Indonesian Journal of Electrical Engineering and Computer Science
Vol. 5, No. 1, January 2017, pp. 214 ~ 219
DOI: 10.11591/ijeecs.v5.i1.pp 214-219

Parallelizing Multi-featured Content Based Search and Retrieval of Videos through High Performance Computing

Azra Nasreen*, Shobha G
Department of Computer Science & Engineering, R V College of Engineering, Bangalore, India
*Corresponding author, e-mail: azranasreen@rvce.edu.in

Abstract

Video Retrieval is an important technology that helps to design video search engines and allow users to browse and retrieve videos of interest from huge databases. Though, there are many existing techniques to search and retrieve videos based on spatial and temporal features but are unable to perform well resulting in high ranking of irrelevant videos leading to poor user satisfaction. In this paper an efficient multi-featured method for matching and extraction is proposed in parallel paradigm to retrieve videos accurately and quickly from the collection. Proposed system is tested on datasets that contains various categories of videos of varying length such as traffic, sports, nature etc. Experimental results show that around 80% of accuracy is achieved in searching and retrieving video. Through the use of high performance computing, the parallel execution performs 5 times faster in locating and retrieving videos of intrest than the sequential execution.

Keywords*: Content Based video retrieval, CUDA, High performance computing, Multi -feature extraction*

1. Introduction

Rapid technological advancement in communication through Internet and social media, sharing of multimedia is considerably increased giving rise to large video collections which in turn requires development of sophisticated video processing applications. Content Based Video retrieval system is one such video processing application that searches for content of interest specified by the user in huge collection of videos. Video retrieval is one of the fastest growing research challenges in the field of multimedia technology. In modern video retrieval systems, the basic need is to handle the faster growing database with the efficiency in retrieving videos. Lot of developments are undergoing to improve the efficiency of video retrieval systems. Content Based Video Retrieval Systems (CBVRs) [1] [14] is an important technology that facilitates efficient search and retrieval of video sequences from large collections. Videos with useful information occupying significant space in the databases are under-utilized unless CBVR systems capable of retrieving desired videos are used to sharply selects relevant videos. Many existing methods are unable to perform well due to inappropriate selection of feature set, improper feature extraction techniques for the selected features leading to retrieval of undesired videos resulting in poor user satisfaction.

The proposed video retrieval system efficiently performs parallel search in video databases with dynamic feature extraction using parallel methods to cope with the large number of videos in a database with high accuracy. Any retreival system based on contents involves fundamental steps such as Video segmentation, Feature Extraction, Similarity Check and Retrieval of best match videos.

As videos have much richer content with lot of raw data and very little prior structure, searching and retrieval of videos is quite difficult. An individual video possesses unique motion features, color histograms, motion histograms, text features, audio features, features extracted from faces and objects existing in it. These features can be extracted and feature vectors can be composed which would then be used to index, classify and retrieve desired and relevant videos while filtering out undesired ones.

Received August 29, 2016; Revised December 3, 2016; Accepted December 23, 2016

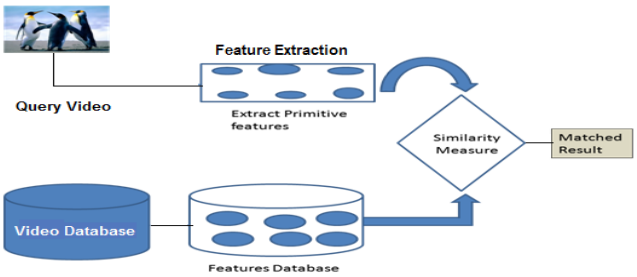

Figure 1. Content based video retrieval system

Figure 1 shows the general structure of content based video retrieval system. Video segmentation is carried out upon the data set to obtain scenes, shots and then frames. These frames need to be processed so that contents of the videos can be used in indexing of videos. Identification and selection of efficient feature extraction technique has to be carried out to compose feature vectors to be stored as feature library for further use. Whenever the query clip is posted, its features are estimated and compared with feature library using a similarity measure. The relevant or closest match videos are displayed as relevant results.

