IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 10, October 2017

Topic Based Video Search and Visualization Techniques of Video Search Results- A Survey

Dr. R. Priya [1], Radhika Krishnan R [2]

Associate Professor, Department of Computer Science, Sree Narayana Guru College, Chavadi,Coimbatore,TN,India[1]
Research Scholar, Department of Computer Science, Sree Narayana Guru College, Chavadi, Coimbatore, TN, India [2]

Abstract: Today's internet contains huge number of videos, due to the technology popularity, everyone using the social network to store and share the videos. This poses a huge challenge on the video search and retrieval. So, an effective technique to perform the video search and retrieval from large video repository is necessary in several applications. Currently, YouTube performs a best video searching process with several unique features. However, the complex query handling to retrieve accurate videos is more difficult. Data mining is the best way to handle such complex queries and visualizing them. This paper gives the survey about the problems associated with the video search and existing video search and visualization techniques. This paper gives the general description about the process of video retrieval and visualization process with hierarchical structures. Finally gives the solutions to perform effective video search.

Keywords: Video Search, Video Retrieval, Visualization, Hierarchical Structure, Visual Analysis.

INTRODUCTION

Nowadays, it is very easy to acquire, store, upload and deliver the video contents because of extreme variations in the development of digital devices, internet infrastructures, and web technologies. The search for video content over the web still seems to be extremely difficult even after the success of different web search engines. Generally, a number of the web search engines index merely the Meta data of videos and search them by texts. In video retrieval, conventional search engines may be restricted, if they are devoid of the capability to comprehend media contents. By means of the usage of the rich media contents for the purpose of video retrieval, there must be ample room in improving conventional search engines.

YouTube is performing search process with various features such as query processing, query suggestion and query expansion techniques for effective video retrieval. However, the video retrieval and visualization techniques are not fully adopted in the current video search websites [1]. Usually it is possible to segment the video retrieval system into two most important elements, they are (i) a phase for the extraction of representative characteristics from video segments and (ii) one defining a fitting similarity model in order to locate identical video clips from the video database. But none performed the effective visualization in the video retrieval. The following fig 1.0 is the sample result from the YouTube for the given query. This shows the summarized results in predefined categories such as update data, type, duration, features and sort by.

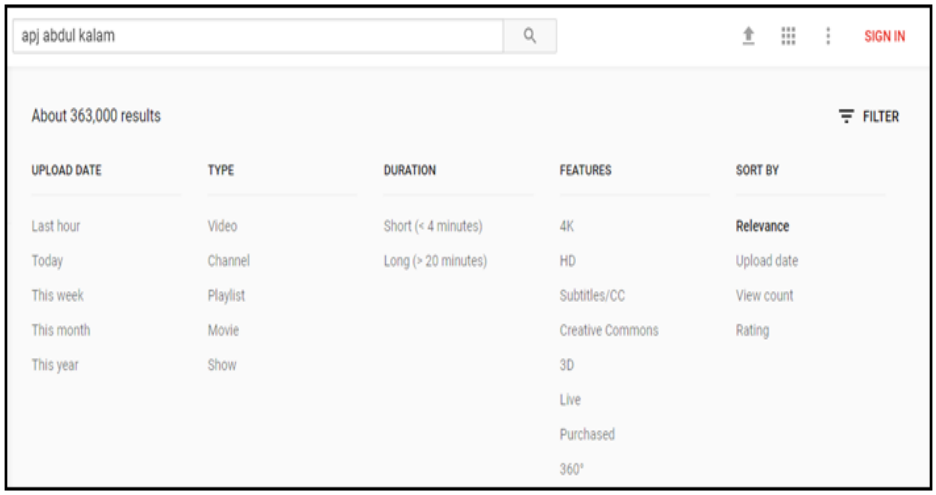

Fig 1.0 search result categorization in YouTube website

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 10, October 2017

LITERATURE REVIEW

Research on the video content retrieval and search has become increasingly active in the past two decades. Considerable progress has been made in video content analysis, video indexing and retrieval, which are the bases for accessing video content. Video content analysis is the basis for later video processing. It includes shot boundary detection and key frame extraction [2]. The two functionalities of video indexing and retrieval are the ultimate goals of a video access system, and they are closely related to video representations.

