International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Textual Description for Annotating Videos

Shama P S[1], Pattan Prakash[2]

[1]PhD Research Scholar, Department of Computer Science and Engineering, PDA College of Engineering Kalaburagi, india

[2]Professor, Department of Computer Science and Engineering, PDA College of Engineering Kalaburagi, India

Abstract: ***In the Present days, large number of multimedia significant resources creates the browsing, retrieval and task. In this survey different types of methods are using, to faster the browsing of large video storage and provides sufficient contents to access. The Video annotation initially cuts the detected video sequence, where the succeeding frames are extracted to show an each shot or frames with a huge difference. The extraction of key frames that shows in an each image is compressing the few images by using standard technologies like HOG, fuzzy clustering and Genetic algorithm regarding the textual description of the scene. Video Annotation improves the sufficient searching and retrieving of the video components. They are inspired and are sometimes directly taken from image techniques. Thus several kinds of images are easily applied to video annotation.***

Keywords: HOG, Fuzzy clustering, GA, fuzzy clustering, video annotating

1. Introduction

Annotation defines as grouping the information of the text along with the particular area of the media. There is a lot of video which are efficiently available like medical, education, news, surveillance, disaster managing video are some different kinds. As the number of videos is increasing, then the complication in searching necessary information is also increased. Video annotation (Video annotation is the task together the graphic components on the screen.) is done by semantic retrieved database. Video contains visual, audio and textual information. Usually the people stored the database through the smart phone, cams, and upload the internet. So video annotation can helps the user to browse the large collection of videos from you tube. Basically the video are in text with particular feature that telling about the important events as well as regarding some video scene to support the respective story. Annotation improves the sufficient finding and retrieving of the video components are often inspired and sometimes directly taken from image techniques. And several kinds of images are also easily applied to video. There are various kind of methods can be used for annotations such as area, period and performance for indexing images which are easily applied to video. Annotation is deprecatory note of noted are attached to text, diagram, document, image or video. Annotation involved data to further segmented information to modify its access. Video technique is very essential technique for analysis and retrieval of video indexing. In previous work they have already done video Annotation as compared to video annotation. Video annotation is difficult because it need more database, memory and respective period.

The identification of video annotation in the text description before text identification is needed to input the image contains text. For instant recording cases, games, book covers. Moreover, the video frames or images regarding text are smaller than the amount of frames due to not text. And the past work on the content is a recognizable proof in picture or video that can be quickly characterized into region based, texture based and edge-based strategies. Region based techniques distinguishes character as the monochrome locales fulfilling certain heuristic limitations. Since the gray scale or shade of the content pixels in information picture are frequently not uniform, picture division or shading grouping pre-process must be performed to lessen the aggregate number of hues or gray scales in picture. The execution of area construct techniques and significantly depend with respect to the monochrome supposition of content characters and are accordingly not hearty to the complex foundation and compacted video.

Video annotation is the process to present the content of video in an effective manner based on two types of summary called static and dynamic video. In static video it select a set of salient images, these images are called key frames and are extracted from actual video to utilise the video content. In the dynamic video, skimming is a technology to divide the video content to several short clips.

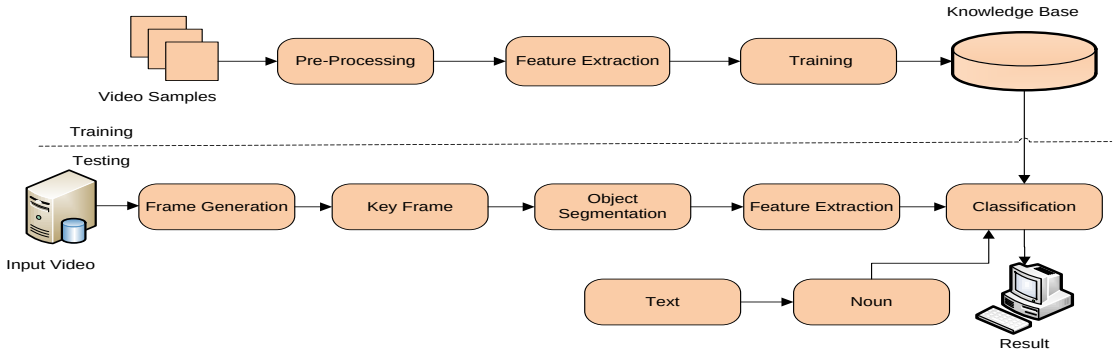

Figure 1: Architecture of Annotating Videos

The proposed system represents in figure1 is of two phases called testing phase and training phase. In the training phase initially the video samples are passed to pre-processing block, this block involved different operation like color conversion, image enhancement, and resizing. Then the feature from all the generated key frames is extracted; after training it will store in the knowledge base. In the testing phase generated key frame block is subjected to segmentation block then segmented images are passed to feature block for feature extracting; regarding images which are classified for comparison with the feature is already stored in knowledge base in the training phase. Matching the text description is the final output.

