International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072

VISUAL CONTENT BASED VIDEO INDEXING USING KEY FRAME EXTRACTION

Hariharan. K[1], Arjun. S. V[2], Nivedha. S[3], Srithar. K[4], Thivaharan. S[5]

[1,2,3,4]*Final Year Students, Dept. of CSE, PSG Institute of Technology and Applied Research, Coimbatore.*
[5]*Asst. Prof (Sel. Gr), Dept. of CSE, PSG Institute of Technology and Applied Research, Coimbatore.*

Abstract:- Generally, videos are structured according to a descending hierarchy of video clips, scenes, shots, and frames. Recording a video is now an everyday activity, as it has become more common to own devices such as video cameras, smartphones and digital cameras. However, videos without any metadata, such as date, creator, keywords, or description, cannot be found using typical search engines. Metadata is usually added manually, and the process is very time consuming. Furthermore, even if a video can be found by its metadata, search engines normally are not capable of finding a specific scene of interest within the video. Its difficult to watch the entire video at the last minute of any preparation. So, our idea is to index the required part of the video choosing its suitable timeframe and playing the video from that time stamp. Visual database systems require efficient indexing to enable quicker access to video sequences in the database. Several content-based indexing methods for video that are based, colour, texture, shape, sketch, object motion, and camera parameters have been reported. The goal of this paper is to provide content-based indexing techniques and to point out the relative advantages and disadvantages that will be useful for the mankind.

Key Words: Indexing, Hashing, Searching, Key frames, Video indexing, Video retrieval, Hash Map.

1. INTRODUCTION

Videos can also be categorized but some clever feature extraction algorithms need to be used. And a particular video can have a large number of frames which individually contain several different subjects. You need to have some key frames from which to extract features as processing all the frames is not practical. Currently, image and video search is mostly based on tags derived from text associated with the images or specified manually. Manually tagging and organizing large volumes of digital media can take hundreds of hours, and important details can be missed. Similarly, associated text is often ambiguous or missing.

1.1.VIDEO INDEXING FLOW CHART:

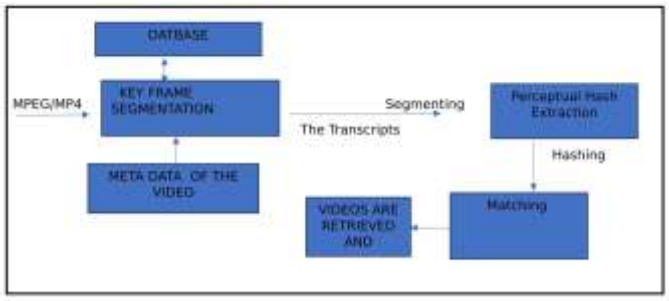

Figure 1: Flow Chart for content based indexing

The technology creates non-semantic metadata (a number versus a pre-loaded word) that represents visual features in a photo or video. This allows the search engine to recognize concepts it has not been trained to recognize. It can be used in a variety of scenarios to sort images by visual similarity from personal photo libraries to real estate listings. In video, it recognizes and allows users to search dynamic content and concepts. For example, users can search for where the sports issue is been told in NEWS. Automated indexing and search technology unlock the visual intelligence contained in large media datasets, providing an effective way to organize and search visual content. Indexing is a data structure technique to efficiently retrieve records from the database files based on some attributes on which the indexing has been done. Indexing is database system is similar to what we see in books.

Content based video browsing and retrieval is the partitioning of a video sequence into shots, where video content analysis is the first phase. A shot is defined as an image sequence that presents continuous action which is captured from a single operation on any suitable devices. The entire video will be made by joining all the shots in a suitable format. Shots can be effectively considered as the smallest indexing unit where no changes in scene content can be perceived and higher level concepts are often constructed by combining and analyzing the inter and intra shot relationships.

2. GENERATION OF KEY FRAMES:

For efficient browsing and retrieval of video from the database extraction of a small number of key frames that can

abstract the content of the video is very important. In this paper, the key frame extraction problem is considered from a set-theoretic point of view, algorithms are developed to find a compact set of key frames that can represent a video segment for a given degree of fidelity. The proposed extraction algorithms can be hierarchically applied to obtain a tree-structured key frame hierarchy that is a multilevel abstract of the video. The key frame hierarchy enables an efficient content-based retrieval by using the depth-first search scheme with pruning., Intensive experiments on a variety of video sequences are presented to demonstrate the improved performance of the proposed algorithms over the existing approaches.

