Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 03, pp.5721-5725, March, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

CONTENT BASED ANALYSIS RETRIEVAL USING AUDIOVISUAL ARCHIVE RETRIEVAL WITH CROSS REFERENCING AND MULTIMODAL RE-RANKING

Nagarajan, G. and *Shankar Ganesh, V.

Department of Computer Science and Engineering, K.S.R. College of Engineering Tiruchengode, Namakkal, India

ARTICLE INFO

Article History:
Received 05th December, 2013
Received in revised form
28th January, 2014
Accepted 20th February, 2014
Published online 25th March, 2014

Key words:

Audio-visual archives,
Evaluation methodology,
Query definitions,
Retrieval tasks,
Keyframes, Re-ranking.

ABSTRACT

In this paper, the system deals with content based retrieval by using three dominant content retrieval techniques and queries used to retrieve the videos. The query sets are given to identify the relevant videos. Media Professionals utilize this audio-visual archives to capture the video material. The primary goal of this project is to investigate how to retrieve the videos by using content based retrieval. The performance of content based video retrieval will improve with the help of cross referencing and multimodal re-ranking method. This project utilizes logged searches, content purchases, session information, and simulators to create realistic query sets and relevance judgments. This paper also guides how to take into account the information needs and retrieves data already present in the audiovisual archive, and demonstrate that retrieval performance can be significantly improved when content-based methods are applied to search. The improvements of the audio visual archives can be established effectively and to produce good results.

INTRODUCTION

Content-based video retrieval (CBVR) (Huurnink *et al.*, 2010) is the technique for audio-visual retrieval problem; it is useful for searching videos in large databases. Content based analysis gives video indexing (VI) for effective audio visual archiving and retrieving. It is a source for media professionals in different field to archive video from database for reusage. Content based analysis provides a solution for the inevitably tedious and incomplete video fragments archive. Fine grained manual and automatic annotation source are helpful for the users to retrieve exact data. Common initial steps for most content based video analysis techniques are to segment a video into elementary shots using video extractor. Based on the description and similarity between the shots result is obtained by query given by the user. Some query input methods are text, image and even video as query for audio visual archive. In images query, the histogram approaches and video query is by shot-by-shot detection. Textual query is the most common query used in the search system to find the relevant data based on the description given in query interface. In this author (Huurnink *et al.*, 2010), the media professionals are finding a new material for audio-visual archive. In cross referencing is introduced here for finding the combination of the similar audio-visual archived material for retrieval process. Finally, multimodal re-ranking is used to ranking the retrieved audio-visual material into the cluster level.

Corresponding author: Shankar Ganesh, V.
Department of Computer Science and Engineering, K.S.R. College of Engineering Tiruchengode, Namakkal, India.

In some approaches, automatic or manual queries are composed using query composer. Finally, all the data are stored in storage server as shown in Figure.1. The file system stores the video data, the indexes and metadata are stored in the storage server. TRECVID is one of the valuable instigators in the advancement of techniques for content based video retrieval, set the base for finding the advancement techniques to archive video data in real world queries. TRECVID started in the year 2001 for evaluation new methods in archiving video; in this study we are going to evaluate some of the optimal approaches in content based video retrieval (CBVR) systems. Due to advancement in technology, the video data are easily captured and stored, we can compress it, transmit and even render it on different platforms, being navigate it (i.e.,) browsing and searching it based in inherent or values of, us is what we are going to guide for new researchers.

Audiovisual archivists aimed to manually describe every shot in every video acquired by their archive. However, due to the rapid increase in the number of videos coming into the archives, it soon became apparent that it was impossible to accomplish this goal through manual labor, given the limited human resources at their disposal. Archivists had to settle instead for describing whole videos, while occasionally providing more detailed within-video descriptions at the archivist's discretion. For example, titles and broadcast dates of videos are likely to be described, but individual shots and scenes are less likely to be described. Such an approach has limited searchers, especially those who do not know which specific video their desired footage appears in. However, 75

