Person-Specific Face Antispoofing With Subject Domain Adaptation

Jianwei Yang, Zhen Lei, *Member, IEEE*, Dong Yi, and Stan Z. Li, *Fellow, IEEE*

***Abstract*— Face antispoofing is important to practical face recognition systems. In previous works, a generic antispoofing classifier is trained to detect spoofing attacks on all subjects. However, due to the individual differences among subjects, the generic classifier cannot generalize well to all subjects. In this paper, we propose a person-specific face antispoofing approach. It recognizes spoofing attacks using a classifier specifically trained for each subject, which dismisses the interferences among subjects. Moreover, considering the scarce or void fake samples for training, we propose a subject domain adaptation method to synthesize virtual features, which makes it tractable to train well-performed individual face antispoofing classifiers. The extensive experiments on two challenging data sets: 1) CASIA and 2) REPLAY-ATTACK demonstrate the prospect of the proposed approach.**

***Index Terms*— Face anti-spoofing, person-specific, subject domain adaptation.**

I. Introduction

NOWADAYS, the security of face recognition systems is challenged. A face image printed on a paper can spoof the systems and then the access is granted to attackers. A practical face recognition system demands not only high recognition performance, but also the capability of anti-spoofing to differentiate faces from real persons (genuine face) and those from attackers (fake face).

Recently, many papers [1]–[5] on face anti-spoofing have been published. Meanwhile, competitions [6], [7] further promoted the development. In all these works, a generic classifier was used to detect spoofing attacks on all subjects. Because samples from different subjects have different distributions, it is difficult to obtain a generic classifier to well detect various spoofing attacks on all subjects. Fig. 1 illustrates the distributions of first three principal components of high quality samples of four subjects from CASIA dataset in the multi-scale local binary pattern (MsLBP) [1] feature space. It is shown that the locations and distributions of samples vary from subject to subject. The genuine samples of one subject overlap the fake samples of another subject. Therefore, it is hard to train a single anti-spoofing classifier which can perform well on all subjects. To address this problem, we propose a person-specific face anti-spoofing framework, in which each subject has a specific classifier for face anti-spoofing.

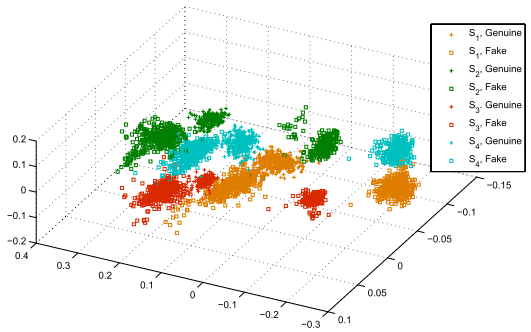


Fig. 1. High quality genuine and fake samples from four subjects in the CASIA dataset shown in the first three principal components of MsLBP feature space (better show when enlarged).

In person-specific anti-spoofing, the most challenging issue is the limited number of samples for training, especially the fake samples. In our framework, we develop anti-spoofing model for each enrolled subject specifically. However, many enrolled subjects have no fake samples in practice. To train face anti-spoofing models for these subjects, we propose a subject domain adaptation method to transfer the information provided by the subjects which have both genuine samples and fake samples (source subjects) to the subjects having no fake samples (target subjects) to synthesize fake samples. The proposed domain adaptation method is based on the assumption that the relation between genuine samples and that between fake samples of two subjects are both caused by the change of identity, and thus be similar mutually. The assumption is derived from our observation on the data. As shown in Fig. 1, one can see the relative location between fake features is similar to that between genuine ones from one subject to another. Based on this observation, we first estimate the relation between genuine samples from two subjects. By applying relation to the fake samples of source subjects, we can synthesize fake samples for target subjects. Once the virtual fake samples are obtained, person-specific anti-spoofing classifiers for the target subjects can be trained. In summary, there are mainly four contributions in our work:

- A person-specific face anti-spoofing approach combining face recognition and anti-spoofing is proposed. It conducts face anti-spoofing in a person-specific manner

Manuscript received June 9, 2014; revised October 28, 2014 and January 27, 2015; accepted January 31, 2015. Date of publication February 12, 2015; date of current version March 20, 2015. This work was supported in part by the National Natural Science Foundation of China under Project 61203267, Project 61375037, and Project 61473291, in part by the National Science and Technology Support Program under Project 2013BAK02B01, in part by the Chinese Academy of Sciences, Beijing, China, under Project KGZD-EW-102-2, and in part by the AuthenMetric Research and Development Funds. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kar-Ann Toh.

The authors are with the National Laboratory of Pattern Recognition, Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: yjwterry@163.com; zlei@cbsr.ia.ac.cn; dyi@cbsr.ia.ac.cn; stan.zq.li@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2015.2403306

according to the prior face identities, which dismisses the interferences among subject domains.
- A brief analysis on the relation among genuine (fake) samples from different subjects is first presented in the perspective of surface reflection model, which provides insights into the textured-based face anti-spoofing approaches.
- A domain adaptation method is proposed to synthesize virtual fake samples, so that the person-specific anti-spoofing classifiers can be trained for the target subjects whose fake samples are unavailable in training stage.
- A new anti-spoofing evaluation criterion is proposed, considering that the person-specific anti-spoofing needs to correctly identify the samples, and detect the spoofing attacks as well. It is different from existing evaluation methods for face anti-spoofing or recognition.

The remain of this paper is organized as follows. Sec. II presents the related works on face anti-spoofing and person-specific model. Sec. III describes the proposed framework briefly. In Sec. IV, we elaborate the proposed subject domain adaptation method. Sec. V illustrates the experimental evaluations, and then we draw the conclusion in Sec. VI.

II. Related Works

A. Face Anti-Spoofing

Existing face anti-spoofing approaches can be mainly categorized into four groups: texture based, motion based, 3D-shape based and multi-spectral reflectance based. Besides, some other works combined two or more of these methods to improve the anti-spoofing performance.

1) Texture-Based Anti-Spoofing: Li et al. [8] proposed a method based on the analysis of Fourier spectra. It assumed the photographs contained fewer high frequency components compared with genuine faces. In [3], Tan et al. used a variational retinex-based method and the difference-of-Gaussian (DoG) filers to extract latent reflectance features on face images, and then a sparse low rank bilinear discriminative model is trained for the classification. Inspired by Tan's work, Peixoto et al. [9] combined the DoG filters and standard Sparse Logistic Regression Model for anti-spoofing under extreme illuminations. After that, Määttä et al. [1] proposed to use LBP features for anti-spoofing, which outperformed previous methods on the NUAA Photograph Imposter Database [3]. Furthermore, the experiments in [10] showed its effectiveness on the REPLAY-ATTACK dataset. In [11], the authors proposed a component-dependent descriptor for face liveness detection, which account for different face parts in different way.

Pereira et al. [12] used a spatio-temporal texture feature for detecting the spoofing attacks. In their method, an operator called Local Binary Patterns from Three Orthogonal Planes (LBP-TOP) was proposed to combine spatial and temporal information into a single descriptor. According to the experimental results on the REPLAY-ATTACK dataset, it outperformed the method in [1]. In [4], it is shown that LBP and LBP-TOP features are applicable in intra-dataset protocol. However, the performance degraded much in a more realistic scenario, i.e., inter-dataset protocol. Komulainen et al. [13] proposed to detect the presence of spoofing medium in the observed scene based on HOG feature. In the 1^{st} competition on 2D face anti-spoofing [7], five out of six teams used textures in their methods. At most recent, seven out of eight teams used textures as clues for anti-spoofing in the 2^{nd} competition [6].

2) Motion-Based Anti-Spoofing: Motion-based approaches can be further divided into two categories: physiological responses based and physical motion based. The physiological responses, such as eye blinking, mouth (lip) movement, are important clues to verify the "liveness". Pan et al. used eye blinking for face anti-spoofing [2], [14]. In their method, a conditional random field was constructed to model different stages of eye blinking. In [15], Kollreider et al. used lip movement classification and lip-reading for the purpose of liveness detection. Furthermore, Chetty et al. [16], [17] proposed a multi-modal approach to aggrandize the difficulty of spoofing attacks. It determined the liveness by verifying the fitness between video and audio signals.

On the other hand, Bao et al. [18] presented a method using optical flow fields to distinguish 2-D planar photography attacks and 3-D real faces. Similarly, Kollreider et al. [15], [19] also relied their method on optical flow analysis. The method is based on the assumption that a 3-D face generates a special 2-D motion which is higher at central face parts (e.g. nose) compared to the outer face regions (e.g. ears). More recently, Anjos et al. proposed to recognize spoofing attacks based on the correlation between optical flows in two regions [20]. At the same time, Yang et al. presented a counter measure to replay attacks based on the correlations among optical magnitude/phase sequences from 11 regions, which won the first place after combining with a texture-based method [6]. Besides, Kollreider et al. [21] used eye-blinking and face movements for detecting liveness in an interaction scenario.

3) 3D Shape-Based Anti-Spoofing: In [22], Marsico et al. proposed a method for moving face anti-spoofing based on 3D projective invariants. However, this method can merely cope with photo attacks without warping, because the coplanar assumption is invalid for warped photos. In [23], the authors proposed to recover sparse 3D shapes for face images to detect various photo attacks. It is showed that the method worked perfectly under both intra-dataset protocols and inter-dataset protocols. However, both methods will fail when coping with 3D mask spoofing, such as the 3D Mask Attack dataset (3DMAD) collected by Erdogmus et al. [24].

4) Multi-Spectral Reflectance-Based Anti-Spoofing: The multi-spectral methods utilize the illuminations beyond visual spectrum to tackle spoofing attacks. In [25] and [26], the authors selected proper working spectrums so that the reflectance differences between genuine and fake faces increased. Different from the methods directly using reflection intensities, a gradient-based multi-spectral method for face anti-spoofing was proposed in [27]. These methods need extra devices to capture face images under the invisible lights, thus it is unreleastic to deploy such devices to the most of recent FR systems, which are merely based on RGB color face images.

