Chapter 5
Face Anti-spoofing: Multi-spectral Approach

Dong Yi, Zhen Lei, Zhiwei Zhang and Stan Z. Li

Abstract With the wide applications of face recognition, spoofing attack is becoming a big threat to their security. Conventional face recognition systems usually adopt behavioral challenge-response or texture analysis methods to resist spoofing attacks, however, these methods require high user cooperation and are sensitive to the imaging quality and environments. In this chapter, we present a multi-spectral face recognition system working in VIS (Visible) and NIR (Near Infrared) spectrums, which is robust to various spoofing attacks and user cooperation free. First, we introduce the structure of the system from several aspects including: imaging device, face landmarking, feature extraction, matching, VIS, and NIR sub-systems. Then the performance of the multi-spectral system and each subsystem is evaluated and analyzed. Finally, we describe the multi-spectral image-based anti-spoofing module, and report its performance under photo attacks. Experiments on a spoofing database show the excellent performance of the proposed system both in recognition rate and anti-spoofing ability. Compared with conventional VIS face recognition system, the multi-spectral system has two advantages: (1) By combining the VIS and NIR spectrums, the system can resist VIS photo and NIR photo attacks easily. And users' cooperation is no longer needed, making the system user friendly and fast. (2) Due to the precise key-point localization, Gabor feature extraction and unsupervised learning, the system is robust to pose, illumination and expression variations. Generally, its recognition rate is higher than the VIS subsystem.

D. Yi (✉) · Z. Lei · Z. Zhang · S.Z. Li
Chinese Academy of Sciences, Institute of Automation,
95 Zhongguancun Donglu, Beijing 100190, China
e-mail: dyi@cbsr.ia.ac.cn

Z. Lei
e-mail: zlei@nlpr.ia.ac.cn

Z. Zhang
e-mail: zhiweiscu@gmail.com

S.Z. Li
e-mail: stan.zq.li@gmail.com

© Springer-Verlag London 2014

S. Marcel et al. (eds.), *Handbook of Biometric Anti-Spoofing*,
Advances in Computer Vision and Pattern Recognition,
DOI: 10.1007/978-1-4471-6524-8_5

5.1 Introduction

Although face recognition has achieved great success during the past decades, little effort has been made to assure its security and reliability in real-world applications. It is now increasingly known that existing face recognition systems are susceptible to fake face attacks, through which unauthorized attackers try to access illegal authorities by exhibiting fake faces of an authorized client. Serious consequences may occur if these attacks succeed, yet there still lack effective anti-spoofing techniques.

Attackers can obtain a client's face images by using portable digital cameras or simply downloading from the Internet, and fake faces can be easily produced, for example, printing photos or showing videos on a laptop. Fake faces like photos and video replays are not only easy to implement but also usually quite effective to attack a face recognition system [1], and has become the main concern [1] in the literature as shown in Sect. 5.2. Actually in real-world applications, face recognition systems may encounter various high quality face attacks and low quality real accesses, therefore an excellent anti-spoofing method should distinguish their difference and perform robust in unpredicted situations.

However, as we review the recent development in Sect. 5.2, we find that current researches mainly concentrate on genuine and fake face with little variations. A shortcoming in variation is the quality of attacks: in [2–4], algorithms extract the high frequency information or micro-texture to detect attacks. But as this high frequency information or micro-texture pattern highly depends on the image quality, how will they perform on good quality and bad quality images? Do their algorithms generalize well? We can see that due to the lack of variational data, many questions remain unanswered. Therefore we cannot predict their performance in real-world applications, because practical attacks are probably not limited to one single type as in previous researches.

Therefore, we think "quality robustness" is a big problem in anti-spoofing research. How to classify high quality attacks and low quality real accesses is the key issue to solve this problem. Multi-modal biometrics may be a practical direction, because we have more information available for anti-spoofing. Likewise, multi-spectral face images are naturally stronger than single VIS face image, by fusing which we have more chance to build a quality robust anti-spoofing module. In this chapter, we propose a high performance face recognition system with anti-spoofing module using multi-spectral imaging. We start by introducing the multi-spectral face recognition system, including spectrums selection, hardware, and algorithms. Then we report the performance of the system under licit transaction and spoofing attack. To resist the attacks, we propose an effective multi-spectral countermeasure and test the performance of the system with countermeasures. The experimental results illustrate that the countermeasures resist the attacks perfectly.

[1] Mask is also a good choice, but usually it is too expensive to produce client-like masks. So the massive usage of masks rarely appears in the literature.

Compared with existing anti-spoofing approaches, the advantages of our system are obvious. First, our system requires no user cooperation, and therefore is user-friendly and fast. Second, our system, by combining multi-spectral information, is more quality robust and can achieve higher recognition rate than traditional single modal face recognition system.

5.2 Related Work

Existing work can be classified as three categories: facial motion detection, facial texture analysis and multi-spectral anti-spoofing methods. The first two are usually applied to VIS face recognition systems. The last one needs extra multi-spectral imaging device to achieve more accurate results.

