ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Model and Extracting Videos Using Intelligent Methods

Satish D Mali
T.P.C.T.'s College of Engineering,
Dr.B.A.M.University, Aurangabad,
Solapur-Osmanabad Road, Osmanabad,
Maharashtra- 413501

Prof. Pravin P. Kalyankar
T.P.C.T.'s College of Engineering,
Dr.B.A.M.University, Aurangabad,
Solapur-Osmanabad Road, Osmanabad,
Maharashtra- 413501

ABSTRACT

Increase in the use of video-based applications has revealed the need for extracting the content in videos. Raw data and low-level features alone are not sufficient to fulfill the user 's needs; that is, a deeper understanding of the content at the semantic level is required. Currently, manual techniques, which are inefficient, subjective and costly in time and limit the querying capabilities, are being used to bridge the gap between low-level representative features and high-level semantic content. Here, we propose a semantic content extraction system that allows the user to query and retrieve objects, events, and concepts that are extracted automatically. We introduce an ontology-based fuzzy video semantic content model that uses spatial/temporal relations in event and concept definitions. This metaontology definition provides a wide-domain applicable rule construction standard that allows the user to construct an ontology for a given domain. In addition to domain ontologies, we use additional rule definitions (without using ontology) to lower spatial relation computation cost and to be able to define some complex situations more effectively.

The proposed framework has been fully implemented and tested on three different domains. We have obtained satisfactory precision and recall rates for object, event and concept extraction.

INTRODUCTION:

THE rapid increase in the available amount of video data has caused an urgent need to develop intelligent methods to model and extract the video content. Typical applications in which modeling and extracting video content are crucial include surveillance, video-on-demand systems, intrusion detection, border monitoring, sport events, criminal investigation systems, and many others. The ultimate goal is to enable users to retrieve some desired content from massive amounts of video data in an efficientand semantically meaningful manner.

Another key issue in semantic content extraction is therepresentation of the semantic content. Many researchers have studied this from different aspects. A simple representation could relate the events with their low-level features (shape, color, etc.) using shots from videos, without any spatial or temporal relations. However, an effective use of spatiotemporal relations is crucial to achieve reliable recognition of events. Employing domain ontologies facilitate use of applicable relations on a domain. There are no studies using both spatial relations between objects, and temporal relations between events together in an ontology-based model to support automatic semantic content extraction. Studies such as BilVideo extended-AVIS, multiView and classView propose methods using spatial/temporal relations but do not have ontology-based models for semantic content representation. Bai et al.present a semantic content analysis framework based on a domain ontology that is used to define semantic events with a temporal description logic where event extraction is done manually and event descriptions only use temporal information. Nevatia and Natarajanpropose an ontology model using spatiotemporal relations to extract complex events where the extraction process is manual. In , each linguistic concept in the domain ontology is associated with a corresponding visual concept with only temporal relations for soccer videos. Nevatia et al.define an event ontology that allows natural representation of complex spatiotemporal events in terms of simpler subevents. A Video Event Recognition Language (VERL) that allows users to define the events without interacting with the lowlevelprocessing is defined. VERL is intended to be a language for representing events for the purpose of designing ontology of the domain, and, Video EventMarkup Language (VEML) is used to manually annotate VERL events in videos. The lack of low-level processing and using manual annotation are the drawbacks of this study. Akdemir et al.present a systematic approach to address the problem of designing ontologies for visual activity recognition. The general ontology design principles are adapted to the specific domain of human activity ontologies using spatial/temporal relations between contextual entities. However, most of the contextual entitieswhich are utilized as critical entities in spatial andtemporal relations must be manually provided for activity recognition. Yildirimprovide a detailed survey of the existing approaches for semantic content representation and extraction.

Considering the above-mentioned needs for content based retrieval and the related studies in the literature, methodologies are required for automatic semantic content extraction applicable in wide-domain videos.

In this study, a new Automatic Semantic ContentExtraction Framework (ASCEF) for videos is proposed for bridging the gap between low-level representative features and high-level semantic content in terms of object, event,concept, spatial and temporal relation extraction. In order to address the modeling need for objects, events and concepts during the extraction process, a wide-domain applicable ontology-based fuzzy VIdeo Semantic Content Model (VISCOM) that uses objects and spatial/temporal relations in event and concept definitions is developed. VISCOM is a metaontology for domain ontologies and provides a domain-independent rule construction standard. It is also possible to give additional rule definitions (without using ontology) for defining some special situations and for speeding up the extraction process. ASCEF performs the extraction process by using these metaontology- based and additional rule definitions, making ASCEF wide-domain applicable

IMAGE FILE FORMATS:

Image file formats are standardized means of organizing and storing images. This entry is about digital image formats used to store photographic and other images. Image files are composed of either pixel or vector (geometric) data that are rasterized to pixels when displayed (with few exceptions) in a vector graphic display. Including proprietary types, there are hundreds of image file types. The PNG, JPEG, and GIF formats are most often used to display images on the Internet.

