KB Video Retrieval at TRECVID 2011

David Etter
November 2011

Abstract

This paper describes KB Video Retrieval's participation in TRECVID 2011. We submitted 4 runs in the automatic Known-Item Search (KIS) task. Our KIS system uses a base text only search for runs KBVR_1_KIS2011 and KBVR_2_KIS2011. Runs KBVR_3_KIS2011 and KBVR_4_KIS2011 extend the text search with pseudo-relevance feedback and semantic concept re-ranking.

Automatic Known-Item Search Runs

- **KBVR_1_KIS2011**
  The first run uses a text only retrieval with a tf-idf ranking model.

- **KBVR_2_KIS2011**
  The second run uses a text only retrieval with a probabilistic ranking model.

- **KBVR_3_KIS2011**
  The third run uses a text only retrieval with probabilistic ranking and pseudo-relevance feedback.

- **KBVR_4_KIS2011**
  The final run uses an initial text retrieval with probabilistic ranking and semantic concept re-ranking.

Automatic Known Item Search

Text Matching

The KB Video Retrieval's system for TRECVID 2011 [1] is based on an initial text only match using the Terrier [5] retrieval system. The text search uses a query constructed from both the visual cues and the query description. A video document is constructed using automatic speech recognition and meta-data from the video. Text used from the meta data included title, subject, keywords, and description.

KBVR_1_KIS2011 and KBVR_2_KIS2011 were text only runs using tf-idf and probabilistic [6] ranking models. These runs were used as a baseline for comparison to the pseudo-relevance feedback and semantic concept re-ranking runs.

The pseudo-relevance run, KBVR_3_KIS2011, uses a probabilistic model to select an initial set of k relevant documents. The original query is then reformulated using key terms from these top rated videos. This model assumes that our initial query has returned relevant documents and that the top terms from these documents provide additional details not found in the original query.

Semantic Re-Ranking

The semantic re-ranking model used in KBVR_4_KIS2011 extends the initial text ranking by considering a set of semantic concepts identified in the retrieved videos. This approach assumes that the relevant video has been returned in our initial text search, but may not be the top ranked video. We map both the query and the initial documents into a semantic concept space. A similarity score is calculated for each video in this semantic space.

The semantic concept space is defined by 130 semantic concepts from the LSCOM [2,3] concept list. We use a Wikipedia [4] knowledge base to enhance the text description of a semantic concept. The Wikipedia knowledge base is constructed by selecting the k most similar articles to our semantic concepts. We use the article title and abstract for the initial concept representation. The knowledge base is then expanded to include related Wiki articles using internal links and category memberships.

Each query and video document is transformed into a 130 dimensional feature vector, where the dimensions represent a similarity score between the video/query text and the Wiki representation of the semantic concept. A second ranked list for the query is calculated using this semantic vector representation. Our final ranking is based on a weighted fusion of the scores from the initial probabilistic model and the semantic concept model.

Conclusions

KB Video Retrieval's run for TRECVID 2011 are based on ASR and meta data features from the video. We extended our baseline text only runs to include a semantic concept re-ranking by mapping both the videos and the queries into a semantic space. The semantic space is defined using 130 LSCOM concepts with the description of each concept enhanced using a Wikipedia knowledge base. The semantic concept re-ranking model uses a weighted fusion of the initial text ranking and the semantic concept ranking.

References

1. Smeaton, Alan F, Over, Paul and Kraaij, Wessel. Evaluation campaigns and TRECVid. New York, NY, USA : ACM, 2006. MIR '06: Proceedings of the 8th ACM international workshop on Multimedia information retrieval. pp. 321-330.

2. LSCOM Lexicon Definitions and Annotations Version 1.0, DTO Challenge Workshop on Large Scale Concept Ontology for Multimedia, Columbia University ADVENT Technical Report #217-2006-3, March 2006.

3. Naphade, M., Smith, J. R., Tesic, J., Chang, S.-F., Hsu, W., Kennedy, L., Hauptmann, A., and Curtis, J. Large-scale concept ontology for multimedia. IEEE MultiMedia 13, 3 (2006), 86–91.

4. Wikimedia Foundation. Wikipedia. [Online] 2010. http://download.wikimedia.org.

5. Iadh Ounis, Gianni Amati, Vassilis Plachouras, Ben He, Craig Macdonald, and Christina Lioma. Terrier: A High Performance and Scalable Information Retrieval Platform. In Proceedings of ACM SIGIR'06 Workshop on Open Source Information Retrieval (OSIR 2006). 10 August, 2006. Seattle, Washington, USA.

6. Robertson, S. E., and Jones, K. S. Relevance weighting of search terms. Journal of the American Society for Information Science. Volume 27, Issue 3, pages 129–146, May/June 1976143–160.