Survey on Different Techniques on Video Annotations

Archana V Potnurwar
Asst .Professor IT Dept
P.I,ET Nagpur

ABSTRACT
The insufficiency of labeled training data for representing the distribution of the entire dataset is a major obstacle in automatic semantic annotation of large-scale video database. Semi-supervised learning algorithms, which attempt to learn from both labeled and unlabeled data, are promising to solve this problem. In this paper **,r**etrieving videos using key words requires obtaining the semantic features of the videos. Most work reported in the literature focuses on annotating a video shot with a fixed number of key words, no matter how much information is contained in the video shot.

General Terms
Algorithms ,Video annotation.

Keywords
Videoannotations, multimodal, automaticvideo annotations, domain specific.

1. INTRODUCTION
Image annotation is an active field of research that serves as a precursor to video annotation in numerous ways. Video features are often inspired and sometimes directly borrowed from image techniques and many methods for image indexing are also easily applied to video. Here we survey some of the most relevant static image annotation literature including modern trends in the field and adaptations of techniques for static image annotation to video. In the following literature the covered topics include emerging and state of the art feature extraction techniques specifically designed for video. We review image features, indexing techniques, and scalable designs that are particularly useful for working with web-scale video collections

2. BACKGROUND
There are three types of image annotation approaches available: manual, automatic and semi automatic. Following table describes difference we can go for semi automatic as per the review. In annotation we can go for multimodality

. **Table 1 Annotation Techniques**

Annotation techniques	Manual	Semi Automatic	Automatic
Initial Human Interaction	Enter some descriptive keyword	Provide initial query at the beginning	No interaction
Machine task	Provide storage for annotation to be saved such as disk space or database	Parse Human's query and extract semantic information to perform annotation	Detect labels semantic keywords automatically using recognition technology

3. MULTIMODAL
Utilizing the available multimodality in video mediums, such as audio and sometimes enclosed text, has received relatively a good attention [2], In spite of that the multimodal features analysis usually increases certainty of video annotation, In this it was preferred to analyze input video's visual features only to keep focusing on wide domain. This was also to accommodate some domains where video clips lacks audio and enclosed text, or they are not so correlated with the visual features such as wild hunts and surveillance [3].

Multimodality. Different types of modality i.e Textual Modality , Visual Modality, Auditory Modality .Further also discussed Content based video indexing compromises of High-level indexing: Index on the basis of high level features e.g. action, time, and space. The main advantages of high-level indexing are that it can give more accurate semantic correct result. In high-level indexing, the high-level and low-level features are map to reduce the semantic gap. Low-level indexing: Index on the basis of low-level features e.g. colour, shape, and texture. Here no semantics is attached. Video can be retrieved by simple pattern matching and similarity measuring techniques. The main advantages of low-level indexing are that it is automatic and fast as compared to high-level indexing.

And Domain specific indexing: These technique uses high-level structure of video to constraints the low-level features extraction and processing. Also the Indexing Techniques of three types Segment-Based Video Indexing, Object-Based Video Indexing, Event-Based Video Indexing are discussed.

And focused on some issues that need to be considered.

a. Need for generalized multimodal video indexing techniques

b. Multimedia data (video) does not have a single unique semantic, so how do we highlight the semantic that will be further used for content based multimedia indexing.

c. The main challenge or complexity in video indexing and searching is that video data is multimodal. There is a need of a system that can decide that which modality is combined or used in order for maximum effectiveness and accurate searching.

d. Need of the framework for indexing that select the most appropriate mode for indexing or using the different modality combination[10].

4. SEMANTIC VIDEO ANNOTATION SYSTEM
a prototype of a video annotation system, called Semantic Video Annotation System (SVAS) in which a three-level annotation architecture and a semantic video search language called Semantic Query Description Language for Video (SQDL-V) is used. SQDL-V engine based on SVAS is able to

return more accurate search results in comparison to the formal video search method.[4].