Facial motion detection techniques expect subjects to exhibit specific facial motion, the detection of which determines the liveness. For methods of this kind, human–computer interaction (HCI) is almost indispensable to detect users' biological motion. The most commonly used motion types include eye blinking [5, 6], head rotation [6, 7], and mouth movement [8], and these motions are mainly detected by adopting optical flow. One main problem of these methods is that users need to be highly cooperative and the duration of liveness detection is relatively long, which will make users feel uncomfortable when using such a system. Another problem is that they cannot deal with some skilled attacks. For example, if the fake face is a photo over a genuine face with eyes and mouth cut out as illustrated in [6], these methods will definitely fail. Therefore, applications of such kind of methods are limited.

Facial texture analysis techniques believe that fake faces probably lack some high frequency information during the reproduction process, and by analyzing and learning the facial texture information, genuine and fake faces can be classified properly. Here the term "texture" represents the high frequency details in face images. In [2] Fourier transform is utilized to extract the high frequency information, and the target face image is judged fake if its energy percentage of high frequency is lower than a certain threshold. In [3] the authors use DoG and LTV algorithms to extract high frequency information from the captured images, and the final model is learned by a complex bilinear sparse low rank logistic regression model. In [4] a more simple but also more powerful LBP+SVM method (named Micro-Texture Analysis, MTA) is proposed, and they achieve very amazing results both on the NUAA [3] and Idiap [9] database. Similarly, [10] also utilize micro-textures and SVM to detect spoofing attacks. However, as noted by [11], most existing spoofing databases did not include enough variations, therefore, the best texture-based methods on the databases may failed in real applications when confronting various imaging environments.

The other class is the multi-spectral methods, which detect the reflectance of object surface under multi-spectrums. To the best of our knowledge, there have been very few papers published in this field, among which two papers are most representative. In [12] Pavlidis and Symosekuses use light at two wavelengths, and a simple threshold method to detect genuine and fake faces. No experiments but only

illustrations were reported in their paper. The second one [13] also selects light at two different wavelengths and then LDA is used to make the final decision. However, this paper requires the distance between the user and the system to be exactly 30 cm, and they utilize users' forehead region to measure reflectance. Not only may the forehead be occluded, but also the exact distance is quite demanding and impossible to execute in practice. Furthermore, the wavelengths they select are actually not as optimal as in [12]. Thermal information is another choice, and we refer readers to a common facial thermal imagery database in [14]. But, the high cost also prevents its usage in real practice.

5.3 The Multi-spectral Face Recognition System

The multi-spectral face recognition system includes the following modules: multi-spectral image acquisition, face detection, key-point localization, feature extraction, subspace learning, matching, and fusion. To achieve good performance and generalization ability in practical applications, the multi-spectral system uses an EBGM like pipeline to process face image, which is mainly composed of precise key-point localization, Gabor filtering, and PCA. Yi et al. [15] have already shown the high performance of the method on FERET (only in VIS spectrum). For each spectrum, face image is processed separately by its corresponding pipeline, and the similarity scores in multi-spectrums are fused by "sum rule." The details of each module are described as follows.

5.3.1 Spectrum Selection and Imaging Device

The use of multiple spectral bands for face biometrics permits to improve the performance and robustness of face recognition in realistic scenarios including uncontrolled illumination conditions. Taking into account the compatibility with normal face recognition system, VIS is selected as the first spectrum for the multi-spectral system. On the other hand, NIR band has become the most used spectral band beyond VIS due to several advantages: radiation which is harmless to health, good-quality images, low-cost cameras, etc. Based on our existing NIR face recognition system [16], we choose the second spectrum as NIR (780 nm). To acquire VIS and NIR images simultaneously, we develop a special imaging device by two CMOS camera modules, which can capture 640×480 images in two bands at 15 fps. The synchronization is controlled by the software system.

In the multi-spectral system, two spectrum bands are available and hence can be exploited for recognition in many ways. A common method is to use all spectral bands simultaneously that match VIS to VIS, and NIR to NIR. Currently, we use multi-spectral images in this way, but heterogeneous matching between VIS and NIR may be added into the system in the future. For example, VIS face images are used for enrollment while NIR face images are used for testing.

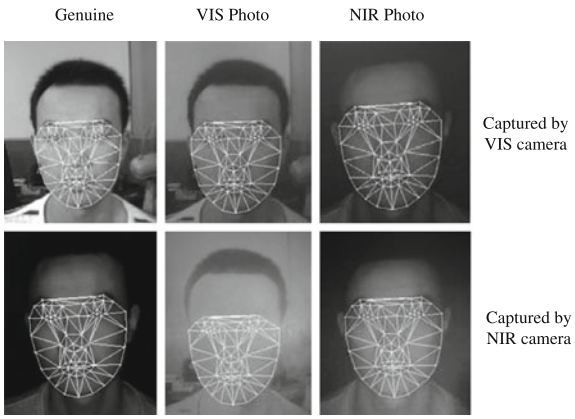

Fig. 5.1 Some genuine and fake face images in different spectrums and their 76 facial landmarks localized by the ASM landmarker. From *left* to *right*, the *three columns* correspond to genuine face, VIS photo, and NIR photo. Images in the *first row* are captured under visible illumination, and the *second row* are captured under NIR illumination. The fake face images are produced by recapture of the VIS and NIR printed photo by the same system