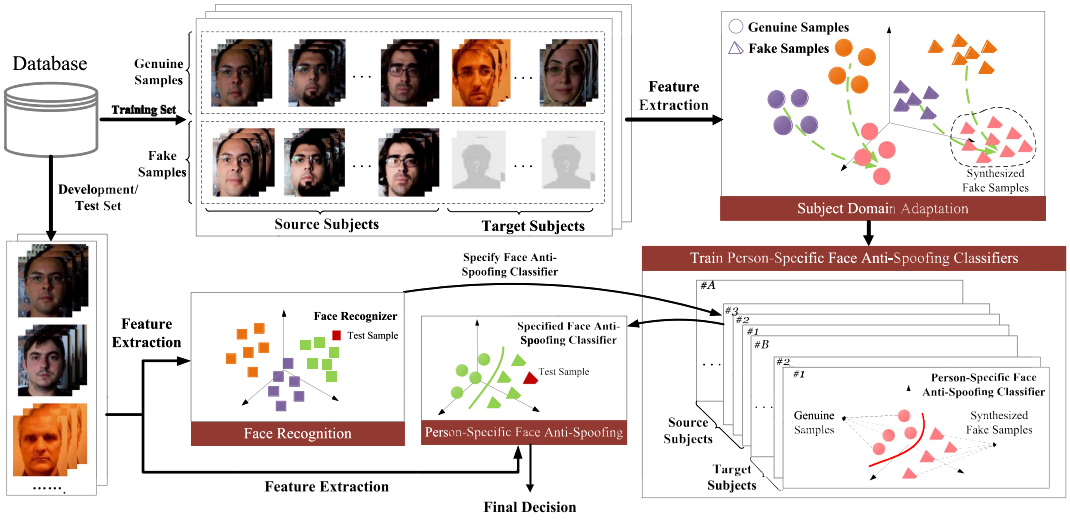


Fig. 2. The flowchart of person-specific face anti-spoofing approach.

Moreover, some works combined two or more of above four kinds of approaches [6], [7], [28]–[30]. Besides, Chingovska et al. proposed to integrate face recognition module into anti-spoofing system in score-level and feature level [31]. Their method is different from ours because we conduct face recognition to obtain prior knowledge for anti-spoofing, rather than fuse them together.

B. Person-Specific Model

Though the person-specific model has not been exploited in face anti-spoofing, it has been applied in many other face-oriented research areas, especially the face recognition issues [32]–[34]. Besides, to train person-specific age estimation model for subjects with limited samples, [35] exploited multi-task warped Gaussian process (MTWGP) to model common features and person-specific features separately. At the most recent, a person-specific expression recognition method based on domain adaptation was proposed in [36]. In their method, expression classifier for each subject was trained by combining the data in the target domain and classifiers trained in source domains. Similarly, Cao et al. [37] proposed to train person-specific face verification classifiers using joint bayesian method for subjects of interests with limited samples in cooperation with a domain adaptation algorithm.

These aforementioned person-specific methods demand target domain possess samples from all categories. In this paper, however, we need to tackle the problem that a number of subjects have *no fake samples*. This situation is a bit similar to the work in [38]. The authors proposed a face sketch synthesis algorithm to improve the performance of sketch recognition, and a more comprehensive version is presented in [39]. Their method is based on the assumption that a new face can be constructed from training samples by using principle component analysis (PCA), and the linear coefficients can be transferred to sketch faces. Though this method achieved good photo-sketch face recognition performance, it is defective when applied in face anti-spoofing. This failure will be illustrated in our experiments.

III. Person-Specific Face Anti-Spoofing

Fig. 2 shows the pipeline of person-specific face anti-spoofing. In the training stage, a unique face anti-spoofing model is trained for each subject. Specifically, for the source subjects, the classifiers are trained with available genuine and fake samples. For the target subjects whose fake samples are unobserved, we first use the proposed subject domain adaptation method to obtain virtual fake samples, and then train the classifiers. To obtain the identity of input face, we use an off-the-shelf face recognizer. Note that face recognition is not the focus of this paper. It is even valid to provide identities manually. In the test stage, the identity of a sample is obtained first and then the corresponding person-specific model is chosen to implement face anti-spoofing. Finally, a decision on its validity (genuine or fake) is given.

IV. Subject Domain Adaptation

The core idea of our domain adaptation method is to adapt fake samples of one subject to another whose fake samples are absolutely unobserved.

A. Assumptions

In our subject domain adaptation method, we make following two assumptions:

1) *Relation between genuine (or fake) samples in two subject domains is formulated as translation and linear transformation*;
2) *The relation between genuine samples is identical to that between fake samples of two subject domains.*

The first assumption provides convenience to derive the relation between two subject domains, and the second one makes it feasible to transfer the relation from genuine to fake samples and vice versa. In the following, we will analyze their reasonability mathematically and experimentally.

1) Mathematical Analysis: According to the Lambertian reflection model, the color intensity of an image is the pixel-wise product of reflectance and shading. Given two images

from two subjects, their intensities can be formulated as

$$\boldsymbol{I}_1 = \boldsymbol{S}_1\boldsymbol{R}_1, \quad \boldsymbol{I}_2 = \boldsymbol{S}_2\boldsymbol{R}_2 \tag{1}$$

where $\boldsymbol{S}$ is the shading, determined by the shape and direction of incident light (we assume it is achromatic); $\boldsymbol{R}$ is the surface reflectance, determined by subject identity. The fake face re-captured from above two genuine face images are formulated as

$$\boldsymbol{I}'_1 = \boldsymbol{S}'_1\boldsymbol{R}'_1, \quad \boldsymbol{I}'_2 = \boldsymbol{S}'_2\boldsymbol{R}'_2 \tag{2}$$

where $\boldsymbol{S}'$ is the shading condition under which the fake face image is captured; $\boldsymbol{R}'$ is reflectance of a fake face image. When displaying a genuine face images on a medium, e.g., paper, photo or screen, degradation is inevitable because of device noises. Here, we assume the genuine face image is polluted by additive noises $\boldsymbol{\eta}$. Then, the captured fake face images are

$$\boldsymbol{I}'_1 = \boldsymbol{S}'_1(\boldsymbol{S}_1\boldsymbol{R}_1 + \boldsymbol{\eta}_1), \quad \boldsymbol{I}'_2 = \boldsymbol{S}'_2(\boldsymbol{S}_2\boldsymbol{R}_2 + \boldsymbol{\eta}_2) \tag{3}$$

In the following, we call $\boldsymbol{S}$ the first-time shading, and $\boldsymbol{S}'$ the second-time shading. In practice, we observe the second-time shading $\boldsymbol{S}'$ is approximately consistent on whole image in the following cases:

- *Attacks by Planar Photography:* The shape over the whole photo is consistent, and thus the same to the shading $\boldsymbol{S}'$;
- *Attacks by Slightly Warped Photography:* In this case, the shape is smoothly changed. We assume the ambient illumination is approximately isotropic, and thus the shading is consistent over the whole face region;
- *Attacks by Electronic Screen:* Because the active light over the whole electronic screen is uniform, the corresponding shading is uniform over the whole face region as well.

Based on above observations, Eq. (3) is simplified to

$$\boldsymbol{I}'_1 = c_1(\boldsymbol{S}_1\boldsymbol{R}_1 + \boldsymbol{\eta}_1), \quad \boldsymbol{I}'_2 = c_2(\boldsymbol{S}_2\boldsymbol{R}_2 + \boldsymbol{\eta}_2) \tag{4}$$

where c_1 and c_2 are two const scalars. In Eq. (4), c_1 and c_2 can be naturally eliminated by normalizing operations. Then,

$$\boldsymbol{I}'_1 = (\boldsymbol{S}_1\boldsymbol{R}_1 + \boldsymbol{\eta}_1), \quad \boldsymbol{I}'_2 = (\boldsymbol{S}_2\boldsymbol{R}_2 + \boldsymbol{\eta}_2) \tag{5}$$

Based on Eq. (1) and 5, we have $\boldsymbol{I}_1 - \boldsymbol{I}_2 = \boldsymbol{I}'_1 - \boldsymbol{I}'_2 = \boldsymbol{S}_1\boldsymbol{R}_1 - \boldsymbol{S}_2\boldsymbol{R}_2$ when $\boldsymbol{\eta}_2$ is similar to $\boldsymbol{\eta}_1$. Fortunately, this condition is valid when fake samples are captured with similar spoofing types. So far, we have shown our assumptions are reasonable under some moderate conditions. In the image space, however, the relation is vulnerable to many factors, such as misalignment, expression, pose, etc. A feasible way to address this problem is representing the face images by extracting more robust texture features. In the following, we will show its experimental validity based on texture representation.

2) Quantitative Analysis: Texture feature extraction is often non-linear, which cannot be formulated explicitly. Fortunately, such non-linear property does not affect the assumed relations much in the used texture feature space. We evaluate our assumptions using MsLBP and HOG features. Firstly, we extract features from all samples, and then compute the centers of genuine and fake samples for each subject. After that, we compute the vectors pointing from the center of one subject to another for genuine and fake samples, respectively. For convenience, we call them directional vectors. For each subject pair, we compute the cosine similarity between the directional vectors. In Fig. 3, we show the histograms of cosine similarities for MsLBP and HOG features on the CASIA and REPLAY-ATTACK datasets. On the CASIA dataset, the average cosine similarities are 0.885 and 0.906 for MsLBP and HOG features, respectively; On the REPLAY-ATTACK dataset, they are 0.833 and 0.797, respectively. High similarities pose great support to our assumptions and are the basis of the proposed subject domain adaptation method. Note that for the CASIA dataset, the directional vectors between subjects are computed using high-quality samples. In this dataset, the fake samples of all quality types (high, normal and low) are captured from photographs, electronic displaying of high quality face images (as stated in [40]). In other words, all the fake samples have similar first-time shading to the high quality genuine samples, rather than the genuine samples with the same quality type. In contrast, the fake samples in the REPLAY-ATTACK dataset are captured from the genuine samples with all kind of illuminations (as stated in [10]), which means that the fake samples have similar first-time shading to the genuine samples. However, the given face locations in this dataset are not aligned well according to the eye positions, which lower the similarities to some extent.

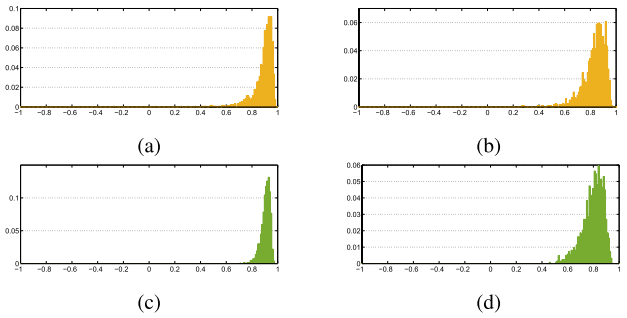


Fig. 3. Histograms of cosine similarities between directional vectors computed for MsLBP and HOG on two datasets. (a) MsLBP on CASIA. (b) MsLBP on REPLAY-ATTACK. (c) HOG on CASIA. (d) HOG on REPLAY-ATTACK.

