Journal of Visual Communication and Image Representation 51 (2018) 112–121

Contents lists available at ScienceDirect

Journal of Visual Communication and Image Representation

journal homepage: www.elsevier.com/locate/jvci

Face sketch-photo synthesis and recognition: Dual-scale Markov Network and multi-information fusion☆

Zhenxue Chen[a,*], Saisai Yao[a], Yunyi Jia[b], Chengyun Liu[a]

[a] School of Control Science & Engineering, Shandong University, Jinan 250061, China
[b] Department of Automotive Engineering, Clemson University, Greenville, SC 29607, USA

ARTICLE INFO

Keywords:
Sketch face synthesis and recognition
Dual-scale Markov Network
Structural information
Feature information
Fusion

ABSTRACT

Sketch face recognition (SFR) has been widely and successfully applied in law enforcement, which attracts a growing number of researchers. In this paper, a face sketch-photo synthesis and recognition method is proposed. Our method has two parts: Firstly, according to the different synthesis results for different scales, a cascade sketch-photo synthesis method via dual-scale Markov Network is utilized for image synthesis; Secondly, structural information and feature information-based data fusion method has been presented for face recognition. It is inspired by the Face Recognition Cognitive Theory, which applies both structural information and feature information for recognition. The experimental results on different databases based on the proposed method, demonstrate the outperformance of our method compared with state-of-the-art methods both in synthesis and recognition processes.

1. Introduction

IN real criminal cases, the photo image of a suspect is not available, because of lack of cameras or the bad quality of camera. Fortunately, the witnesses who are found in most cases can provide certain assist for the police. Professional forensic artist creates sketch drawing of suspect by assisting the witness to recall the appearance characteristics of the criminal suspect. Then the sketch is disseminated to law enforcement officers and media outlets with the hope that someone will recognize the individual and provide relevant information of the suspect. Due to the great help of sketch in law enforcement, a growing number of researchers have been devoted to explore the sketch face recognition methods. But the sketches and photos are two different modalities from different sources that bring great difficulty for identification study. Traditional manual method to find the corresponding photo of probe sketch is feasible, but it costs too much time and manpower. In order to solve the above problems, a face sketch-photo synthesis and recognition method is proposed which is based on dual-scale Markov Network and Multi-information fusion.

Face sketch-photo synthesis is to transform the photo (sketch) into sketch (photo). A typical method is searching for the neighbor photo patches of the test photo and creating the target sketch by a linear combination of the corresponding neighbor sketch patches. The selection of the patch scale is a significant factor in the synthesis procedure. A large scale can lead to the coarse synthesis image with distortion and mosaic effect, while a small scale neglects the connection of face structures. Most existing methods use single patch scale on the sketch-photo synthesis process, which ignore the influence of patch scale. In view of the above problem, this paper proposes a sketch-photo synthesis method based on dual-scale Markov Network.

After synthesis process, sketches and photos are in the same modality, which are needed to do further recognition. According to the famous "Thatcher illusion" [1], cognitive psychologists believe that two kinds of information (structural information and feature information) are utilized for face recognition. Structural information refers to the spatial relations between facial features, and feature information refers to the feature of single facial components. Structural information and feature information are different from each other, but are essential and complementary in face recognition process. Therefore, this paper proposes a data fusion recognition method based on human cognition, which takes both the structural information and feature information into consideration for face recognition.

The main contributions of this paper are summarized as follows:

(1) Dual-scale Markov Network-based sketch photo synthesis method is presented to synthesize sketch and photo, which uses larger and smaller scales Markov Network sequentially.
(2) Multi-information fusion-based sketch face recognition approach is

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author.
E-mail address: chenzhenxue@sdu.edu.cn (Z. Chen).

https://doi.org/10.1016/j.jvcir.2017.12.010
Received 6 July 2017; Received in revised form 9 December 2017; Accepted 19 December 2017

presented, which utilizes structural information and feature information for sketch face recognition according to Face Recognition Cognitive Theory.

(3) Leading accuracy rates are achieved on multiple sketch face databases, which prove the effectiveness of the proposed method.

The arrangement of the rest of this paper is as follows. Section 2 introduces several representative methods on SFR. The proposed sketch face synthesis and recognition method via dual-scale Markov Network and Multi-information is presented in Section 3. Section 4 shows experimental results and analyses on different sketch face databases.

