KB Video Retrieval at TRECVID 2010

David Etter

Oct 2010

Abstract

This paper describes KB Video Retrieval's participation in the TREC Video Retrieval Evaluation for 2010. This year we submitted results for the Semantic Indexing, Known-item Search, Instance Search, and Event Detection in Internet Multimedia tasks. Our goal this year was to evaluate ranking strategies and expand our knowledge based approach to a variety of data sets and tasks.

1. Introduction

This paper describes KB Video Retrieval's participation in the TREC Video Retrieval Evaluation for 2010 (1). The KBVR system for 2010 maintains its foundation in knowledge based search through query expansion techniques (8). We expand that approach this year by evaluating ranking strategies in both supervised and unsupervised settings. Our participation included the Semantic Indexing and Know-item Search tasks, where we submitted four runs for each task. We also participated in the pilot tasks Instance Search and Event Detection in Internet Multimedia, where we submitted four runs and one run. Our overall retrieval system and task submissions are described in the next section.

2. KB Video Retrieval System

Our KB Video Retrieval System was constructed with over 400 semantic classifiers based on the LSCOM concept list (9). These classifiers were trained using both positive and negative training data from an internet archive of image and video data (3,4,6).

A visual feature vector was constructed using low-level features which included edge, color, texture, and local keypoint. The visual feature vectors were then clustered to create sub-semantic topics within each concept. A final visual ranking function for each concept was then constructed using a boosting algorithm to learn the contribution for each sub-semantic ranker (7). In the internet multimedia task, we model an event by creating an event feature vector which consists of up to 10 concatenated shot feature vectors.

We also constructed an ontology of concepts and their relationships using a meta-data extract from Wikipedia (5). Relationships between concepts were identified by both reference links and category membership.

Our test data approach in each task was fundamentally the same. A set of key frames was extracted from each potential shot or video clip. We then extracted our low-level edge, color, texture, and local keypoint features to create a visual feature vector. Each key frame was then tagged by our semantic models with a ranking score. For each task where speech text or meta data is provided and allowed, we tagged the key frames with our semantic topics and their relationships. The query topics from each task followed a similar process of visual concept tagging and text semantic tagging. The system produces a final ranking using a similarity comparison of the combined concept vectors.

2.1 Semantic Indexing (SIN)

We submitted four runs in the Semantic Indexing task at TREC Video Retrieval Evaluation for 2010. Our submitted runs are listed below:

Semantic Indexing (SIN)

- *KBVR_1_SIN2010 – This is a run using a feature vector of 5 related concepts.*

- *KBVR_2_SIN2010 – This is a run using a feature vector of 10 related concepts.*

- *KBVR_3_SIN2010 – This is a run using a feature vector of 25 related concepts.*

- *KBVR_4_SIN2010 – This is a run using a feature vector of 50 related concepts.*

2.2 Known-item Search (KIS)

We submitted four runs in the Known-item Search task at TREC Video Retrieval Evaluation for 2010. Our submitted runs are listed below:

Known-item Search (KIS)

- *KBVR_1_KIS2010 – This is a run using a feature vector of 3 related concepts.*

- *KBVR_2_KIS2010 – This is a run using a feature vector of 5 related concepts.*

- *KBVR_3_KIS2010 – This is a run using a feature vector of 10 related concepts.*

- *KBVR_4_KIS2010 – This is a run using a feature vector of 15 related concepts.*

2.3 Instance Search (INS)

We submitted four runs in the Instance Search pilot task at TREC Video Retrieval Evaluation for 2010. Our submitted runs are listed below:

Instance Search (INS)

- *KBVR_INS_1 – This is a run using a feature vector of 1 concept.*

- *KBVR_INS_2 – This is a run using a feature vector of 10 related concepts.*

- *KBVR_INS_3 – This is a run using a feature vector of 15 related concepts.*

- *KBVR_INS_4 – This is a run using a feature vector of 5 related concepts.*

2.4 Event Detection in Internet Multimedia (MED)

We submitted one run, which included each of the three events, in the Event Detection in Internet Multimedia pilot task for TREC Video Retrieval Evaluation for 2010. Our submitted runs are listed below:

Event Detection in Internet Multimedia (MED)

- *KBVR_2010_MED_EVAL_p-baseline_1 – This run includes the three internet events.*

3. Conclusions

The KBVR system for 2010 was expanded from previous years to perform in four different tasks. Our system remains focused on a knowledge based approach, where we incorporate both specific and generic knowledge to expand our queries and enhance our understanding of the relationships between semantic concepts. This years system began to look at ways in which to successful combine rankings from a variety of feature based rankers. In the future, we will plan to continue our work on the learning to rank problem in video search .

4. References

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

Recognition. Huijbregts, Marijn, Ordelman, Roeland and Jong, Franciska de. [ed.] Bianca Falcidieno, et al. s.l. : Springer, 2007. SAMT. Vol. 4816, pp. 78-90.

3. Efficient use of local edge histogram descriptor. Park, Dong Kwon, Jeon, Yoon Seok and Won, Chee Sun. New York, NY, USA : ACM, 2000. MULTIMEDIA '00: Proceedings of the 2000 ACM workshops on Multimedia. pp. 51-54.

4. k-means++: the advantages of careful seeding. Arthur, David and Vassilvitskii, Sergei. Philadelphia, PA, USA : Society for Industrial and Applied Mathematics, 2007. SODA '07: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. pp. 1027-1035.

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

6. Google. Google Image Search. [Online] 2010. http://images.google.com.

7. Massih-Reza Amini, Vinh Truong, Cyril Goutte, A Boosting Algorithm for Learning Bipartite Ranking Functions with Partially Labeled Data, International ACM SIGIR conference, 2008.

8. KB Video Retrieval at TRECVID 2009. Etter, David. Gaitherburg, MD : NIST, 2009. Online Proceedings of the TRECVID Workshops.

9. 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.