Proposed system is mainly built to achieve better performance and high accuracy of video retrieval through high performance computing platform such as CUDA (Compute Unified Device Architecture) along with multiple feature extraction. The performance and accuracy factors of the approach are the motivating characteristics that lead to carry on the development of the work. It extracts both texture and color features using LBP operator and Mean Variance and Skewness model respectively and stores those values in the vector database. Performs classification and clustering of videos followed by video indexing based on the vector values obtained. Finally computes the similarity measure between the values of query video and videos in the database using KNN algorithm with Euclidean distance method.

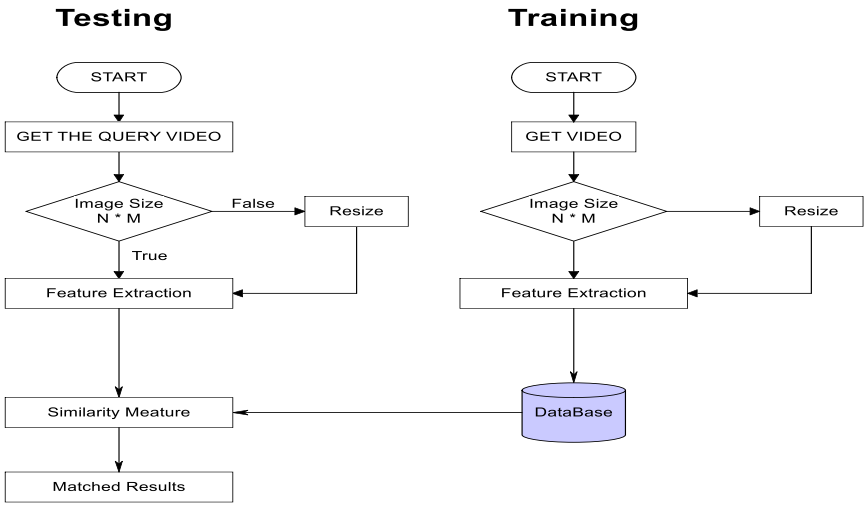

Figure 2. Steps involved in the proposed content-based video retrieval system

Figure 2 shows the representation of the actual working of proposed system. Parallel version of the proposed method is developed using CUDA, which is a parallel computing environment and a programming pattern developed by NVIDIA [15]. Basically, CUDA programming architecture consists of components namely, main memory, CPU, Graphics Processing Unit (GPU) and the GPU Memory. The process flow of CUDA programming is as shown in the Figure 3.

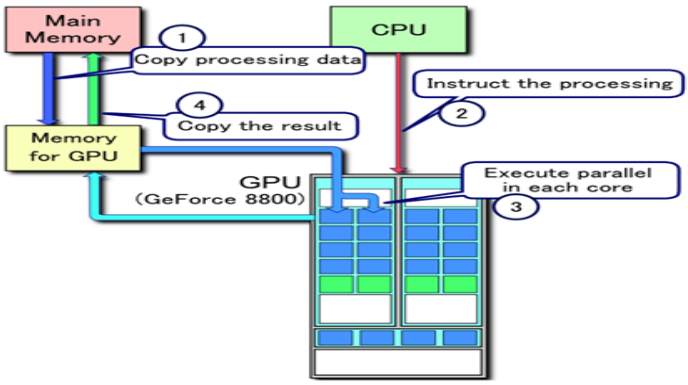

Figure 3. Process flow of Compute Unified Device Architecture

The processing steps are as follows:

1. Data is copied from the main memory to the GPU memory.
2. GPU is instructed by the CPU to process the instructions.
3. Parallel execution in each core at the GPU memory.
4. Result from GPU memory is copied back to main memory.

