



INTRODUCTION

The volume of video data is growing exponentially. This data need to be annotated to facilitate search and retrieval, so that we can quickly find a video whenever needed.

Manual Annotation, especially for such volume, is time consuming and would be expensive. Hence, automated annotation systems are required.

AIM

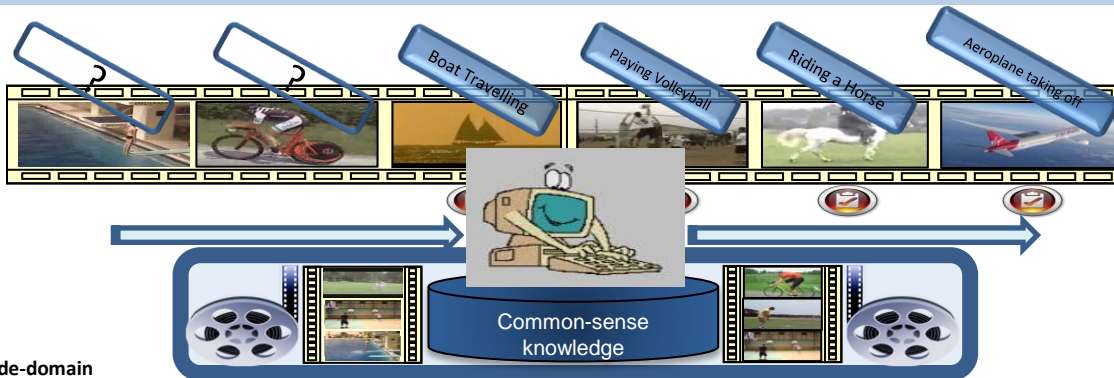
Automated Semantic Annotation of wide-domain videos (i.e. no domain restrictions). This is an important step towards bridging the "Semantic Gap" in video understanding.

METHOD

1. Extracting "Video Signature" for each video.
2. Match signatures to find most similar videos, with annotations
3. Analyse and process obtained annotations, in consultation with Common-sense knowledge-bases
4. Produce the suggested annotation.

EVALUATION

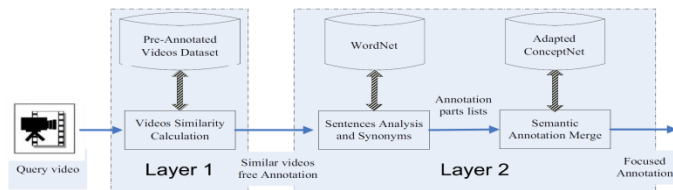
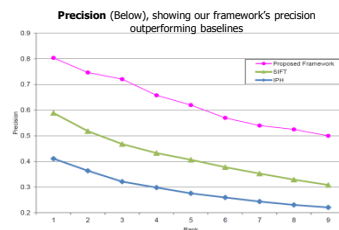
- Two standard, and challenging Datasets were used. TRECVID BBC Rush and UCF.
- Black-box and White-box testing carried out.
- Measures include: Precision, Confusion Matrix.



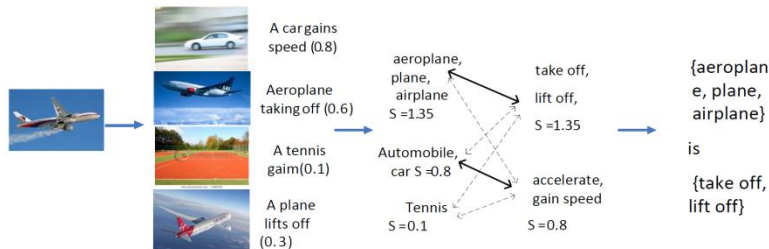
RESULTS

	airplane	boat	car	horse	jumping	tennis	volleyball
airplane	0.88	0.04	0.28	0.00	0.00	0.00	0.00
boat	0.04	0.88	0.32	0.00	0.00	0.04	0.00
car	0.28	0.32	0.88	0.00	0.00	0.00	0.00
horse	0.00	0.00	0.00	0.88	0.00	0.00	0.00
jumping	0.00	0.00	0.00	0.00	0.88	0.00	0.00
tennis	0.00	0.04	0.00	0.00	0.00	0.88	0.00
volleyball	0.00	0.00	0.00	0.00	0.00	0.00	0.88

Confusion Matrix (Above), showing most of videos recognised correctly. Jumping has confusion with others



Our Annotation Framework (Above), and a working example (below) with intermediate outputs



CONCLUSION

- Developed an Automatic Semantic Video Annotation framework.
- Not restricted to a specific domain videos.
- Utilising Common-sense Knowledge enhances scene understanding and improve semantic annotation.

Publications

- A Framework for Automatic Semantic Annotation for Domain-Independent Uncontrolled Videos Utilizing Similarity and Commonsense, Journal of Multimedia Tools And Applications (MTAP), Vol 64, No 2., Springer, (2012)
- Semantic levels of domain-independent commonsense knowledgebase for visual indexing and retrieval applications. Neural Information Processing. Lecture Notes in Computer Science, 7663. pp. 640-647. (2012)
- VisualNet: Commonsense Knowledgebase for Video and Image Indexing and Retrieval Application. In Proceedings of the IEEE International Conference on Intelligent Computing Systems (ICICS'09), China, (2009)
- Video databases annotation enhancing using commonsense knowledgebases for indexing and retrieval. In: The 13th IASTED International Conference on Artificial Intelligence and Soft Computing, Spain. (2009)