2. Literature Survey

In this section, we are representing the literature survey of texture description video annotation. Video annotation refers the annotation tool that lets you easily layer text, links, and spotlights to your videos. The key objective is to highlight the advantages and limitations of these methods. There are different techniques of segmentation and extraction. As the segmentation means in simple words its grouping or clustering of similar parameter pixels grouped together for forming a segment. Extraction defines that the Feature extraction involves reducing the amount of resources required to describe a large set of data. In this paper we represent the textual description for particular video.

3. Feature Extraction

Feature extraction defines developing the feature vector for given images or dataset using respective algorithm is done in feature. Khushboo Khurana et al. [1] the author worked on Image annotation is a well performed field of research that serves as precursor video annotation. Now a day's increasing the number of video and specific information present in them. There is a need to annotate the video. The video annotation refers to the extraction of the information about video, adding this information to the video which helps video browsing, detecting, and dividing etc. this survey performance view of entire video within a short period of time using HOG technique. HOG feature have been fame and effective selection for different kind of groups participating in Pascal visual objects classes risks. There are several features can be integrate to form a feature vector which is our input image that is machine learning systems to final annotation.

Roberto Di Salvo et al. [2] have worked on video object segmentation Method. This paper describes a simple and effective with interactive video object segmentation method to utilise the huge feature reduction data grouped from crowd by using HOG and SVM technique. According to this technique the video acquisition and storage methods have driven and developing the object detection for tracking and recognition. The main constrains to collect the necessary ground truth of human effort; annotation of a particular frame supposed for getting 30 minutes. It depends on video while taking time. This method performed in such a manner that it achieved a desire result for different complex scenarios, and also the clutters scenes and multiple objects. Here the paper is utilised for generating high quality video than annotation of images. The method can be easily generalised any video scenarios. The enhancement of this paper is for collecting the video object segmentation ground truth. And planning enriches the game.

Cheng Cai et al.[3] have worked dramatically by increasing of video resources. The human strength need to manually semantic the video annotation and the automatic annotation is an efficient and suitable solution. This paper has variety of resources schemes using random forest and is applied on video shot semantic annotation. Here the extracted using k-mean clustring technique. The results Indicate that the video shot annotation is based on random forest along with this it achieves good performance

4. Segmentation

Narasimha Murthy K N et al [4] have worked on different video methods. This paper approaches a method of extracting text from different type of video inputs. The important information within the text can be estimated by text storing the data and referred products. This system proposes uses contour based protocol like SUSAN algorithm for estimating the contour detection. And the system can travel or unfamiliar areas and refines the edge by fuzzy c mean clustering the non text portions are removing using morphological operation like dilation. This paper results in obtained the traditional algorithm used by the good evaluating results.

Lynn Wilcox et al [5] This survey provides description on a retrieval of segments of audio and video suitable for insertion in several media documents. This paper described a method for indexing and retrieving of multimedia data based on annotation and segmentation. annotation defines to grouping of text information along with certain time locations of the media segmentation is the dividing of continuous media into homogenous regions retrieval is activated over segments of the media Using the grouped annotation with the segments. This paper presents in two frameworks. Which are helps how to apply these techniques. First framework is approaches how extract frame videotaped usage of a new device are placed insertion a report on the analysis of the device. Second framework represents how audio bits recorded regard for use in authoring a summary of discussion.

