Information Fusion 33 (2017) 86–99

Contents lists available at ScienceDirect

Information Fusion

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

Composite sketch recognition using saliency and attribute feedback

Paritosh Mittal, Aishwarya Jain, Gaurav Goswami, Mayank Vatsa*, Richa Singh

IIIT Delhi, India

ARTICLE INFO

Article history:
Received 26 June 2015
Revised 14 April 2016
Accepted 17 April 2016
Available online 19 April 2016

Keywords:
Biometrics
Sketch recognition
Composite sketch
Face recognition
Saliency

ABSTRACT

Recent interest and requirement of law enforcement agencies in matching composite sketches with digital images has instigated research in this important face recognition problem. In this paper, we propose feature extraction and matching algorithm using visual saliency and combination of texture features for matching composite sketches with digital photos. The attributes such as gender, ethnicity, and skin color are utilized for re-ordering the ranked list. Further, information from multiple experts such as multiple composite sketch generation tools or artists is combined for improving the matching performance. The results obtained on the extended PRIP database show that the proposed algorithm improves the state-of-art in matching composite sketch and digital face images and yields the rank 50 identification accuracy of 70.3% on a database of 1500 subjects.

1. Introduction

Forensic investigations may involve generating facial sketches of a suspect by an expert artist based on the description of one or more eye-witness. These sketches not only perform a crucial role in the immediate notification to the general public, they also serve as a key component in criminal investigations. While face images captured using a digital camera retain fine characteristics of the face accurately (depending on camera's resolution and other features), hand-drawn forensic sketches are approximate reproductions of witness descriptions and capture basic features such as shape along with some salient and distinguishable facial characteristics such as scars and moles. In the process, not only texture information and other minute details are left out, reproduction of basic features is also subjective and depends on the witness and the artist. This leads to challenges which are specific to the problem of matching sketches with digital images. Due to these challenges, state-of-the-art face recognition approaches cannot be applied directly in this forensic application. In the past few years, several discriminative and generative algorithms have been proposed for matching hand drawn sketches with digital images [1–9].

With technological advancements, law enforcement agencies have started utilizing software solutions to produce composite sketches as a replacement of hand drawn sketches. These solutions provide multiple templates of facial features, such as nose, eyes, and mouth to generate composite sketches with different shapes and sizes. A face composite is created based on the selected feature template. Composite sketches have assisted in successful arrest of several criminals such as the Lexington armed robber [10], the Manchester serial rapist [11], and the Virginia Highland robber [12]. Some real world examples of composite and digital images are shown in Fig. 1.

Composite sketch generation is a relatively well researched area. Some of the popular solutions include E-FIT [20], which interprets descriptions of facial features to create a prototype which can then be refined manually. It also provides the functionality, where a user can select a minimalistic face template and then search and add facial features to it. EvoFIT [21] is another state-of-the-art solution for creating composite sketches that utilizes an attribute based method to shortlist a gallery of face templates from its database. Witnesses must provide gender, ethnicity, and age after which they repeatedly select the most similar faces from a few face template groups presented via the interface. The selected templates are merged together and then refined using complex attributes such as masculinity, weight, friendliness, and honesty represented using sliders that affect the template in real time. It also offers a variety of animated spatial transformations to help account for inaccuracies in the generated face composite. Faces by IQ Biometrics [10] is another widely used composite sketch generation software that creates sketches by combining facial sketch templates (chosen by eye-witness). A recent research explores the possibility of reconstructing a 3D face using attributes and genetic markers [22]. This research has sparked the exploration of generating 3D facial composites utilizing DNA obtained from a crime scene without the requirement for witness description.

* Corresponding author. Tel.: +919654653404.
E-mail addresses: paritosh10059@iiitd.ac.in (P. Mittal), aishwarya10007@iiitd.ac.in (A. Jain), gauravgs@iiitd.ac.in (G. Goswami), mayank@iiitd.ac.in (M. Vatsa), rsingh@iiitd.ac.in (R. Singh).

http://dx.doi.org/10.1016/j.inffus.2016.04.003
1566-2535/

Fig. 1. Composite and digital face images corresponding to some real world cases. Images are taken from the Internet.

Table 1
A review of algorithms proposed for matching composite sketches to digital face images.