5.3.2 Key-Point Localization

Because face detection is relatively mature than other steps, we skip it. Interested readers can refer to [17] for details. After face detection, we localize the facial landmarks by Active Shape Model (ASM) [18]. ASM is composed of three parts: shape model, local experts, and optimization strategy. In most ASM variants, shape model is usually PCA [19], and we follow this model. For local experts, we use LBP [20] feature and Boosting classifier for each landmark, which is similar to the method in [21]. Based on the output of Boosting classifiers, we can get a confidence map for each landmark. These confidence maps are fed to a Landmark Mean-Shift procedure [22]. Then we can get the final positions of all facial landmarks. For robustness and efficiency, the optimization process is repeated several times on two scales.

The training set of our landmarker is constructed from the MUCT database [23]. Three views (a, d, and e) with small pose variations are used for training. Because the backgrounds of images in the MUCT are almost uniform, we replace them with some random backgrounds and mirror all images to augment the dataset (see Fig. 5.2). The uniform backgrounds of the face images are segmented by GrabCut [24], which is initialized by the results of face detection. Figure 5.1 shows two example images in the FERET database [25] and their 76 facial landmarks localized by the landmarker, from which we can see that the landmarks are robust to small pose variations and have good precision for the next steps.

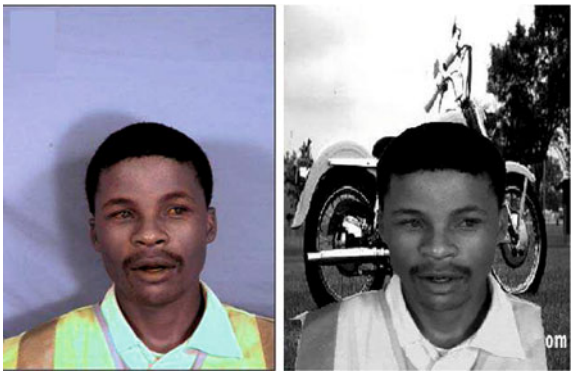

Fig. 5.2 Sample images in the MUCT training set for the ASM landmarker. *Left* A color face image in the MUCT database, which has uniform background. *Right* A face image is converted to *gray* scale and the background is replaced by a random image from the Internet

5.3.3 Gabor Feature and Subspace Learning

Given an aligned face image and the 76 landmarks, we extract local features on the landmarks by a Gabor wavelet, which is described in [26].

$$\psi_{\mathbf{k},\sigma}(\mathbf{x}) = \frac{k^2}{\sigma^2} e^{\frac{k^2}{-2\sigma^2}x^2} \{e^{i\mathbf{kx}} - e^{-\frac{\sigma^2}{2}}\} \tag{5.1}$$

The wavelet is a plane wave with wave vector $\mathbf{k}$, restricted by a Gaussian envelope, the size of which relative to the wavelength is parameterized by σ. The second term in the brace removes the DC component. Following the popular way, we sample the space of wave vectors $\mathbf{k}$ and scale σ in a discrete hierarchy of 5 resolutions (differing by half-octaves) and 8 orientations at each resolution (See Fig. 5.3), thus giving $5 \times 8 = 40$ complex values for each landmark. Because the phase information is sensitive to image shift or misalignment, we drop the phase and use the amplitude as feature for face recognition.

Merging the feature values at all landmarks together, we get a feature vector with $76 \times 40 = 3{,}040$ dimensions. To reduce the dimensionality of feature and remove the redundant information, Principle Component Analysis (PCA) [27] is used to learn a low-dimensional subspace. To remove the large variations caused by extrinsic factors such as illumination and expression, we discard the first several principal components. In the reduced PCA subspace, the similarity of feature vectors are evaluated by Cosine metric.

$$s(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^T \mathbf{y}}{\sqrt{\mathbf{x}^T \mathbf{x} \mathbf{y}^T \mathbf{y}}} \tag{5.2}$$

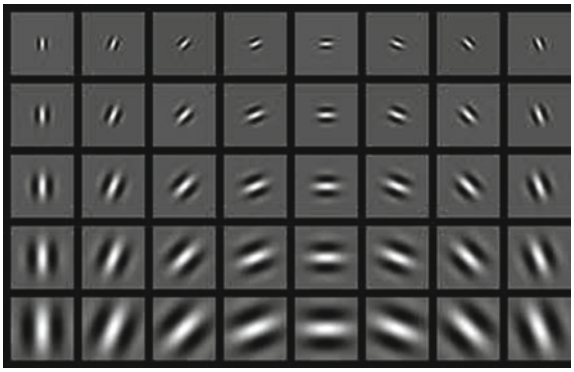

Fig. 5.3 The real part of the Gabor wavelet in 5 resolutions and 8 orientations

In practice, we usually normalize the feature vector $\mathbf{x}$ and $\mathbf{y}$ to unit length as $\mathbf{x}'$ and $\mathbf{y}'$. Then Eq. (5.2) can be written as

$$s(\mathbf{x}, \mathbf{y}) = s(\mathbf{x}', \mathbf{y}') = \mathbf{x}'^T \mathbf{y}'. \tag{5.3}$$

5.4 Performance Under Spoofing Attack

In this section, we build a database to evaluate the performance of the multi-spectral system and analyze the vulnerabilities of the multi-spectral system when confronted to spoofing attacks. Because making 3D mask is expensive, currently we use printed photos (VIS and NIR) to attack the system. In order to obtain comprehensive results, the VIS subsystem and NIR subsystem are also evaluated separately.