The above said platform is selected as video retrieval is a computationally complex task. The serial and parallel runs are compared to appreciate the usefulness of parallel paradigm.

2. Research Method

The search and retrieval method [2] portrays the data set using Video Segmentation and Optimal Key Frame. The feature vectors are generated from VSR using texture analysis using Local Binary Pattern (LBP) technique and for color moments YCbCr model is used. Then based on the calculation of probability of occurrence of pixel intensity values with pixel location among frames, optical frames are obtained. Similarity check is done using Weighted Distance Measure. An optimal key frame representation scheme for video segmentation is proposed [3] uses video shot detection and retrieval considering both spatial and global color statistics of the objects in the frames. A technique called Temporally Maximum Occurrence Frame (TMOF) is proposed as a new method for video shot representation. Later at each pixel position of a video shot, k pixel values with highest probability of occurrence are considered to enhance the performance. The resultant methods are named as k-TMOF and k-pTMOF. Finally a comparison is made between these new schemes. The experimental results prove that the methods achieve better performance of retrieval. A method [4] for automated video indexing and video search in large lecture video database is proposed, wherein automatic video segmentation and key-frame detection is done. Textual metadata is extracted by OCR on key-frames. Automatic speech recognition is used on lecture videos. This method is dependent on the text contents of the video. Another method to gain performance over retrieving the videos using intermediate block truncation coding technique [5] divides each frame into blocks based on dimensions row wise and column wise. Then Block Truncation Coding is used for feature extraction. For each block seven different color spaces are applied to calculate the pixel intensity values. From each block the average or mean is calculated upon all seven color spaces and each color index is used to get the final vector. The same step is carried out for all other blocks and all other frames. Later all feature vectors are stored in a database. Then query feature vectors are compared with these feature vectors using absolute difference. Same method is followed [6] but for odd and even videos. An approach called mesh- local binary pattern (LBP) applied for 2D LBP variants adopted for 3D texture patterns of triangular mesh surfaces is proposed in [7]. A method for addressing rotation invariance is also given which confirms the repeatability factor. The work offers effectiveness, generalization, adaptability and simplicity factors. The experimental results prove the uniformity for various types of scalar functions and allow extension on all variants. Techniques for texture feature includes statistical

analysis and structural analysis are discussed [9] and comparative analysis is made on each of the technique. There are many existing techniques and mechanisms to perform video retrieval on content basis. The default text based [10] and color based [11] [12]. video retrieval techniques and also techniques involving multiple features such as color and texture, texture color and shape are proposed. Techniques for shape feature include shot boundary detection and edge boundary detection and for non-visual features: captions, annotations, relational attributes and structural descriptions can be considered. There are many existing techniques and mechanisms to perform video retrieval on content basis [15]. However, most of them are unable to perform well resulting in high ranking of irrelevant videos and also long response time leading to poor user satisfaction. Therefore, a new system is developed making use of multiple features along with parallel programming[16] to achieve better performance and high accuracy in retrieving the videos.

The proposed system analyzes the structure of the videos and performs video segmentation process to obtain shots, scenes and frames. From each video shot, a set of frames that best matches the shot contents are extracted as key frames. Later the Texture and Color features of the videos are extracted using Local Binary Pattern and Mean Variance Skewness techniques respectively. Feature vectors are composed and stored in the database. Using KNN along with Euclidean distance measure, similarity check is done between the query video and the videos from the data set. The relevant or closest match videos are displayed as results. Then the complete framework is parallelized using CUDA to achieve high performance.

3. Results and Analysis

3.1. Evaluation Metrics

Proposed video retrieval system is measured based on the following metrics:

1. **Accuracy of retrieval**: It is the ratio of count of exact matching videos that are retrieved to the total number of videos multiplied by 100.
2. **Precision**: it is the ratio of count of videos retrieved similar to the query clip to the total count of retrieved videos
3. **Recall**: it is the ratio of count of retrieved videos similar to the query clip to the total count of similar videos available in the database.