years later, the original dream of individually describing every shot in every video in an archive has now come within grasp. Where manual description of shots is not realistic, machines may fill the gap with the automatic shot descriptions associated with content-based video retrieval. Though these descriptions are not flawless, they may be helpful when searching through the archive. The most important contribution of this paper, then, is a detailed investigation of how content-based video retrieval can improve audiovisual archive search. We develop an experimental methodology that allows us to quantitatively evaluate how retrieval performance for professional searches is affected. To enable replication of our experiments, we provide a publicly available evaluation collection that includes manually created program annotations from the archive, queries based on the information needs of users from the audiovisual archive, and their associated relevance judgments.1We present methods for reconciling retrieval on information sources at different granularities (shot-level and program-level), and use weighted fusion to combine results. This allows us to provide an extensive analysis of the potential impact of content-based video retrieval in an archive.

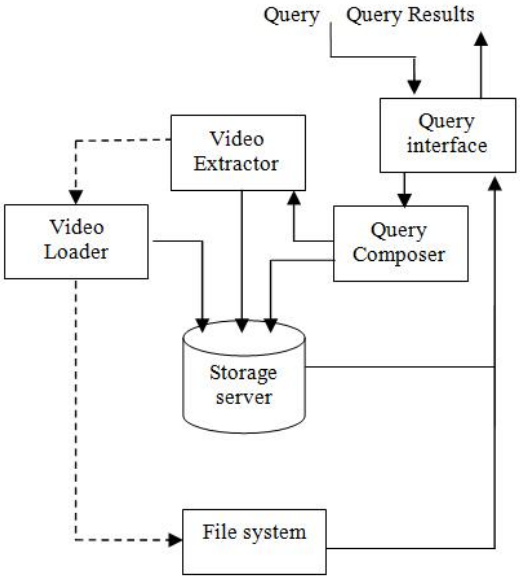

Figure 1.1. System Architecture

Related works

In this paper (vanGemert *et al.,* 2010), the color descriptors are used to recognizing the object and scenes in the video. The distinctiveness of color descriptors can be accessed from image domain and video domain. The Extracted images can be describes in Bags-of-visual-Words (BoW) in (Chen *et al.,* 2010). The extracted images based on the keypoints. By mapping the keypoints of an image back into visual words of the vocabulary, so the representation of an image is a histogram of visual words and uses it for classification. The first VideOlympics (Snoek *et al.,* 2008) brings content-based analysis to the archive and allows for many-to-many communication. In this paper (Wactlar *et al.,* 1999), deals with extracting information automatically from digitized Video. Creating interfaces that allowed users to search for and retrieve videos based on extracted information. The system validation can be through user testbeds. The author (Huijbregts *et al.,* 2007) describes about the construction and utilization of two novel semantic spaces consists of Ontology-enriched Semantic Space (OSS) and Ontology-enriched Orthogonal Semantic Space (OS^2).

EVALUATION METHODOLOGY

Dominant search methods

In content based analysis (Huurnink *et al.,* 2010), there are three dominant search content based video retrieval methods such as, transcript based search, low-level feature based search and detector based search. The first method, transcript based search method uses automatic speech recognition and given textual query (22) for retrieve high quality videos are retrieved. The second method, low-level feature based search visual information of image keyframes are matched to query image (Wactlar *et al.,* 1999), which is done by similarity metrics between global image histograms. The third method, detector-based search utilizes shot based detection scores by human defined concepts to retrieve video is shown in Figure 1.1.