B. Subject Domain Adaptation

In this part, we will elaborate how to adapt fake samples from one subject domain to another in the feature space. Denote n_a and n_b features extracted from genuine samples of subject a and b by $\boldsymbol{G}_a = \{\boldsymbol{g}_a^1, \boldsymbol{g}_a^2, \ldots, \boldsymbol{g}_a^{n_a}\} \in \mathbb{R}^{d\times n_a}$ and $\boldsymbol{G}_b = \{\boldsymbol{g}_b^1, \boldsymbol{g}_b^2, \ldots, \boldsymbol{g}_b^{n_b}\} \in \mathbb{R}^{d\times n_b}$, respectively. The fake features extracted from subject a are denoted by $\boldsymbol{F}_a = \{\boldsymbol{f}_a^1, \boldsymbol{f}_a^2, \ldots, \boldsymbol{f}_a^{n_a}\} \in \mathbb{R}^{d\times n_a}$. Based on our assumptions, given features $\boldsymbol{g}_a^i$, $\boldsymbol{g}_b^j$, and $\boldsymbol{f}_a^i$ captured from $\boldsymbol{g}_a^i$, we have

$$\boldsymbol{f}_b^j = \boldsymbol{f}_a^i + (\boldsymbol{g}_b^j - \boldsymbol{g}_a^i) \tag{6}$$

Ideally, the synthesized feature $\boldsymbol{f}_b^j$ has the same first-time shading to $\boldsymbol{g}_b^j$. To synthesize more reliable feature, we take all

samples from subject a into account:

$$\boldsymbol{f}_b^j = \frac{1}{n_a}\sum_{i=1}^{n_a}(\boldsymbol{f}_a^i + (\boldsymbol{g}_b^j - \boldsymbol{g}_a^i)) \tag{7}$$

Further averaging the virtual features synthesized for subject b

$$\frac{1}{n_b}\sum_{j=1}^{n_b}\boldsymbol{f}_b^j = \frac{1}{n_b}\sum_{j=1}^{n_b}\frac{1}{n_a}\sum_{i=1}^{n_a}(\boldsymbol{f}_a^i + (\boldsymbol{g}_b^j - \boldsymbol{g}_a^i)) \tag{8}$$

Eq. (8) can be re-formulated to

$$\hat{\boldsymbol{F}}_b = \hat{\boldsymbol{F}}_a + (\hat{\boldsymbol{G}}_b - \hat{\boldsymbol{G}}_a) \tag{9}$$

where $\hat{\boldsymbol{G}}_a$, $\hat{\boldsymbol{G}}_b$, $\hat{\boldsymbol{F}}_a$ and $\hat{\boldsymbol{F}}_b$ are the centers of feature groups. Eq. (9) expresses a very simple domain adaptation algorithm. However, such a simple translation may not suffice to model the complicated relation between two subject domains in practice. In Eq. (9), all genuine samples of one subject are averaged for the calculation of translation. In this case, some outlier samples caused by unexpected factors, e.g., inaccurate face locations and extreme illumination conditions may affect the final domain adaptation result. Therefore, we attempt to find and match those non-outlier features of two subjects before domain adaptation. Moreover, to cope with more complicated situations, the domain adaptation should consider not only translation, but also the transformation as well. With this goal in mind, we model the relation from $\boldsymbol{G}_a$ to $\boldsymbol{G}_b$ in a more general formula

$$\boldsymbol{G}_b = \boldsymbol{H}_{ab}\boldsymbol{G}_a\boldsymbol{P}_{ab} + \boldsymbol{T}_{ab}, \quad if\ n_a \geq n_b \tag{10}$$

$$\boldsymbol{G}_b\boldsymbol{P}_{ba} = \boldsymbol{H}_{ab}\boldsymbol{G}_a + \boldsymbol{T}_{ab}, \quad if\ n_a < n_b \tag{11}$$

where $\boldsymbol{H} \in \mathbb{R}^{d\times d}$ is a transformation matrix; $\boldsymbol{T}$ is a bias matrix modelling the translation, whose columns are all equal to $\boldsymbol{t}_{ab}$; $\boldsymbol{P}$ is the correspondence matrix to label matchings among samples. For the case $n_a \geq n_b$ in Eq. (10), $\boldsymbol{T}_{ab} \in \mathbb{R}^{d\times n_b}$ and $\boldsymbol{P}_{ab} \in \{0,1\}^{n_a\times n_b}$; For the case $n_a < n_b$ in Eq. (11), $\boldsymbol{T}_{ab} \in \mathbb{R}^{d\times n_a}$ and $\boldsymbol{P}_{ba} \in \{0,1\}^{n_b\times n_a}$. In above two equations, $\boldsymbol{H}$ is used to model transformation between $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$, and can further reduce the residual. $\boldsymbol{P}$ plays two roles: (1) removing the outliers in $\boldsymbol{G}_a$ or $\boldsymbol{G}_b$; (2) matching features of one subject to another. To synthesize virtual fake features, we first derive $\boldsymbol{H}_{ab}$, $\boldsymbol{T}_{ab}$ given $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$ based on Eq. (10) or Eq. (11), and then derive $\boldsymbol{F}_b$ by transferring the relation hold by genuine samples to fake samples

$$\boldsymbol{F}_b = \boldsymbol{H}_{ab}\boldsymbol{F}_a + \boldsymbol{T}_{ab} \tag{12}$$

Note that the column size of $\boldsymbol{T}_{ab}$ in above equation is equal to that of $\boldsymbol{F}_a$. If not mentioned, we consider the case $n_a \geq n_b$ for the sake of clarification in the following.

1) Domain Adaptation Algorithm: We update the correspondences and relations between subjects in an iterative manner, which is shown in Alg. 1. In the algorithm, there are two procedures need to be clarified: (1) updating correspondence $\boldsymbol{P}_{ab}$ between $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$; (2) deriving the optimal relation (translation and transformation) from $\boldsymbol{G}_a$ to $\boldsymbol{G}_b$ given $\boldsymbol{P}_{ab}$.

Update Correspondence: To get the optimal correspondences means to match features of one subject to another so that the outliers are excluded. At this point, we define the loss function $D(\boldsymbol{P}_{ab})$ as

$$D(\boldsymbol{P}_{ab}) = \sum_i\sum_j P_{ab}^{ij}E_{ab}^{ij} \tag{13}$$

where $\boldsymbol{E}$ is a proximity matrix, whose entity is the similarity of two feature sample from two subject domains and is formulated as

$$E_{ab}^{ij} = \exp(-\frac{\|g_a^i - g_b^j\|_2^2}{2\sigma^2}) \tag{14}$$

The optimal correspondences can be obtained by maximizing D given current $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$. Obviously, above maximization can be achieved by singular-value-decomposition (SVD), which is earlier used in [41]. Briefly, $\boldsymbol{P}$ is first decomposed into $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$. Replacing $\boldsymbol{\Sigma}$ by a identity matrix $\boldsymbol{\Sigma}_I$, we can obtain the optimal correspondences by finding the largest entries in rows and columns in matrix $\boldsymbol{U}\boldsymbol{\Sigma}_I\boldsymbol{V}^T$.

Estimating Relation: We use three methods to estimate $\boldsymbol{H}_{ab}$ and $\boldsymbol{T}_{ab}$. Given the correspondence matrix $\boldsymbol{P}^*$, we formulate the loss function as

$$L(\boldsymbol{H}_{ab}, \boldsymbol{T}_{ab}, \boldsymbol{P}_{ab}^*) = \frac{1}{2}||\boldsymbol{G}_b - \boldsymbol{H}_{ab}\boldsymbol{G}_a\boldsymbol{P}_{ab}^* - \boldsymbol{T}_{ab}||_F^2 \tag{15}$$

which is minimized at line 5 in Alg. 1.

Center Shift (CS): It assumes that the transformation $\boldsymbol{H}$ to be identity matrix. In this case, Eq. (15) becomes

$$L(\boldsymbol{H}_{ab}, \boldsymbol{T}_{ab}, \boldsymbol{P}_{ab}^*) = \frac{1}{2}||\boldsymbol{G}_b - \boldsymbol{G}_a\boldsymbol{P}_{ab}^* - \boldsymbol{T}_{ab}||_F^2 \tag{16}$$

In Eq. (16), we only need to estimate $\boldsymbol{T}_{ab}$ given $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$. The optimal solution for $\boldsymbol{T}_{ab}$ in current iteration is obtained by computing the center shift vector from the center of $\boldsymbol{G}_a\boldsymbol{P}_{ab}^*$ to that of $\boldsymbol{G}_b$.

As mentioned earlier, the CS method cannot cope with complicated transformation between subjects. In the following, we propose two other methods to estimate both translation and transformation between subjects.

Algorithm 1 Estimating Relation Between Two Subject Domains

Input:
$\boldsymbol{G}_a$: = feature collection of genuine samples of subject a;
$\boldsymbol{G}_b$: = feature collection of genuine samples of subject b;
$\boldsymbol{P}_{ab}^0$: = random correspondence between $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$;
Output:
$(\dot{\boldsymbol{H}}_{ab}, \dot{\boldsymbol{T}}_{ab})$: = final domain relation;
1: $\dot{\boldsymbol{H}}_{ab} \leftarrow \boldsymbol{I}$, $\dot{\boldsymbol{T}}_{ab} \leftarrow \boldsymbol{0}$
2: $L_p \leftarrow L(\dot{\boldsymbol{H}}_{ab}, \dot{\boldsymbol{T}}_{ab}, \boldsymbol{P}_{ab}^0)$
3: **repeat**
4: $\quad\boldsymbol{P}_{ab}^* \leftarrow \arg_{\boldsymbol{P}_{ab}}\max D(\boldsymbol{P}_{ab})$
5: $\quad(\boldsymbol{H}_{ab}^*, \boldsymbol{T}_{ab}^*) \leftarrow \arg_{(\boldsymbol{H}_{ab}, \boldsymbol{T}_{ab})}\min L(\boldsymbol{H}_{ab}, \boldsymbol{T}_{ab}, \boldsymbol{P}_{ab}^*)$
6: $\quad L_c \leftarrow L(\boldsymbol{H}_{ab}^*, \boldsymbol{T}_{ab}^*, \boldsymbol{P}_{ab}^*)$
7: $\quad\dot{\boldsymbol{H}}_{ab} \leftarrow \boldsymbol{H}_{ab}^*\dot{\boldsymbol{H}}_{ab}$
8: $\quad\dot{\boldsymbol{T}}_{ab} \leftarrow \boldsymbol{H}_{ab}^*\dot{\boldsymbol{T}}_{ab} + \boldsymbol{T}_{ab}^*$
9: $\quad\boldsymbol{G}_a \leftarrow \boldsymbol{H}_{ab}^*\boldsymbol{G}_a + \boldsymbol{T}_{ab}^*$
10: $\quad\Delta \leftarrow |L_c - L_p|, L_p \leftarrow L_c$
11: **until** $\Delta < \varepsilon$ or iteration times exceed a threshold

Ordinary Least Square (OLS) Regression: Firstly, $\boldsymbol{G}_a$ and $\boldsymbol{G}_b$ are centralized to $\bar{\boldsymbol{G}}_a$ and $\bar{\boldsymbol{G}}_b$, respectively. Then, we reformulate the loss in Eq. (15) to

$$L = \frac{1}{2}||\bar{\boldsymbol{G}}_b - \boldsymbol{H}'_{ab}\boldsymbol{X}||_F^2 \tag{17}$$

where $\boldsymbol{H}'_{ab} = [\boldsymbol{H}_{ab}, \boldsymbol{t}_{ab}]$; $\boldsymbol{X} = \left[\bar{\boldsymbol{G}}_a \boldsymbol{P}^*_{ab}; \boldsymbol{1}^T\right]$. Eq. (17) is a typical least square error problem. It has a closed-form solution $\bar{\boldsymbol{G}}_b \boldsymbol{X}^T(\boldsymbol{X}\boldsymbol{X}^T)^{-1}$. In some cases, the feature dimension may be larger than the number of matched feature samples, and $\boldsymbol{X}\boldsymbol{X}^T$ is thus not invertible. To solve this problem, we compute the pseudo-inverse of $\boldsymbol{X}\boldsymbol{X}^T$ alternatively.