2. Related works

At present, the studies of sketch face recognition mainly divided into two categories: sketch/photo transformation and matching sketches and photos directly. For the aspect of sketch/photo transformation, Tang and Wang [2–7] proposed some methods. One is that synthesized a pseudo sketch by applying Eigen-transformation on entire face photo which used for matching in sketch modality [2,3] and they improve the synthesis framework by applying Eigen transformation on local patches [4]. Then, a nonlinear approach for face sketch synthesis and recognition is proposed in Ref. [5], in which the pseudo-sketch generation method is based on local linear preserving of geometry between photo and sketch images and the nonlinear discriminate analysis is used to recognize the probe sketch from the synthesized pseudo-sketches. Based on references proposed above, Wang and Tang synthesized and recognized sketch/photo images using a multi-scale Markov Random Fields (MRF) model [6] in which extensive experiments are conducted on a face sketch database including 606 faces. In 2011, Zhang and Wang [7] proposed a new inter-modality face recognition approach by reducing the modality gap at the feature extraction stage. A new face descriptor based on coupled information-theoretic encoding is used to capture discriminative local face structures and to effectively match photos and sketches. Guided by maximizing the mutual information between photos and sketches in the quantized feature spaces, the coupled encoding is achieved by the proposed coupled information-theoretic projection tree, which is extended to the randomized forest to further boost the performance. At the same time, Xiao [8] studied the mapping relationship between the sketch and pictures about different information expression patterns, focusing on fundamental problems of the sketch-photo identification in his doctoral degree thesis, and proposed the sketch-photo synthesis methods on the basis of constructing the mapping relation models, so as to transform sketch-photo identification to face recognition within the same mode. Obviously, the performance of those methods is highly dependent on the result of photo/sketch synthesis. Chen et al. [9] presented an improved sketch-photo transformation algorithm which uses the effective characteristics of the photo image more reasonably during transforming a photo image into sketch. Gao et al. [10] proposed an automatic sketch-photo synthesis and retrieval algorithm based on sparse representation [35–37]. A novel multiple representations based face sketch photo synthesis method that adaptively combines multiple representations to represent an image patch is proposed by Peng et al. [38]. Then, Peng et al. [39] proposed a super pixel based face sketch photo synthesis method by estimating the face structures through image segmentation. In [40], Liang et al. synthesized images from a simple line drawing that contains only a few lines without any texture. Song et al. [41] present a facial sketch to photo synthesis method via online coupled dictionary learning.

On the other hand, Klare and Jain [11] represented face photo and sketch using SIFT descriptors [12] with common representation matching. Later, Klare et al. [13] proposed local feature based discriminative analysis (LFDA) to match forensic sketches to mug shot photos with individually represent both sketches and photos using SIFT feature descriptors and multi-scale local binary patterns (MLBP) [14]. Li et al. matched sketches and photos using an algorithm similar to the Eigen-transform, focusing on converting sketches to photos [15]. Zhang [16] compared the performances of different sketch image recognition to automated sketch image recognition and showed the contributions of using multiple sketches drawn by different artists. Klum et al. [17] researched the problem of automatic facial sketch to mugshot matching and compare the effectiveness of forensic sketches and composite sketches. The main contribution of this paper is that a comparison of holistic facial representations versus component based representations for sketch to mugshot matching. Klum [18] has researched the FaceSketchID System, a scalable and operationally deployable software system that achieves state-of-the-art matching accuracy on facial composites using two complementary algorithms (holistic and component based). Sen Bekir [19] presented a novel framework for face recognition from sketch which is based on Principle Component Analysis (PCA) and Canonical Correlation Analysis (CCA). Wang et al. [20] presented a novel transductive face sketch-photo synthesis method that incorporates the given test samples into the learning process and optimizes the performance on these test samples. Han et al. [21] proposed a component-based representation (CBR) approach to measure the similarity between a composite sketch and mugshot photograph. In Ref. [22], an automatic sketch synthesis algorithm is proposed based on embedded hidden Markov model (E-HMM) and selective ensemble strategy. Xiao et al. [23] devoted to synthesizing a photo from the sketch and transforming the sketch-photo recognition to photo-photo recognition to achieve better performance in mixture pattern recognition. Marco and Guillermo1 [24] proposed a new approach for photo/sketch recognition based on the Local Feature-based Discriminant Analysis (LFDA) method. Kukharev et al. [25] presented the state of the art problem of comparing photo portrait and the corresponding hand-drawn portrait (sketch). It is shown that for sketch recognition you can use simple system. Wang [26] represented sketches and photos using HOG feature descriptors and used Null space-based LDA (NLDA) to get the feature space for recognition. Klare et al. [27] extended the prototype-based method into heterogeneous face recognition and present a kernel prototype random subspace to deal with the small sample problem.

In this paper, a face sketch-photo synthesis and recognition method is proposed. Firstly, a dual-scale Markov Network based face sketch-photo synthesis method is presented to transform the sketches and photos into the same modality. Then, according to Face Recognition Cognitive Theory, a data fusion method which applies both structural information and feature information for recognition is proposed. The HOG feature of the whole face image is extracted to represent the structural information. And the sparse kernel prototype representation is utilized to describe the feature information on local face components. More details of the proposed sketch face synthesis and recognition method are provided in the next Section.

3. Proposed synthesis method

3.1. Dual-scale Markov Network synthesis

Given a test photo t, training sets p and s, we evenly divide the images into N overlapped patches based on large scale. For the ith test photo patch t_i, we firstly search for the nearest K neighbor photo patches $\{p_{i,1},p_{i,2},...,p_{i,k}\}$ of test photo patch among the training set. And the corresponding neighbor sketch patches $\{s_{i,1},s_{i,2},...,s_{i,k}\}$ are obtained. Then, their weight vector $w_i = [w_{i,1},w_{i,2},...,w_{i,k}]$ is estimated using Markov Network.

The joint probability of the test photo patches t_i and weight vectors w_i ,$\forall\ i \in \{1,2,...,N\}$ is described as follows:

$$p(t_1,...,t_N,w_1,...,w_N) \propto \prod_{i=1}^{N} \Phi(t_i,w_i) \prod_{(i,j)\in\Xi} \Psi(w_i,w_j) \quad (1)$$