The evaluation methodology includes manual catalog annotation for aggregating different fields by free text, tags and technical metadata. In multimedia content analysis dominant search methods above mentioned are used. The query sets such as archive queries are concatenation of text queries in various sessions by users the query is defined. In lab queries, the query (Wactlar *et al.,* 1999) is defined based on if the program video contains shots relevant to query, then entire video is considered relevant to query, video collections varies from year to year for finding relevant shots. In future queries, the analyzing search sessions and use them to formulate multimedia queries by using transaction logs. In simulation query by using simulation framework of (Huurnink *et al.,* 2010) using logs archive users to generate simulated query. In concept-based video retrieval (Snoek *et al.,* 2008), two retrieval tasks shot retrieval and program retrieval combining both annotation (Jiang *et al.,* 2010) phrase and multimedia contents to retrieve high quality data as result. In the retrieval data sources can be divided into two categories such as, manual catalog annotation and multimedia content analysis. In the manual catalog annotation have source of retrieval data that can be used as a collection of manually created catalog entries that describe each program. It consists of three different entries namely, free text, tags and technical metadata. Free text describes and summarizes the content of a program by natural language descriptions. Tags structured thesaurus terms that describe the people, locations, named entities, and subject areas that appear in or are the topic of a program. Technical metadata have technical information about a program such as identification codes, copyright owners, available formats, and the program title. In the Multimedia content analysis, it creates the catalog entries using transcript based search, feature based search and detector based search.

Query Definitions

Archive Queries

In this method, the user's previous query words and the search result returned are stored in the database. In future, if the new query matches with any of these queries, then the search result

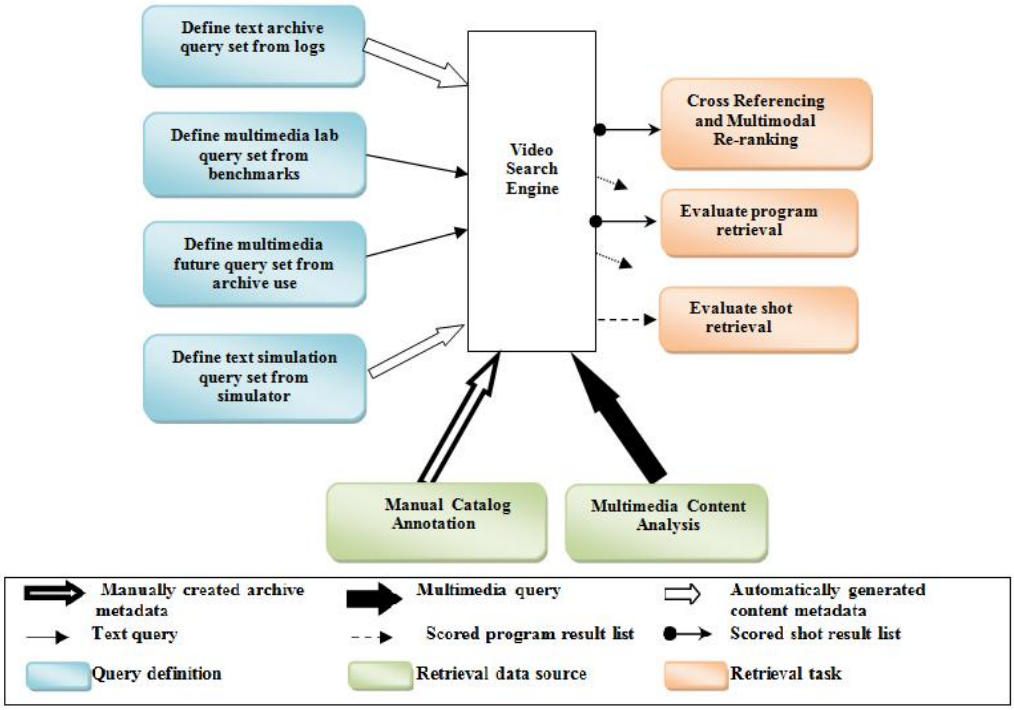

Figure 1.1.Architecture design for content based video retrieval

is taken from those logs. To create a set of archive queries based directly on today's user needs, user makes use of the archive's transaction logs. This approach is different because this module also includes purchase data, in addition to click data. An archive query can be defined by first identifying all logged search sessions that resulted in a purchase from the archive's video collection.

