Image and Vision Computing 77 (2018) 1–9

Contents lists available at ScienceDirect

Image and Vision Computing

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

On the generalization of color texture-based face anti-spoofing☆

Zinelabidine Boulkenafet[a,*], Jukka Komulainen[a], Abdenour Hadid[a,b]

[a] Center for Machine Vision and Signal Analysis, University of Oulu, Finland
[b] School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China

ARTICLE INFO

Article history:
Received 13 July 2017
Received in revised form 19 January 2018
Accepted 27 April 2018
Available online xxxx

Keywords:
Face recognition
Presentation attack detection
Spoofing
Color texture analysis
Cross-database
Generalization

ABSTRACT

Despite the significant attention given to the problem of face spoofing, we still lack generalized presentation attack detection (PAD) methods performing robustly in practical face recognition systems. The existing face anti-spoofing techniques have indeed achieved impressive results when trained and evaluated on the same database (i.e. intra-test protocols). Cross-database experiments have, however, revealed that the performance of the state-of-the-art methods drops drastically as they fail to cope with new attacks scenarios and other operating conditions that have not been seen during training and development phases. So far, even the popular convolutional neural networks (CNN) have failed to derive well-generalizing features for face anti-spoofing. In this work, we explore the effect of different factors, such as acquisition conditions and presentation attack instrument (PAI) variation, on the generalization of color texture-based face anti-spoofing. Our extensive cross-database evaluation of seven color texture-based methods demonstrates that most of the methods are unable to generalize to unseen spoofing attack scenarios. More importantly, the experiments show that some facial color texture representations are more robust to particular PAIs than others. From this observation, we propose a face PAD solution of attack-specific countermeasures based solely on color texture analysis and investigate how well it generalizes under display and print attacks in different conditions. The evaluation of the method combining attack-specific detectors on three benchmark face anti-spoofing databases showed remarkable generalization ability against display attacks while print attacks require still further attention.

1. Introduction

The vulnerability of biometric systems against the learned or forged biometric traits has been the subject of many recent studies, including [1-5]. These works have concluded that most of the biometric systems, even those presenting a high recognition performance, are vulnerable to spoofing attacks (or a presentation attack as defined in the current ISO/IEC 30107-3 standard [6]). Face recognition systems are particularly easy to be deceived. With the increase of the social networks' popularity and the improvement of the camera resolution, it is easy to spoof the identity of a target person by using his/her images published in the web or captured from distance without permission. For instance, in a recent study [5], six commercial face recognition systems, namely Face Unlock, Facelock Pro, Visidon, Veriface, Luxand Blink and FastAccess, were easily fooled with crude photo attacks using images of the targeted person downloaded from social networks. Even worse, also their dedicated challenge-response based liveness detection mechanisms were circumvented using simple photo manipulation to imitate the requested facial motion (liveness cues), including eye blinking and head rotation.

To discriminate between real and fake face images, many face presentation attack detection (PAD) methods have been proposed in the literature (see [7-9] for extensive surveys). The existing face anti-spoofing techniques analyzing motion, facial texture content and image quality have already achieved impressive results particularly when trained and evaluated on the same database (i.e. intra-test protocols). As all the existing benchmark publicly available datasets lack variations in the collected data, e.g. user demographics, application scenarios, illumination conditions and input cameras, the reported anti-spoofing results may unfortunately not reflect the real uncontrolled operating conditions that will be definitely faced in real-world applications, such as mobile authentication. For instance, in the widely used Replay-Attack Database [10], the video samples of the training, development and test sets have been collected using a single camera.

To gain insight into the generalization performance of face anti-spoofing techniques, de Freitas Pereira et al. [11] suggested a cross-database evaluation in which the anti-spoofing models are trained

☆ This paper has been recommended for acceptance by Vitomir ?truc.

* Corresponding author.
E-mail address: zinelabidine.boulkenafet@oulu.fi (Z. Boulkenafet).

https://doi.org/10.1016/j.imavis.2018.04.007
0262-8856/

and tuned on one database and then tested on other databases. The experiments have revealed that the performance of the state-of-the-art methods drastically drops as they failed to cope with new spoofing scenarios that have not been seen during training and development phases. So far, even the popular convolutional neural networks (CNN) have failed to derive well-generalizing features for face anti-spoofing [12,13].

Cross-database testing has been increasingly applied in face PAD research [12-20] to overcome the shortcomings of the public datasets since the generalization issue was pointed out by de Freitas Pereira et al. [11]. This has been a nice trend but the main limitation with these preliminary studies has been that, in general, the generalization performance has been only broadly evaluated on the plain overall protocol (i.e. combining all types of spoofing scenarios) without any deep analysis on the effect of different factors such as input sensor or presentation attack instrument (PAI) variation on the generalization capability. Since the overall performance of the state of the art has been far from satisfying the strict security demands of biometric systems, one can even question the meaningfulness of this kind of benchmarking.

