International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 4, April 2014)

A Study on Digital Image Processing Techniques for Content Extraction from Video

PREMILA SHANKER RAO
Lecturer, Department of Computer Science
V.G.Womens.College, Gulbarga
toranrajlaxmi@rediffmail.com

Sharanabasappa
Lecturer, Govt.First Grade for Womens Degree College
Jewargi Colony Gulbarga,
smaisalgi.1983@rediffmail.com

***Abstract*- In this paper extracting the contents from the video is described by using the various Digital Image Processing Techniques such as some image detection and image segmentation techniques because the increase in the diversity and availability of electronic information led to additional processing requirements, in order to retrieve relevant and useful data: the accessibility problem. This problem is even more relevant for audiovisual information, where huge amounts of data have to be searched, indexed and processed. Most of the solutions for this type of problems point towards a common need: to extract relevant information features for a given content domain. A process which underlies two difficult tasks: deciding what is relevant and extracting it. In fact, while content extraction techniques are reasonably developed for text, video data still is essentially opaque. Despite its obvious advantages as a communication medium, the lack of suitable processing and communication supporting platforms has delayed its introduction in a generalized way. This situation is changing and new video based applications are being developed.**

Keywords: **Digital Image Processing, Image Detection, Image Segmentation, video data.**

I. INTRODUCTION

The image or video is stored only as a set of pixels with RGB values in computer. The computer knows nothing about the meaning of these pixel values. The content of an image is quite clear for a person. However, it is not so easy for a computer. For example, it is a piece of cake to recognize yourself in an image or video, even in a crowd. But this is extremely difficult for computer. The preprocessing is to help the computer to understand the content of image or video. What is the so-called content of image or video? Here content means features of image or video or their objects such as color, texture, resolution, and motion. Object can be viewed as a meaningful component in an image or video picture. For example, a moving car, a flying bird, a person are all objects. There are a lot of techniques for image and video processing. This chapter starts with an introduction to general image processing techniques and then talks about video processing techniques. The reason we want to introduce image processing first is that image processing techniques can be used on video if we treat each picture of a video as a still image.

II. BACKGROUND

A few years ago, the problems of representation and retrieval of visual media were confined to specialized image databases (geographical, medical, pilot experiments in computerized slide libraries), in the professional applications of the audiovisual industries (production, broadcasting and archives), and in computerized training or education. The present development of multimedia technology and information highways has put content processing of visual media at the core of key application domains: digital and interactive video, large distributed digital libraries, multimedia publishing. Though the most important investments have been targeted at the information infrastructure (networks, servers, coding and compression, delivery models, multimedia systems architecture), a growing number of researchers have realized that content processing will be a key asset in putting together successful applications. The need for content processing techniques has been made evident from a variety of angles, ranging from achieving better quality in compression, allowing user choice of programs in video-on-demand, achieving better productivity in video production, providing access to large still image databases or integrating still images and video in multimedia publishing and cooperative work.

Content of image includes resolution, color, intensity and texture. Image resolution is just the size of image in term of display pixels. Color is represented using RGB color model in computer. For each pixel on the screen, there are three bytes (R,G,B color component) to represent its color. Each color component is in the range of 0 to 255. Intensity is the gray level information of pixels represented by one byte. The intensity value is in the range of 0 to 255. Texture characterizes local variations of image color or intensity. Although texture-based methods have been widely used in computer vision and graphics, there is no single commonly accepted definition of texture. Each texture analysis method defines texture according to its own model. We consider texture as a symbol of local color or intensity variation. Image regions that are detected to have a similar texture have similar pattern of local variation of color or intensity.

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 4, April 2014)

III. BASIS OF VIDEO PROCESSING

A. *Content of Digital Video*

Generally speaking, there is much similarity between digital video and image. Each picture of video can be treated as a still image. All the techniques applicable to images can also be applied to video pictures. However, there are still different. The most significant difference is that video has temporal information and uses motion estimation for compression. Video is a meaningful group of pictures that tells a story or something else. Video pictures can be grouped as a shot. A video shot is a set of pictures taken in one camera break. Within each shot, there can be one or more key pictures. Key picture is a representative of the content of a video shot. For a long video shot, there may be multiple key pictures. Usually video processing segments video into separate shots, selects key pictures from these shots, and then generates features of these key pictures. The features (color, texture, object) of key pictures are searched in video query.

Video processing includes shot detection, key picture selection, feature generation, and object extraction.