Partial Least Squares (PLS) Regression [42]: Besides the OLS algorithm, we also use PLS to minimize the loss function in Eq. (17). Different from OLS regression, the PLS regression algorithm is implemented in a subspace, rather than the original feature space. Similarly, we obtain the centralized features $\bar{\boldsymbol{G}}_a$ and $\bar{\boldsymbol{G}}_b$ before regression. According to [42], to model the relation between $\bar{\boldsymbol{G}}_a \boldsymbol{P}^*_{ab}$ and $\bar{\boldsymbol{G}}_b$, we decompose them into

$$\begin{aligned}(\boldsymbol{P}^*_{ab})^T \bar{\boldsymbol{G}}_a^T &= \boldsymbol{U}_a \boldsymbol{R}_a^T + \boldsymbol{E}_a \\ \bar{\boldsymbol{G}}_b^T &= \boldsymbol{U}_b \boldsymbol{R}_b^T + \boldsymbol{E}_b\end{aligned} \tag{18}$$

where $\boldsymbol{U}_a \in \mathbb{R}^{K\times d'}$ and $\boldsymbol{U}_b \in \mathbb{R}^{K\times d'}$ are the score matrices; $\boldsymbol{R}_a$ and $\boldsymbol{R}_b$ are $d \times d'$ are loadings; K is the number of matched feature samples. For regression, we use the PLS2 form introduced in [42]. Specifically, it assumes there is a linear relation between score matrices $\boldsymbol{U}_a$ and $\boldsymbol{U}_b$, i.e., $\boldsymbol{U}_b = \boldsymbol{U}_a \boldsymbol{D}_{ab} + \boldsymbol{L}_{ab}$. Based on this linear relation, we yield the transformation

$$\bar{\boldsymbol{G}}_b^T = \boldsymbol{U}_a \boldsymbol{D}_{ab} \boldsymbol{R}_b^T + \boldsymbol{L}_{ab} \boldsymbol{R}_b^T + \boldsymbol{E}_b \tag{19}$$

Based on the deductions in [42] and [43], above relation can be further formulated as

$$\bar{\boldsymbol{G}}_b^T = (\bar{\boldsymbol{G}}_a^*)^T \boldsymbol{C}_{ab} + \boldsymbol{F} \tag{20}$$

where

$$\boldsymbol{C}_{ab} = \bar{\boldsymbol{G}}_a^* \boldsymbol{U}_b \left(\boldsymbol{U}_a^T (\bar{\boldsymbol{G}}_a^*)^T \bar{\boldsymbol{G}}_a^* \boldsymbol{U}_b\right)^{-1} \boldsymbol{U}_a^T \bar{\boldsymbol{G}}_b^T \tag{21}$$

$$\boldsymbol{F}^* = \boldsymbol{L}_{ab} \boldsymbol{R}_b^T + \boldsymbol{E}_b \tag{22}$$

and $\boldsymbol{G}_a^* = \bar{\boldsymbol{G}}_a \boldsymbol{P}^*_{ab}$. In Eq. (21) and (22), the intermediate variables are obtained in an iterative manner. More details are presented in [42]. In the iteration process, we set the dimension of subspace $d' = \min(50, K/3)$.

Upon above three domain adaptation methods, we estimate the relation between subjects iteratively. As presented in Alg. 1, the final transformation matrix $\dot{\boldsymbol{H}}_{ab}$ and $\dot{\boldsymbol{T}}_{ab}$ are first initialized to be identity matrix and zero matrix, respectively. Then, in each iteration, the matched feature samples are obtained via $\max D(\boldsymbol{P})$. Given the optimal correspondence in current iteration, the corresponding optimal transformation $\boldsymbol{H}^*_{ab}$ and translation $\boldsymbol{T}^*_{ab}$ are obtained by minimizing $L(\boldsymbol{H}_{ab}, \boldsymbol{T}_{ab}, \boldsymbol{P}^*_{ab})$, which are then used to update $\dot{\boldsymbol{H}}_{ab}$ and $\dot{\boldsymbol{T}}_{ab}$ incrementally (line 7 and 8). Meanwhile, the features from subject a are updated iteratively (line 9). The iteration suspends when the change of error Δ (computed at line 10) is smaller than a predefined value ϵ or iteration times is larger than a value (line 11). In our experiment, we set ϵ to be $1e^{-5}$ and the maximal iteration times as 5.

Algorithm 2 Synthesize Fake Features for Target Subjects

Input:
$\boldsymbol{F}^{Sr}_{\{1,..,A\}}$: = fake features from A source subjects;
$\boldsymbol{G}^{Tr}_{\{1,..,B\}}$: = genuine feature from B target subjects;
$\{(\dot{\boldsymbol{H}}_{ab}, \dot{\boldsymbol{T}}_{ab})\}$: = domain relation from A source subjects to B target subjects;
$\{\dot{\boldsymbol{T}}^{CS}_{ab}\}$: = translation from A source subjects to B target subjects learned from CS algorithm.

Output:
$\boldsymbol{F}^{Tr}_{\{1,..,B\}}$: = synthesized fake features for B target subjects;

$\boldsymbol{F}^{Tr}_b \leftarrow \emptyset, b \in \{1, .., B\}$
1: **for** b from 1 to B **do**
2: **for** a from 1 to A **do**
3: $\boldsymbol{D}^{Tr}_{ab} = (\dot{\boldsymbol{H}}_{ab}\boldsymbol{F}^{Sr}_a + \dot{\boldsymbol{T}}_{ab}) + \dot{\boldsymbol{T}}^{CS}_{ab}$
4: **if** $\Phi(\boldsymbol{G}^{Tr}_b, \boldsymbol{D}^{Tr}_{ab}) > \tau$ **then**
5: $\boldsymbol{F}^{Tr}_b \leftarrow \boldsymbol{F}^{Tr}_b \cup \boldsymbol{D}^{Tr}_{ab}$
6: **end if**
7: **end for**
8: **end for**

2) Feature Synthesis: After obtaining the relation between source subjects and target subjects, we can synthesize fake features for target subjects based on our assumption. If more than one source subjects exist, we combine the synthesized fake features from all source subjects. Assume we have A source subjects, and B target subjects. $\boldsymbol{D}^{Tr}_{ab}$ is the synthesized fake features transferred from $a^{th}, a \in \{1, \ldots, A\}$ source subject for $b^{th}, b \in \{1, \ldots, B\}$ target subject. $\boldsymbol{T}^0_{ab}$ is the center-shift vector from $\boldsymbol{G}^{Sr}_a$ to $\boldsymbol{G}^{Tr}_b$. Due to the noises in practice, we add the synthesized fake features only if the minimal distance from features in $\boldsymbol{G}^{Tr}_b$ to those in $\boldsymbol{D}^{Tr}_b$ is larger than a pre-determined value τ, which is formulated by $\Phi(\boldsymbol{G}^{Tr}_b, \boldsymbol{D}^{Tr}_{ab}) > \tau$. In our experiment, τ is tuned using development set. Finally, to retain the balance between genuine and fake features, we uniformly select $1/A$ number of features from $\boldsymbol{F}^{Tr}$ for each target subject. The synthesis algorithm is shown in Alg. 2. In the algorithm, $\dot{\boldsymbol{H}}$ is an identity matrix, and $\dot{\boldsymbol{T}} = \boldsymbol{0}$ for CS method.

V. Experiment

In this section, we evaluate the performance of person-specific face anti-spoofing method, compared with generic one. To train the person-specific models, we use the aforementioned three domain adaptation methods for feature synthesis. We also implement the PCA-based algorithm proposed in [38] for comparison. In the PCA-based feature synthesis, 98% energy is preserved in PCA dimensionality reduction. Moreover, we train generic face ant-spoofing models with synthesized features for comparison. In this paper, the MsLBP are extracted in the same way as in [1], whose dimension is 833. For HOG feature extraction, we divide the face image into 6×7 cells with the size of 20×20. Without local

TABLE I
PERFORMANCE EVALUATION CRITERION FOR rFAS

		Output		
	Validity	Genuine		Fake
Input	Genuine	$k_{pred} = k_{gt}$	True Positive	False Negative
		$k_{pred} \neq k_{gt}$	False Negative	
	Fake	False Positive		True Negative

normalization, the dimension is $42 \times 9 = 378$. We use linear support vector machine (SVM) to train the classifiers.

For person-specific face anti-spoofing, we consider two implementation scenarios:

- *Face Anti-Spoofing in Ideal Situation:* In this scenario, we assume an ideal situation where all test samples are correctly identified by an user or a perfect face recognizer. We denote it by **PS-iFAS**.
- *Face Anti-Spoofing in Realistic Situation:* We consider a realistic situation where the face recognizer is defective somewhat for person-specific face anti-spoofing, which means that some samples may be wrongly identified. In our experiments, we train a generic face recognition engine using the method in [44]. For simplicity, it is denoted by **PS-rFAS**.

A. Performance Metric

In this paper, we set up a new criterion to evaluate the performance of person-specific face anti-spoofing. The new criterion takes both face anti-spoofing and recognition results into account. In Table I, k_{gt} is the ground-truth identity of input face image, and k_{pred} is the predicted identity. Accordingly, false negative occurs when a genuine face image with identity k_{gt} is wrongly recognized, or detected as a fake face. The false positive refers to the case where a fake face image is predicted as a genuine sample.

