International Journal of Trend in Research and Development, Volume 2(4), ISSN 2394-9333
www.ijtrd.com

Video Search Reranking Via Cross Reference Based Fusion Strategy

P. Perumal [1], D. Anandhu [2]
[1,2]Sri Ramakrishna Engineering College, Department of CSE, Coimbatore, India.

Abstract: In this paper the video retrieval process is evaluated to produce the top ranked search results to the query relevance. i.e., every large search engine log shows that the users are really interested in top ranked result according to their query. Therefore, it is essential to achieve high accuracy in video search retrieval. Generally, the search query for video retrieval is converted to text query and then searches for the relevant results (videos). While many methods exist for improving video search performance, they pay less attention to the above factor or encounter difficulties in practical applications. To overcome the limitations of the existing reranking methods we present a flexible and effective method called Cross Reference Reranking (CR- Reranking), to improve the retrieval effectiveness. To provide high accuracy in video retrieval, CR-Reranking involves a cross reference method, to fuse multimodal features. Particularly, multimodal features are first taken separately to rerank the initial search results at the cluster level with cluster number, and then all the ranked clusters from different modalities are fused together and produce the top ranked search results with high relevance to user query.

Keywords: Clustering, Fusion, Pseudo relevance feedback, Reranking, Relevance feedback.

I. INTRODUCTION

As an emerging research field, content-based video retrieval (CBVR) has attracted a great deal of attention in recent years. While various retrieval models have been developed to improve video search quality, most of them implement search procedure by implicitly or explicitly measuring the similarity between the query and database shots in some low-level feature spaces. However, such similarity is not usually consistent with human perception due to the limitation of current image/video understanding techniques. That is, the semantic gap exists between the low-level features and high-level semantics. For example, although a scene with red flags and a scene with red buildings share similar color features, they have completely different semantic meanings. The semantic gap will enlarge linearly with the increase of data set size since a larger data set means more confusion, which thereby leads to rapid deterioration of search performance.

Performance comparison between TRECVID'05 and TRECVID'06 evaluation on all the three search types, i.e., automatic, manual, and interactive, also reveals it. Consequently, it is more attainable for low-level features to reliably distinguish different shots in a relatively small collection, which is the basis of proposed reranking scheme. If we consider that the final aim of search engines is to meet users' information needs, it is reasonable to take user satisfaction and user behavior into account when designing a search engine. According to the analysis in, users are rarely patient to go through the entire result list. Instead, they usually check the top-ranked documents. Analysis on click-through data from a very large Web search engine log also reflects such preference. Therefore, it is more crucial to offer high accuracy on the top-ranked documents than to improve the whole search performance on the entire result list.

II. RELATED WORK

There are many methods, proposed for improving the retrieval performance of video search engines. The earlier work which is based on relevance feedback (RF) strategy focuses mainly on the refinement of the initial search results in an interactive fashion. However, RF-based methods require users' labeling for updating the query model, which is usually time-consuming and even impractical in some search scenarios. In contrast, pseudorelevance feedback (PRF)-based methods assume that the top-ranked documents are relevant and use them to automatically refine the search process.

Pseudo relevance feedback (PRF)-based methods assume that the top-ranked documents are relevant and use them to automatically refine the search process. For instance, the coretrieval algorithm treats the top-ranked results as positive examples and others as negative ones. Using these noisy training samples, a retrained retrieval model is then built via an Adaboost-based ensemble learning method. Although both RF- and PRF-based methods have achieved precision improvement on the entire result list by returning more relevant shots, no mechanism guarantees that these relevant shots will be top positioned.

Recently, the metasearch strategy, which is originally put forward in the field of information retrieval, is imported to CBVR for improving video retrieval effectiveness. The key idea of metasearch is that multiple result lists returned by several different search engines in response to a given query are aggregated into a single list in an optimal way. Metasearch is generally based on the "unequal overlap property": different search models retrieve many of the same relevant documents, but different irrelevant documents. Using this property, the combination of the returned lists is performed by simply giving higher ranks to the documents that are contained simultaneously in multiple result lists. Similar schemes include the PageRank-like graph-based approach and the model-based reranking algorithm. As a kind of multimodal fusion method, metasearch can simultaneously leverage multiple ranked lists from several engines based on various modalities. However, a general problem with metasearch is that it is usually hard to expect users to provide query examples with multimodal representations. In addition, it is not easy in practice to get access to multiple search engines based on different modalities.