In this work, we show that the blind overall assessment might actually lead to overly pessimistic conclusions on the contrary as a method might be able to generalize under some conditions even if its plain overall performance is poor. We argue that careful breakdown analysis across different covariates, especially attack scenarios, is very crucial to gain better insights into the performance and importantly the generalization of different face anti-spoofing methods. The recently standardized ISO/IEC 30107-3 metrics [6] are an important step to the right direction because the attack potential is taken into account as the overall PAD performance corresponds to the most successful PAI. However, this indicates how easy a biometric system is to fool on average by exploiting its (possible) vulnerability, which suits well for evaluating the robustness of complete biometric solutions. Since it is reasonable to assume that no single superior technique is able to detect all known, let alone unseen, attacks types, it is also important to find out the operating conditions of different PAD methods and how complementary countermeasures could be combined to achieve more robust overall performance [21].

Based on the above observations, we present in this work an in-depth analysis on the generalization of color texture-based face anti-spoofing. This is motivated by our recent works [16-18] showing that color texture features extracted from both luminance and chrominance color channels provide the state-of-the-art performance and very promising generalization abilities in face PAD. We perform extensive cross-database tests which measure the robustness of seven different facial color texture descriptions across different covariates, like acquisition conditions and attack scenarios. Our experiments depict that most of the methods are unable to generalize to unseen spoofing attack scenarios but some of the methods are more robust to particular PAIs. Inspired by this, we propose an attack-specific approach to cope with the problem of generalized face PAD. Compared to the state of the art, we obtained very competitive intra-database and inter-database results on three benchmark face spoofing databases. More importantly, the color texture based method can generalize extremely well against display attacks (digital photo and video-replay attacks) launched at short distance, while further work or other complementary countermeasures is needed for tackling print attacks.

The rest of the article is organized as follows. First, in Section 2, we give a brief overview on the different approaches for face PAD proposed in the literature. Section 3 presents the different color texture descriptors studied in this work. The experimental setup is described in Section 4. Section 5 is devoted to the in-depth analysis, exploring the generalization problem across different conditions, and describing the newly proposed scheme along with a fair comparison against state of the art. Concluding remarks are drawn in Section 6.

2. Related work

There exists no unified taxonomy for the different face PAD approaches. In this article, we categorized the methods into two groups: hardware-based and software-based methods.

Hardware-based methods are probably the most robust ones for anti-spoofing because the dedicated sensors are able to directly capture or emphasize specific intrinsic differences between genuine and artificial faces in 3D structure [22,23] and (multi-spectral) reflectance [23-26] properties. For instance, planar PAI detection becomes rather trivial if depth information is available [22], whereas near-infrared or thermal cameras are efficient in display attack detection as most of the displays in consumer electronics emit only visible light. On the other hand, these kinds of unconventional sensors are usually expensive and not compact, thus not (yet) available in mobile devices, which prevents their wide deployment.

It would be rather appealing to perform face PAD by further analyzing only the same data that is used for the actual biometric purposes or additional data captured with the standard acquisition device. These kinds of software-based methods can be broadly divided into active (requiring user collaboration) and passive approaches. Additional user interaction can be very effectively used for face anti-spoofing because we humans tend to be interactive, whereas a photo or video-replay attack cannot respond to randomly specified action requirements. Furthermore, it is almost impossible to perform liveness detection or facial 3D structure estimation by relying only on spontaneous facial motion. Challenge-response based methods aim at performing face PAD detection based on whether the required action (challenge), e.g. facial expression [27,28], mouth movement 27,29] or head rotation (3D structure) [30-32], was observed within a predefined time window (response). Also, active software-based methods are able to generalize well across different acquisition conditions and attack scenarios but at the cost of usability due to increased authentication time and system complexity.

Ideally, passive software-based methods would be preferable for face PAD because they are faster and less intrusive than their active counterparts. Due to the increasing number of public benchmark databases, numerous passive software-based approaches have been proposed for face anti-spoofing. In general, passive methods based on analyzing different facial properties, like frequency content [33,34], texture [10,35-39] and quality [40-42], or motion cues, like eye blinking [43-46], facial expression changes [27,44-46], mouth movements [27,44-46], or even color variation due to blood circulation (pulse) [47], to discriminate face artifacts from genuine ones. Passive software-based methods have shown impressive results on the publicly available datasets but the preliminary cross-database tests, like [11,32], revealed that the performance is likely to degrade drastically when operating in unknown conditions.

Recently, the research focus on software-based face PAD has been gradually moving into assessing and improving the generalization capabilities of the proposed and existing methods in a cross-database setup instead of operating solely on single databases. Among handcrafted feature based approaches, image distortion analysis [14], combination of texture and image quality analysis with interpupillary distance (IPD) based reject option [19], dynamic spectral domain analysis [15] and pulse detection [48] have been applied in the context of generalized face anti-spoofing but with only moderate results.

The initial studies using deep CNNs have resulted in excellent intra-test performance but the cross-database results have still been unsatisfactory [12,13]. This is mainly due to the fact that the current publicly available dataset may not probably provide enough data for training well-known deep neural network architectures from scratch or even for fine-tuning pre-trained networks, thus the CNN models have been suffering from overfitting. In [20], deep dictionary learning based formulation was proposed to mitigate the requirement