

A Novel Framework for Video Search

¹Sunil Parihar, ²Kshitij Pathak

^{1,2}Dept. of IT, MIT, RGPV University Bhopal, India

Abstract

The technologies are growing continually and due to this there are a large network of social networking and media collection library is increased. Particularly searching for a video from media libraries on the network has become a challenging task. Presently video search is mainly based on text, titles, descriptions and image features associated with it. There are many methods have been developed to improve the video search performance but they don't provide high accuracy on top ranked documents. In this paper we present a novel framework that integrates multiple features and help us to improve the video search performance in terms of relatedness of documents. We use semantic mapping and feedback policy to gain high accuracy on top ranked result. The proposed framework may be the most promising framework to gain high accuracy on the top ranked documents. We will provide the result on the basis of two performance parameters namely lost query result and ghost query result.

Keywords

Ghost Results, Media Library, Semantic Mapping, Video Search, WWW

I. Introduction

Due to increasing demand of media over the internet leads to design media library for collecting different kinds of videos. There is various popular video collection libraries are available i.e. YouTube and others. But the problem arises where we required fetching the correct video form this library. Video search has developed slowly through several basic search formats which exist today and uses keywords. The keywords for each search can be found in the title of the media, any text attached to the media and content linked web pages, also defined by authors and users of video hosted resources.

Some video search is performed using human powered search, others create technological systems that work automatically to detect what is in the video and match the searchers needs. Many efforts to improve video search including both human powered search as well as writing algorithm that recognize what's inside the video have meant complete redevelopment of search efforts.

The purpose of the search engine is to fulfill user's need and to provide accurate search results to the user. We should also take care of user behavior while designing search engine. The earlier work [1-2] show that users are highly interested in visiting top ranked documents they hardly check the entire result list. Thus it is more important to provide high accuracy on the top ranked documents.

In this paper, we propose a searching technique that not only improves the search performance but also fetch the most relevant documents and place it on the top of the result list.

II. Related Work

Several methods have been presented to improve the retrieval performance of video search engines. For Example [3-8], which are mainly deals with relevance feedback strategy. Relevance feedback technique is an interactive technique which improves the initial search result but take much time in labeling for updating the query model. On the contrary pseudo relevance feedback method

considers that the top ranked documents are relevant and use them to automatically refine the search process [9]. Relevance feedback and Pseudo relevance feedback techniques improve the retrieval performance of video search but do not promise that these relevant shots will be top positioned. Metasearch strategy [10-11] is also used to improve the performance of video search. Meta-search is mainly based on the unequal overlap property in which result obtained by several search engines is combined to form a single list in an optimum way and assign a highest rank to the document which are simultaneously occur in multiple result list. But the main problem associated with metasearch is that in multimodal representations it is very difficult for a user to provide query examples and also it is very hard to access multiple search engines for the same query example.

Recently the reranking methods are used to improve the video search performance. In reranking the initial search results are rearranged so that highly relevance documents can be reached on the top of the result list. Usually reranking methods are used in web search for ranking the web pages [12-13]. In multimedia search community reranking technique is successfully implemented in IB-Reranking [14], and CR-Reranking [15]. IB-Reranking uses bottleneck principle to reorder the initial search result and it works only for single feature space while CR-Reranking based on cross reference strategy to hierarchically combine all ranked clusters from various modalities and provides better performance over IB-Reranking. In CR-Reranking firstly initial search results are clustered separately in diversified feature space then ranking the clusters with relevance to the query, and then hierarchically fusing all the ranked clusters using a cross-reference strategy. This overall process of CR-Reranking makes it more time consuming.

III. Research Gap

In order to get the effective and efficient video search technique following key factors should be taken into consideration which are:

- Required to enhance the way of input query: use the relevance query keywords, and example image as input.
- Required to extract the feature space to make effective methods for extracting the most optimal results.
- Required to implement the semantics and text query optimization that understand the user query and perform visual mapping for the query text.

IV. Video Search Techniques

These are some techniques which we have used in our proposed framework.