As an alternative scheme, the reranking method can improve search quality by reordering the initial result list. Although the total number of relevant documents remains fixed after reranking, the precision improvement at the low depth of the result list can be expected by forcing true relevant documents to move forward. Traditionally, this kind of technique is used in the field of Web search. The predominant work includes PageRank and HITS. In the multimedia search community, the idea of reranking has been extended to develop advanced video search engines. As a successful attempt, IB-Reranking based on the Information Bottleneck (IB) principle, explores multimodal cues to reorder the initial search results. It finds some relevance-consistent clusters first and then ranks shots within the resulting clusters. In this method, however, multiple modalities are integrated in a unique feature space, that is, multimodal features are fused by concatenating them into a single representation. This fusion strategy is called early fusion. As a consequence, IBReranking is carried out only in a single feature space by which the accuracy on the top-ranked documents receives relatively less attention.

A) CR-Reranking

The reranking method, called CRReranking, which combines multimodal features in the manner of cross reference. The fundamental idea of CRReranking lies in the fact that the semantic understanding of video content from different modalities Multiview learning strategy, a semisupervised method in machine learning. Multiview 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. 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: Specifically, the initial search results are first divided into several clusters individually in different feature spaces. Then, the clusters from each space are mapped to the predefined ranks according to their relevance to the query. Given the ranked clusters from all the feature spaces, the cross-reference strategy can hierarchically fuse them into a unique and improved result ranking. Experimental results show that the search effectiveness, especially on the topranked results, is improved significantly.

As analyzed, the reranking method is sensitive to the number of clusters due to the limitation of cluster ranking.

III. PROPOSED WORK

In the proposed work, four key contributions are made to the video Search reranking. The first contribution is that multiple modalities are considered individually during clustering and cluster ranking processes. It means that reranking at the cluster level is conducted separately in distinct feature spaces, which provides a possibility for offering higher accuracy on the top-ranked documents. In contrast, in previous work multimodal features are first concatenated into a unique feature, and the subsequent clustering and cluster ranking are then implemented once in the above unique feature space. The second contribution is adaptively giving the cluster numbers and improves efficiency. The third contribution is defining a strategy for selecting some query-relevant shots to convey users' query intent. Instead of directly treating the top-ranked results as relevant examples like PRF, further filter out some irrelevant shots using some properties existing in the initial rankings. Reliably selecting a query-relevant shot set has a beneficial effect on cluster ranking. The third contribution is presenting a cross-reference strategy to hierarchically combine all the ranked clusters from various modalities. We assume that the shot with high relevance should be the one that simultaneously exists in multiple high-ranked clusters from different modalities. Based on this assumption, the shots with high relevance can be inferred cooperatively using the cross-reference strategy and then be brought up to the top of the result list. As a result, the accuracy on the top-ranked documents is given more consideration. Because the "unequal overlap property" is employed implicitly, this fusion strategy is similar to the metasearch methods to a certain extent. However, our cross reference strategy differs in two ways from metasearch. The first difference is that, instead of combining multiple ranked lists from different search engines, we integrate multiple reordered variants of the same result list obtained from only one text-based video search engine. The second one is that, instead of fusing multiple lists at the shot level, we first coarsely rank each list at the cluster level, and then integrate all the resulting clusters hierarchically. Experimental results indicate that CR-Reranking method indeed achieves higher accuracy on the top-ranked shots.

IV. PROBLEM ANALYSIS

Currently, text information associated with video content is the main source used in successful semantic video search engines. In those search engines, researchers give much consideration to feature extraction and similarity measurement. Before presenting the proposed reranking scheme, in this section, we first analyze the weakness in those search engines and then judge whether it is possible to alleviate the weakness using the reranking technique.

A) Weakness of Current Search Engines

As a well-recognized community for video search, NIST TRECVID provides 24 query topics for all participants to test their video search systems. In annual competition, all participants are required to return a ranking of 1,000 shots for each query topic and to submit at least one run (including 24 rankings where one ranking corresponds to one topic) for performance evaluation. In TRECVID'06, 76 runs, which are obtained mainly from text-based video search engines, are submitted, including the run (named as BJTU) from our developed video search system. Analyzing the retrieval effectiveness of these runs, we can reveal the weakness of current video search engines. Here, the average numbers of the relevant shots at different depths of the result list are used as the evaluation criterion. Given a depth X, the average number at depth X can be obtained by averaging the numbers of relevant shots in the top-X results over all 24 rankings.