Lab Queries

In this method, annotation phrase is generated and saved at regular intervals say in hours or days. Then if the query contains the annotation phrase, then the videos matching with the phrase are returned. Each annotator was able to browse through the video using transcript-based search, feature-based search, and detector-based search. To create Lab queries that are representative of those used in content retrieval research by adopting them from several existing evaluation initiatives. It performed additional relevance judging to identify relevant shots in the experimental collection used in this paper; a group of annotators manually labeled shots from the video collection as relevant or non-relevant using an interactive annotation tool. Each annotator was given a minimum of half an hour and a maximum of one-and-a-half hours per query to find as many relevant shots as possible.

Future Queries

Turning back to the needs of archive users, create a set of future queries. These are based on logged user needs, but reformulated in terms of an archive retrieval system that includes content-based video retrieval capabilities. The video owner enter the 'N' query phrases for a single video and the user's query phrase matches with these 'N' query phrases, the videos are returned. The information contained in the sessions included searches, result clicks, and purchases. An independent query creator from the archive was given the information from each session, and was asked to develop queries that the user felt reflected the underlying information need of the broadcast professional. To be precise, the query creator was asked to:

1) scan the session to get an idea of the general information needs of the searcher;
2) view the video fragments that were ordered;
3) note down the visual information needs that the user may possibly have had; and
4) rank the noted information needs according to the confidence that they reflect the actual information need of the user.

Simulation Queries

This method is focusing to generate a set of simulated queries using the simulation framework, on the basis of query logs from the same archive. A given catalog entry is used to generate a simulated query by this simulation approach. The related program is then considered relevant to that query. Using a simulator to create a set of queries for evaluation, it gives the advantage of being able to create as many queries. The simulators create relevance judgments at the level of an entire program, and are therefore not suitable for evaluating shot-level retrieval. In addition, the simulated queries do not necessarily reflect the needs of real users. Keeping these limitations in mind, to generate simulated queries for each of the programs in the archive footage collection, resulting in a set of simulated purchase query pairs.

Video Retrieval Tasks

In the video retrieval have two different tasks namely, shot retrieval and program retrieval.

Shot Retrieval

Shot retrieval describes the shots for each program in order of appearance with the program-level catalog annotations. This shot based video retrieval allows users to search the shots much more effectively. In this shot retrieval depends on three dominant content based video retrieval search methods and three different queries set i.e., (3x3).

Program Retrieval

In the program based video retrieval allows to search entire program and it is based on the shot-level annotations. In this program retrieval depends on three dominant content based video retrieval search methods and four queries set i.e., (3x4).

Cross References and Multimodal Re-ranking

After retrieve the video samples for output, the proposed system also presents a flexible and effective re-ranking method, called Cross Reference-Re-ranking, to improve the retrieval effectiveness. To offer high accuracy on the top-ranked results, CR-Re-ranking employs a cross-reference (CR) strategy to fuse multimodal cue. Specifically, multimodal features are first utilized separately to re-rank the initial returned results at the cluster level, and then all the ranked clusters from different modalities are cooperatively used to infer the shots with high relevance.

Performance analysis

In the description of the above evaluation methodology, then performance will improve. Experimental results show that the search quality, especially on the top-ranked results, is improved significantly. The new system is being to develop to eliminate the drawbacks in the existing system. We summarize the statistics of our three query sets and their associated relevance judgments.

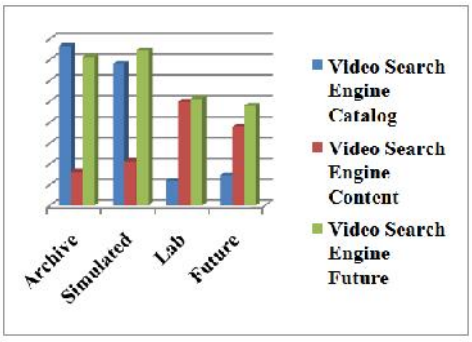

Figure 1.2 Four query sets with three video search engines

RESULTS

In the above figure shows that the three video searches using four different query sets can be measured appropriately. Cross referencing and multimodal re-ranking can be implemented in this system. These techniques will produce better results.