5.4.1 Database

All face images in the database are acquired by using the self-developed device (5.3), which includes a VIS camera, an NIR camera with some NIR LEDs. VIS and NIR images are synchronized by the system. The imaging device works at a rate of 15 fps for 640 × 480 images. Figure 5.4 shows the scenario of database collection.

The database comprises genuine face images of 100 subjects and their corresponding fake face samples. All subjects are imaged under VIS and NIR illuminations (5 images per subject per spectrum). For the fake faces, photos are printed using both the visible and NIR face image, named as VIS photo and NIR photo respectively. The printed photos are acquired by the same system described above. We use a kind

Fig. 5.4 Setup used for the acquisition of real-accesses for the multi-spectral face spoofing database

of coarse paper as printing material, because its relatively rough surface makes the reflectance weak and the fake face more vivid. Some examples are shown in Fig. 5.5.

In summary, the information about the database is: We denote these kinds of face images as:

- G_{VIS}: Genuine subjects captured by VIS camera;
- G_{NIR}: Genuine subjects captured by NIR camera;
- VP_{VIS}: VIS photo captured by VIS camera;
- VP_{NIR}: VIS photo captured by NIR camera;
- NP_{VIS}: NIR photo captured by VIS camera;
- NP_{NIR}: NIR photo captured by NIR camera.

By observing the face images in Fig. 5.5 we give some conjectures: (1) VP_{VIS} is easy to attack G_{VIS}; (2) NP_{VIS} is harder to attack G_{VIS} than VP_{VIS}. And by using the color information, we can easily detect the fake samples belonging to NP_{VIS}; (3) Due to the strong specular reflectance, VP_{NIR} is hard to attack G_{NIR}; (4) G_{NIR} is easily attacked by NP_{NIR}. These conjectures will be verified in the following experiments and in the next Sect. 5.5.

5.4.2 Protocol for Licit Transactions

The genuine face images in the database could be used to evaluate the performance of the multi-spectral face recognition system. Here we describe how to use the database to do training, enrollment, and recognition. To illustrate the performance of individual

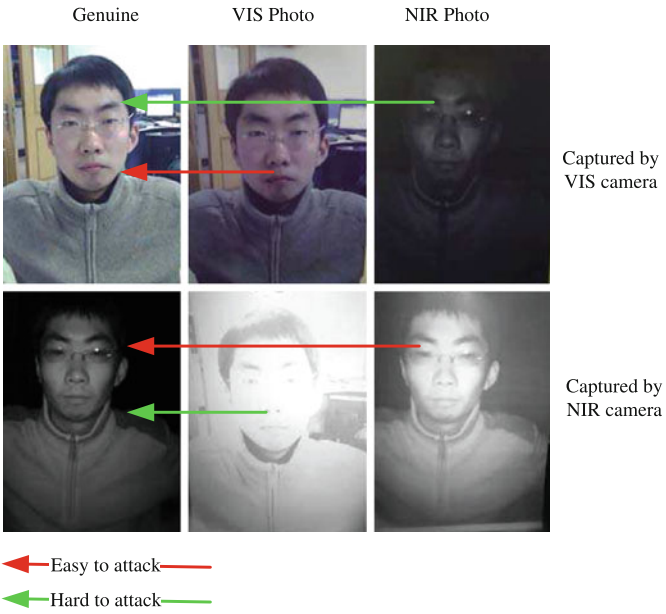

Fig. 5.5 Illustration of multi-spectral face images in the database. From *left* to *right*, the *three columns* correspond to genuine face, VIS photo and NIR photo. Images in the *first row* are captured under visible illumination, and the *second row* are captured under NIR illumination. The fake face images are produced by recapture of the VIS and NIR photo by the same system

modality (VIS and NIR) and the improvement of the multi-spectral fusion, three sub-experiments were conducted:

- VIS vs. VIS;
- NIR vs. NIR;
- VIS + NIR vs. VIS + NIR.

For the three experiments, the database is split into four subsets including:

- Training set: genuine face images of the first 30 subjects, $30 \times 5 = 150$ pairs;
- Development set: genuine face images of the following 30 subjects, $30 \times 5 = 150$ pairs;
- Licit Gallery set: two genuine face images of the other 40 subjects, $40 \times 2 = 80$ pairs;
- Licit Probe set: three genuine face images of the other 40 subjects, $40 \times 3 = 120$ pairs.

