Chapter XV

Video Summarization Based on Human Face Detection and Recognition

Hong-Mo Je, Pohang University of Science and Technology, Korea
Daijin Kim, Pohang University of Science and Technology, Korea
Sung-Yang Bang, Pohang University of Science and Technology, Korea

ABSTRACT

In this chapter, we deal with video summarization using human facial information by face detection and recognition. Many methods of face detection and face recognition are introduced as both theoretical and practical aspects. Also, we describe the real implementation of the video summarization system based on face detection and recognition.

INTRODUCTION

The growing availability of multimedia data such as video and home equipment creates a strong requirement for efficient tools to manipulate this type of data. Automatic summarization is one such tool that automatically creates a short version or subset of key frames that contains as much information as possible from the original video. Summaries are important because they can rapidly provide users with some information.
about the content of a large video or set of videos. From a summary, the user should be able to evaluate if a video is interesting or not, for example, if a documentary contains a certain topic, or a film takes place partly in a certain location.

Automatic summarization is the subject of very active research, and several approaches have been proposed to define and identify what is the most important content in a video. Many researchers and engineers have been actively developing technologies for video summarization in order to facilitate efficient management, exchange, and consumption of digital videos. The goal of video summarization is to obtain a compact representation of the original video that usually contains large volume of data. Oh and Hua (2000) stated that there are two methods of video summarization. One is to extract key frames from each shot or scene and present them as the summary of the video. This approach is good for quick browsing of lengthy videos. The other way is to extract its “interesting” or “important” scenes using content-based features such as caption on shot, audio, or visual information. The human face can play an important role in indexing key information for video summarization because it is a unique feature of human beings and it is ubiquitous in TV news, dramas, documentaries, and movie videos. So it can also be salient feature to represent importance of video shot.

In this chapter, we introduce an application of video summarization to extract interesting/important scenes using human face detection and recognition and show how to implement the system. The target person for the system can be a particular person in video sequences, such as an anchor of news or the main actor in a drama or movie. The first step of this application, face detection, is presented in the second section. In the third section, face recognition is reviewed. Then, we describe implementation of the proposed system. Experimental results are shown in the fifth section. Finally, concluding remarks and future works are summarized in the last section.

FACE DETECTION

A face detection problem can be defined as follows: an arbitrary image can be a digitized video signal or a scanned photograph as input, so we must determine whether there are any human faces in the image or not, and if so, report their location. A first step of any face detection system is detecting the locations in images where faces are present. However, face detection from a single image is a challenging task because of variability in scale, location, orientation (upright, rotated), and pose (frontal, profile). Facial expression, occlusion, and lighting conditions also change the overall appearance of faces. The challenges associated with face detection can be attributed to the following factors:

- **Pose:** The images of a face vary due to the relative camera-face pose (frontal, 45 degrees, profile, upside-down), and some facial features (such as an eye or the nose) may become partially or wholly occluded.
- **Presence or absence of structural components:** Facial features such as beards, mustaches, and glasses may or may not be present, and there is a great deal of variability among these components, including shape, color, and size.
- **Facial expression:** The appearance of faces are directly affected by a person’s facial expression.
Related Content

Hybrid Recommender System Using Emotional Fingerprints Model
[www.igi-global.com/article/hybrid-recommender-system-using-emotional-fingerprints-model/230326?camid=4v1a](www.igi-global.com/article/hybrid-recommender-system-using-emotional-fingerprints-model/230326?camid=4v1a)

A Technique for Resolution of the Assignment Problem Containers in a Container Terminal
www.igi-global.com/chapter/a-technique-for-resolution-of-the-assignment-problem-containers-in-a-container-terminal/197699?camid=4v1a

A Discrete Black Hole Optimization Algorithm for Efficient Community Detection in Social Networks
www.igi-global.com/chapter/a-discrete-black-hole-optimization-algorithm-for-efficient-community-detection-in-social-networks/197700?camid=4v1a