4.1 Video semantic annotation using graph diffusion technique

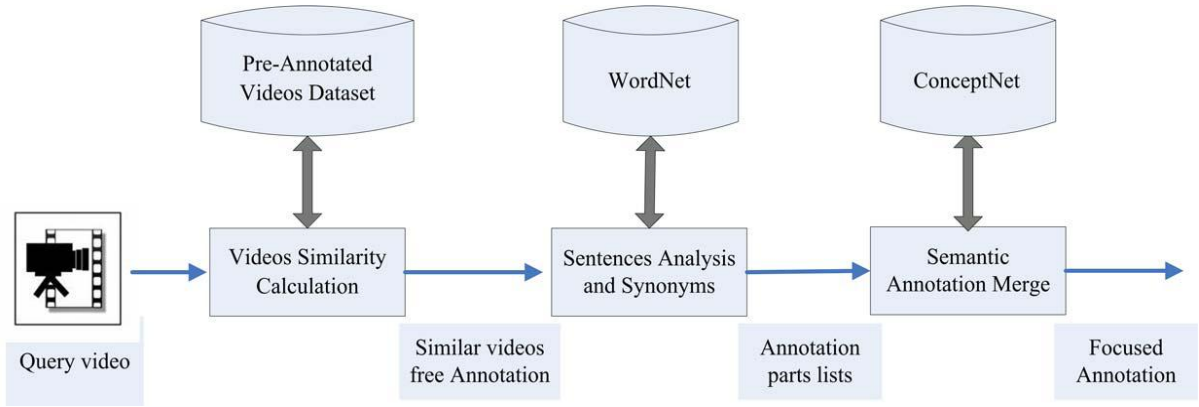
A novel and efficient approach for scalable to large data sets where only a couple of minutes improving large scale video semantic annotation using graph diffusion technique. The main concentration in this paper was on. Firstly, it allowed the online update of semantic context for addressing the problem of domain shift .Second , it was

required to complete approach implemented over hundreds of concepts for thousands of video shots[6].

4.2 Automatic video annotation method

The former was done using consensus foreground object template (CFOT) for moving object detection, and the later was achieved by the integration of heterogeneous features from different domains. In this work, the focus is on the challenging task of Web video annotation, in which most existing Web videos were captured under uncontrolled environments, with insufficient quality or limited tag information available[5].

The System has collected a complex, uncontrolled, and challenging Web video dataset from YouTube for the experiments carried out. The video data were captured by moving or shaky cameras and the moving object of interest were present in cluttered background. Significant scale and viewpoint variations of the objects were observed, and the resolution of a large portion of videos in dataset was low. The system considered six different moving object categories: Airplane, Ambulance, Car, Fire Engine, Helicopter, and Motorbike. Each object category had 25 to 30 video sequences, and each sequence has one moving foreground object present in it. Randomly select 10 from each class for training, and the remaining for testing



5. USES OF VIDEO ANNOTATIONS

1. Broadcasters generally annotate material that will be used later for either immediate "highlights" purposes, or for archiving
2. "Production Logging" in which producers will mark up an event live, to note shots to be edited into highlights packages and "Posterity Logging" in which librarians make detailed annotation of video tape for long term reuse, where depth and historical context is also noted.

6. OBJECTIVE

- The enhanced annotations resulting can be used directly in improving existing text-based search engines.
- Automated video annotation must explicitly address the issue of scalability, both in terms of the quantity of video and the expansiveness of the annotation vocabulary.
- 3.Research in video search and mining techniques is progressing rapidly yet most works are limited by small vocabularies and dataset sizes we can develop a prototype system to enhance web scale video search with automated Video annotation .Testing the model on a portion of YouTube can demonstrates the scalability and efficacy of our approach that will be used.

6.1 GAPS

It is well-known that analyzing and reasoning about video data are not easy due to

- the difficulty of approaching and simulating human being's perception by computers, and
- the lack of semantically meaningful annotations and technologies in understanding complex audio/visual data This is often referred to as the "semantic gap" in the multimedia retrieval community which limits the retrieval effectiveness.

7. CONCLUSION

In this paper, we have recalled some problems related with different techniques. retrieval. The state of the art of existing approaches in each major issue has been described with the focus on the following tasks: video structure analysis including shot boundary detection, key frame extraction and scene segmentation, extraction of features of static key frames, objects and motions, video data mining, video classification and annotation, video search including interface, similarity measure and relevance feedback, and video summarization and browsing.In this paper uses,gaps and objectives of video annotation is been given.

8. REFERENCES

[1] Amjad Altadmri and Amr Ahmed "Automatic Semantic Video Annotation in Wide Domain Videos Based on Similarity and Commonsense Knowledgebase's ,"in IEEE International Conference on Signal and Image Processing Applications ,2009.