A. Video search, indexing and retrieval techniques:

In text based video retrieval method, users use keyword or description to the videos as query. This description is, then, applied to a database in order to retrieve the relevant videos to the description or keyword. In this approach, it is necessary that the video content has annotated with some metadata standards, such as MPEG-7 (Moving Picture Experts Group) [3]. This approach is more efficient from a semantic point of view. However, it is very much dependent on manual notes previously made.

According to the authors in [4] text based video retrieval has several shortcomings: subjective, inaccurate, and incomplete. Moreover, the manual production of video content description is even more time-consuming and thus more costly and is almost impossible to generate it for vast amount of video data available.

Authors in [5] indicated that some images (banners or specific logos) could not be annotated due to difficulty in describing their content with words. In addition, the advance of the World Wide Web (WWW) carries some new challenges to conventional text based information retrieval. Nowadays web based digital libraries can be accessed around the world. Users may have different languages and cultural backgrounds, different understandings of or interpretations of the same object and may not be able to use effective keyword searches of these libraries. Visual-based query techniques use visual features of the images and videos in the search process will increase the accessibility of these digital libraries greatly, and this is probably a main reason it has become a hot research area in the past decade. In image based video retrieval method, users use image to the videos as query so that they can use the retrieved videos, which are relevant to the query image. This method is also not successful in reducing the gap between the image and user's interest. Visual-based video retrieval is intended as one of the most practical solutions to improve the retrieval quality.

One of the earliest examples is the VideoQ system developed in the literature [6], which allows the user to search video based on a rich set of visual features and spatio-temporal relationships. The video clips in the database are stored as MPEG files. Through a Web interface, the user can formulate a query scene as a collection of objects with different attributes, including motion, shape, color and texture. Once the user has formulated the query, it is sent to a query server, which contains several databases for different content features. On the query server, the similarities between the features of each object specified in query and those of the objects in the database are computed; a list of video clips is then retrieved based on their similarity values. For each of these video clips, keyframes are dynamically extracted from the video database and returned to browser. The matched objects are highlighted in the returned keyframe. The user can interactively view these matched video clips by simply clicking on the keyframe. Meanwhile, the video clip corresponding to that keyframe is extracted from the video database.

In information retrieval, the search techniques are very important to the quality of the results gained, and depend on the scope of results to obtain the relevant answer. In the paper [7], some typical query methods by which information may be retrieved are:

Query by Semantics: The simplest way to achieve basic content-based retrieval is to use keywords, or free text associated with the keyframes or video shots in the annotation process.

Query by Sketch (QbS): A user draws a visual sketch of an image or video clip he/she has in mind. By drawing, the user specifies low-level features that are used in the similarity search process. For example, an image query involves the specification of a desirable shape, color etc.

Browsing: This can be random browsing or category based navigation (CBN). CBN is similar to operations on hypertext scripts in browsing of web pages, which is usually applied in prompt or hierarchical querying of semantics.

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

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 10, October 2017

Query by Example : Query results are determined by the similarity matching between user specified example and candidate samples and the similarity here are calculated by selected features, such as color, texture or tone.

Authors in [8] have presented a novel scheme to match a video clip against a large database of videos. Unlike previous schemes that match videos based on image similarity, this scheme matches videos based on similarity of temporal activity, i.e., it finds similar "actions". Furthermore, it provides precise temporal localization of the actions in the matched videos. The fingerprint of the query video is matched against the fingerprints of videos in a database using sequential matching. The fingerprints are computed directly from compressed MPEG videos. The matching is much faster than real time. This scheme is used to find similar actions in sporting events, such as diving and baseball.

Authors in [9] have introduced a novel spatio-tempral model for video matching. This model uses information on image object spatial relationship in frames as well as transition between frames to generate a vector representing that frame. The analysis of video content, which fully considers temporality of video, has been a very active research area among researchers. The authors have proposed a video retrieval algorithm on the basis of the integration of a number of visual cues. On the contrary to keyframe based representation of shot, their approach has examined all frames within a shot in order to build a compact representation of video shot. In the video matching step, by combining the color and motion features, a similarity measure has been defined to spot out the occurrence of related video clips in the database. For that reason, their approach has been able to fully make use of the spatio-temporal information contained in video. Experimental results have been revealed that their approach is efficient and outperforms a few of the existing techniques. The common drawback of these frame sequence matching technique is that they lack in searching accuracy of retrieved videos.

