Bindhu. N, Bala Murugan. C / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 3, Issue 2, March -April 2013, pp.416-419

An Adaptive Novel Approach for Detection of Text and Caption in Videos

Bindhu. N[1], Bala Murugan. C[2]

[1]PG Scholar, Department of Computer Science and Engineering, Vel Tech Dr.RR & Dr.SR Technical University,
#42 Avadi-Vel Tech Road, Avadi, Chennai-62

[2]Asst. Professor, Department of Computer Science and Engineering, Vel Tech Dr.RR & Dr.SR Technical University, #42 Avadi-Vel Tech Road, Avadi, Chennai-62

ABSTRACT

The video image spitted into number of frames, each frame maintains the text. Then the Image is converted into Gray Scale to avoid the text color variation. A single value is corresponding to gray value and detecting the edge. Detecting the edge process is the boundary between two regions with relatively distinct gray-level properties. One is the horizontal direction of the image. Another is the vertical direction of the image. The features to describe text regions are area, saturation, orientation, aspect ratio and position. Then convert into binary image. The comer detected we can using the new techniques for neural network. To apply the Harris corner algorithm to implementing the Feature description by selecting the text area. The Orientation is defined as the angle (ranging from to -90 degree to 90 degree) between the x-axis and the major axis of the ellipse that has the same second-moments as the region. The input image is given equalization based extracting the text. After finding out the text, it is checked with the database. If tracked image is matched with the existing database. From that we can identify the text of the particular image. Next input image is spited into frames to extract text for sequences, and compare to previous text. Finally convert into voice.

Key words: Edge detection, text, video retrieval

Introduction

Efficient Management of large collection of data in multimedia like Video, Satellite Imagery, are important issues that needs great attention and research work, as the amount of data collected by Digital Libraries as Video, Satellite Imagery, tend to grow exponentially, in order to reuse this data effectively, we need to organize this data such that it helps provide an effective access. Archival of Video clips to enable reuse is a time consuming, tedious and inefficient process [4].

In Video images, index corresponds to an event occurred in video image, On the other hand [2] table of contents corresponds to a hierarchical structure of topic.

A straightforward video analysis is shot boundary detection. Boundaries are typically found by computing an image-based distance adjacent (or regularly) frames of the video and noting when this distance exceeds a certain threshold. The distance between frames can be based on statistical properties of pixels, histogram differences [8], or motion detection [11]

Other forms of scientific data including Time Series, Medical Images(such as MRI, CT, and PET) and Seismic Data, have arching and search requirements similar to that for Satellite imagery [10].

Content Based Retrieval can be used more effective by organizing the data. Cataloging is performed to help this task achieved. It is performed by relying on minimal human input [4]. In order to simplify and speed up this process, human input in Video content can be used as index for Video Browsing, this can include visual material that helps browsing, such as Key Frames, Story Board, Audio Clips [4], while Texture, Color Histogram can help in image retrieval.

Design Challenges and search Strategies

Content Based Retrieval of Images and Video Databases involve comparing a query object with the objects stored in the data repository. The search is usually based on similarity rather than on exact match, and the retrieved results are then ranked according to a similarity index [6, 13].

The nature of video content as in a indoor video with a single speaker user verses an outdoor sports video impacts the retrieval [4].

Video segmentation is a commonly used as a first step to automatically analyze content [4]. Shot Boundary Detection Algorithms [1] are used to partition the video into elemental units called Shots. “Shot” is a fundamental unit of processing (analyzing, indexing representation) of video. Shots are annotated with text and Keywords one or more frames are extracted and treated as till images to apply visual search technology [4].

Presentation of video. Shots are annotated with text and Keywords one or more frames are extracted and treated as till images to apply visual search technology [4].

Data Organization and Content Based Retrieval

Significant images and words form a paragraph are extracted to produce a short summary of the Video. The cataloger supports time based and content based key frame simplifying text based annotation and metadata.

Objects extracted at image ingestion time can be indexed much more efficiently. However, it is usually very difficult to anticipate all the types of objects in which a user might be, and thus systems allowing only search based on pre-extracted objects are severely limiting. On the other hand, recognizing objects entirely at query time will limit the scalability of a system, due to the high expense of such computing [10].

These problems were alleviated by object-oriented framework which allows flexible composition of queries relying on both types of objects. Within this frameworks, objects can be specified at multiple abstraction levels namely, the raw data level, the feature level, and the semantic level as shown in figure:1 [10].

Raw Data: At the lowest abstraction level, objects are simply aggregations of raw pixels from the image. Comparison between objects or regions is done pixel-by-pixel Comparison at the pixel level is very specific, and is therefore only used when a relatively precise match is required [10].

Feature: The next higher abstraction level for representing images is the feature level. An image feature is a distinguishing primitive characteristics or attribute of an image. Some features such as luminance, shape descriptor, and gray scale texture are natural since they correspond to visual appearance of an image.

Semantic: This is the highest abstraction level at which a content-based search can be performed. An object-oriented definition of a semantic object also involves prescribing a set of pertinent features or pixels as well as a method (such as a classification algorithm with the appropriate training data). For satellite images, examples of semantic objects include the type of land cover for a specific area such as water, forest, or urban. For medical images, examples include the type of organ such as liver, stomach, or colon. A semantic network can be constructed which groups similar semantic terms into categories. For example, pine and maple are grouped into trees, rose and sunflower are grouped into flowers, corn and wheat are grouped into crops, etc. The purpose of constructing such a semantic network is to allow the generalization of retrieval at the semantic level [10].