Table 1: Survey on feature extraction Techniques

Title	Year	Algorithm	Performance
Study of Various Video Annotation Technique	2013	HOG	View of entire video within a short period of time.
Generating Reliable Video Annotations by Exploiting the Crowd	2016	HOG	The method can easily generalise to any videos/scenarios, thus providing a flexible and powerful mean to generate Video annotations. Accuracy 85.1%
Video Shot Annoting Using Random Forest	2015	K-Mean Clustering Method, HOG	HOG and SURF get the best annotation and the cumulative correct probability is 0.6 which show the effectiveness of scheme.
Framework For Group Based Image Retrieval And Video Annotation	2013	GIR, ASVA, SIFFT	Improve the annotation process in terms of all parameters except for speed. The speed of the proposed algorithm increased on average by 1.8 extract while combining the algorithms

Amjad Altadmri et al [6] have worked a novel framework for automatic Semantic Video Annotation. This framework finding the possible events happening in video clips, it forms the annotating based video search engine. The proposed system achieves to operate on unconstrained huge domain videos through the generic features. This scheme performed to bridge the ''semantic gap'', it defines the contrast between the low-level visual features and the humans ability, through the detecting same like visual events, after words utilising their text annotation to detect a general location and good description for this new video using knowledge bases The experimental results shows integrity between those two layers in order to find expressing annotations for the input video.

Dr. V. Radha et al [7] have worked on annotation techniques. This paper proposes two automatic video annotation techniques that are examined. The first method uses ontology to moderate the semantic gap while video retrieval and other activities together through the image retrieval using video files. This paper proposes uses GIR algorithm to generate same image group towards this pure Set of images, SIFT features are extracted and the usage is ASVA (Advanced Video Information System). This algorithm is performed to annotate the video in a semantic fashion. The video annotation algorithm activate in three steps: first step is to evaluate the video comparison using SIFT features and sentence utilisation. Second step is comparison of the same meaning annotations and then finally combining the sentence and increasing the each annotation using concept Net.

5. Classification

Vandana Gupta et al [8] have worked on detection and extraction of text information. This paper proposes a new approach for detection and extraction of text information from both scanned document pictures and scene images. This paper focused on text detection and extraction. Both are performed in four steps: pre-processing the pre-processing regards of binarization, noise removal of an image, image segmentation using particular features. And feature extracted by different technique these extraction is finally classified by the selected threshold value of conflict property.

Ms. Amanpreet Kaur et al [9] have work on data of indexing and retrieving. This paper describes the feature extraction from video data for indexing and retrieval.

In the present year the several media storage grows and the cost for storing multimedia data is less expensive because of that reason the more videos are available in video repositories. There is a huge value of multimedia data types; accessible bandwidth is huge with the demand of video retrieval systems. They are consisting the users shift from text based to content based video retrieval system.

Zeng Cheng et al [10] author worked on this paper to give detail explanation of a cross media annotation and traditional information annotation technology is commonly based on text description methods. In now a day's multimedia are increasing fast and rapidly developed cross media technology. Here the annotation between several procedures of media information is becoming possible. This technology is proved to be feasible and effectual in the arched type system called CMA (Cross Media Annotation).

Table 2: Survey on Segmentation

Title	Year	Algorithm	Performance
Robust Model for Text Extraction from Complex Video Inputs based on SUSAN Contour Detection and Fuzzy C Mean Clustering	2011	Fuzzy C Mean Clustering	Accuracy 97%
Text Driven Temporal Segmentation of Cricket	2006	Temporal Segmentation Video, Visual Feature based scene segmentation	Automatic annotation of multimedia with text
Automatic Semantic Video Annotation in Wide Domain Videos Based on Similarity and Commonsense Knowledgebase	2009	SIFT features	Videos to form the annotating part of video search engine.
Video Summarization Using Fuzzy Descriptors And A Temporal Segmentation	2013	Gabor filters, fuzzy histogram, heterogeneous features	Accuracy 91%

Young Chol Song et al [11] have worked on the video technology and data storage have made large scale video collections of complex activities readily available and conclude temporal segments in a video is to group the spatiotemporal segments in a video with its actual language description. This is using the different algorithms. And it is

having a different model that presents to be effective for unexamined alignment of objects in video with language.

Meng Wang, Xian et al [11] have worked on optimised multi graphs based OMG-SSL. This paper proposes a method named optimised multi graph-based semi-supervised learning (OMG-SSL), the goal is to expects all while handling these challenges in an infield together plan. We demonstrate that different critical considerations on video explanation, including various modalities, numerous separation capacities, and fleeting consistency, all compare to various connections among video units, and consequently they can be spoken to various diagrams.

