Image and Vision Computing 58 (2017) 181–192

Contents lists available at ScienceDirect

Image and Vision Computing

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

Facial attributes for active authentication on mobile devices☆

Pouya Samangouei[a], Vishal M. Patel[b], Rama Chellappa[a]

[a] *Electrical and Computer Engineering Department, University of Maryland, College Park, MD, USA*
[b] *Electrical and Computer Engineering Department, Rutgers University, 94 Brett Road, Piscataway, NJ, USA*

ARTICLE INFO

Article history:
Received 8 October 2015
Received in revised form 20 April 2016
Accepted 5 May 2016
Available online 13 May 2016

Keywords:
Attributes
Security
Active authentication
Face
Mobile

ABSTRACT

We present a method using facial attributes for continuous authentication of smartphone users. We train a bunch of binary attribute classifiers which provide compact visual descriptions of faces. The learned classifiers are applied to the image of the current user of a mobile device to extract the attributes and then authentication is done by simply comparing the calculated attributes with the enrolled attributes of the original user. Extensive experiments on two publicly available unconstrained mobile face video datasets show that our method is able to capture meaningful attributes of faces and performs better than the previously proposed LBP-based authentication method. We also provide a practical variant of our method for efficient continuous authentication on an actual mobile device by doing extensive platform evaluations of memory usage, power consumption, and authentication speed.

Published by Elsevier B.V.

1. Introduction

Most probably the first time that a password was used for authenticating users of a computer was in 1961 on the famous Compatible Time-Sharing System (CTSS) which was developed at MIT's computing center [1]. However, passwords turned out to be hard to remember and maintain, and they need to be kept secure. Soon researchers began investigating less vulnerable and easier to maintain forms of authentication. One of the most studied type of methods is biometric identification [2], such as fingerprints, retinal scans or facial image matching. These methods are appealing because while they are unique for each person and hence more secure, they are more intuitive, hence requiring less user effort either for remembering or entering them.

Mobile devices are one of many categories of devices in which the password-based approaches are widely used as the sole authentication method. Smartphones, tablets, and wearable devices fall into this category. These devices have become an inseparable part of people's lives. They contain a lot of valuable information, from bank account details to emails and other private content. Therefore, these devices are being increasingly targeted by different kinds of attacks [3]. Typical devices incorporate no mechanisms to verify that the user originally authenticated is still the user in control of the mobile device. Thus, unauthorized individuals may improperly obtain access to personal information of the user if a password is compromised or if a user does not exercise adequate vigilance after initial authentication on a device. Biometrics-based algorithms have emerged as a solution for continuous authentication on these devices [4–6]. See [7] for a comprehensive survey of recent mobile continuous authentication systems.

Modern mobile devices come with a variety of built-in sensors and accessories such as cameras, microphone, gyroscope, accelerometer and pressure sensor. These sensors can be used to extract the biometric data for the user [8–10], [6]. For instance, faces can be captured using the front-facing camera of a mobile device and can be used to continuously authenticate a mobile device user [6]. Also, the gyroscope touchscreen and accelerometer can be used to measure biometrics such as gait, touch gestures and hand movement.

One of the most popular active authentication methods is based on the face biometric. Authentication methods such as the ones reviewed in [9], [10], [6] use camera sensor images to detect the face of the user, extract low-level features and apply patter recognition algorithms on these features to authenticate the user. The common drawbacks of these approaches are that the low-level features can vary significantly in different environmental conditions and head pose changes. Also as emphasized in [10], the tradeoff between verification accuracy and mobile performance is an important challenge of active authentication. Many of the methods with good accuracies explored in [9] and [10] have either costly enrollment phase or test phase, in terms of computation or memory.

We propose to use a large number of facial attributes, like gender, race, ethnicity, etc. as intermediate representations to authenticate

☆ This paper has been recommended for acceptance by Patrick Flynn.
E-mail addresses: pouya@umiacs.umd.edu (P. Samangouei), vishal.m.patel@rutgers.edu (V. Patel), rama@umiacs.umd.edu (R. Chellappa).

http://dx.doi.org/10.1016/j.imavis.2016.05.004
0262-8856/Published by Elsevier B.V.