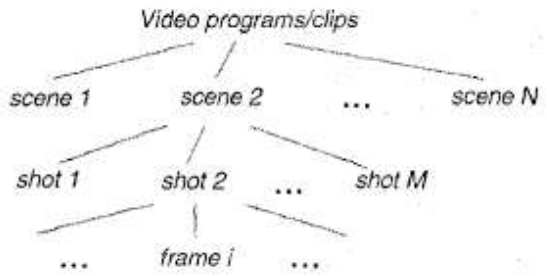

Figure 2: Separation of Key frames

The examination in two types of video abstraction: the key frame set. For the extraction of key frames, these components are the size of key frame set, portrayal scope, base unit, basic computational instruments with every component being additionally sorted. Likewise, for the generation of video skims, five imperative components are recognized and depicted, including the general procedure, the length of video skim, working information area; fundamental calculation system and highlights are utilized.

2.1. SEPARATION OF VIDEO ON THE KEY FRAMES:

One of the powerful tool by which video content can be done is by Key Frame Extraction by Clustering methods. Most of the existing key frames extraction methods are not suitable for video indexing, as they do not meet specific requirements.

Generally, key frame extraction techniques can be roughly categorized into four types [1], based on shot boundary, visual information, movement analysis [2]. Nowadays, cluster-based methods are mostly applied in video content analysis research. In these methods, key frame extraction is usually modelled as a typical clustering process that divides one video shot into several clusters and then one or several frames are extracted based on low or high level features[3]. The methods in compressed domain usually are not suitable for diverse formats of videos from the Internet. Transcoding may increase time complexity and inaccuracy.

2.2. ACHIEVE A MEANINGFULL KEY FRAME:

How to achieve a meaningful key frame is an important problem in various communities. The focus of the work is to represent the video content adequately and fast. In this paper, an active detection method is proposed. First, the key frame is defined for video copyright protection. And then, a key frame extraction algorithm based on two-step method with low level features is proposed.

The distinct features of our algorithm are as follows. (1) The definition of key frame is specific for video copyright protection. (2) Convert the video into audio file.(3) Obtain the subtitle i.e., Transcripts from the audio with the right time stamp. (4) Search the required content (5) It presents the keyframe of the video in which the suitable content is been told.

Based on the number of key frames, color feature extraction method for video sequence obvious video content conversion has a good ability

Figure 3: Obtaining of meaningful key frames.

2.3. HASH MAP FOR RETRIEVAL OF KEY FRAMES

Hashing is a technique that is used to uniquely identify a specific object from a group of similar objects. Similarly hash map has been used to find the similar key frames at the required time stamp. Once the required time stamp has been obtained, it indexes the suitable value at the key frame and indicates which point the required content is been searched.

By using Hash Functions the Large key values are converted into smaller keys. The values are then stored in a data structure called hash table. Across an array the key values are evenly distributed. Each element is assigned a key (converted key). By using that key you can access the element in O(1) time. Using the key, the algorithm (hash function) computes an index that suggests where an entry can be found or inserted.

Hashing is implemented in two steps:

1. From the Hash Function the element is converted into integer, which is stored as an index in the hash table.

2. Using hash key the elements can be quickly retrieved from the database.

   Hash-hashfunction(key)

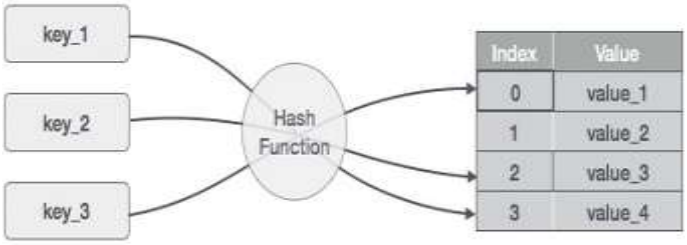

Figure 4: Working of a hash map

2.3.1. HASH FUNCTION – WORKING MODEL:

In this method, the hash is independent of the array size and it is then reduced to an index (a number between 0 and array size − 1) by using the modulo operator (%).

Once the video is sent as input, its respective transcripts are obtained with its respective time stamp. Once the respective time stamp is obtained the required content to be searched can be easily identified. Using the applications of hashing and applying the function of hashing the suitable search algorithm the required content within the video can be obtained with its respective time-stamp. If the required content is present more than one time in the video then it indicates the time-stamp in the position where the content has been told.

3. CONCLUSION

A key frame extraction method based on frame difference with low level features is proposed for video indexing. Exactly, an algorithm to extract accurate key frames to cover the content for the entire video sequence has been implemented. Firstly, the transcripts are obtained from the video that is given as input. Secondly, the transcripts that are obtained are maintained and stored in the database with its suitable timestamp. And thirdly, using Hash Map the required content is been searched and indexed in the position where it has been told in the video. Tested with several television videos with different content, formats, and resolutions, it is shown that the proposed method has advantages in computation complexity and robustness on several video formats, video resolution, and so on.