Conclusion

This paper includes three query set definitions, three state-of-the-art content and archive-based video search engines, and two challenging retrieval tasks that are grounded in a real-world audiovisual archive. By using query definitions, the queries can be used effectively. This search engine combines manually created archive metadata and automatically generated content metadata by using the simulated queries. Thus the search engine derives queries from logged searches of the media professionals. It is found that the queries are taken directly from search logs. The cross referencing is used to retrieve the videos with the combination of the results. In addition, the re-ranking approach is providing to get the result which is efficient in terms of more than two feature spaces. Options are provided to filter the search by various features such as file size, duration, category and quality aspects.

REFERENCES

Chaisorn L., K.-W. Wan, Y.-T. Zheng, Y. Zhu, T.-S. Kok, H.-L. Tan, Z. Fu, and S. Bolling, "TRECVID 2010 known-item search (KIS) task by I2R," in *Proc. TRECVID*, 2010.

Chen X., J.Yuan, L. Nie,Z.-J.Zha,S.Yan, andT.-S.Chua, "TRECVID 2010 known-item search by NUS," in *Proc. TRECVID*, 2010.

Hofmann K., B. Huurnink, M. Bron, and M. de Rijke, "Comparing click-through data to purchase decisions for retrieval evaluation," in *Proc. SIGIR*, 2010, pp. 761–762.

Huijbregts M., R. Ordelman, and F. de Jong, "Annotation of heterogeneous multimedia content using automatic speech recognition," in *Proc. SAMT*, Berlin, Germany, 2007, LNCS, Springer Verlag.

Huurnink B., C. G. M. Snoek, M. de Rijke, and A. W. M. Smeulders , "Today's and tomorrow's retrieval practice in the audiovisual archive," in *Proc. CIVR*, 2010, ACM.

Huurnink B., K.Hofmann,M. de Rijke, andM. Bron, "Validating query simulators: An experiment using commercial searches and purchases," in *Proc. CLEF*, Padova, Italy, 2010, Springer.

Huurnink B., L. Hollink, W. van den Heuvel, and M. de Rijke, "The search behavior of media professionals at an audiovisual archive: A transaction log analysis," *JASIST*, vol. 61, no. 6, 2010.

Jiang Y.-G., J. Yang, C.-W. Ngo, and A. G. Hauptmann, "Representations of keypoint-based semantic concept detection: A comprehensive study," *IEEE Trans. Multimedia*, vol. 12, pp. 42–53, 2010.

Oomen J. and R. Ordelman, "Accessing audiovisual heritage: A roadmap for collaborative innovation," *IEEE Multimedia*, vol. 18, no.4, pp. 4–10, 2011.

Snoek C. G. M. and A. W. M. Smeulders, "Visual-concept search solved?," *IEEE Comput.*, vol. 43, no. 6, pp. 76–78, 2010.

Snoek C. G. M. and M. Worring, "Concept-based video retrieval," *Found. Trends Inf. Retriev.*, vol. 4, no. 2, pp. 215–322, 2009.

Snoek C. G. M., M. Worring, O. de Rooij, K. E. A. van de Sande, R. Yan, and A. G. Hauptmann, "VideOlympics: Real-time evaluation of multimedia retrieval systems," *IEEE Multimedia*, vol. 15, no. 1, pp. 86–91, 2008.

Tan H. and C. Ngo, "Fusing heterogeneous modalities for video and image re-ranking," in *Proc. ICMR*, 2011, p. 15, ACM.

Van de Sande K. E. A., T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene

recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, pp. 1582–1596, 2010.

Van Gemert J. C., C. J.Veenman, A.W. M. Smeulders, and J.M. Geusebroek, "Visual word ambiguity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1271–1283, Jul. 2010.

Wactlar H. D., M. G. Christel, Y. Gong, and A. G. Hauptmann, "Lessons learned from building a terabyte digital video library," *IEEE Comput.*, vol. 32, no. 2, pp. 66–73, 1999.

Wei X.-Y., C.-W. Ngo, andY.-G. Jiang, "Selection of concept detectors for video search by ontology-enriched semantic spaces," *IEEE Trans.Multimedia*, vol. 10, no. 6, pp. 1085–1096, 2008.

****