Dr.M.B.Chandak et al [12] have works on video internet offline world. This paper describes led to flooding of video consist of internet and offline world. For each huge quantity of content the access of similar video in moderate period is a matter of concern. So the system will helps in content based access of video arises and it is mandatory. This paper introduced a novel and are analysed. The complete video frames are extracted by videos are analysed, including whole videos. Here only one key frame is utilised to identify the present objects. The object detectors are trained for finding the object. The detected objects are then added in the annotation file. This annotation based ontology which eases the semantic retrieval of videos. Its performance significantly accelerates the retrieval systems.

Anastasis Kounoudes et al [13] have worked on a multi level; video annotation tool based on XML directories. It consists of 3 different panels and provides a friendly user interface that seems very powerful user profiles. This paper proposes video technology have outcomes in a digital video accessible based on the web or in private natural resources. That created various video data levels of conditions that are movies and video annotation tool.

6. Conclusion

This survey concluded that the Video annotation process is complicated because it requires a large database, memory & time for processing. In this paper we are doing a survey on different video annotation techniques. Different technique works in different environment with different level of accuracy and timing constraints. There is a need to improve the existing methods in terms of accuracy & time. Further it needs to handle a large amount of database. Training for labelling can be done efficiently by using the fuzzy logic. According to our survey we have improved to provide a better performance using different techniques like HOG, Fuzzy c mean, SUSAN.

Table 3: Text Based Algorithm

Title	Year	Algorithm	Performance
Robust Model for Text Extraction from Complex Video Inputs based on SUSAN Contour Detection and Fuzzy C Mean Clustering	2011	SUSAN, Morphological Operations	The achieved result is much in Contrast compared to majority of prior significant research in this area. Accuracy 97%
Video Annotation Methodology Based on Ontology for Transportation Domain	2013	SIFT features, SVM	Performance of the system can be made error free by increasing the positive and negative images for training.
Visualization-Based Active Learning for Video Annotation	2016	iso-contour-based scatter plot, semi supervised metric learning method, sampling method	Can effectively enhance sample selection in active learning.

References

[1] Khushboo Khurana and M. B. Chandak "Study of Various Video Annotation Techniques", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, No. 1, 2013.
[2] Roberto Di Salvo, Concetto Spampinato and Daniela Giordano, "Generating Reliable Video Annotations by Exploiting the Crowd", IEEE, pp. 7-10, 2016
[3] Cheng Cai, "Video Shots Annotation using Random Forest", IEEE, pp. 1 - 4, 2015.
[4] Narasimha Murthy K N and Dr. Y S Kumaraswamy," Robust Model for Text Extraction from Complex Video Inputs Based on SUSAN Contour Detection and Fuzzy C Means Clustering", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No. 3, 2011.
[5] Amjad Altadmri and Amr Ahmed, "A framework for automatic semantic video annotation Utilizing similarity and commonsense knowledge bases", Springer, pp. 1167 – 1191, 2013.
[6] Arnon Amira, "A multi-modal system for the retrieval of semantic video events", Vol. 96, Issue 2, pp. 216 – 236, 2004.
[7] Dr. V. Radha and K.Tamil Selvi "Framework for Group Based Image Retrieval and Video Annotation", Journal of Global Research in Computer Science, Vol. 4, No. 12, 2013.
[8] Vandana Gupta, Kanchan Singh, "A Novel Approach for Detection and Extraction of Textual Information from Scanned Document Images and Scene Images", International Journal of Advanced
[9] Ms. Amanpreet Kaur, "Feature Extraction from Video Data for Indexing and Retrieval", International Research Journal of Engineering and Technology (IRJET), Vol. 02, Issue 01, 2015.
[10] Zeng Cheng; Liu Yaqin, "Cross-Media Annotation Based on Semantic Network", IEEE, pp. 15 – 20, 2008.
[11] Young Chol Song, "Unsupervised Alignment of Actions in Video with Text Descriptions", IEEE, Vol. 13, Issue 2, pp. 206 - 215, 2011

[12] Dr.M.B. Chandak, Domain Khushboo Khurana, "Video Annotation Methodology Based on Ontology for Transportation", IEEE, pp. 227 – 234, 2008
[13] Anastasis Kounoudes and Nicolas Tsapatsoulis ,Kyprianos Str, "A multi-level Video Annotation Tool based on XML-dictionaries", pp. 455 - 460, 2008.