To keep consistent to previous works, we use HTER as the metric in our experiments. Specifically, we first find the operating point where the false rejection rate (FRR) is equal to false acceptance rate (FAR) on the development set. Then, the threshold θ corresponding to the operating point is used to calculate the HTER on the test set. By far, another difference arises between the person-specific and generic models. In the person-specific model, a threshold is associated with each subject, and thus the HTER is calculated for each subject separately. However, for generic face anti-spoofing, a single threshold is obtained from the development set and then applied to the test set to calculate a single HTER. To compare the subject-wise performance with person-specific model, we additionally compute the HTERs for subjects separately using the single threshold for generic anti-spoofing model.

B. Protocol

1) Dataset: To prove the efficiency of the proposed approach, we conduct the experiments on two up-to-date publicly available datasets, CASIA [40] and REPLAY-ATTACK [10]. Followings are the brief introductions of two datasets:

TABLE II
DATA ORGANIZATION FOR CASIA DATASET

Domain	Train		Devel/Test	
	Genuine	Fake	Genuine	Fake
Source	Nor-Q (1-20)	Nor-Q (1-20)	Low-Q (1-20)	Low-Q (1-20)
Target	Nor-Q (21-50)	N/A	Low-Q (21-50)	Low-Q (21-50)

TABLE III
DATA ORGANIZATION FOR REPLAY-ATTACK DATASET

Domain	Train		Devel/Test	
	Genuine	Fake	Genuine	Fake
Source	En-Set (1-15)	FixedN (1-15)	Tr-Set (1-15)	HandN (1-15)
Target	En-Set (16-50)	N/A	Dev-Set Te-Set (16-50)	HandN (16-50)

- *CASIA Dataset [40]:* This dataset contains 50 subjects in total. It covers three kinds of imaging qualities (high, normal and low qualities) and three kinds of attacks. For each subject, the genuine faces are captured under three quality conditions. The spoofing faces are fabricated by implementing three kind of spoofing attacks, i.e., warped photo attack, cut photo attack and electronic screen attack in three qualities, respectively. As a result, for each subject, the dataset contains 12 sequences (3 genuine and 9 fake ones). The overall number of sequences in the dataset is 600.
- *REPLAY-ATTACK Dataset [10]:* It also contains 50 subjects. For each subject, four genuine video sequences are collected in front of adverse and controlled backgrounds. As for fake samples, three spoofing types are used, including print attack, digital photo attack, and video attack. The spoofing sequences are captured from hand-hold and fixed support mediums. Besides, two extra real-access sequences are collected as enrollment data for each subject. As a result, the overall number of sequences in the dataset is 1300.

2) Organization: To make it compatible to person-specific models, we re-organize the data in two datasets:

- *CASIA:* We divide the data into three parts according to the image quality: high, normal and low. The high quality genuine video sequences are used to estimate the relations between source and target subjects; The normal quality sequences are used to train person specific models and low quality sequences are equally divided into two parts for development and test, respectively. As shown in Table II, 'Nor-Q' means normal quality samples, and 'Low-Q' means low quality samples. To evaluate the domain adaptation performance, we use the first 20 subjects in the dataset as source subjects, and the remaining 30 subjects as target subjects, whose fake samples are assumed to be unavailable during model training.
- *REPLAY-ATTACK:* As shown in Table III, we use the genuine samples in the enrollment set to learn the relations between source subjects and target subjects, and train face anti-spoofing models by combining them with

TABLE IV
EERs OF G-FAS AND PS-iFAS ON THE DEVELOPMENT SET OF THE CASIA DATASET

Feature	Set	G-FAS					PS-iFAS			
		Orig	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS
MsLBP	*Devel-S*	6.82	6.83	**6.59**	6.72	7.58	**4.38**	**4.38**	**4.38**	**4.38**
	Devel-T	17.66	17.51	17.46	15.68	**15.40**	2.25	**1.18**	1.66	1.69
	Devel	13.01	12.84	12.70	**11.73**	11.77	3.10	**2.46**	2.74	2.76
HOG	*Devel-S*	**2.10**	2.19	3.12	2.92	2.79	**0.07**	**0.07**	**0.07**	**0.07**
	Devel-T	16.87	16.71	**15.65**	16.25	17.06	6.96	2.04	2.13	**1.69**
	Devel	11.63	11.57	**11.20**	11.27	11.88	4.20	1.25	1.31	**1.04**

TABLE V
HTERs OF G-FAS, PS-iFAS AND PS-rFAS ON THE TEST SET OF THE CASIA DATASET

Feature	Set	FR	G-FAS					PS-iFAS				PS-rFAS			
			Orig	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS
MsLBP	*Test-S*	99.63	7.54	**7.31**	7.99	8.02	8.55	**5.60**	**5.60**	**5.60**	**5.60**	5.62	**5.59**	5.63	5.63
	Test-T	93.40	17.49	17.54	17.77	16.46	**16.13**	2.52	**2.00**	2.72	2.76	5.42	**4.52**	5.18	5.25
	Test	95.86	13.03	12.90	13.29	12.61	**12.50**	3.75	**3.44**	3.87	3.90	5.50	**4.95**	5.36	5.40
HOG	*Test-S*	99.63	**3.52**	3.54	4.37	4.71	4.40	**0.82**	**0.82**	**0.82**	**0.82**	0.97	**0.89**	0.91	0.90
	Test-T	93.40	16.00	15.77	**14.95**	15.45	15.77	6.83	3.26	3.36	**2.80**	10.80	7.17	7.24	**6.66**
	Test	95.86	11.42	11.30	**11.16**	11.34	11.54	4.43	2.28	2.34	**2.01**	6.87	4.66	4.71	**4.35**

fake samples captured via fixed support. The remaining genuine samples in the dataset, and fake samples captured with hand-hold medium are divided into two parts for development and test, respectively. For domain adaptation, 15 subjects are used as source subjects, and the other 35 subjects as target subjects.

C. Experiments on the CASIA Dataset

In this part, we compare the performance of generic and person-specific face anti-spoofing on the CASIA dataset. The feature synthesis is based on Alg. 1 and 2 introduced in Sec. IV. Specifically, we first estimate the relations between extracted features from genuine samples of 20 source subjects to 30 target subjects based on Alg. 1. Then, such relations are applied to synthesize fake features based on Alg. 2. As mentioned in Sec. IV, the relations are obtained from high quality genuine samples considering they have similar first-time reflection to fake samples. Then we transfer them to normal quality to synthesize normal quality features for fake samples of target subjects, followed by the training of anti-spoofing models. For convenience, we denote three domain adaptation methods proposed in this paper by CS, OLS and PLS, and the one in [38] by PCA. The generic model is trained using all genuine and fake samples available in the training set, which is defined as *original generic model*. Besides, we train four other generic models by augmenting the training set with synthesized features of fake samples, which we call *augmented generic model* in the following. As for person-specific model, we obtain a unique face anti-spoofing classifier for each subject. Among them, 20 classifiers are trained with available genuine and fake samples for source subjects, and the remaining 30 classifiers are trained with genuine features and synthesized features for fake samples of target subjects separately. In the following, the generic anti-spoofing model is denoted by *G-FAS*, and the ideal and realistic person-specific models are denoted by *PS-iFAS*, *PS-rFAS*, respectively.

With the anti-spoofing classifiers obtained in the training stage, we use the development set to compute the equal error rate (EER) and the corresponding thresholds. To show the effects of person-specific model and subject domain adaptation separately, we divide the development set *Devel* into two subsets, *Devel-S* and *Devel-T*, which contains samples from source subjects and target subjects, respectively. For the generic model, we first obtain the EER and threshold on the whole *Devel* set, and then use the threshold to compute the HTERs for all subjects separately. Then, these HTERs are averaged on *Devel-S* and *Devel-T*. For consistency, we also call them EERs. As to the person-specific model, the EER is computed on each subject, and then averaged on *Devel-S*, *Devel-T*, and *Devel* sets, respectively.

1) Results on the Development Set: The EERs on the development set are reported in Table IV. For good presentation, we bold the minimal errors for different models. In the table, "Orig" represents the original generic models. As we can see, the proposed PS-iFAS methods achieve lower EERs than G-FAS methods consistently over two feature types and feature synthesis methods, indicating the effectiveness of person-specific methods. On the other hand, the superior performance of PS-iFAS and augmented generic models on the *Devel-T* set also proves the effectiveness of domain adaptation methods for feature synthesis. Among all domain adaptation methods, we can observe that the proposed subject domain adaptation methods achieve better performance for both G-FAS and PS-iFAS models compared with the PCA-based feature synthesis.

2) Results on the Test Set: After obtaining the thresholds on the development set, we evaluate the performance on test set. Similarly, the test set *Test* is also split into *Test-S* and *Test-T*. The HTERs are reported in Table V. In the table, we compare the performances of generic models, PS-iFAS models and PS-rFAS models. In PS-rFAS, we use the genuine samples in the training set as gallery set. The third column in Table V

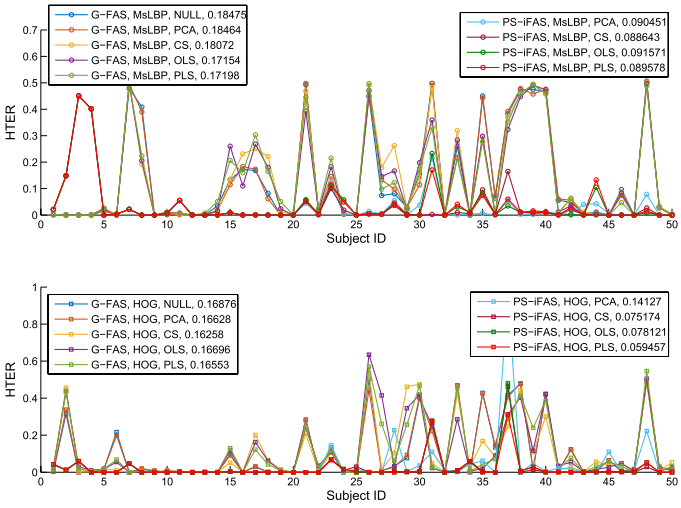
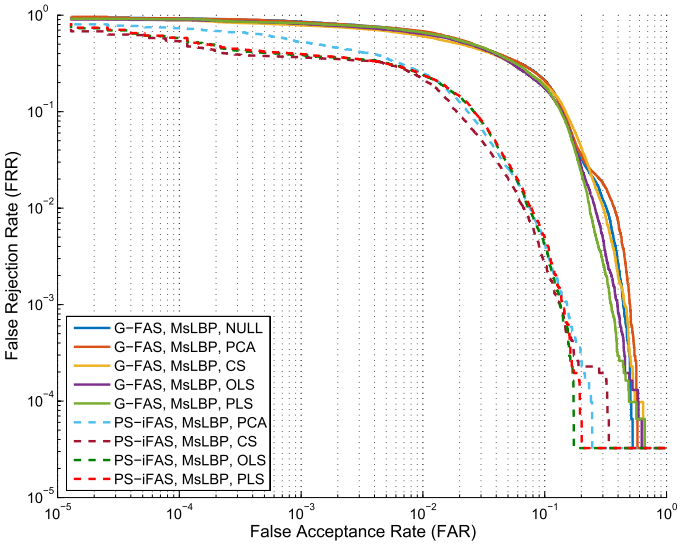


Fig. 4. HTER of each subject on the test set of the CASIA dataset.