The PCA subspace is trained on the training set and the parameters are tuned on the development set. During the testing phase, the parameters of model should remain fixed. The Detection-Error Trade-off (DET) curve is used to illustrate the performance of the system.

5.4.3 Protocol for Spoofing Attacks

Similar to the protocols for Licit Transactions, six scenarios are evaluated to illustrate the influence of spoofing attacks to the performance of the system. There are:

- S1.1: using the VIS photos in VP_{VIS} to attack the VIS subsystem;
- S1.2: using the NIR photos in NP_{VIS} to attack the VIS subsystem;
- S2.1: using the VIS photos in VP_{NIR} to attack the NIR subsystem;
- S2.2: using the NIR photos in NP_{NIR} to attack the NIR subsystem;
- S3.1: using the VP_{VIS} and VP_{NIR} photo pairs to attack the multi-spectral system;
- S3.2: using the NP_{VIS} and NP_{NIR} photo pairs to attack the multi-spectral system.

In the attacking scenarios, the training sets, testing sets, and gallery sets are as same as those in licit transaction, and the probe sets are augmented by the printed photos. In these cases the genuine user enrolls with their faces and the attacker tries to access the system with the corresponding printed VIS or NIR photos. A successful attack is accomplished when the system confuses a genuine face image with its corresponding printed photo. The protocol is shown as follows:

- S1.1 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in VP_{VIS};
- S1.2 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in NP_{VIS};
- S2.1 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in VP_{NIR};
- S2.2 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in NP_{NIR};
- S3.1 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in VP_{VIS} and VP_{NIR};
- S3.2 Probe set: Licit Probe set described in Sect. 5.4.2 and their corresponding printed photos in NP_{VIS} and NP_{NIR};

The DET curves are also utilized for performance reporting, which describes the relationship between false detection rate and false rejection rate.

5.4.4 Results

5.4.4.1 Attacking the VIS Subsystem

The experimental results of S1.1 and S1.2 are shown in Figs. 5.6 and 5.7. From the results we can see that the VIS subsystem has good performance when it does not confront spoofing attacks. The FAR@FRR = 3% is about 2%. Furthermore, the score distributions of imposters and true claimants are well separated. These observations indicate that the VIS subsystem performs well on this database. Because the baseline

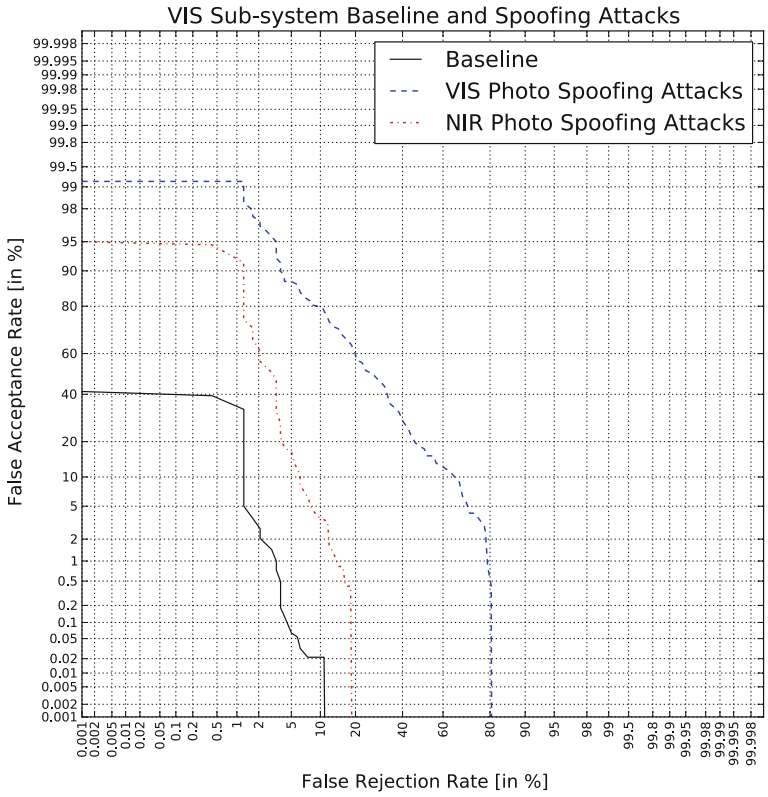

Fig. 5.6 DET curves for the VIS subsystem baseline and spoofing attacks

is a full unsupervised method, based on local feature + PCA, it is expected to have good generalization ability.

Under the VIS and NIR photo attacks, the performance of the VIS subsystem drops drastically. For example, when FRR = 2% the FAR increases from 3 to 97 % and 60 % respectively. This indicates that the VIS subsystem is attacked by the VIS and NIR photos, and the VIS photo can attack the system more easily than the NIR photo. Figure 5.7 shows the same phenomenon, the score distribution of the VIS photo is heavily overlapped with the true claimants.