B. Shot Detection

Shot detection is a process to detect camera shots. A camera shot consists of one or more pictures taken in one camera break. The general approach to shot detection has been the definition of a difference metric. If the differences between two pictures are above the metric, then there is a shot between them. An algorithm can be proposed for this. This algorithm uses binary search to detect shot which makes it very fast and achieve good performance as well.

C. Key Picture Selection

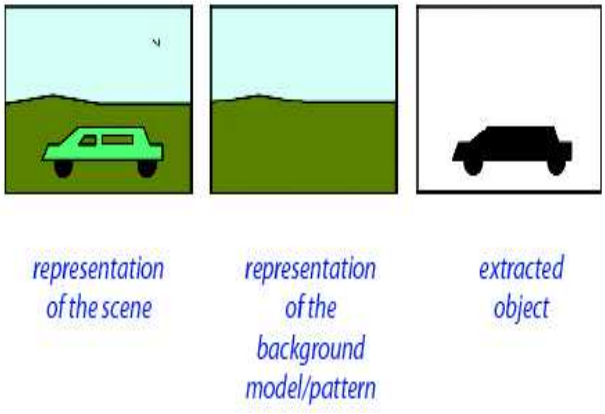
After shot detection, each shot is represented by at least one key picture. The choice of key picture could be as simple as a particular picture in the shot: the first, the last, or the middle. However, in situations such as long shot, no single picture can represent the content of the entire shot. QBIC (Query by Image Content) uses a synthesized key picture created by seamlessly mosaicking all the pictures in a given shot using the computed motion transformation of the dominant background. This picture is an authentic depiction of all background captured in the whole shot. In CBIR (Content based Image Retrieval) system, key picture selection is a simple process that usually chooses the first and last pictures of a shot as key pictures.

D. Feature Generation

After key picture selection, features of key pictures such as color, texture, intensity are stored as indexes of the video shot. Users can perform traditional search by using keyword querying and content-based query by specifying a color, intensity, or texture pattern. Only the generated features will be searched against and the retrieval can be in real time.

E. Object Extraction

During the process of shot detection and key picture selection, the objects in the video are also extracted using image segmentation techniques or motion information. Segmentation-based techniques are mainly based on image segmentation. And objects are recognized and tracked by segmentation projection. Motion-based techniques make use of motion vectors to distinguish objects from background and keep track of their motion. It is a very difficult problem. And the new standard being developed will talk about how to get objects in the video and encode them separately into different layers. Hopefully this process is not manual and it is also unrealistic to expect it to be full automatical.



IV. COMMON IMAGE PROCESSING TECHNIQUES

A. Dithering

Dithering is a process of using a pattern of solid dots to simulate shades of gray. Different shapes and patterns of dots have been employed in this process, but the effect is the same. When viewed from a great enough distance that the dots are not discernible, the pattern appears as a solid shade of gray.

B. Erosion

Erosion is the process of eliminating all the boundary points from an object, leaving the object smaller in area by one pixel all around its perimeter. If it narrows to less than three pixels thick at any point, it will become disconnected (into two objects) at that point. It is useful for removing from a segmented image objects that are too small to be of interest. Shrinking is a special kind of erosion in that single-pixel objects are left intact. This is useful when the total object count must be preserved. Thinning is another special kind of erosion. It is implemented in a two-step process. The first step will mark all candidate pixels for removal. The second step actually removes those candidates that can be removed without destroying object connectivity.

C. Dilation

Dilation is the process of incorporating into the object all the background pixels that touch it, leaving it larger in area by that amount. If two objects are separated by less than three pixels at any point, they will become connected (merged into one object) at that point. It is useful for filling small holes in segmented objects. Thickening is a special kind of dilation. It

International Journal of Ethics in Engineering & Management Education
Website: www.ijeee.in (ISSN: 2348-4748, Volume 1, Issue 4, April 2014)

is implemented in a two-step process. The first step marks all the candidate pixels for addition. The second step adds those candidates that can be added without merging objects.

D. Opening

The process of erosion followed by dilation is called *opening*. It has the effect of eliminating small and thin objects, breaking objects at thin points, and generally smoothing the boundaries of larger objects without significantly changing their area.

E. Closing

The process of dilation followed by erosion is called *closing*. It has the effect of filling small and thin holes in objects, connecting nearby objects, and generally smoothing the boundaries of objects without significantly changing their area.