Authors	Algorithm	Database and Results
Han et al. [13]	A component based representation (CBR) approach is proposed. After detecting facial landmarks using active shape models, features are extracted using MLBP and per component similarity scores are fused to obtain the similarity score between a face and a composite sketch.	PRIP database: 123 composite sketches pertaining to digital face images from AR database. Rank 50 accuracy - 55% on a gallery of 10,000.
Mittal et al. [14]	Multiple circular patches are extracted from digital faces and composite sketches. DAISY features are extracted and recognition is performed by matching these features using boosting.	PRIP database. Rank 10 accuracy of 46%.
Chugh et al. [15]	Shape, orientation, and texture features are extracted locally using image moments and histogram of oriented gradients (HOG). Match score level fusion is performed for efficient matching.	Composite sketch database of 59 subjects, 49.2% rank 50 accuracy against a gallery of 4904 subjects.
Klum et al. [16]	Holistic and component based facial representations are compared and the effect of demographic filtering on digital gallery is analyzed.	Composite sketch database of 75 subjects, 62.67% rank 200 accuracy against a gallery of 10,075 subjects.
Klum et al. [17]	The FaceSketchID System: Matching Facial Composites to Mugshots	2356 composite sketch pairs, 61.3% true accept rate at 1.0% false accept rate against a gallery of size 100,000.
Mittal et al. [18]	Dictionary based matching using bag-of-words SSD features	60.0% rank 50 accuracy against a gallery of 1500 subjects.
Mittal et al. [19]	Deep learning based transfer learning approach	58.0% rank 40 accuracy against extended gallery of 2400 subjects.

Table 1 provides a brief summary of algorithms proposed for composite sketch recognition. Han *et al.* [13] have proposed a component based representation which extracts facial landmarks using active shape models. These features are encoded for every component using multiscale local binary patterns. The features extracted from corresponding components pertaining to both composite and digital images are matched followed by score normalization and fusion to generate the matching result. The algorithm also uses gender information as an indexing/ancillary parameter. Mittal *et al.* [14] have proposed a boosting based approach that combines various weak feature descriptors extracted from fiducial features and multiple circular patches. Chugh *et al.* [15] have proposed an approach that combines multiple discriminative features based on shape, texture, and orientation at match score level for recognition. Klum *et al.* [16] presented a comparison of different facial representations for sketch to digital image matching and also demonstrated the effect of filtering digital mugshot images on the basis of demographic information. Klum *et al.* have proposed the FaceSketchID [17] system which provides a unified framework for composite sketch recognition by combining a holistic and a component based approach. Mittal *et al.* [18] presented a Self Similarity Descriptor (SSD) based dictionary learning approach for matching composite images with digital face images. Mittal *et al.* [19] recently proposed a deep learning based architecture using autoencoders and deep belief networks to implement a transfer learning approach for composite sketch matching. The algorithm utilizes 30,000 frontal face images from the CMU Multi-PIE database to train the deep network which is then fine-tuned for sketch matching.

In this paper, we propose a face recognition algorithm to match composite sketches and digital images using saliency and attribute feedback. In the proposed algorithm, visual saliency determines a subset of facial patches that are utilized for texture feature extraction. These features are extracted at both global and local levels; and then combined with attribute information such as gender, ethnicity, and age, via attribute feedback. We also demonstrate the results of combining information from sketches obtained from multiple sources. Further, a new set of composite sketches generated by an Indian user has been created in accordance with the faces included in the PRIP database. In addition, attributes such as age and skin color are also incorporated to improve the results.

The performance of the proposed algorithm has been evaluated using the extended PRIP database. The results are also compared with a commercial face recognition system FaceVACS and one state-of-the-art algorithm in digital image to hand-drawn sketch matching [9]. Since matching composite sketches with digital images involve “human in the loop”, we perform a user analysis to understand the matching capabilities of humans' for matching composite sketches with digital face images. To the best of our knowledge, this is the first extensive study on the impact of various factors in sketch recognition.

2. Proposed approach

Generally, a witness sees the subject for only a couple of seconds or minutes. The description provided by the witness based on the memory of such a short duration may not be completely accurate and therefore, designing a matching algorithm based on a single global facial feature may not provide the best results. Along with facial characteristics, a witness description also consists of attributes such as ethnicity, gender, and age. The proposed algorithm is designed based on the hypothesis that several of these facial features along with attributes can be combined to create a strong classifier. The steps involved in the proposed algorithm are shown in Fig. 2. First, the salient regions in a face are identified. The local facial features of the most salient regions are encoded using DAISY