5.4.4.2 Attacking the NIR Subsystem

The experimental results of S2.1 and S2.2 are shown in Figs. 5.8 and 5.9. From the figures we can see similar results as the VIS subsystem. When not confronting spoofing attacks, the NIR subsystem performs well too, and even better than the VIS

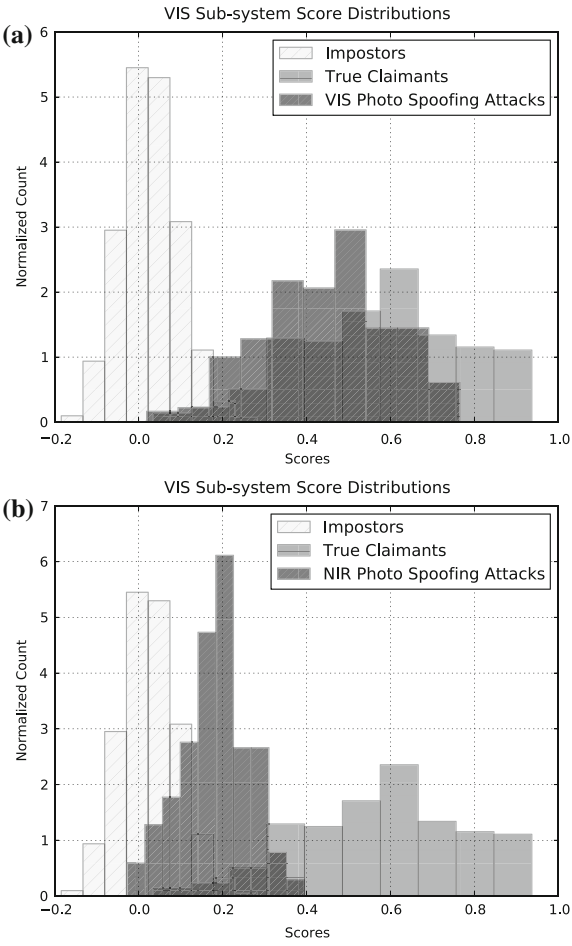

Fig. 5.7 Score distributions for the VIS subsystem baseline and spoofing attacks. **a** Attacked by VIS Photo. **b** Attacked by NIR Photo

subsystem due to its illumination invariant property. Its EER (equal error rate) is about 2 %, slightly lower than the VIS subsystem.

Under the VIS and NIR photo attacks, the performance of the NIR subsystem drops too, but decline is not as sharp as the VIS subsystem. For example, when FRR = 2 % the FAR increases from 2 to 10 and 78 %, which shows that the NIR face modality is inherently better than the VIS face modality in terms of anti-spoofing. On the contrary with VIS, the NIR subsystem is more easily attacked by the NIR photo, which verifies the conjectures in Sect. 5.4.1.

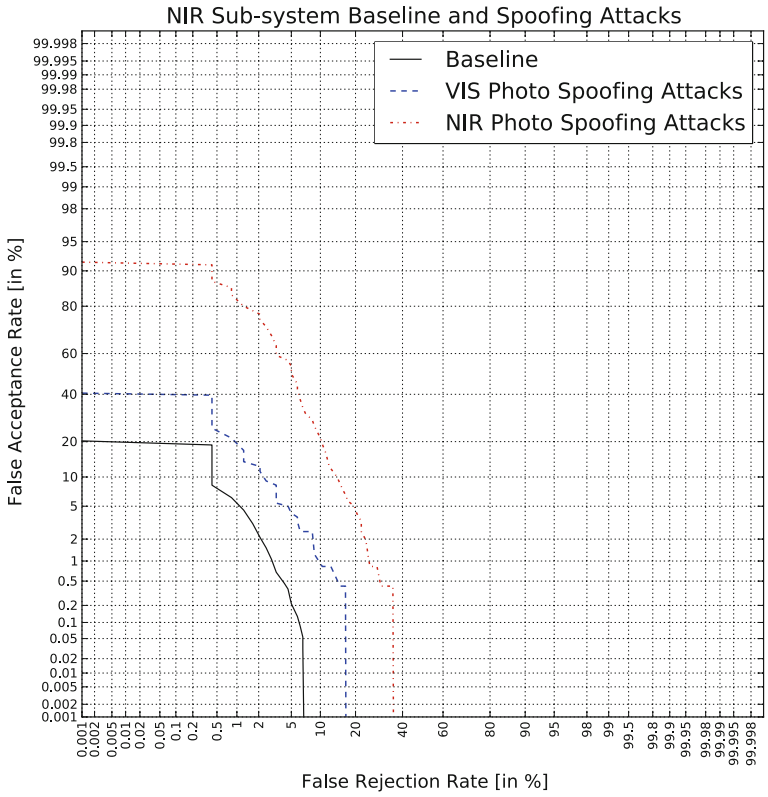

Fig. 5.8 DET curves for the NIR subsystem baseline and spoofing attacks

5.4.4.3 Attacking the Multi-spectral System

By fusing the scores of the VIS and NIR subsystems, we show the experimental results of S3.1 and S3.2 in Figs. 5.10 and 5.11. Compared with the VIS and NIR subsystems, the performance of the multi-spectral system is improved a little, EER from 3 % and 2 to 1.8 %. However, an interesting thing is that the multi-spectral system is vulnerable both to the VIS and NIR photo attack. By the effect of "sum rule," the DET curves and score distributions all achieve a balance between the VIS and NIR subsystems, which result in the multi-spectral system that is weaker than the VIS subsystem to resist the NIR photo attack and is weaker than the NIR subsystem to resist the VIS photo attack.