A. Semantic Mapping

Our aims is to find a set of concepts that have the highest linguistic relatedness to the text queries, To measure such relatedness, required to adopt previously proposed work given in [16] by first building a Semantic Space (SS) and then performing concept reasoning. SS is an orthogonal linear space encapsulating the semantic relationship (mainly is-a relation) among text words. The relationship is learnt from Word Net using an ontology-based measure named WUP [16]. In SS, a concept or word, when projected to this space, is represented as a vector. The semantic

similarity (the quantitative linguistic relatedness) between two concepts C_i and C_j is measured with cosine similarity as:

$$Semantics(C_i, C_j) = \frac{Sc_i \cdot Sc_j}{|Sc_i| |Sc_j|} \quad (1)$$

Where Sc_i and Sc_j are the feature vectors of C_i and C_j in SS respectively. Different from conventional ontology-based measures, SS is a computable space and provides a global view of concepts in determining semantic relatedness between concepts.

Denote $Q = \{q_1, q_2, \dots\}$ as a text query, and $V = \{C_1, C_2, \dots\}$ as a concept set. Through projecting the query terms in Q into the SS, the relatedness between a text term $q_i \in Q$ and a concept $C_j \in V$ can be computed by Equation (1). By considering all the query terms and selecting one concept with the highest relatedness for each term, we have

$$S = \bigcup_{q_i \in Q} \operatorname{argmax}_{C_j \in V} \{Semantics(q_i, C_j)\} \quad (2)$$

Where S is the set of semantic concepts selected for query answering, and Semantic (q_i, C_j).

B. CR-Reranking

CR-Reranking [15], combines multimodal features in the manner of cross reference. The fundamental idea of CR- Reranking lies in the fact that the semantic understanding of video content from different modalities can reach an agreement. Actually, this idea is derived from the multi- view learning strategy [17-20] a semi-supervised method in machine learning. Multi-view learning first partitions available attributes into disjointed subsets (or views), and then cooperatively uses the information from various views to learn the target model. Its theoretical foundation depends on the assumption that different views are compatible and uncorrelated. In our context, the assumption means that various modalities should be comparable in effectiveness and independent of each other. Multiview strategy has been successfully applied to various research fields, such as concept detection. However, this strategy, here, is utilized for inferring the most relevant shots in the initial search results, which is different from its original role. CR-Reranking method contains three main stages: clustering the initial search results separately in diverse feature spaces, ranking the clusters by their relevance to the query, and hierarchically fusing all the ranked clusters using a cross-reference strategy.

V. Proposed Work

As we discussed in section II there are various approaches and work available to optimize video search results and we found that few techniques fulfill the user requirement and still there is a scope for betterment.

A. Problem Domain

Our concluded problem is related to design a high performance video search engine that works with the user search query to identify the most relevant search results. For this purpose we need to deal with following problems:

1. User query understanding.
2. More exact definition creation by the end user.
3. Able to search with examples.
4. Reducing the search complexity for end user
5. Reducing the search time and training time for improving the search results.

B. Solution Domain

Our framework is based on two different models for video search first concept based semantic and second multimodal fusion for defining constraints of video search. Here we provide some suggestion for implementation of a hybrid model that is based on CR- Reranking and Concept based reigning. The suggested solutions are as follows:

1. Apply the offline training method to extract images from video.
2. Building text mapping from this image objects.
3. Introducing semantics by defining user query image and more than one keyword applications.
4. Define the results on the basis of success and failure ratio.

C. Proposed Framework

The proposed framework is shown in fig. 1. Here we are planning to implement the complete search scenario in some small modules these are:

1. Offline Training

In offline training input video is provide to system and extract some images from video, use the keyword database to introduce image definition and used to get relativity of image keywords and video definition. Here a feedback is collected by the system to improve semantics.

2. Search Query

Introduces the AJAX based schemes to help for writing query for video using text and images.