F. Filtering

Image filtering can be used for noise reduction, image sharpening, and image smoothing. By applying a low-pass or high-pass filter to the image, the image can be smoothed or sharpened respectively. Lowpass filter is used to reduce the amplitude of high-frequency components. Simple lowpass filters apply local averaging. The gray level at each pixel is replaced with the average of the gray levels in a square or rectangular neighborhood. *Gaussian Lowpass Filter* applies *Fourier* transform to the image. Highpass filter is used to increase the amplitude of high-frquency components.

G. Segmentation

It is useful for detecting a set in which all the pixels are adjacent or touching. Within each region, there are some common features among the pixels, such as color, intensity, or texture. When a human observer views a scene, his visual system will automatically segment the scene for him or her. The process is so fast and efficient that one sees not a complex scene, but rather a collection of objects. However, computer must laboriously isolate the objects in an image by breaking the image into sets of pixels, each of which is the image of one object.

Image segmentation can be approached from three ways. The first approach is called *region approach*, in which each pixel is assigned to a particular object or region. In the *boundary approach*, only the boundaries that exist between the regions are located. The third is called *edge approach*, where people try to identify edge pixels and then link them together to form the required boundaries.

H. Object Recognition

The most difficult part of image processing is object recognition. Although there are many image segmentation algorithms that can segment image into regions with some continuous feature, it is still very difficult to recognize objects from these regions. There are several reasons for this. First, image segmentation is an ill-posed task and there is always some degree of uncertainty in the segmentation result. Second, an object may contain several regions and how to connect different regions is another problem. At present, no algorithm can segment general images into objects automatically with high accuracy. In the case that there is some prior knowledge about the foreground objects or background scene, the accuracy of object recognition could be pretty good. Usually the image is first segmented into regions according to the pattern of color or texture. Then separate regions will be grouped to form objects. The grouping process is important for the success of object recognition. Full automatical grouping only occurs when the prior knowledge about the foreground objects or background scene exists. In the other cases, human interaction may be required to achieve good accuracy of object recognition

V. APPLICATIONS

Videocel applications

Video browser:

This application is used to visualize video streams. The browser can load a stream and split it in its shot segments using cut detection algorithms. Each shot is then represented in the browser main window by an icon that is a reduced form of its first frame the shots can be played using several view objects.

Weather Digest

The Weather Digest application generates HTML documents from TV weather forecasts. The temporal sequence of maps, presented on the TV, is mapped to a sequence of images in the HTML page. This application illustrates the importance of information models.

News analysis

News analysis developed a set of applications to be used by social scientists in content analysis of TV news. The analysis was in filling forms including news items duration, subjects, etc., which our attempts to automate. The system generates HTML pages with the images and CSV (Comma Separated Values) tables suitable for use in spreadsheets such as Excel. Additionally, these HTML pages can be also used for news browsing and there also is a java based tool for accessing this information.

VI. CONCLUSION

Visual information has always been always an important source of knowledge. With the advances in information computing & communication technology, this information in the type of digital images & digital video, is highly available also through the computer. To be able to cope with the explosion of visual information, an organization of material which allows for fast search & retrieval is required. These calls for the system which is in some way can provide content-based handling of visual information. In this seminar I have tried to give the basic image processing techniques,

status of content based access to images & video databases, some applications regarding to video content extraction.

An image extraction system is necessarily for users that have large collection of images like digital library. During the last few years, some content based techniques for image retrieval system are commercially available. These systems offer retrieval by color; texture or shape & smart combinations of these images help users in finding the image he is looking for. A video retrieval system is useful for video archiving, video editing, production etc.

REFERENCES

[1]. Basis of Video and Image Processing by Jian Wang
[2]. A survey on content based access to image & video databases by Kjersti Aas & Line Eikvil
[3]. Content-based representation and retrieval of visual Media: a state-of-the-art review 1 Philippe Aigrain, Hongjiang Zhang Dragutin Petkovic
[4]. Advanced content-based semantic scene analysis and information retrieval: the schema project E. Izquierdo, J.R. Casas, R. Leonardi, P. Migliorati, Noel e. O'connor, I. Kompatsiaris and M. G. Strintzis.
[5]. Relevance feedback techniques in content-based image retrieval By Yong Rui, Thomas U Huang, Sharad Mehrotra
[6]. IEEE conference on Advanced Video and Signal Based Surveillance Video Extraction in Compressed Domain
[7]. IEEE paper on Image Segmentation; block based feature extraction; large Video/image databases;
[8]. IEEE paper on Face location in wavelet-based video compression for high perpetual quality Feature extraction