Quality Based Frame Selection For Video Face Recognition

Kaneswaran Anantharajah, Simon Denmon, Sridha Sridharan, Clinton Fookes and Dian Tjondronegoro
Science and Engineering Faculty,
Queensland University of Technology,
GPO Box 2434, 2 George St., Brisbane, Queensland 4001.
{*k.anantharajah*, *s.denman*, *s.sridharan*, *c.fookes*, *dian*}@qut.edu.au

***Abstract*—Quality based frame selection is a crucial task in video face recognition, to both improve the recognition rate and to reduce the computational cost. In this paper we present a framework that uses a variety of cues (face symmetry, sharpness, contrast, closeness of mouth, brightness and openness of the eye) to select the highest quality facial images available in a video sequence for recognition. Normalized feature scores are fused using a neural network and frames with high quality scores are used in a Local Gabor Binary Pattern Histogram Sequence based face recognition system. Experiments on the Honda/UCSD database shows that the proposed method selects the best quality face images in the video sequence, resulting in improved recognition performance.**

I. Introduction

Compared to other biometrics, video based face recognition has attracted significant attention in recent years due to its unique advantages and the wide range of applications including security, law enforcement and entertainment [1]. Video based face recognition techniques can be classified as either extensions of still-image-based techniques, or multi modal cue and spatial-temporal information based techniques [1]. Video contains redundant information and using whole video sequences for face recognition contributes to high computational costs and does not guarantee optimal performance [2]. This is due to facial images that may be non-frontal, contain out of plane rotations, or be poor resolution, degrading the recognition rate. Thus, selecting high quality frames for video based face recognition is crucial. In the past, quality frame selection has been explored by the multimedia and image processing research community and has been used in many applications including video summarization [3], [4] and face recognition [5], [6], [2]. We have been motivated by the work reported in [7] to summarize faces in surveillance video sequences, where head pose, tilt, brightness, sharpness, resolution, openness of the eye, direction of the eyes and closeness of the mouth are combined to select good quality frames. Normalized feature scores are combined using a neural network and frames with higher scores are selected to represent the video sequence. However, this method is not validated using a face recognition experiment and is only used as a means to summarize a sequence for a human operator.

Argones Rua *et al*. [5] proposed a quality based score normalization and frame selection technique for video based person authentication, where sharpness of the image and face symmetry features are used to select good quality frames. Other researchers have used clustering techniques to extract the exemplar, including spatio temporal hierarchical agglomerative clustering (STHAC) [8] and locally linear embedding combined with k-means clustering [9]. Wong et al. [2] considered image characteristics such as sharpness, head pose, geometric alignment and cast shadows to select quality frames. Facial symmetry analysis including pose and lighting, distance from the user to the camera, illumination intensity and sharpness properties are considered in [10] for standardization of face image quality. This method was tested to asses the face image quality and has not been used for face recognition. In this paper, we evaluate quality frame selection for video based face recognition, using high level and low level image features. In our proposed approach, first, frames are extracted from a video sequence and faces are normalized using eye position. Then we extract image features such as closeness of mouth, sharpness, brightness, openness of the eye, contrast and face symmetry; and determine a quality score for each individual feature. Following this, we fuse all normalized feature scores using neural network and images with a high score are used in a face recognition system. In this paper, we used the local gabor binary pattern histogram sequence (LGBPHS) [11] technique for face recognition.

The remainder of the paper is organized as follows. In Section II we describe the face detection and facial feature extraction techniques used; Section III outlines the local gabor binary pattern histogram sequence framework for face recognition. In Section IV, we summarize the experimental results and we conclude the paper in Section V.

II. Face Detection and Facial Feature Extraction

A. Face Normalization

We use a Haar-like feature based cascade classifier to detect eye position [12], [13]. Then based on the eye position we normalize the face image to a consistent size (130×150 pixels). In the normalization process we also normalize the image intensity using histogram equalisation. Eye position is also used to remove severely in-plane rotated face images. When the detected eyes Y coordinate is higher than a pre determined threshold, we set that face as an outlier in our

978-1-4673-2393-2/12/$31.00