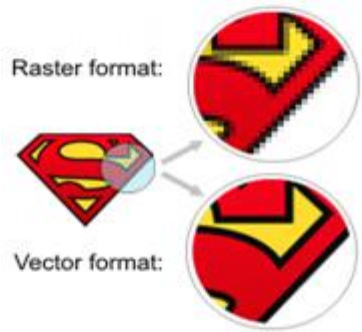

Fig 1: Raster And Vector Format

In addition to straight image formats, Metafile formats are portable formats which can include both raster and vector information.The metafile format is an intermediate format. Most Windows applications open metafiles and then save them in their own native format.

RASTER FORMATS:

These formats store images as bitmaps (also known as pixmaps).

JPEG/JFIF:

JPEG (Joint Photographic Experts Group) is a compression method. JPEG compressed images are

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

usually stored in the JFIF (JPEG File Interchange Format) file format. JPEG compression is lossy compression. Nearly every digital camera can save images in the JPEG/JFIF format, which supports 8 bits per color (red, green, blue) for a 24-bit total, producing relatively small files. Photographic images may be better stored in a lossless non-JPEG format if they will be re-edited, or if small "artifacts" are unacceptable.

The JPEG/JFIF format also is used as the image compression algorithm in many Adobe PDF files.

EXIF:

The EXIF (Exchangeable image file format) format is a file standard similar to the JFIF format with TIFF extensions. It is incorporated in the JPEG writing software used in most cameras. Its purpose is to record and to standardize the exchange of images with image metadata between digital cameras and editing and viewing software. The metadata are recorded for individual images and include such things as camera settings, time and date, shutter speed, exposure, image size, compression, name of camera, color information, etc. When images are viewed or edited by image editing software, all of this image information can be displayed.

TIFF:

The TIFF (Tagged Image File Format) format is a flexible format that normally saves 8 bits or 16 bits per color (red, green, blue) for 24-bit and 48-bit totals, respectively, usually using either the TIFF or TIF filename extension. TIFFs are lossy and lossless. Some offer relatively good lossless compression for bi-level (black & white) images. Some digital cameras can save in TIFF format, using the LZW compression algorithm for lossless storage. TIFF image format is not widely supported by web browsers. TIFF remains widely accepted as a photograph file standard in the printing business. TIFF can handle device-specific color spaces, such as the CMYK defined by a particular set of printing press inks.

PNG:

The PNG (Portable Network Graphics) file format was created as the free, open-source successor to the GIF.

The PNG file format supports true color (16 million colors) while the GIF supports only 256 colors. The PNG file excels when the image has large, uniformly colored areas. The lossless PNG format is best suited for editing pictures, and the lossy formats, like JPG, are best for the final distribution of photographic images, because JPG files are smaller than PNG files.

PNG, an extensible file format for the lossless, portable, well-compressed storage of raster images. PNG provides a patent-free replacement for GIF and can also replace many common uses of TIFF. Indexed-color, grayscale, and true color images are supported, plus an optional alpha channel. PNG is designed to work well in online viewing applications, such as the World Wide Web. PNG is robust, providing both full file integrity checking and simple detection of common transmission errors.

GIF:

GIF (Graphics Interchange Format) is limited to an 8-bit palette, or 256 colors. This makes the GIF format suitable for storing graphics with relatively few colors such as simple diagrams, shapes, logos and cartoon style images. The GIF format supports animation and is still widely used to provide image animation effects.

It also uses a lossless compression that is more effective when large areas have a single color, and ineffective for detailed images or dithered images.

BMP:

The BMP file format (Windows bitmap) handles graphics files within the Microsoft Windows OS.

Typically, BMP files are uncompressed, hence they are large. The advantage is their simplicity and wide acceptance in Windows programs.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

FUNDAMENTAL STEPS IN DIGITAL IMAGE PROCESSING:

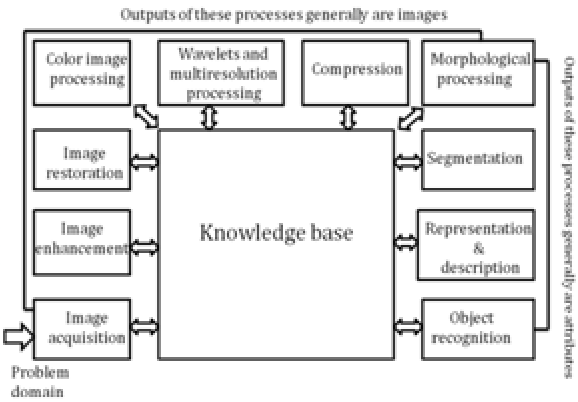

Image Acquisition:

Image Acquisition is to acquire a digital image. To do so requires an image sensor and the capability to digitize the signal produced by the sensor. The sensor could be monochrome or color TV camera that produces an entire image of the problem domain every 1/30 sec. the image sensor could also be line scan camera that produces a single image line at a time. In this case, the objects motion past the line.