In order to improve the searching accuracy, algorithms [10] have been developed based on an integrated approach for video shot matching to retrieve exact and similar video shots. In contrast to keyframe based representation of shot, this approach analyzes all the frames within a shot to construct compact representation of video shot. In all the cases, the user has to provide the sample query video shot. The algorithms are designed and developed based on low level visual primitives for effective CBVR system. Effort has also been made to retrieve high level visual features (object location identification, movement detection) for CBVR System.

With the number of accessible video programs increasing every day, video indexing has become an important research topic. It is not only essential for efficient video retrieval, but also essential for locating desired access points within a long video sequence. After obtaining the video structure from the shot boundary detection process, low and/or high-level features can be subsequently extracted from the basic video units to facilitate the content indexing. The indexing phase includes the process of feature extraction and content analysis. For instance, authors in the paper [11] extracted four image features (color, texture, shape and edge) and two temporal features (camera operations and temporal brightness variation) based on the segmented shot sequence and generated keyframes.

In indexing or storage process, images and video are processed to extract features which describe their semantics. The extracted features are then represented, organized, and stored in the database [12]. The indexing phase starts with segmentation of video data into smaller pieces. This shot boundary detection approach is done using video processing and pattern analysis tools.

A comprehensive video indexing and browsing environment (VIBE) was discussed in [13] for compressed video database. Specifically, given a video sequence, the VIBE system would first extract its shot structure, and then represent each shot with a hierarchical structure called shot contents. Next, all shots were classified into pseudo-semantic classes based on their contents. Finally, these results were presented to end users in an active browsing environment based on similarity pyramid data structure.

Authors in [14] presented a generic framework of integrating existing low and high-level indexing features, where the low-level features might include color, texture, motion and shape, and the high-level features could be video scenes, events and hyperlinks.

In paper [15], authors have presented an interactive multimedia based e-Learning environment that enables users to interact with it to obtain knowledge in the form of logically segmented video clips. They have developed a two-phase approach to conduct content-based video indexing and retrieval to identify video clips appropriate to addressing user's interests. Their approach integrates natural language processing, named entity extraction, text and frame based video indexing and information retrieval techniques. The relevance of video clips to questions were measured based on the similarity between generated templates of questions and clip content. Their research explores a way to access instructional videos in interactive e-Learning. Some results have shown that their approach has achieved more advanced precision than the traditional keyword based approach.

In paper [16], authors have proposed an innovative method to achieve the high quality of content-based video retrieval by discovering the temporal patterns in the video contents. On basis of the discovered temporal patterns, an efficient indexing technique and an effective sequence matching technique are integrated to reduce the computation cost and to raise the retrieval accuracy, respectively. In this thesis, videos are processed to extract low-level and high-level features. The extracted features are then stored in the feature database or library.

Video search summarization Techniques:

Several researches have been made for web content searching and categorization. Extracting appropriated contents from the web and resulting them into more user friendly manner is more important for the application. This paper gives the recent techniques, which provides the video search and visualization in effective manner with its merits and demerits.

In the paper [17] authors proposed topic based video browsing and video search result organization using semantic hierarchic, this is more relevant to the approach proposed in [18][19]. In the papers, a regular clustering is performed on textual data's, which contains the information about the videos. This is performed to determine the hidden topic structures in the video search results. This also used to adapt the hierarchy extracted from Wikipedia website, which crawls and summarizes the results into categories. After performing the hierarchical process, a simple optimization problem is formulated to assign the videos to each node of the hierarchy considering three important criteria. However, the techniques were not fully adopted for the high dimensional data sources, which may have many duplicate and multiple meaning tags.

CONCLUSION

Over the past decade, Content-based indexing and retrieval of visual information have received attention from the research community. Numerous reviews of the literature on video retrieval have been published, from a variety of viewpoints. Many authors reviewed methods for providing subject access to video data depending on three category framework. They discussed the strengths and limitations both of conventional methods based on text based video retrieval and image based video retrieval for both indexing and searching and summarization with effective visualization techniques, and experimental systems use content based video retrieval for one or both of these. Their conclusions are that, while there are serious limitations in current text based and image based techniques for subject access to video data, it is essential to develop a video retrieval system, which can automatically extract the indexing structure and categories, from a compact yet meaningful data representation, to facilitate effective video retrieval. This paper concentrates on video retrieval, indexing and visualization techniques for video shot representation and a video similarity measure to achieve video retrieval task with more accurate and effective.