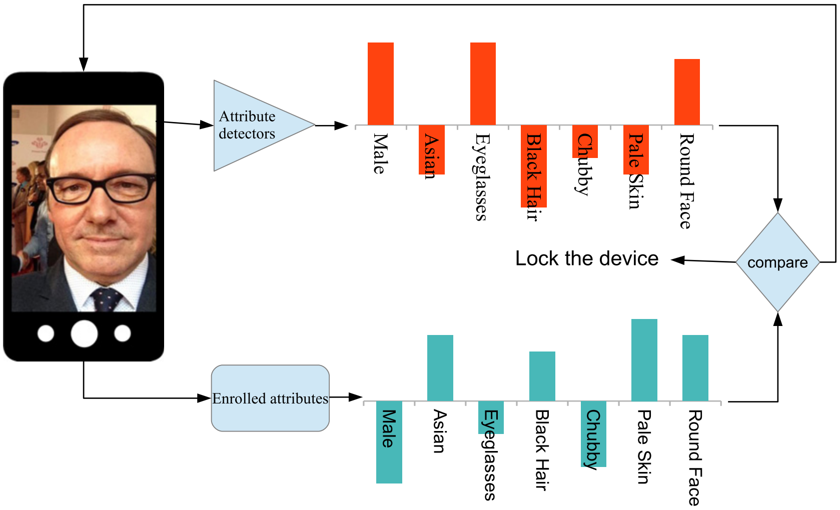

Fig. 1. Overview of our attribute-based authentication method.

the user of the mobile device. The overview of our method is shown in Fig. 1. These attributes give a compact and discriminative representation for the task of continuous authentication.

There are several benefits for using facial attributes as features. First, the training of attribute classifiers can be done offline on large datasets of images which can embody various conditions. As a result, we achieve robustness to changes that can make low-level features ineffective for authentication. Secondly, as we show later in Section 5, the attribute models can be run efficiently, authenticating more than four frames per second on an average mobile device with low power and memory consumption. This is very important since the algorithm must seamlessly run continuously alongside other applications. Furthermore, the attributes are compact representations. Suppose we have n binary attributes, the probability that two people having the same attribute is bounded above by $1/2^n$. This probability can be very low if attribute scores are continuous like age or skin tone. Therefore, just by comparing the derived attributes with the enrolled ones, one can detect with a high probability whether the current user of the phone is the registered user or not. Furthermore, enrollment of attributes can be done in different ways, they either can be asked directly from the user or they can be captured on one device and be used on some other device within the same network. Finally, if stricter security is needed, we show that traditional low-level features can be fused with attribute features to give better performance.

The contributions of this paper are three fold. First, we introduce the use of facial attributes to the task of active authentication and show with extensive experiments on MOBIO [11] and AA01[9] datasets that this approach produces promising results. Second, we present evaluations based on an implementation of our algorithm on an actual mobile device. These evaluations are necessary for all mobile continuous authentication systems to prove their feasibility. Lastly, we also present the labels for UMDAA dataset of 50 subjects each having 44 attributes.

2. Related work

2.1. Attributes

In computer vision, almost in all problems, the very first step is to extract features from a given visual signal. The first use of attributes as higher order features was introduced in content-based image retrieval where they are presented as a solution to decrease the semantic gap [12–14]. Attributes were also referred to as a kind of "intermediate features". This term initially appeared in [14] referred to the features that are "low-level" semantic features but "high level" image features.

Subsequent applications of attributes were in object recognition domain and human identification. Ferrari et al. [15] learned visual attributes for objects such as "dotted" or "striped". In [16] Farhadi et al. used L1-regularized logistic regression to learn object attributes such as "has wheels" or "metallic" from images of PASCAL VOC 2008 [17] and then used them to describe objects in the image. In [18] Lampert et al. learned object attributes via kernel Support Vector Machines (SVMs)[19] in two learning paradigms, Direct and Indirect Attribute Prediction and then used these to perform object recognition. They demonstrated good results on their Animal with Attributes dataset. There are several other areas where attribute features have been shown to be useful: zero-shot learning [20], scene classification [21], and action recognition [22].

Human attributes or "soft biometrics" such as age and gender suggested in [23], have been successfully used for identity recognition/verification in many applications. In [23], Jain et al. combined height, race and gender information with fingerprint to improve the recognition accuracy on an in-house dataset. Face image retrieval solely based on attributes was investigated in [24] by Kumar et al., and in [25] by Park et al. For face verification, Kumar et al. in [26] extracted attribute feature vectors. Zhang et al. [27] used attributes to improve face clustering in the presence of pose and illumination variations. Klare et al.[28] defined 46 facial attributes to perform suspect identification task. In [29] Layn et al. showed that attributes such as "jeans", "headphones" and "sunglasses" can help re-identifying people seen on different cameras of a distributed camera network. Vaquero et al. [30] developed a method for searching with attributes in surveillance environments using Viola–Jones attribute detectors. In [31] Thornton et al. used attribute profiles to search in large datasets of surveillance video data. In [32] Jain et al. fused fingerprints, a few soft biometrics, and low-level features for face recognition on an in-house dataset of 263 users having 10 images each.

Detecting the presence of each attribute has been the focus of many researchers. These algorithms can be roughly divided into two groups, those which learn a specific model per attribute and those which present a general framework to learn all the target attributes together at once. Our focus in this paper is on the second group of