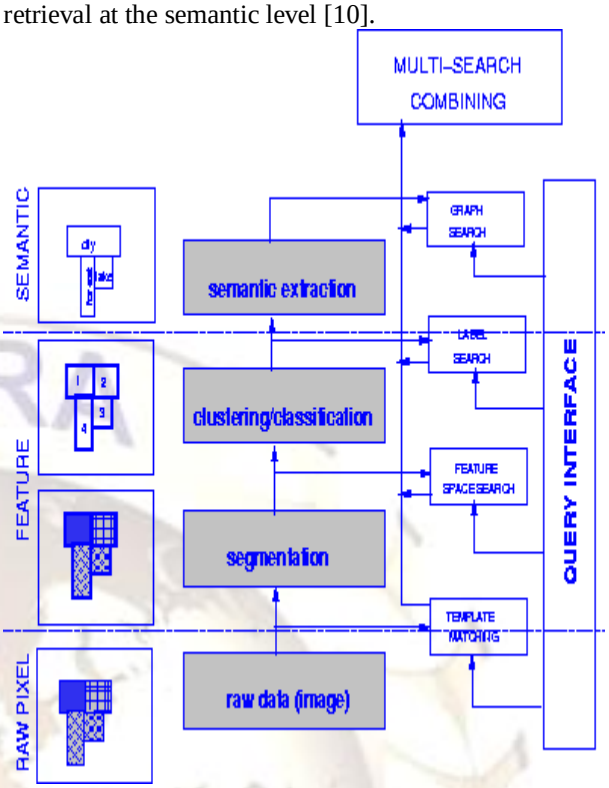

Figure 1: Abstraction levels of an image.

Data Organization and Retrieval

The techniques to syntactically organize the video data are summarized as follows

1. Cut detection: The difference between consecutive frames is computed based on a histogram method and if it is greater than some threshold, it is regarded as a cut point.
   The shot is extracted as the section ended at the cut points [2].

2. Extraction of common scene: Common scenes included in long video data give information about the repetition [2].

3. Extraction of camera works: Camera work indicates the intention of a cameraman or director. Therefore it is possible to segment the video data in terms of homogeneous work within a shot. The camera work can be extracted by a projection method or Affined parameters[2].

Let us now have a look at the scene change detection in video and analyze using the "Cut measure" and "dissolve measure" which were discussed by Changick Kim in "Automated Shot Change Detection" CIDIL, Digital Video Library Project.

"Cut": In editing, an immediate switch from one image to anther, without the aid of transitions such as dissolves or wipe.

"Dissolve": A video transition in which the existing image is partially or totally replaced by superimposing another image, one image fades in as the other fades out.

Here are the Frames taken from a Video Clip to study the Cut and Dissolve measures.

Frame #1

Frame # 35, Frame #45,

Frame # 52 Frame # 60 faden)

Cut Frame : # 90 and #91

Cut frame: Frame # 928 and # 929.

Conclusion

Content Based video retrieval and Scene change detection continue to pose challenging problems, approaches that integrate different indexing techniques like cataloging, Organization of data that produce automatic table of contents, indices and defining better ways to analyze the cut and dissolve measures, Analyzing images at various levels of abstractions and finding methods which improve recognizing the objects entirely at query time remain as promising avenues to be explored. Summarizing video is a challenging user interface problem. Thus proper organization and making use of metadata efficiently can improve the performance of Content Based Retrieval, which can be Further applied to Educational Technical and Medical fields.

References:

[1] Arman, F., Depommier, R., Hsu, A., and Chiu, M.Y. *Content based browsing of video sequences,* in *ACM Multimedia 94* pp.97-103, Aug. 1994

[2] Ariki Yauso, *Organization and Retrieval of Continues media,* ACM *Multimedia 2000,* NOV 2000.

[3] C. Kim, *Automated Shot change Detection*, CIDIL, Digital Video Library Project.

[4] D. Ponceleon, S. Srinivasan, A. Amir D.Petkovic, *Key to Effective Video Retrieval: Effective Cataloging and Browsing,* ACM *Multimedia 98,* Sep 1998.

[5] E. Deardroff, T.D.C Little, J.D, Marshall and Venketasj, *Video Scene Decomposition with the Motion Picture Paser,* Proc, SPIE, vol. 2187, Digital Video Compression on Personal Computers: Algorithms and Technologies, Feb 1994, pp 44-55.

[6] F. Arman, A. Hsu, M. Y. Chiu, *Feature Management of large video Database*, Proc. SPIE vol 1908, 1993, pp 271-298.

[7] J. M. Gauch, S. Gauch, *Scene Change Detection in Video Images,* Design Lab, Digital Video Library System.

[8] J. Foote, J. Boreczky, A. Girensohn, and Lynn Wilcox, *An Intelligent Media Browser using Automatic Multimodal Analysis,* ACM *Multimedia 98,* Sep 1998.

[9] L. Rutledge, Lhardman, J.van Ossenburggen, D.C.A Bulterman, CWI, Amsterdam, Vrije, University, Amsterdam, Netherlands *Structural Distinction between Hypermedia Storage and Presentation* in ACM Multimedia 98, Sep 1998.

[10] L. D. Bergman, V. Castelli, C. Li, *Progressive Content Based Retrieval from Satellite Image Archives,* D-Lib Magazine, Oct 97.

[11] Shahrary, B., *Scene change Detection and Content Based Sampling Video Sequences*, in Digital Video Compression: Algorithms and Technologies, Rodriguez, Safranek, Delp, Eds., Proc APIE 2419, Feb, pp 2-13.

[12] S.W. Smoliar and H. Zhang , *Content based Video indexing and Retrieval, IEEE Multimedia,* vol 1, no. 2, Summer 1994, pp 62-72.