V. MULTIMODALITIES SCHEME

Current web video search results rely exclusively on text keywords or user-supplied tags. A search on typical popular video often returns many duplicate and near-duplicate videos in the top results. We aim to present technique that gives the client with the apt result they need. the video itself is generally endowed with multiple information sources. Hence, fusing information from multiple modalities, i.e., multimodal fusion for short, is a popular way currently to enhance the understanding of video content, which thereby helps to develop excellent video search engines. Likewise, video search reranking can also benefit from multimodal fusion, especially when the size of the returned result set is relatively small. Based on the idea, a multimodal reranking scheme called CR-Reranking is proposed.

A) Over view

The framework of CR-Reranking shown in Fig.1, where $\{d_1; d_2; \ldots ; d_7\}$ denotes the initial result list ranked according to text-based search scores. The initial result list is processed individually in two distinct feature spaces, i.e., feature spaces A and B. In each feature space, all the results are first clustered into three clusters, and then the resulting clusters are mapped to three predefined rank levels, i.e., High, Median, and Low, in terms of their relevance to the query. Finally, a unique and improved shot ranking is formed by hierarchically combining all the ranked clusters from two different spaces. Note that only two modalities (or features) are considered here; however, the system can be easily extended to more modalities (or features).

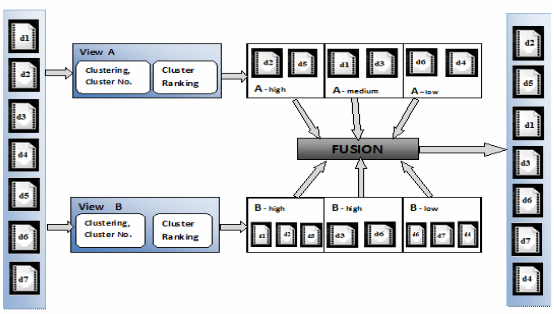

Fig.1 Framework of proposed CR - Reranking

B) Multispace Clustering

We handle the initial search results by performing clustering and cluster ranking operations separately in two feature spaces. Clustering the initial search results, we can obtain three clusters from each feature space, which are needed for the hierarchical fusion. In our case, the initial result list of 1,000 shots used for reranking is a relatively small shot set. Hence, it is possible to nicely partition the initial list into several clusters in certain low-level feature spaces. Specifically, after extracting multiple features for each shot, we carry out clustering independently in these feature spaces. As a result, we can obtain a certain number of clusters from each feature space, which paves the way for implementing our cross-reference strategy. In our scheme, NCuts clustering algorithm, one of the popular spectral clustering algorithms, is employed for clustering.

C) Ranking at the Cluster Level

After several clusters are obtained from one feature space, the next step in our scheme is to coarsely rank them by their relevance to the query. To this end, some query-relevant shots should be selected in advance to convey the query intent. Similar to our selecting approach is also inspired by the PRF method. That is, the top-ranked initial results are considered as the informative shots. Here, the limited results are selected. Compared with directly treating these shots as relevant shots or adopting "soft" pseudolabels strategy, the proposed scheme only chooses K most informative shots from them by exploiting the centralization and decentralization properties. By doing this, some irrelevant shots (i.e., noisy points) can be filtered out effectively.

Note that only the visual feature of grid color moment is utilized here. As we have analyzed, the relevant results in the limited shots $\{ A_1;A_2; \ldots;A_N \}$ usually group together in visual feature space, yet the irrelevant shots are scattered. It means that the distances between relevant shots are smaller than those distances between irrelevant shots or between relevant shots and irrelevant shots. Therefore, K shots with the smallest distances are more possible to be the shots conveying the query intent, which can be selected to form the query-relevant shot set E. The value of K is selected empirically and fixed to 10.

Therefore, the implementation of cluster ranking is equivalent to measuring the similarity between the E is the query-relevant set and the clusters C. For measuring the relevance between shot sets, we employ the modified Hausdorff distance

$$hd(E,C) = \underset{e\in E}{mean}\{\min_{c\in C}\{d(e,c)\}\}$$

We can assign corresponding ranks to the clusters in each modality space.