In addition to this as a part of performance evaluation, the serial and parallel processing time of the system is compared.

3.2. Experimental Dataset

The experimental dataset consists of various categories of video clips of different length that includes set of traffic videos, sports videos, nature videos, ocean videos and so on and are of .avi format. The sample dataset is shown in Figure 4.

Figure 4. Sample Dataset

A set of videos are collected and stored in the database for each category. There are totally four such categories. Videos from each of these categories are compared with that of the query clip posted and the best matches are retrieved and returned to the user.

3.3. Performance Analysis

Performance is the factor that judges the efficiency of the system. Performance of the proposed system is estimated by comparing the sequential and parallel execution time of video retrieval on different processor speed. Proposed system was tested successfully and was found to be working accordingly in retrieving the relevant videos from the database that matches the query clip. The accuracy of video retrieval system was found to be 80% in retrieving the relevant videos. The performance of the proposed parallel system was found to be 5 times faster than the sequential processing. Figure 5 depicts the comparison between serial and parallel execution of the proposed system of core i3 processor.

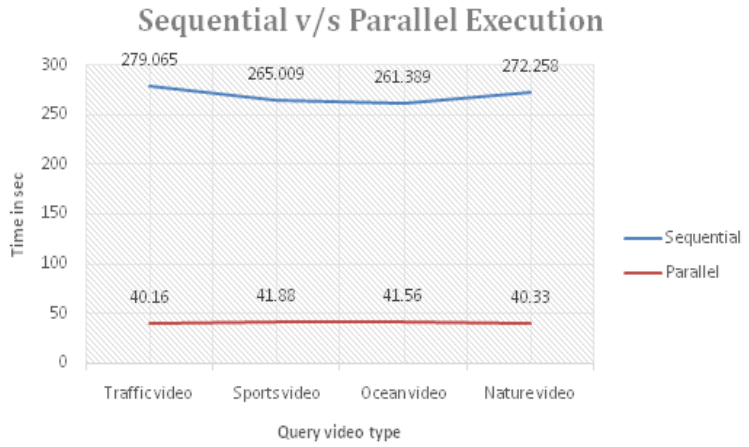

Figure 5. Proposed system performance comparison on i3 processor

Figure 6 depicts the comparison between serial and parallel execution of the proposed system on core i5 processor. The graph is plotted by considering query video type on x-axis and execution time in sec on y-axis.

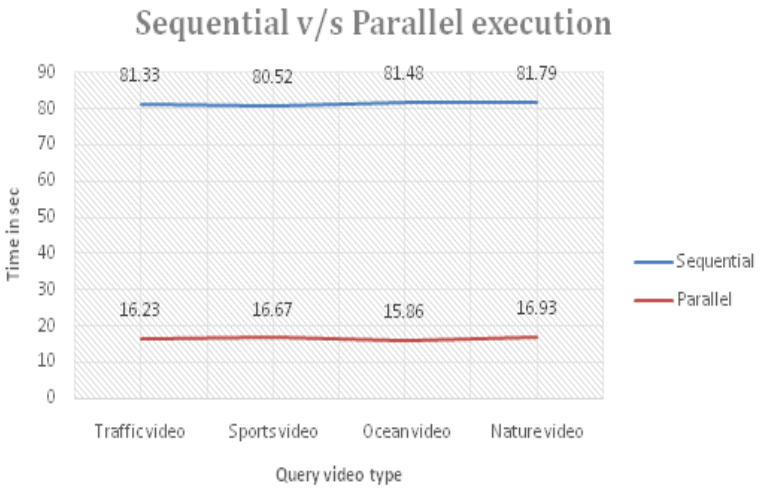

Figure 6. Proposed system performance comparison on i5 processor

From the above two graphs it is evident that execution time reduces to 5 times lesser than the sequential execution as processing on GPU is more efficient for image processing applications than normal CPU execution.