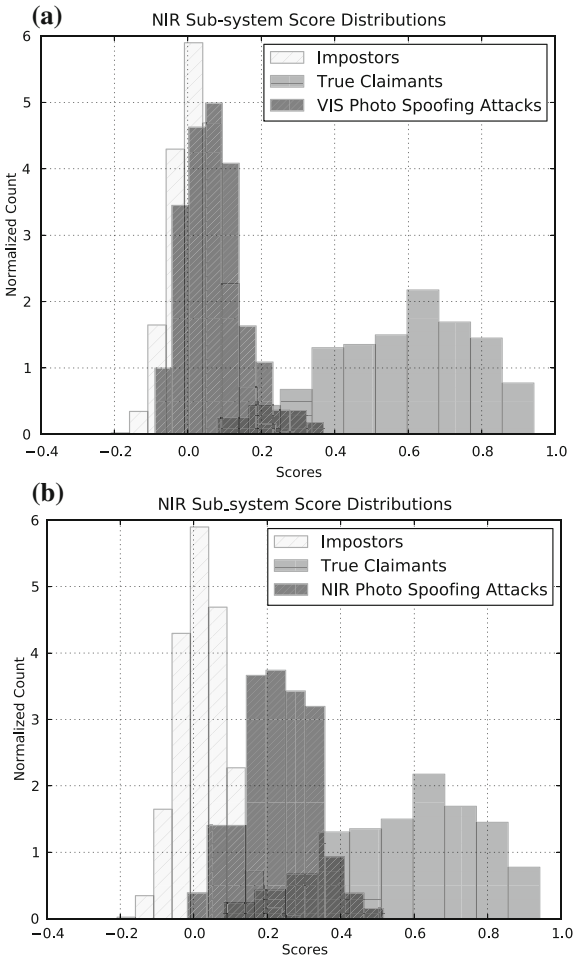

Fig. 5.9 Score distributions for the NIR subsystem baseline and spoofing attacks. **a** Attacked by VIS Photo. **b** Attacked by NIR Photo

5.5 Countermeasure Integration

5.5.1 Color and Texture-Based Countermeasure

From the experiments in the previous section, we can see "sum rule" is good to improve the performance of face recognition but not robust to resist the spoofing attacks. Because "sum rule" prefers to get a trade-off between the VIS and NIR subsystems, it makes the multi-spectral system neither robust to VIS photo attack nor to NIR photo attack Table 5.1.

From the results in the previous section, we can see the VIS subsystem is robust to the NIR photo attack, and the NIR subsystem is robust to the VIS photo attack. By

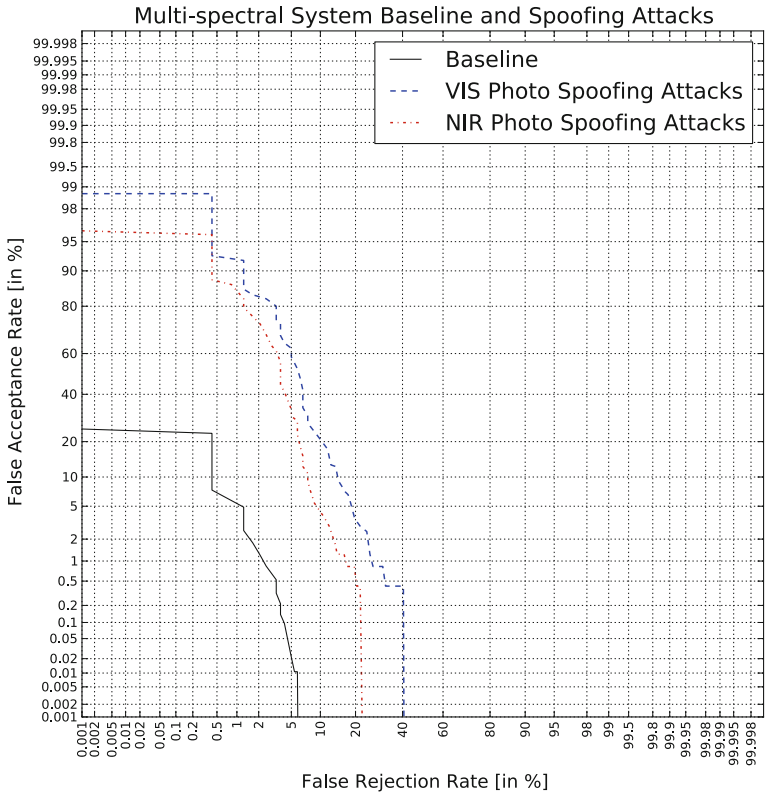

Fig. 5.10 DET curves for the multi-spectral system baseline and spoofing attacks

Table 5.1 The information of the multi-spectral face spoofing database

	Genuine subjects	VIS photo attack	NIR photo attack
Captured by VIS camera	100 subjects × 5	100 subjects × 5	100 subjects × 5
Captured by NIR camera	100 subjects × 5	100 subjects × 5	100 subjects × 5

combining their advantages, we propose a two-step countermeasure based on color and texture analysis. The color analysis is used to resist the NIR photo attack, because NIR photo captured by the VIS camera has no color. By setting a threshold based on color information, we could easily reject NIR photos. VIS photo usually can pass the color analysis, but it will be rejected by the following texture analysis due to its strong specular reflectance in the NIR spectrum. The process of the countermeasure is shown in Fig. 5.12.

