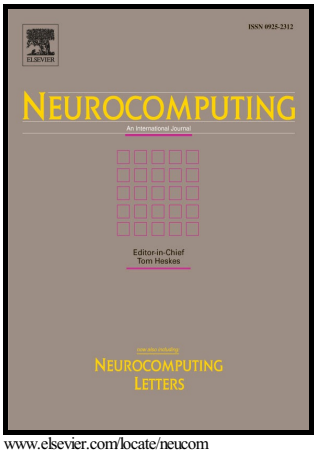

www.elsevier.com/locate/neucom

Surveillance Video Face Recognition with Single Sample per Person Based on 3D Modeling and Blurring

Xiao Hu, Shaohu Peng, Li Wang, Zhao Yang, Zhaowen Li

PII: S0925-2312(17)30001-2
DOI: http://dx.doi.org/10.1016/j.neucom.2016.12.059
Reference: NEUCOM17890

To appear in: *Neurocomputing*

Received date: 1 September 2015
Revised date: 26 August 2016
Accepted date: 26 December 2016

Cite this article as: Xiao Hu, Shaohu Peng, Li Wang, Zhao Yang and Zhaowen Li, Surveillance Video Face Recognition with Single Sample per Person Based on 3D Modeling and Blurring, *Neurocomputing*, http://dx.doi.org/10.1016/j.neucom.2016.12.059

Surveillance Video Face Recognition with Single Sample per Person Based on 3D Modeling and Blurring

Xiao Hu, Shaohu Peng, Li Wang, Zhao Yang, Zhaowen Li

School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China.

Abstract

Video surveillance has attracted more and more interests in the last decade, video-based Face Recognition (FR) therefore became an important task. However, the surveillance videos include many vague non-frontal faces especially the view of faces looking down and up. As a result, most FR algorithms would perform worse when they were applied in surveillance videos. On the other hand, it was common at video monitoring field that only Single training Sample Per Person (SSPP) is available from their identification card. In order to effectively improve FR for both the SSPP problem and the low-quality problem, this paper proposed an approach to synthesis face images-based on 3D face modeling and blurring. In the proposed algorithm, firstly a 2D frontal face with high-resolution was used to build a 3D face model, then several virtual faces with different poses were synthesized from the 3D model, and finally some degraded face images were constructed from the original and the virtual faces through blurring process. At last multiple face images could be chosen from frontal, virtual and degraded faces to build a training set. Both SCface and LFW databases were employed to evaluate the proposed algorithm by using PCA, FLDA, scale invariant feature transform, compressive sensing and deep learning. The results on both datasets showed that the performance of these methods could be improved when virtual faces were generated to train the classifiers. Furthermore, in SCface database the average recognition rates increased up to 10%, 16.62%, 13.03%, 19.44% and 23.28% respectively for the above-mentioned methods when virtual view and blurred faces were taken to train their classifiers. Experimental results indicated that the proposed method for generating more train samples was effective and could be considered to be applied in intelligent video monitoring system.

Keywords: single training sample per person; video surveillance; scale invariant feature transform; compressive sensing; deep learning.