REFERENCES

[1] Abu-El-Haija, S., Kothari, N., Lee, J., Natsev, P., Toderici, G., Varadarajan, B., & Vijayanarasimhan, S. (2016). YouTube-8M: A large-scale video classification benchmark. *arXiv preprint arXiv:1609.08675*.

[2] Dimitrova, N., Zhang, H. J., Shahraray, B., Sezan, I., Huang, T., & Zakhor, A. (2002). Applications of video-content analysis and retrieval. *IEEE multimedia*, *9*(3), 42-55.

[3] Chen, Billy, Eyal Ofek, Bo Zhang, and Ying-Qing Xu. "Calibration and annotation of video content." U.S. Patent 8,769,396, issued July 1, 2014.

[4] Priya, R., T. N. Shanmugam, and R. Baskaran. "A Content Based Video Retrieval Analysis System with Extensive Features by Using Kullback-Leibler." *International Journal of Computational Intelligence Systems* 7, no. 2 (2014): 242-263.

[5] Bai, Qinxun, Henry Lam, and Stan Sclaroff. "A Bayesian framework for online classifier ensemble." In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pp. 1584-1592. 2014.

[6] Chang, Shih-Fu, and Horace J. Meng. "Methods and architecture for indexing and editing compressed video over the world wide web." U.S. Patent No. 9,330,722. 3 May 2016.

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

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 10, October 2017

[7] Ren, Reede, and Joemon M. Jose. "Football video segmentation based on video production strategy." *European Conference on Information Retrieval*. Springer, Berlin, Heidelberg, 2005.
[8] Apostolidis, Evlampios, et al. "Automatic fine-grained hyperlinking of videos within a closed collection using scene segmentation." *Proceedings of the 22nd ACM international conference on Multimedia*. ACM, 2014.
[9] Petkovic, Milan, and Willem Jonker. "Content-based video retrieval by integrating spatio-temporal and stochastic recognition of events." *Detection and Recognition of Events in Video, 2001. Proceedings. IEEE Workshop on*. IEEE, 2001.
[10] Yang, Yan, Brian C. Lovell, and Farhad Dadgostar. "Content-based video retrieval (cbvr) system for cctv surveillance videos." *Digital Image Computing: Techniques and Applications, 2009. DICTA'09.*. IEEE, 2009.
[11] Chakraborty, S., Tickoo, O., & Iyer, R. (2015, January). Adaptive keyframe selection for video summarization. In *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on* (pp. 702-709). IEEE.
[12] Rao, L. Koteswara, D. Venkata Rao, and L. Pratap Reddy. "Local mesh quantized extrema patterns for image retrieval." *SpringerPlus* 5.1 (2016): 1-15.
[13] C. Taskiran, J.-Y. Chen, A. Albiol, L. Torres, C. A. Bouman, and E. J. Delp, "VIBE: A compressed video database structured for active browsing and search," IEEE Trans. Multimedia, vol. 6, no. 1, pp. 103–118, Feb. 2004.
[14] Tusch, R., Kosch, H., & Böszörményi, L. (2000, October). VIDEX: an integrated generic video indexing approach. In *Proceedings of the eighth ACM international conference on Multimedia* (pp. 448-451). ACM.
[15] Zhang, D., & Nunamaker, J. F. (2004). A natural language approach to content-based video indexing and retrieval for interactive e-learning. *IEEE Transactions on multimedia*, *6*(3), 450-458.
[16]Su, Ja-Hwung, et al. "Effective content-based video retrieval using pattern-indexing and matching techniques." *Expert Systems with Applications* 37.7 (2010): 5068-5085.
[17] J. Wang, Y.-G. Jiang, Q. Wang, K. Yang, and C.-W. Ngo, "Organizing video search results to adapted semantic hierarchies for topic-based browsing," in Proc. ACM Int. Conf. Multimedia, 2014, pp. 845–848.
[18] Z.-Y. Ming, K. Wang, and T.-S. Chua, "Prototype hierarchy based clustering for the categorization and navigation of web collections," in Proc. Prototype Hierarchy Based Clustering Categorization Navigation Web Collections, 2010, pp. 845–848.
[19] Jiang, Y. G., Wang, J., Wang, Q., Liu, W., & Ngo, C. W. (2016). Hierarchical visualization of video search results for topic-based browsing. *IEEE Transactions on Multimedia*, *18*(11), 2161-2170.