Fig 2: color TV camera

Scanner produces a two-dimensional image. If the output of the camera or other imaging sensor is not in digital form, an analog to digital converter digitizes it. The nature of the sensor and the image it produces are determined by the application.

Fig 3: Scanner produce

CONCLUSION

The primary aim of this research is to develop a framework for an automatic semantic content extraction system for videos which can be utilized in various areas, such as surveillance, sport events, and news video applications. The novel idea here is to utilize domain ontologies generated with a domain-independent ontology-based semantic content met ontology model and a set of special rule definitions. Automatic Semantic Content Extraction Framework contributes in several ways to semantic video modeling and semantic content extraction research areas. First of all, the semantic content extraction process is done automatically. In addition, a generic ontology-based semantic metaontologymodel for videos (VISCOM) is proposed. Moreover, the semantic content representation capability and extractionsuccess are improved by adding fuzziness in class, relation, and rule definitions. An automatic Genetic Algorithm-based object extraction method is integrated to the proposed system to capture semantic content. In every component of the framework, ontology-based modeling and extraction capabilities are used. The test results clearly show the success of the developed system. As a further study, one can improve the model and the extraction capabilities of the framework for spatial relation extraction by considering the viewing angle of camera and the motions in the depth dimension.

REFERENCES

[1] A.D. Bagdanov, M. Bertini, A. Del Bimbo, C. Torniai, and G. Serra, “Semantic Annotation and Retrieval of Video Events Using Multimedia Ontologies,” Proc. IEEE Int’l Conf. Semantic Computing (ICSC), Sept. 2007.

[2] R. Nevatia, J. Hobbs, and B. Bolles, “An Ontology for Video Event Representation,” Proc. Conf. Computer Vision and Pattern Recognition Workshop, p. 119, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1384914, 2004.

[3] U. Akdemir, P.K. Turaga, and R. Chellappa, "An Ontology Based Approach for Activity Recognition from Video," Proc. ACM Int'l Conf. Multimedia, A. El-Saddik, S. Vuong, C. Griwodz, A.D. Bimbo, K.S. Candan, and A. Jaimes, eds., pp. 709-712, http://dblp.unitrier.de/db/conf/mm/mm2008.html#AkdemirTC08, 2008.

[4] Y. Yildirim, "Automatic Semantic Content Extraction in Video Using a Spatio-Temporal Ontology Model," PhD dissertation, Computer Eng. Dept., METU, Turkey, 2009.

[5] T. Yilmaz, "Object Extraction from Images/Videos Using a Genetic Algorithm Based Approach," master's thesis, Computer Eng. Dept., METU, Turkey, 2008.

[6] Y. Yildirim and A. Yazici, "Ontology-Supported Video Modeling and Retrieval," Proc. Fourth Int'l Conf. Adaptive Multimedia Retrieval: User, Context, and Feedback (AMR), pp. 28- 41, 2006.

[7] Y. Yildirim, T. Yilmaz, and A. Yazici, "Ontology-Supported Object and Event Extraction with a Genetic Algorithms Approach for Object Classification," Proc. Sixth ACM Int'l Conf. Image and Video Retrieval (CIVR '07), pp. 202-209, 2007.

[8] V. Mezaris, I. Kompatsiaris, N.V. Boulgouris, and M.G. Strintzis, "Real-Time Compressed-Domain Spatiotemporal Segmentation and Ontologies for Video Indexing and Retrieval," IEEE Trans. Circuits Systems Video Technology, vol. 14, no. 5, pp. 606-621, May 2004.

[9] D. Song, H.T. Liu, M. Cho, H. Kim, and P. Kim, "Domain Knowledge Ontology Building for Semantic Video Event Description," Proc. Int'l Conf. Image and Video Retrieval (CIVR), pp. 267-275, 2005.

[10] W. Chen and D.S. Warren, "C-logic of Complex Objects," PODS '89: Proc. Eighth ACM SIGACT-SIGMOD-SIGART Symp. Principles of Database Systems, pp. 369-378, 1989.

[11] S. Hongeng, R. Nevatia, and F. Bre´mond, "Video-Based Event Recognition: Activity Representation and Probabilistic Recognition Methods," Computer Vision and Image Understanding, vol. 96, no. 2, pp. 129-162, 2004.

[12] A. Hakeem and M. Shah, "Multiple Agent Event Detection and Representation in Videos," Proc. 20th Nat'l Conf. Artificial Intelligence (AAAI), pp. 89-94, 2005.

[13] J.F. Allen, "Maintaining Knowledge about Temporal Intervals," Comm. ACM, vol. 26, no. 11, pp. 832-843, 1983.

[14] M.J. Egenhofer and J.R. Herring, "A Mathematical Framework for the Definition of Topological Relationships," Proc. Fourth Int'l Symp.Spatial Data Handling, pp. 803-813, 1990.

[15] M. Vazirgiannis, "Uncertainty Handling in Spatial Relationships," SAC '00: Proc. ACM Symp. Applied Computing, pp.494-500,2000.