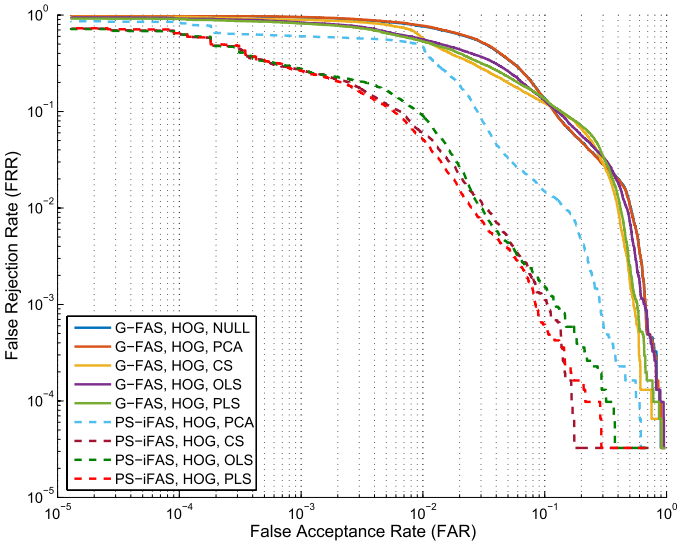


Fig. 5. ROC curves for different face anti-spoofing methods on the CASIA dataset.

presents the face recognition accuracies. Accordingly, the proposed person-specific model also outperforms the generic one consistently as on the development set. Meanwhile, the synthesized features for fake samples make sense to the generic models. On the *Test-T* set, the HTERs of augmented generic models decrease in almost all cases compared with original generic model. However, due to interferences among subjects, the synthesized features for fake samples impose negative effects more or less on the source subjects. Furthermore, we can find that the one based on CS, OLS and PLS outperform the PCA-based one in most situations. Ideally, if the center shift vectors between features of genuine samples are strictly equal to those between fake samples, the PCA-based feature synthesis can also generate plausible features as the proposed methods. However, it performs poorly in practice because the PCA-based method is more vulnerable to outliers. Specifically, because it sums the weighted features from all subjects to synthesize feature vector for a fake sample, samples violating our assumptions may lead to a bias.

Compared with PS-iFAS, the face recognition results affect the performance of PS-rFAS. Face images misclassified by the face recognizer are conveyed to an unmatched anti-spoofing classifier, and thus are more likely to be wrongly recognized. As a result, the performance degrades to some extent generally compared with the PS-iFAS. Fortunately, recent techniques on face recognition suffice to obtain satisfactory results in our case. We can find that the PS-rFAS methods still outperform G-FAS methods consistently, validating that the proposed person-specific methods are able to achieve better performance than generic models in a realistic scenario.

In Fig. 4, we show the HTERs of PS-iFAS for all subjects separately. As we can see, the generic models generally have higher HTERs with more fluctuations. Meanwhile, the PCA-based person-specific model also has poor performance on the target subjects. In contrast, the person-specific model based on the proposed domain adaptation methods have lower and slightly changed HTERs over subjects. For quantitative comparison, the standard deviations are shown in the legend. Note that we can only see one curve of PS-iFAS models on source subjects because the HTERs in all test scenarios are the same. In Fig. 5, we show the ROC curves for different methods and features. For person-specific method, we compute its overall ROC curve by the following way. Assume the person-specific thresholds obtained from the development set are $\{th\}_{i=1}^{B}$. On the test set, we separately compute the number of false accepted and false rejected samples for all subjects with their threshold be $\{th\}_{i=1}^{B} + \eta$. With the change of η, the overall number of false acceptance and false rejection are obtained.

3) Effect of the Number of Source Subjects: To further prove the effectiveness of our domain adaptation methods, we evaluate the face anti-spoofing performance of person-specific and generic models with the increasing number of source subjects. To train original generic models, the genuine and fake samples from a certain number of source subjects and genuine samples from all target subjects are merged to train a single SVM. As for augmented generic models, the

TABLE VI

EERs OF G-FAS AND PS-iFAS ON THE DEVELOPMENT SET OF THE REPLAY-ATTACK DATASET

Feature	Set	G-FAS					PS-iFAS			
		Orig	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS
MsLBP	*Devel-S*	6.51	6.23	2.42	**2.22**	2.45	**0.87**	**0.87**	**0.87**	**0.87**
	Devel-T	6.93	6.70	4.73	**4.44**	4.67	14.41	2.04	**1.93**	2.56
	Devel	6.81	6.57	4.03	**3.78**	4.00	10.34	1.69	**1.61**	2.05
HOG	*Devel-S*	18.35	19.04	9.28	**7.68**	9.07	**0.24**	**0.24**	**0.24**	**0.24**
	Devel-T	20.42	21.40	16.57	**15.77**	16.58	23.01	**8.67**	10.25	11.55
	Devel	19.84	20.73	14.39	13.34	**14.33**	16.18	**6.15**	7.25	8.15

TABLE VII

HTERs OF G-FAS, PS-iFAS AND PS-rFAS ON THE TEST SET OF THE REPLAY-ATTACK DATASET

Feature	Set	FR	G-FAS					PS-iFAS				PS-rFAS			
			Orig	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS
MsLBP	*Test-S*	99.59	8.65	8.50	**3.34**	3.56	3.44	**1.45**	**1.45**	**1.45**	**1.45**	**1.45**	**1.45**	**1.45**	**1.45**
	Test-T	99.50	6.42	6.42	4.11	**3.94**	4.07	14.60	**2.97**	3.24	3.57	14.79	**2.97**	3.24	3.57
	Test	99.52	7.11	7.07	3.89	**3.83**	3.89	10.66	**2.51**	2.70	2.94	10.79	**2.51**	2.70	2.94
HOG	*Test-S*	99.59	22.29	23.01	14.32	**12.74**	14.28	**3.58**	**3.58**	**3.58**	**3.58**	**3.58**	**3.58**	**3.58**	**3.58**
	Test-T	99.50	20.89	21.52	16.41	**15.60**	16.70	25.51	**9.40**	11.74	13.00	25.63	**9.40**	11.74	13.00
	Test	99.52	21.31	21.98	15.78	**14.72**	15.97	18.93	**7.65**	9.29	10.18	19.02	**7.65**	9.29	10.18

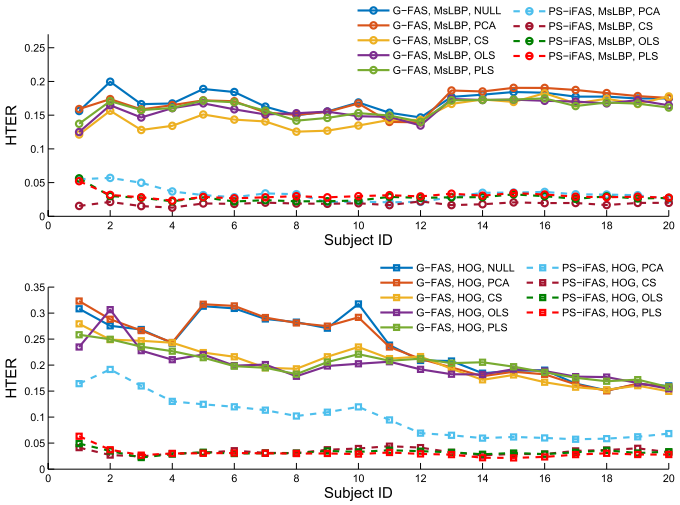


Fig. 6. HTERs on the target subjects in the CASIA dataset with the increased number of source subjects used for feature synthesis.

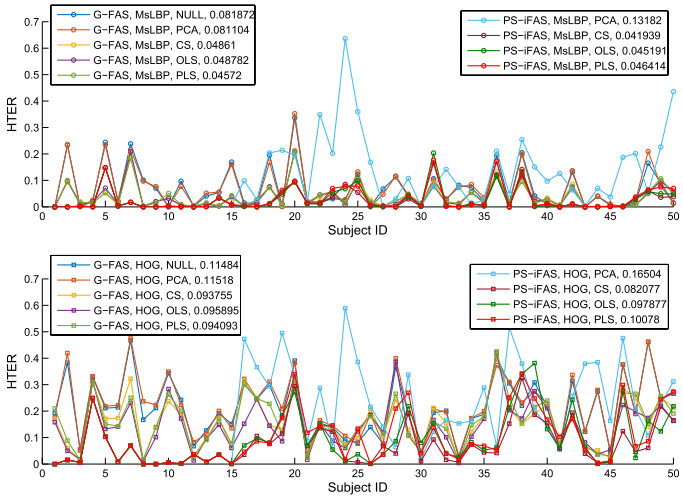


Fig. 7. HTER for each subject in the test set of the REPLAY-ATTACK dataset.

synthesized features for fake samples of all target subjects are added into the training set. Differently, each person-specific model is trained using merely the given genuine samples of one subject and the synthesized features of fake samples for it. The HTERs on the *Test-T* set of CASIA dataset are shown in Fig. 6. Accordingly, the HTERs associated with the proposed PS-iFAS methods are lower than the generic models consistently, and the HTERs decrease to a stable level lower than 5% when only four source subjects are supplied. These trend occurs on both MsLBP and HOG features. Moreover, with the assistance of synthesized features via our domain adaptation methods, the augmented generic models achieve slightly better performance than original generic model.

D. Experiments on the REPLAY-ATTACK Dataset

Similar to the experiments on the CASIA dataset, we compare G-FAS methods with PS-FAS methods on the REPLAY-ATTACK dataset. Before training classifiers, we estimate relations between 15 source subjects and 35 target subjects, and then synthesize features of fake samples for target subjects. To preserve the consistency of reflection factors, we split the genuine samples in enrollment set into two groups, which are captured in front of adverse and controlled backgrounds, respectively. Then, the relations of each group are transferred to the fake samples captured under the same condition.