3. Search Results

Introduces the search results in steps to filter unwanted or unmatched search results and to bring the most relevant documents on the top of the result list.

4. Performance Analysis

The system performance will be measured in the form of success and failure ratio, where failure ratio represents the relevant shots that are lost and irrelevant shots that are appeared after our reranking scheme. These two parameters are known as lost query results and ghost query results. Additionally the success of the system is calculated as amount of correct results/total results in the category.

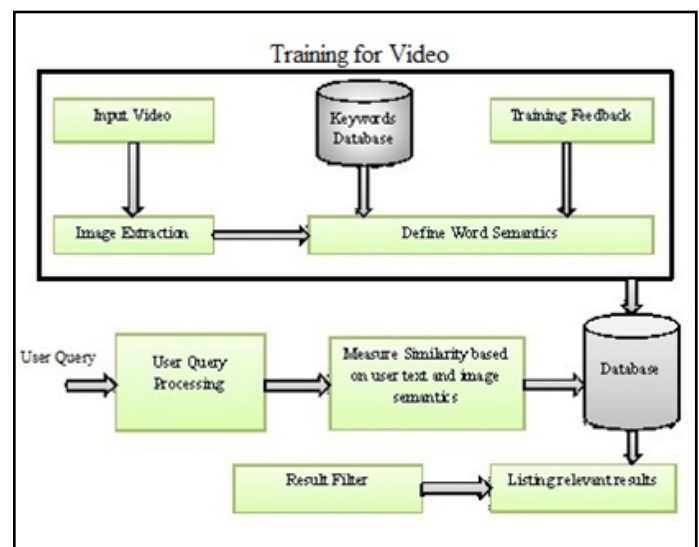


Fig. 1: Proposed Framework

VI. Conclusion and Future Work

In this paper our main goal is to analyze and review the traditional methods of video search and collect the facts by which we design and implement newer approach of the video search. After analysis we found the following conclusion over the traditional methods:

1. We get the two most popular and effective search techniques that produces accurate results.
2. We propose a new hybrid technique to get a newer model for video search.

In future we will implement this proposed Framework and simulate the results.

References

- [1] T. Joachims, "Optimizing Search Engines Using Clickthrough Data", Proc. ACM SIGKDD, pp. 133-142, 2002.
- [2] C. Silverstein, H. Marais, M. Henzinger, M. Moricz, "Analysis of a Very Large Web Search Engine Query Log", ACM SIGIR Forum, Vol. 33, pp. 6-12, 1999.
- [3] C.G.M. Snoek, J.C. van Gemert, J.M. Geusebroek, B. Huurnink, D.C. Koelma, G.P. Nguyen, O. de Rooij, F.J. Seinstra, A.W.M. Smeulders, C.J. Veenman, M. Worring, "The MediaMill TRECVID 2005 Semantic Video Search Engine", TREC Video Retrieval Evaluation Online Proc., 2005.
- [4] A. Amir, J. Argillander, M. Campbell, A. Haubold, G. Iyengar, S.Ebadollahi, F. Kang, M.R. Naphade, A. Natsev, J.R. Smith, J. Te_si_c, T. Volkmer, "IBM Research TRECVID-2005 Video Retrieval System", TREC Video Retrieval Evaluation Online Proc., 2005.
- [5] S.F. Chang, W.H. Hsu, L. Kennedy, L. Xie, A. Yanagawa, E. Zavesky, D.-Q. Zhang, "Columbia University TRECVID-2005 Video Search and High-Level Feature Extraction", TREC Video Retrieval Evaluation Online Proc., 2005.
- [6] A.G. Hauptmann, M. Christel, R. Concescu, J. Gao, Q. Jin, W.-H. Lin, J.-Y. Pan, S.M. Stevens, R. Yan, J. Yang, Y. Zhang, "CMU Informedia's TRECVID 2005 Skirmishes", TREC Video Retrieval Evaluation Online Proc., 2005.
- [7] J.H. Yuan, W.J. Zheng, L. Chen, D.Y. Ding, D. Wang, Z.J. Tong, H.Y. Wang, J. Wu, J.M. Lin, B. Zhang, "Tsinghua University at TRECVID 2005", TREC Video Retrieval Evaluation Online Proc., 2005.
- [8] S.K. Wei, Y. Zhao, Z.F. Zhu, N. Liu, Y.F. Zhao, L. Zhang, F. Wang, "BJTU TRECVID 2006 Video Retrieval System", TREC Video Retrieval Evaluation Online Proc., 2006.
- [9] W.H. Hsu, L.S. Kennedy, S.-F. Chang, "Reranking Methods for Visual Search", IEEE Trans. Multimedia, Vol. 14, No. 3, pp. 14-22, July-Sept. 2007.
- [10] J.H. Lee, "Analyses of Multiple Evidence Combination", ACM SIGIR Forum, Vol. 31, pp. 267-276, 1997.
- [11] J.A. Aslam, M. Montague, "Models for Metasearch," Proc. 24th Ann., Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 276-284, 2001.
- [12] L. Page, S. Brin, R. Motwani, T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web", Stanford Digital Libraries Working Paper, 1998.
- [13] T. H. Haveliwala, "Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search", IEEE Trans. Knowledge and Data Eng., Vol. 15, No. 4, pp. 784-796, July/Aug. 2003.
- [14] W.H. Hsu, L.S. Kennedy, S.-F. Chang, "Reranking Methods for Visual Search," IEEE Trans. Multimedia, Vol. 14, No. 3, pp. 14-22, July-Sept. 2007.