D) Cross-Reference-Based Fusion Strategy

Our final goal is to obtain a unique and improved reranking of the initial results, especially paying more attention to the accuracy on the top-ranked results the user is interested. In order to move vigorously toward this goal, we hierarchically fuse all the ranked clusters from different modalities using a cross-reference strategy. Our fusion approach is composed of three main components: combining these ranked clusters using cross-reference strategy, ranking subsets with the same rank level, and ranking shots within the same subset. Note that the rank levels are denoted numerically in the following formulas for the convenience of expression. The rank levels High, Median, and Low are equivalent to the rank levels 1, 2, and 3, respectively. We assume that a shot has a high rank if it exists simultaneously in multiple high-ranked clusters from different modalities. Based on this assumption, we put forward a cross-reference strategy to hierarchically combine all the ranked clusters, leading to a coarsely ranked subset list. Specifically, let { $A_1;A_2; \ldots;A_N$ } and { $B_1;B_2; \ldots;A_N$ } be the sets of the ranked clusters from feature spaces A and B, respectively, and Rank be the operation of measuring the rank level of a cluster or shot. The ranked clusters in each set are arranged from high-rank level to low-rank level in ascending order of their subscripts, that is, Rank(A_i) is greater than Rank(A_{i+1}). Then, two ranked cluster sets can be integrated into a unique and coarsely ranked subset list according to the following inference rule:

$$\mathrm{Rank}(A_i \cap B_j) \succ \mathrm{Rank}(Am \cap Bn);$$
$$\text{if } (i+j) \prec (m+n), i, j, m, n=1, \ldots, N,$$

where N is the number of clusters, and Ai Bj stands for the intersection of clusters Ai and Bj. As a matter of fact, the rank levels of subsets cannot be compared using merely the above criteria if (i + j) is equal to (m n), just like the intersections (A1 ⋂ B2) and (A2 ⋂ B1). To address this issue, we employ the method used in the cluster ranking step to order those subsets, which can be formulized as follows:

$$\mathrm{Rank}(A_i \cap B_j) \succ \mathrm{Rank}(Am \cap Bn);$$
$$\text{if } (i+j)=(m+n), \mathrm{hd}(E, A_i \cap B_j) \prec \mathrm{hd}(E, A_m \cap B_n),$$

where the distance can be computed in any of the feature spaces. So far, an ordered subset list has been formed. Although the ranks of shots in different subsets can be compared by the ranks of their corresponding subsets, we do not know which shot within the same subset is more relevant to the query. Hence, we need to find a method to order the shots within the same subset, i.e., ranking at the shot level. Here, the score or rank information of the initial ranking is used to order these shots. The ranking rule is defined as follows:

$$\mathrm{Rank}(d_m) \succ \mathrm{Rand}(d_n), \text{ if } S_m \succ S_n$$

where d_m and d_n denote shots m and n within the same subset, respectively, S_m and S_n correspond to the scores or ranks of shots m and n, respectively.

VI EXPERIMENTS

Evaluation on Different Ranking Methods

The proposed scheme is compared with several available methods for video search reranking in this section. All these reranking methods are conducted using only the TEXT feature and MM visual feature, which are constructed as follows:

A) Single-Reranking: This kind of reranking method is constructed by performing clustering and cluster ranking once in only one modality space. Here two systems are built individually in the TEXT and MM feature spaces, namely, Single-TEXT and Single-MM.

B) Early-Fusion Reranking: We construct this scheme by clustering and cluster ranking once in a single feature space. The main difference from Single-Reranking is that, instead of using only one modality, the feature vector used in Early-Fusion is formed by concatenating the vectors of multiple modalities. Here, we only concatenate the TEXT feature vector and MM visual feature vector.

C) Late-Fusion Reranking: The clustering results from two feature spaces (i.e., TEXT and MM spaces) are directly fused by randomly intersecting any two clusters from different modalities and then ranking the newly formed subset list. Compared with the proposed method, the Late-Fusion scheme skips the cluster ranking step before combination.

Table-1 Comparisons of different ranking schemes with text-baseline search

System	MAP (gain)	Prec_5	Prec_10	Prec_15	Prec_20	Prec_30	Prec_100
Text-only baseline	0.0333(0%)	0.1167	0.1375	0.1222	0.125	0.1264	0.0987
Single-TEXT	0.0398(19.5%)	0.1583	0.1708	0.1472	0.1375	0.1347	0.0908
Single-MM	0.0461(38.4%)	0.175	0.1792	0.1611	0.1437	0.1361	0.1062
Single-GCM	0.0489(46.87%)	0.15	0.1708	0.1611	0.1458	0.1458	0.1087
Single-EDH	0.0376(12.9%)	0.1167	0.1292	0.1333	0.1312	0.1236	0.1029

VII PERFORMANCE ANALYSIS ON QUERIES

Here, we evaluate the performance of our proposed system with all the query topics. This show that the proposed reranking scheme works well for named persons and named objects, such as “D. Cheney” and “Boats,” as the search quality on these topics can benefit from the TEXT feature used in our scheme which is shown in Fig.2.