4. Conclusion

The proposed system retrieves relevant videos from the database based on the user input effectively using multiple features such as texture and color moments using Local Binary Pattern and Mean Variance and Skewness respectively. Distance between the query and the video in the database is computed using K-nearest neighbor algorithm with Euclidean distance measure. The system is achieving more than 80% accuracy in retrieving the best match videos for a given query. The sequential and parallel versions of the system are then compared for performance analysis in retrieving videos. By the evidence of the results it is clear that parallel computation performs 5 times faster than serial computation. Hence the system is achieving promising performance by exploiting parallel architecture of CUDA.

References

[1] P Geetha, Vasumathi Narayanan. A Survey of Content-Based Video Retrieval. *Journal of Computer Science*. 2008; 4(6): 474-486.
[2] S Padmakala, Dr GS Anandha Mala, M Shalini. An Effective Content Based Video Retrieval Utilizing Texture, Color and Optimal Key Frame Features. *IEEE International Conference on Image Processing*. 2011. 978-1-61284-861-7/11.
[3] Kin-Wai Sze, Kin-Man Lam, and Guoping Qiu. A New Key Frame Representation for Video Segment Retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*. 2005; 15(9).
[4] Haojin Yang, Christoph Meinel. Content Based Lecture Video Retrieval using Speech and Video Text Information. *IEEE Transactions on Learning Technologies*. 2014; 7(2).
[5] Dr Sudeep D Thepade, Krishnasagar S Subhedarpage, Ankur A Mali and Tushar S Vaidya. *Performance Gain of Content Based Video Retrieval Technique using Intermediate Block Truncation Coding on Different Color Spaces.* International conference on Communication and Signal Processing, India. 2013.
[6] 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. 2013.
[7] Naoufel Werghi, Stefano Berretti, and Alberto del Bimbo. *The Mesh-LBP: A Framework for Extracting Local Binary Patterns From Discrete Manifolds*. IEEE Transactions on Image Processing. January 2015; 24(1).
[8] Madhav Gitte, Harshal Bawaskar, Sourabh Sethi, Ajinkya Shinde. Content Based Video Retrieval System. *International Journal of Research in Engineering and Technology*. 2014; 3(6.2).
[9] Ojala T, Pietikainen M, Harwood D. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*. 1996; 29(1): 51–59.
[10] BV Pateland B Bmeshram. Content Based Video Retrieval Systems. *International Journal of UbiComp (IJU)*. 2012; 3(2).
[11] Hamdy K. Elminir and Mohamed Abu ElSoud. Multi feature content based video retrieval using high level semantic concept. *IJCSI International Journal of Computer Science Issues*. 2012; 9(4, 2).
[12] Aasif Ansari and Muzammil H Mohammed. Content based Video Retrieval Systems - Methods, Techniques, Trends and Challenges. *International Journal of Computer Applications* (0975 – 8887). 2015; 112(7).
[13] NianhuaXie, Li Li, Xianglin Zeng, and Stephen Maybank. A Survey on Visual Content-Based Video Indexing and Retrieval. *IEEE Transactions On Systems, Man, And Cybernetics.* 2011; 41(6).
[14] M Petkovic, W Jonker. Content-Based Video Retrieval by IntegratingSpatio- Temporal and Stochastic Recognition of Events. In proceedingsof *IEEE Intl. Workshop on Detection and Recognition of Events in Video*. 2001: 75-82.
[15] Azra Nasreen and Dr Shobha G. A comparative study of content based search and retrieval of video sequences. *International Journal of Computer Science and Information Technology & Security (IJCSITS).* 2016; 6(4): 1-3. ISSN: 2249-9555.
[16] Ashok Ghatol. Implementation of parallel image processing using NVIDIA GPU framework. *Advances in computing communication and control, Springer Berlin Heidelberg*. 2011: 457-464.