1) Results on the Development Set: In Table VI, the EERs on the development set are presented. As we can see, the person-specific model achieves much lower EERs on source subjects. Because of the nearly perfect face recognition, the PS-rFAS has approximately identical performance to the

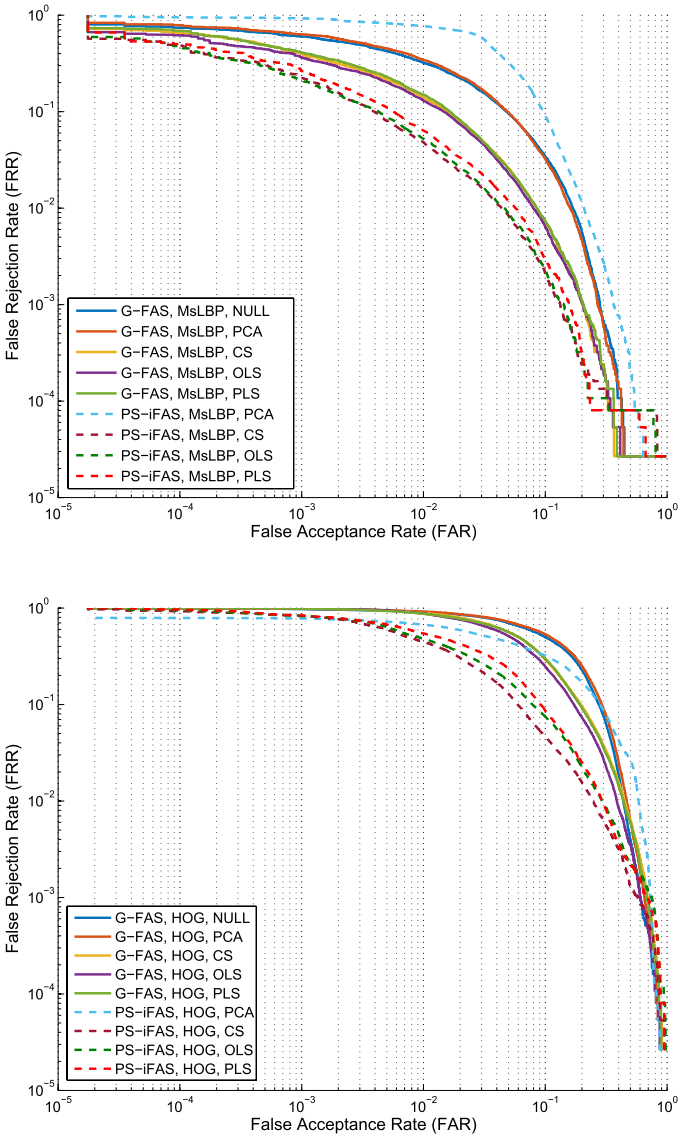


Fig. 8. ROC curves of G-FAS and PS-iFAS models on the test set of the REPLAY-ATTACK dataset.

PS-iFAS, still outperforming the generic models in all cases. On the target subjects, the person-specific models with proposed domain adaptation methods beat both original and augmented generic models. Compared with the proposed domain adaptation methods, the PCA-based one fails to synthesize effective fake features for target subjects, resulting in poor performance as on the CASIA dataset.

2) Results on the Test Set: Having obtained the thresholds corresponding to EERs, we evaluate the performance on the test set. In Table VII, the HTERs of different face anti-spoofing methods on test set are reported. As we can see, the person-specific model performs much better on source subjects than generic models, especially the original one. On the target subjects, the improvements are also noticeable. Given a nearly perfect face recognizer, the PS-rFAS performs identically to PS-iFAS when using the proposed three feature synthesis methods. In Fig. 7, the HTER for each subject is shown. As we can see, the original generic model has an intense fluctuation on both source and target subjects, which is suppressed by

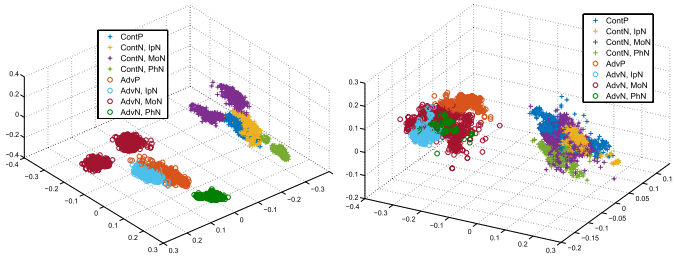


Fig. 9. Samples of the third subject from the REPLAY-ATTACK dataset projected into the 3D MsLBP (left) and HOG (right) feature subspace. "ContP" and "AdvP" mean genuine samples captured from controlled and adverse background, respectively. "ContN" and "AdvN" are fake samples. The fake samples are further categorized into three groups: IpN, MoN and PhN.

using the augmented generic and person-specific models with proposed domain adaptation methods. For details, we present ROC curves for different anti-spoofing methods on the test set in Fig. 8.

3) Results of Unknown Capturing Condition: In above experiments, we assume the information on capturing condition of samples are known. In this case, the domain adaptation can be implemented for a single capturing condition specifically. To evaluate the performance of proposed domain adaptation methods in more realistic situations, we assume no prior information on capturing condition is provided during training. Under this condition, we combine genuine samples captured in front of adverse and controlled backgrounds together, and then derive the relation between source and target subjects globally. After obtaining the relations between subjects, we transfer them to features of fake samples captured from fixed medium. Also, all the fake samples are transferred together to the target domain. In Table VIII, we report the HTERs with this new testing protocol on the test set. As we can see, the person-specific model achieves much lower EERs on source subjects. On target subjects, the person-specific model also beat the generic models when using CS method. However, the other domain adaptation methods do not synthesize proper features of fake samples for target subjects, resulting in even bad performance compared with generic models. These degradations are mainly caused by the poor regression results between genuine samples with large intra variations. As shown in Fig. 9, in both MsLBP and HOG feature spaces, the genuine samples captured in front of controlled and adverse backgrounds deviate much from each other. In this case, the OLS and PLS regression algorithms are very susceptible to the estimated correspondences of samples between two subjects. In our experiments, we observe that some samples with adverse background from one subject are matched to the samples with controlled background from another subject, then the estimated transformations will distort the synthesized fake samples much. Moreover, fake samples are captured in various conditions. Transferring them all together in one time based on OLS or PLS are also infeasible because the transformation are not applicable to multi-cluster samples under our assumptions, unlike the translation. Therefore, when no prior information is provided for domain adaptation, we recommend the most robust CS method.

TABLE VIII
HTERs ON THE TEST SET OF THE REPLAY-ATTACK DATASET WHEN NO PRIOR INFORMATION IS PROVIDED FOR DOMAIN ADAPTATION

Feature	Set	G-FAS					PS-iFAS			
		Orig	PCA	CS	OLS	PLS	PCA	CS	OLS	PLS
MsLBP	*Test-S*	8.63	8.43	**3.28**	6.04	5.11	**1.45**	**1.45**	**1.45**	**1.45**
	Test-T	6.43	6.48	**3.24**	5.22	4.33	21.16	**3.18**	14.41	18.27
	Test	7.10	7.08	**3.26**	5.48	4.58	15.24	**2.66**	10.52	13.22
HOG	*Test-S*	22.38	22.44	**12.56**	18.35	14.33	**3.58**	**3.58**	**3.58**	**3.58**
	Test-T	20.89	20.97	**16.43**	17.87	16.61	25.85	**12.95**	18.81	18.70
	Test	21.33	21.42	**15.27**	17.98	15.93	19.17	**10.14**	14.24	14.16

E. Discussion

According to the experimental results on two datasets, the person-specific models has consistently better performance on the source subjects compared with generic models. One difference between generic and person-specific models is that the latter needs face recognition. The performance of face recognition affects partially the final person-specific anti-spoofing performance. However, as shown in our experiments, the face recognition algorithm we used suffice to cope with the frontal face images in two datasets, regardless of the illumination and image quality.

Different domain adaptation methods perform differently on two dataset. On the CASIA dataset, the proposed three methods have no dominant superiority to each others. Different methods achieve the best under different conditions. On the REPLAY-ATTACK dataset, however, it is obvious that the CS method performs better than the other two methods. Remarkably, the CS method surpass much the OLS and PLS methods when no prior information is provided for domain adaptation. The reason for that CS method obtains better performance may be two-fold. Firstly, OLS and PLS algorithms need correct correspondences between features, which are actually difficult to obtain for high dimensional features without any auxiliary. On the other hand, OLS and PLS regressions may be under-constrained when the training samples are scarce or similar.

Another issue we want to discuss is the effect of number of source subjects for domain adaptation. In the experiments on CASIA dataset, we can see a few source subjects can boost the person-specific anti-spoofing performance much. This effectiveness makes it feasible to deploy person-specific anti-spoofing classifiers in an off-the-shelf face recognition system where most of the enrolled subjects have no fake faces for training. Moreover, we can observe that the lowest HTER occurs when not all source subjects are used. This indicates that the performance can be further improved by adapting the fake samples in source subject domains selectively.

Finally, we stress that the assumptions in Sec. IV-A are based on the condition that the fake samples have similar first-time reflection factors to the genuine samples from the same subject. When this condition is violated, the synthesized fake samples may not be helpful to tackle spoofing attacks. However, this problem can be solved by firstly estimating the illuminations on the samples and then also conducting a domain adaptation algorithm to eliminate the differences of reflection factors among those samples. Alternatively, we can narrow the gap by collecting genuine face images under various illuminations, and then choose the genuine samples with similar first-time reflection factors to the fake samples for subject domain adaptation.

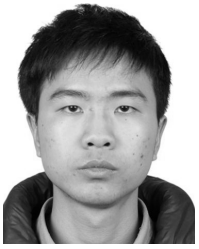
VI. CONCLUSION

In this paper, we have proposed a person-specific face anti-spoofing framework, which can be naturally integrated into face recognition systems. The new framework recognizes faces' identities first, and then feed the faces into specified spoofing classifiers, which avoids the interferences among subjects. Furthermore, to train person-specific anti-spoofing classifiers for the subjects whose fake faces are void, a subject domain adaptation method was proposed. Experiments on two dataset indicate it is promising to integrate face recognition and anti-spoofing in a person-specific manner. Upon this paper, one of the future work is exploiting a supervised subject domain adaptation method to estimate and transfer the relation between subjects, which may further improve the performance.

