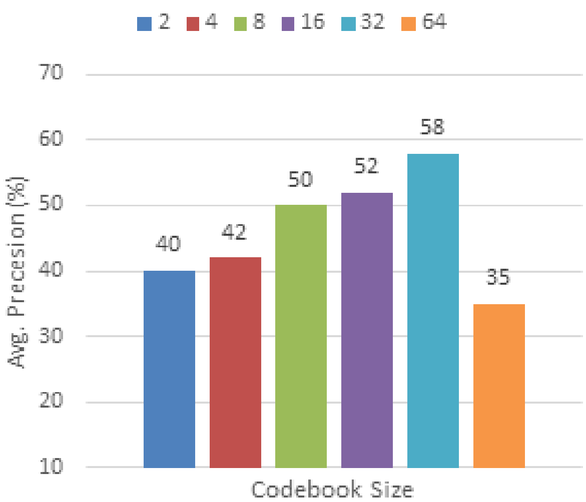

Fig. 5. Performance of the algorithm by Euclidean

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 7, July 2016

Distance similarity measure for different code-book sizes Adding to that, from this chart of the results as given in fig. 4, code-book size 32 (i.e. 25) has maximum average precision of 58%. Which is the highest precision for Euclidean Distance similarity measure. We thought code-book size 64 will give higher results, but now results are far below unlike as expected for Euclidean Distance similarity measure. Code-book size 64 has only 35% mean precision. This may true for different similarity measure.

VI. CONCLUSION

This paper gives light on the use of Vector Quantization with the content based video retrieval systems. As this is not used before, we have tried to present this concept with some results after huge experiments over 500 test bed videos with 10 different categories.

Among different similarity measures, we have evaluated performance only for Euclidean Distance similarity measure. The highest percentage of average precision is shown for code-book size of 32. In other words, best performance is given by code-book size of 32 with Euclidean Distance similarity measure which is 58%.

Moving ahead it will definitely be more remarkable to explore the evaluation range with ensembles of different similarity measures for the betterment of Content based video retrieval.

ACKNOWLEDGMENT

Thanks Rahul Gaikwad sir for the active interest in this research area. We would like to extend our thanks to Swapnil Gore, Rahul Somwanshi, Umesh Patil, Pravin Bhole, Dilip Patil and Pravin Patil for their help in sharing their systems to experiment, without which we could not complete the video analysis within time. Special thanks to **Sanjay Patil** (sir), **Kiran Patil** and **Pritam Patil** for their help, co-operation and moral support. All of them were there for us even at midnight for any kind of help. Finally, we would like to express our deep sense of gratitude towards our parents and friends who have supported us in all possible ways for successful completion of this paper.

REFERENCES

[1] Dipak Pardhi and Jitesh Neve, “Performance Rise in Novel Content Based Video Retrieval Using Vector Quantization,” International Conference on Electrical, Electronics and Optimization Technique (ICEEOT), Chennai, 2016.

[2] Dr. Sudeep D. Thepade and Nalini B. Yadav, “Assessment of Similarity Measurement Criteria in Thepade’s Sorted Ternary Block Truncation Coding (TSTBTC) for Content Based Video Retrieval,” International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, 2015, pp. 1-6.

[3] Dr. Sudeep D. Thepade, Krishnasagar. S. Subhedarpage and Ankur A. Mali, “Performance Rise in Content Based Video Retrieval using Multi-level Thepade's sorted ternary Block Truncation Coding with intermediate block videos and even-odd videos,” International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore, 2013, pp. 962-966.

[4] Sudeep D. Thepade and Pritam H. Patil, “Novel Keyframe Extraction for Video Content Summarization using LBG Codebook Generation Technique of Vector Quantization,” J. International Journal of Computer Applications (0975 – 8887), vol-111, no-9, pp. 49-53, Feb, 2015.

[5] Dr. Sudeep D. Thepade and Nalini B. Yadav, “Novel Efficient Content Based Video Retrieval method using Cosine-Haar Hybrid Wavelet Transform and Fractional Energy,” International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, 2015, pp. 615-619.

[6] Sudeep.D.Thepade, Krishnasagar Subhedarpage, Ankur.A.Mali and Tushar.S.Vaidya, “Color Content based Video Retrieval using Block Truncation Coding with Different Color Spaces,” J. International Journal of Computer Applications (0975 – 8887), vol-64, no-3, pp. 35-38, Feb, 2013.

[7] Sudeep D. Thepade and Pritam H. Patil, “Novel Visual Content Summarization in Videos using Keyframe Extraction with Thepade's Sorted Ternary Block Truncation Coding and Assorted Similarity Measures,” International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, 2015, pp. 1-5.

[8] Aasif Ansari and Muzammil H Mohammed, “Content based Video Retrieval Systems - Methods, Techniques, Trends and Challenges,” J. International Journal of Computer Applications (0975 – 8887), vol-112, no-7, pp. 13-22, Feb, 2015.

[9] Dr P. N Chatur and R. M. Shende, “A Simple Review On Content Based Video Images Retrieval”, J. International Journal of Engineering Research & Technology (IJERT) (ISSN: 2278-0181), vol-2, no-3, pp. 1-6, Mar, 2013.

[10] W. H. Equitz, “IEEE Transactions on Acoustics, Speech, and Signal Processing”, IEEE Journal (ISSN: 0096-3518), vol-37, Issue-10, pp- 1568 – 1575, Oct, 1989.

[11] Y. Linde, A. Buzo, and R. M. Gray, “An algorithm for vector quantizer design,” IEEE Trans.Commun., vol.COM-28, no. 1, pp.: 84-95, 1980.