

ROBUST FACE TRACKING BASED ON REGION CORRESPONDENCE AND ITS APPLICATION FOR PERSON BASED INDEXING SYSTEM

HANJIN RYU, MYUNGHOON KIM, VIETCUONG DINH
SEONGSOO CHUN AND SANGHOON SULL¹

Department of Electronics and Computer Engineering
Korea University
1, 5-Ga, Anam-Dong, Sungbuk-Ku, Seoul 136-701, Korea
{hanjin; mhkim; cuongdv; sschun; sull}@mpeg.korea.ac.kr

Received September 2007; revised March 2008

ABSTRACT. *In this paper, a face tracking approach is proposed based on template matching for applying a video indexing application. Typically, this method is simpler and faster than other methods, but the main drawback is the poor performance in the case of face scaling changes and template drifts. To overcome these problems, the information about the facial features is incorporated into a template matching process. First, the face template is represented by two projection histograms of the face region and matching methods are used to determine the candidate face region. The matching is done this way as it is faster to match two 1-dimensional projection histograms than it is to match one 2-dimensional image. Next, the facial features are extracted from the candidate face region and a dissimilarity measure, based on the proximity of the facial features, is used to verify the face region. Finally, an anthropometric model, based on geometrical relationship between facial features, is used to refine the face region. The template to determine the face region in the next frame is dynamically updated using the refined face region. Thus, the proposed method can adapt to scale any changes with less computational cost. Experimental results are provided to demonstrate the efficiency of the proposed method.*

Keywords: Face tracking, Projection histogram, Template matching, Video indexing

1. Introduction. Face detection in images and video sequences has been one of the most important research issues in the area of computer vision. As the amount of multimedia content that can be accessed by the public has recently exploded with the phenomenal growth of digital media, the need for an automatic video content indexing has been also increased. For example, the positions or the number of times that actors appear in a movie provide good information for organizing and presenting the video content. A wide variety of face detection methods have been proposed in recent research publications for static images as well as video sequences [1-3]. Among these, the image based method is known to be more robust than most of the alternatives.

The image based methods [4-8] make direct use of all available image intensity information, thereby eliminating the need to determine a special set of local features, such as in the feature based methods [9-12]. Some methods, to characterize face patterns using statistical learning algorithms such as neural networks [7] and support vector machines [8], have been explored, but they lack the ability to track faces directly. This is because these methods are focused on detecting faces that have been trained for each frame of an input video.

¹Corresponding author