REFERENCES

[1] G.-J. Qi, X.-S. Hua, Y. Rui, J. Tang, T. Mei, and H.-J. Zhang, "Correlative multi-label video annotation," in *Proc. ACMInt. Conf. Multimedia*,

[2] Adcock, J., Cooper, M., Denoue, L., Pirsiavash, H.: Talkminer: A lecture webcast search engine. In: Proc. of the ACM International Conference on Multimedia, MM 2010, Firenze, Italy, pp. 241–250. ACM (2010).

[3] M.-Y. Chen and A. Hauptmann, "Active learning in multiple modalities for semantic feature extraction from video," in *Proc. 20th National Conf. Artificial Intelligence Workshop Learning in Computer Vision*, 2005.

[4] M.Wang, X.-S. Hua, R. Hong, J. Tang, G.-J. Qi, and Y. Song, "Unified video annotation via multi-graph learning," *IEEE Trans. Circuits Syst.Video Technol.*, vol. 19, no. 5, pp. 733–746, May 2009.

[5] E. Parzen, "On the estimation of a probability density function and the model," *Ann. Math. Statist.*, vol. 33, no. 2, 1962.

[6] Z.-J. Zha, Y.-T. Zheng, M. Wang, F, Chang, and T.-S. Chua, "Locally

[7] J. Jeon , V. Lavrenko , R. Manmatha, Automatic image annotation and retrieval using cross-media relevance models, Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, July 28-August 01, 2003, Toronto, Canada .

[8] L. Liu and G. Fan, "Combined key-frame extraction and object-based video segmentation," IEEE Transactions on Circuits & Systems for Video Technology, vol. 15, no. 7, pp. 869–884, 2005.

[9] J. Li, Y. Ding, Y. Shi, and Q. Zeng, "DWT-based shot boundary detection using Support Vector Machine," in Proceedings of the 5th International Conference on Information Assurance and Security (IAS '09), pp. 435–438, September 2009.

[10]Z. Sun, K. Jia, and H. Chen, "Video key frame extraction based on spatial-temporal color distribution," in Proceedings of the 4th International Conference on Intelligent Information Hiding and Multimedia Signal (IIH-MSP '08), pp. 196–199, August 2008.

[11]J. Zhang, X. Jiang, G. Li, and L. Jiang, "Key frame extraction based on particle swarm optimization," Journal of Computer Applications, vol. 31, no. 2, pp. 358–361, 2011.

[12] Steven C.H. Hoi, Michael R. Lyu, "A multimodal and multilevel ranking frame work for content-based video retrieval", ICASSP, 2007.

[13] M. Petkovic, W. Jonker, "Content-Based Video Retrieval by Integrating Spatio-Temporal and Stochastic Recognition of Events", Proceedings of IEEE Workshop on Detection and Recognition of Events in Video, pp. 75-82, 2001.

[14] Ara V. Nefian, Monson H. Hayes III, "Hidden Markov Models for Face Recognition", IEEE International Conference on Acoustics, Speech and Signal Processing, 1998.

[15] M.Ravinder and T.Venugopal, "Content-Based Video Indexing and Retrieval using Key Frames Texture, Edge and Motion Features" International Journal of Current Engineering and Technology, Vol.6, No.2, April 2016.

[16] Pentland AP, Picard RW, Davenport G, Welsh, B (1993) The BT/MIT project on advanced image tools for telecommunications: an overview, Technical Report No. 212, Media Lab, Massachusetts Institute of Technology, Cambridge, Mass.

[17] Zhang HJ, Smoliar SW, Wu JH (1995) Content-based video browsing tools, Proceedings IS&T/SPIE Conference Multimedia Computing and Networking, to be held in San Jose, Calif.

[18] Pentland AP, Picard RW, Davenport G, Welsh, B (1993) The BT/MIT project on advanced image tools for telecommunications: an overview, Technical Report No. 212, Media Lab, Massachusetts Institute of Technology, Cambridge, Mass

AUTHOR BIOGRAPHIES

Mr. Hariharan. K, currently pursuing Final Year Computer Science in PSG Institute of Technology and Applied Research. He is currently the Chairman of Students' Union and he has published many Papers and the Project Lead in Final Year Academic Project.

Mr. Arjun. S. V, currently pursuing Final Year Computer Science in PSG Institute of Technology and Applied Research. He is currently working in Zoho and his area of interest is Operating Systems and Content Mining.

Ms. Nivetha. S, currently pursuing Final Year Computer Science in PSG Institute of Technology and Applied Research. Her area of interest is Data Structures and Video Indexing.

Mr. Srithar. K, currently pursuing Final Year Computer Science in PSG Institute of Technology and Applied Research. His area of interest is Content Mining.

Mr. S. Thivaharan, currently working as Asst. Prof (Sel. Gr) in the department of Computer Science and Engineering in PSG Institute of Technology and Applied Research, Coimbatore. He is having a total experience of 13 years both in industry and Academia. His areas of interest include Language processing and content mining.