[2] M. L. Shyu, Z. Xie, M. Chen, and S. C. Chen, "Video semantic event/concept detection using a subspace-based multimedia data mining framework,"

IEEE Transactions on Multimedia, vol. 10, no. 2, pp. 252–259, 2008.

[3] N. Haering, R. J. Qian, and M. I. Sezan, "A semantic event-detection approach and its application to detecting hunts in wildlife video," IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, no. 6, pp. 857–868, 2000

[4] ChengZhi Xu, HuiChuan Wu1, Bo Xiao1, Philip C-Y Sheu1, Shu-Ching Chen," A Hierarchical Video Annotation System".

[5] Shih-Wei Suna, Yu-Chiang Frank Wanga, Yao-Ling Hung "AUTOMATIC ANNOTATION OF WEB VIDEOS" a Institute of Information Science and Research Center for Information Technology InnovationAcademia Sinica, Taipei, Taiwan

[6] Yu-Gang Jiang,Jun Wang, Shih-Fu Changand Chong-Wah Ngo ,"Domain Adaptive Semantic Diffusion for Large ScaleContext-Based Video Annotation",by ieee ICCV sept29-oct 2,2009

[7] Vincenzo Lombardo & Rossana Damiano "An intelligent tool for narrative-based video annotation and editing", in 2010 International Conference on Complex, Intelligent and Software Intensive Systems

[8] Meng Wang, Xian-Sheng Hua Jinhui Tang and Richang Hong "Beyond Distance Measurement: Constructing Neighborhood Similarity for Video Annotation" IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 11, NO. 3, APRIL 2009 465.

[9] Jenny Yuen, Bryan Russell& Ce Liu1 Antonio Torralba, " LabelMe video: Building a Video Database with Human Annotations " in International Journal of Research and Reviews in Information Sciences

[10] Nida Aslam, Irfanullah, Kok-Keong Loo, Roohullah,"Growing Trend from Uni-to-Multimodal Video Indexing ", in nternational Journal of Digital Content Technology and its Applications Volume 3, Number 2, June 2009

[11] Khasfariyati Razikin, Dion Hoe-Lian Goh, Ee-Peng Lim, Aixin Sun, Yin-Leng Theng, Thi Nhu QuynhKim, Kalyani Chatterjea, Chew-Hung Chang, "Managing Media Rich Geo-spatial Annotations for a map-based Mobile Application using Clustering"in 2010 IEEE

[12] S. Dasiopoulou, E. Giannakidou, G. Litos, P. Malasioti, and I. Kompatsiaris,"A Survey of Semantic Image"

[13] Jinhui Tang, Xian-Sheng Hua, Guo-Jun Qi , Zhiwei Gu , Xiuqing Wu ,"Beyond Accuracy: Typically Ranking For Video Annotation " ICME 2007 IEEE

[14] Emily Moxley, Tao Mei, and B. S. Manjunath Video Annotation Through Search and Graph Reinforcement Mining,", ieee transactions on multimedia, vol. 12, no. 3, april 2010

[15] Matthew Butler, Tim Zapart, and Raymond Li," Video Annotation – Improving Assessment of Transient Educational Events", in Proceedings of the 2006 Informing Science and IT Education Joint Conference Salford, UK – June 25-28.

[16] T. Sumathi1, C.Lakshmi Devasena2, and M Hemalath "An Overview of Automated Image Annotation Approaches"International Journal of Research and Reviews Information Sciences Vol. 1, No. 1, March 2011

[17] Guo-Jun Qi , Xian-Sheng Hua, Yan Song , Jinhui Tang, Hong-Jiang Zhang,"Transductive Inference with Hierarchical Clustering for Video Annotation",IEEE ICME 2007

[18] Helmut Neuschmied " MPEG-7 Video Annotation Tool" ,inProceedings of I-MEDIA '07 and I-SEMANTICS '07Graz, Austria, September 5-7, 2007

[19] Mario Cordeiro, Cristina Ribeiro ,"Reuse Of Video Annotations Based On Low-level Descriptor Similarity", 2009 ieee.

[20] Pector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender Video Annotation Tools"decisions", Journal of Systems and Software, 2005, in press.