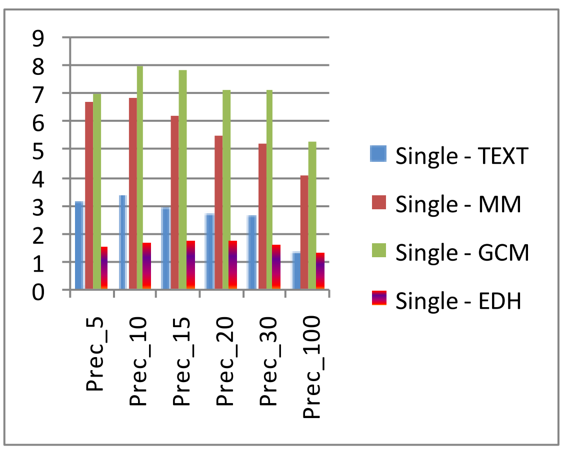

Fig.2 Performance Evaluation On Different Reranking Schemes

However, the search performance after reranking is even below the performance of text-only baseline for some topics with motion properties, like “leaving a vehicle.” The reason is that features used in our scheme lack the capability to capture motion properties in video. Hence, new research fruits in precise representation of shot will provide much more room for performance improvement. In addition, our proposed method also fails in some query topics with very few relevant shots within the limited results, such as “meeting” and “people with uniform.” It is because cluster ranking is based essentially on the relevant shots within the limited result sets.

CONCLUSION & FUTURE EHANCEMENTS

This system presents a new reranking method that combines multimodal features via a cross-reference strategy. It can handle the initial search results independently in various modality spaces. Specifically, the initial search results are first divided into several clusters individually in different feature spaces using Ncut clustering algorithm. Then, the clusters from each space are mapped to the predefined ranks according to their relevance to the query in advance by PRF method. Thus, the ranked clusters from all the feature spaces, the cross-reference strategy can hierarchically fuse them into a unique and improved result ranking. Finally the fused and top ranked result list is produced relevant to the query.

As analyzed that existing reranking method is sensitive to the number of clusters due to the limitation of cluster ranking. Thus the proposed method adaptively chooses cluster number for different feature spaces.

In the future, the proposed system will develop this method to produce the exact top ranked result list for all the feature spaces available in the video search results. In addition, new strategies are to be investigated for selecting query-relevant shots, e.g., using pseudonegative samples to exclude irrelevant shots.

ACKNOWLEDGEMENT

With sincere and heartfelt thanks to our Principal, Director of Sri Ramakrishna Engineering College, Coimbatore for helpful discussions and support to finish this work successfully.