- [15] Shikui Wei, Yao Zhao, Zhenfeng Zhu, Nan Liu, "Multimodal fusion for Video Search Reranking", IEEE Trans. Knowledge and Data Eng., Vol. 22, No. 8, pp. 1191-1199, August 2010.
- [16] X.-Y. Wei, C.-W. Ngo, Y.-G. Jiang, "Selection of concept detectors for video search by ontology-enriched semantic spaces", IEEE Trans. on Multimedia, Vol. 10, No. 6, pp. 1085-1096, 2008.
- [17] A. Blum, T. Mitchell, "Combining Labeled and Unlabeled Data with Co-Training", Proc. 11th Ann. Conf. Computational Learning Theory, pp. 92-100, 1998.
- [18] K. Nigam, R. Ghani, "Understanding the Behavior of Co-Training", Proc. KDD-2000 Workshop Text Mining, 2000.
- [19] Z.H. Zhou, M. Li, "Tri-Training: Exploiting Unlabeled Data Using Three Classifiers", IEEE Trans. Knowledge and Data Eng., Vol. 17, No. 11, pp. 1529-1541, Nov. 2005.
- [20] I. Muslea, S. Minton, C. Knoblock, "Active+ Semi-Supervised Learning $\frac{1}{4}$ Robust Multi-View Learning", Proc. Int'l Conf. Machine Learning, pp. 435-442, 2002.



Sunil Parihar has received his B.E. degree in Computer Science & Engineering from RGPV University Bhopal, Madhya Pradesh, India, in 2006 and pursuing M. Tech from RGPV University Bhopal, Madhya Pradesh, India. He is active member of IEEE Society. His main research interest includes Data Mining and Multimedia Computing.



Kshitij Pathak is working as a Faculty in the Department of Information Technology, Mahakal Institute of Technology Ujjain since August 2005. He is also heading the Software Cell of the institute. He served as a faculty in Ujjain Engineering College Ujjain during 2003 to 2005. He has a rich academic and research experience of more than 7 years. Mr. Pathak has a meritorious academic career and did his B.E. (CSE) and M. Tech (IT) with distinction. He is actively involved in Management Committee of Computer Society of India –Ujjain Chapter for the duration of 2011-12. He is instrumental in developing an ERP package for MIT group of Institutions. He has participated and contributed in various National/International conferences / Seminars/Workshops/FDPs. His area of interest and research are: Data Mining, Algorithms, Computational Learning and Software Engineering.