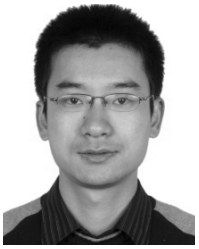
REFERENCES

[1] J. Määttä, A. Hadid, and M. Pietikainen, "Face spoofing detection from single images using micro-texture analysis," in *Proc. Int. Joint Conf. Biometrics (IJCB)*, Oct. 2011, pp. 1–7.
[2] G. Pan, L. Sun, Z. Wu, and S. Lao, "Eyeblink-based anti-spoofing in face recognition from a generic webcamera," in *Proc. IEEE 11th ICCV*, Oct. 2007, pp. 1–8.
[3] X. Tan, Y. Li, J. Liu, and L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in *Proc. 11th ECCV*, 2010, pp. 504–517.
[4] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "Can face anti-spoofing countermeasures work in a real world scenario?" in *Proc. ICB*, Jun. 2013, pp. 1–8.
[5] J. Komulainen, A. Hadid, M. Pietikainen, A. Anjos, and S. Marcel, "Complementary countermeasures for detecting scenic face spoofing attacks," in *Proc. ICB*, Jun. 2013, pp. 1–7.
[6] I. Chingovska *et al.*, "The 2nd competition on counter measures to 2D face spoofing attacks," in *Proc. ICB*, Jun. 2013, pp. 1–6.
[7] M. M. Chakka *et al.*, "Competition on counter measures to 2-D facial spoofing attacks," in *Proc. IJCB*, 2011, pp. 1–6.
[8] J. Li, Y. Wang, T. Tan, and A. K. Jain, "Live face detection based on the analysis of Fourier spectra," *Proc. SPIE*, vol. 5404, pp. 296–303, Aug. 2004.
[9] B. Peixoto, C. Michelassi, and A. Rocha, "Face liveness detection under bad illumination conditions," in *Proc. 18th IEEE ICIP*, Sep. 2011, pp. 3557–3560.
[10] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2012, pp. 1–7.
[11] J. Yang, Z. Lei, S. Liao, and S. Z. Li, "Face liveness detection with component dependent descriptor," in *Proc. ICB*, Jun. 2013, pp. 1–6.
[12] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "$LBP - TOP$ based countermeasure against face spoofing attacks," in *Computer Vision*. Berlin, Germany: Springer-Verlag, 2013, pp. 121–132.

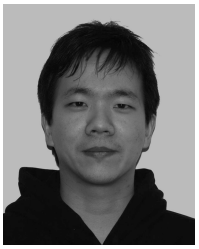
[13] J. Komulainen, A. Hadid, and M. Pietikainen, "Context based face anti-spoofing," in *Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl., Syst. (BTAS)*, Sep./Oct. 2013, pp. 1–8.
[14] L. Sun, G. Pan, Z. Wu, and S. Lao, "Blinking-based live face detection using conditional random fields," in *Proc. ICB*, 2007, pp. 252–260.
[15] K. Kollreider, H. Fronthaler, M. I. Faraj, and J. Bigün, "Real-time face detection and motion analysis with application in 'liveness' assessment," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 548–558, Sep. 2007.
[16] G. Chetty, "Biometric liveness checking using multimodal fuzzy fusion," in *Proc. IEEE Int. Conf. FUZZ*, Jul. 2010, pp. 1–8.
[17] G. Chetty and M. Wagner, "Audio-visual multimodal fusion for biometric person authentication and liveness verification," in *Proc. NICTA-HCSNet Multimodal User Interact. Workshop (MMUI)*, vol. 57. Sydney, NSW, Australia, 2005, pp. 17–24.
[18] W. Bao, H. Li, N. Li, and W. Jiang, "A liveness detection method for face recognition based on optical flow field," in *Proc. Int. Conf. Image Anal. Signal Process.*, 2009, pp. 233–236.
[19] K. Kollreider, H. Fronthaler, and J. Bigün, "Evaluating liveness by face images and the structure tensor," in *Proc. 4th IEEE Workshop AutoID*, Oct. 2005, pp. 75–80.
[20] A. Anjos, M. M. Chakka, and S. Marcel, "Motion-based counter-measures to photo attacks in face recognition," *IET Biometrics*, vol. 3, no. 3, pp. 147–158, Sep. 2014.
[21] K. Kollreider, H. Fronthaler, and J. Bigün, "Verifying liveness by multiple experts in face biometrics," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2008, pp. 1–6.
[22] M. De Marsico, M. Nappi, D. Riccio, and J. Dugelay, "Moving face spoofing detection via 3D projective invariants," in *Proc. 5th IAPR Int. Conf. Biometrics (ICB)*, Mar./Apr. 2012, pp. 73–78.
[23] T. Wang, J. Yang, Z. Lei, S. Liao, and S. Z. Li, "Face liveness detection using 3d structure recovered from a single camera," in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2013, pp. 1–6.
[24] N. Erdogmus and S. Marcel, "Spoofing 2D face recognition systems with 3D masks," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2013, pp. 1–8.
[25] I. Pavlidis and P. Symosek, "The imaging issue in an automatic face/disguise detection system," in *Proc. IEEE Workshop Comput. Vis. Beyond Vis. Spectr., Methods Appl.*, Jun. 2000, pp. 15–24.
[26] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li, "Face liveness detection by learning multispectral reflectance distributions," in *Proc. IEEE Int. Conf. FG*, Mar. 2011, pp. 436–441.
[27] Y.-L. Hou, X. Hao, Y. Wang, and C. Guo, "Multispectral face liveness detection method based on gradient features," *Opt. Eng.*, vol. 52, no. 11, p. 113102, 2013.
[28] G. Pan, L. Sun, Z. Wu, and Y. Wang, "Monocular camera-based face liveness detection by combining eyeblink and scene context," *Telecommun. Syst.*, vol. 47, nos. 3–4, pp. 215–225, 2011.
[29] R. Tronci *et al.*, "Fusion of multiple clues for photo-attack detection in face recognition systems," in *Proc. Int. Joint Conf. Biometrics (IJCB)*, 2011, pp. 1–6.
[30] J. Yan, Z. Zhang, Z. Lei, D. Yi, and S. Z. Li, "Face liveness detection by exploring multiple scenic clues," in *Proc. 12th Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Dec. 2012, pp. 188–193.
[31] I. Chingovska, A. Anjos, and S. Marcel, "Anti-spoofing in action: Joint operation with a verification system," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2013, pp. 98–104.
[32] B. Yao, H. Ai, and S. Lao, "Person-specific face recognition in unconstrained environments: A combination of offline and online learning," in *Proc. 8th IEEE Int. Conf. FG*, Sep. 2008, pp. 1–8.
[33] G. Chiachia, A. X. Falcao, and A. Rocha, "Person-specific face representation for recognition," in *Proc. IEEE Int. Joint Conf. Biometrics Compendium Biometrics (IJCB)*, Washington, DC, USA, 2011.
[34] G. Chiachia, N. Pinto, W. R. Schwartz, A. Rocha, A. X. Falcão, and D. Cox, "Person-specific subspace analysis for unconstrained familiar face identification," in *Proc. BMVC*, 2012, pp. 1–12.
[35] Y. Zhang and D.-Y. Yeung, "Multi-task warped Gaussian process for personalized age estimation," in *Proc. IEEE Conf. CVPR*, Jun. 2010, pp. 2622–2629.
[36] J. Chen, X. Liu, P. Tu, and A. Aragones, "Learning person-specific models for facial expression and action unit recognition," *Pattern Recognit. Lett.*, vol. 34, no. 15, pp. 1964–1970, 2013.
[37] X. Cao, D. Wipf, F. Wen, G. Duan, and J. Sun, "A practical transfer learning algorithm for face verification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3208–3215.
[38] X. Tang and X. Wang, "Face sketch synthesis and recognition," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 687–694.
[39] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009.
[40] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face antispoofing database with diverse attacks," in *Proc. 5th IAPR ICB*, 2012, pp. 26–31.
[41] M. Pilu, "A direct method for stereo correspondence based on singular value decomposition," in *Proc. IEEE Comput. Soc. CVPR*, Jun. 1997, pp. 261–266.
[42] R. Rosipal and N. Krämer, "Overview and recent advances in partial least squares," in *Subspace, Latent Structure and Feature Selection*. Berlin, Germany: Springer-Verlag, 2006, pp. 34–51.
[43] R. Manne, "Analysis of two partial-least-squares algorithms for multivariate calibration," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 187–197, 1987.
[44] D. Yi, Z. Lei, and S. Z. Li, "Towards pose robust face recognition," in *Proc. IEEE Conf. CVPR*, Jun. 2013, pp. 3539–3545.



Jianwei Yang received the B.S. degree from the School of Information Science and Engineering, Central South University, Changsha, China, in 2011, and the M.E. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2014. During the master's study, his research was mainly focused on face anti-spoofing, under the supervision of Prof. S. Z. Li. He won the first place at the second competition on counter measures to 2-D facial spoofing attacks in 2013. His current research interests include face anti-spoofing, face recognition, and image recognition.



Zhen Lei received the B.S. degree in automation from the University of Science and Technology of China, in 2005, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, in 2010, where he is currently an Associate Professor. He has published over 90 papers in international journals and conferences. His research interests are in computer vision, pattern recognition, image processing, and face recognition in particular. He serves as an Area Chair of the International Joint Conference on Biometrics in 2014, the IAPR/IEEE International Conference on Biometric in 2015, and the IEEE International Conference on Automatic Face and Gesture Recognition in 2015.



Dong Yi received the B.S. degree in electronics engineering, in 2003, the M.S. degree in communication and information systems from Wuhan University, in 2006, and the Ph.D. degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences. He developed the face biometric modules and systems for the immigration control projects and 2008 Beijing Olympic Games. His research areas are unconstrained face recognition, heterogeneous face recognition, and deep learning.

Stan Z. Li (F'09) received the B.Eng. degree from Hunan University, Changsha, China, the M.Eng. degree from the National University of Defense Technology, Changsha, and the Ph.D. degree from Surrey University, Guildford, U.K. He was an Associate Professor with Nanyang Technological University, Singapore. He was with Microsoft Research Asia as a Researcher from 2000 to 2004. He is currently a Professor and the Director of the Center for Biometrics and Security Research with the Institute of Automation, Chinese Academy of Sciences. He has published over 200 papers in international journals and conferences, and authored and edited eight books.

His research interests include pattern recognition and machine learning, image and vision processing, face recognition, biometrics, and intelligent video surveillance. He was an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, and serves as the Editor-in-Chief of the *Encyclopedia of Biometrics*. He has served as the Program Co-Chair of the International Conference on Biometrics in 2007, 2009, and 2015, and has been involved in organizing other international conferences and workshops in the fields of his research interest.