REFERENCES

1. M.S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-Based Multimedia Information Retrieval: State of the Art and Challenges," ACM Trans. Multimedia Computing, Comm., and Applications, vol. 2, pp. 1-19, 2006.
2. [A. Smeaton and T. Ianeva, "TRECVID-2006 Search Task," TREC Video Retrieval Evaluation Online Proc., 2006.
3. W. H. Hsu and S.-F. Chang. Visual cue cluster construction via information bottleneck principle and kernel density estimation. In CIVR, Singapore, 2005.
4. Natsev, M. R. Naphade, and J. Tesic. Learning the semantics of multimedia queries and concepts from a small number of examples. In ACM Multimedia, pages 598–607, Singapore, 2005.
5. R. Yan, A. Hauptmann, and R. Jin. Multimedia search with pseudo-relevance feedback. In CIVR,Urbana- Champaign, IL, 2003.
6. A.Amir et al. IBM Research TRECVID-2005 video retrieval system. In TRECVID Workshop, Washington DC, 2005.
7. R. Yan, A. Hauptmann, and R. Jin, "Multimedia search with pseudo-relevance feedback," *ACM CIVR*, 2003.
8. T.-S. Chua et al, "TRECVID 2004 search and featureextraction task by NUS PRIS,"*NIST TRECVID workshop*, 2004.
9. S.-F. Chang et al, "Columbia University TRECVID-2006video search and high-level feature extraction", *NIST TRECVID workshop*, 2006.
10. J. Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture*, Portfolio Trade, 2006.
11. W. Hsu et al "Video search reranking through random walk over document-level context graph," *ACM Multimedia*, 2007.
12. L. Kennedy and S.-F. Chang, "A reranking approach for context-based concept fusion in video indexing and retrieval,"*ACM CIVR*, pp. 333–340, 2007.
13. R. Herbrich, T. Graepel, and K. Obermayer, "Support vector learning for ordinal regression," *ICANN*, pp. 97– 102, 1999.
14. Z. Cao et al, "Learning to rank: from pair wise approach to list wise approach," *IEEE ICML*, pp. 129–136, 2007.
15. C. Snoek, M. Worring, J. van Gemert, J. Geusebroek, and A. Smeulders.The challenge problem for automated detection of 101 semantic concepts in multimedia. In *Proc. 14th Annual ACM Intl. Conference on Multimedia*, pp. 421–430, Santa Barbara, CA, 2006.
16. J. Battelle. The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture. Portfolio Trade, 2006.
17. L. Page and et al. The page rank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1998.
18. A. N. Langville and C. D. Meyer. A survey of eigenvector methods for web information retrieval. SIAM Review, 47(1):135–161, 2005.
19. H. D.Wactlar, T. Kanade, M. A. Smith, and S. M. Stevens, "Intelligent access to digital video: Informedia project," *IEEE Computer*, vol. 29, no. 5, pp. 46–53, 1996.
20. M. R. Lyu, E. Yau, and K. S. Sze, "iview: An intelligent video over internet and wireless access system," in *Proc. 11th Int. World Wide Web Conf. (WWW2002), Practice and Experience Track*, Honolulu, HI, 2002.
21. S. C. H. Hoi, W. Liu, M. R. Lyu, and W.-Y. Ma, "Learning distancemetrics with contextual constraints for image retrieval," in *Proc. IEEEConf. Computer Vision and Pattern Recognition (CVPR'06)*, New York, Jun. 17–22, 2006.
22. Muslea, S. Minton, and C. Knoblock, "Active+ Semi-Supervised Learning ¼ Robust Multi-View Learning," Proc. Int'l Conf. Machine Learning, pp. 435-442, 2002.
23. R. Yan and M. Naphade, "Multi-Modal Video Concept Extraction Using Co-Training," Proc. IEEE Int'l Conf. Multimedia and Expo, pp. 514-517, 2005.
24. J. Shi and J. Malik, "Normalized Cuts and Image Segmentation,"IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 888-905, Aug. 2000.
25. W.H. Hsu, L.S. Kennedy, and S.-F. Chang, "Video Search Reranking via Information Bottleneck Principle," Proc. 14th Ann. Int'l Conf. Multimedia, pp. 35-44, 2006.

26. R. Yan and A.G. Hauptmann, "Co-Retrieval: A Boosted Reranking Approach for Video Retrieval," IEE Proc.Vision, Image and Signal Processing, vol. 152, pp. 888-895, 2005.
27. D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge, "Comparing Images Using the Hausdorff
28. Distance," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 9, pp. 850-863, Sept. 1993.
29. W.H. Hsu, L.S. Kennedy, and S.-F. Chang, "Video Search Reranking via Information Bottleneck Principle," Proc. 14th Ann. Int'l Conf. Multimedia, pp. 35- 44, 2006.
30. *Yi-Hsuan Yang and Winston H. Hsu, "*Video Search Reranking Via Online Ordinal Reranking", National Taiwan University, 2000.
31. Winston H. Hsu, Lyndon S. Kennedy, Shih-Fu Chang, "Video Search Reranking through Random Walk over Document-Level Context Graph", ACM-*MM'07,* September 23–28, 2007, Augsburg, Bavaria, Germany.
32. Steven C. H. Hoi, *Member*, *IEEE*, and Michael R. Lyu, *Fellow*, *IEEE*, "A Multimodal and Multilevel Ranking Scheme for Large-Scale Video Retrieval", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 10, NO. 4, JUNE 2008.
33. Shikui Wei, Yao Zhao, Member, IEEE, Zhenfeng Zhu, and Nan Liu, "Multimodal Fusion for Video Search Reranking", *IEEE Transactions on Knowledge and Data Engineering, Vol. 22, No. 8, August.2010.*