For color analysis, we first crop the face region from image and then use the histogram of chroma (HoC) as feature, where the chroma of each pixel is calculated

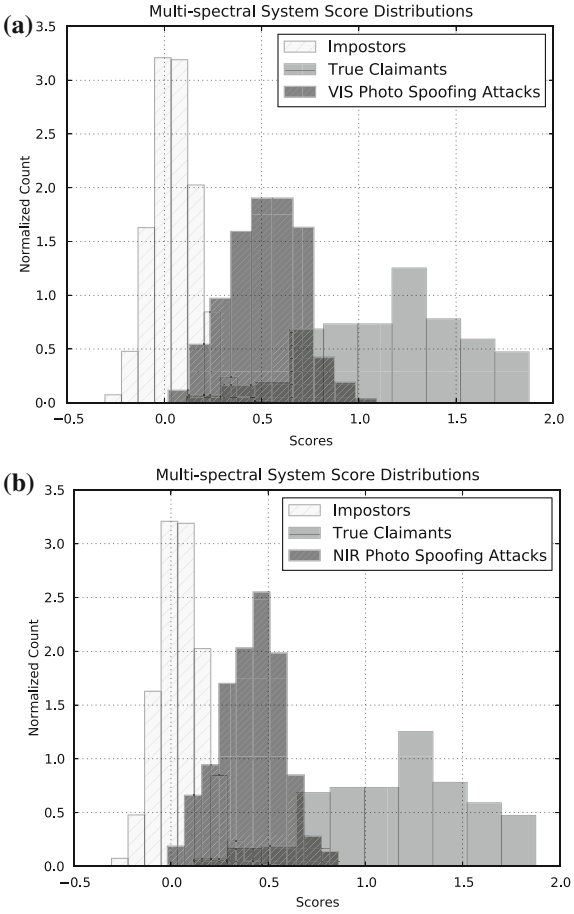

Fig. 5.11 Score distributions for the multi-spectral system baseline and spoofing attacks. **a** Attacked by VIS Photo. **b** Attacked by NIR Photo

by $\max(R, G, B)$–$\min(R, G, B)$. For texture analysis, the Gabor features, the same as face recognition (see Sect. 5.3), are used. Finally, the linear SVM is used to train the classifiers based on these two kinds of features respectively.

5.5.2 Protocol for Countermeasure

To train the classifiers for our countermeasure and evaluate the performance, we construct three subsets from the database as well:

- Training set: genuine and fake face images of the the first 30 subjects;

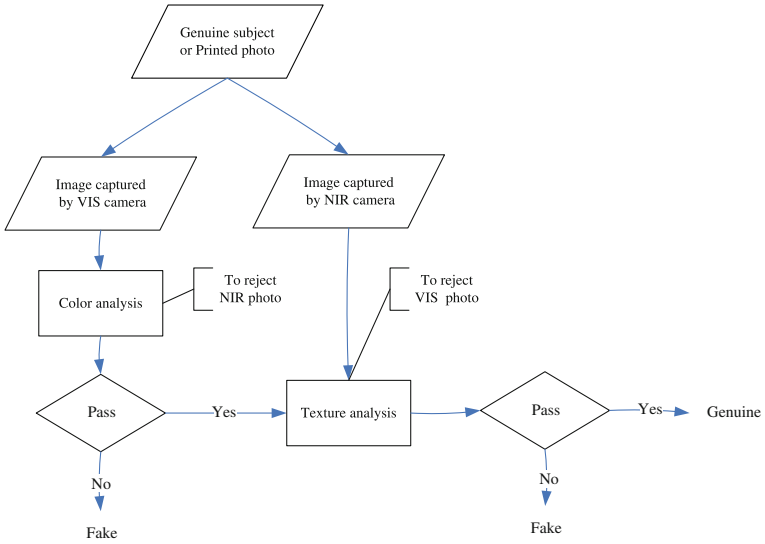

Fig. 5.12 The process of our countermeasure for the multi-spectral system

- Development set: genuine and fake face images of the following 30 subjects;
- Testing set: genuine and fake face images of the other 40 subjects.

First, we train the color-based classifier using G_{VIS} and NP_{VIS} in the training set. Second, we train the texture-based classifier using G_{NIR} and VP_{NIR}. The final countermeasure is constructed according to the structure shown in Fig. 5.12. The performance will be evaluated with respect to two kinds of errors: FLR (False Living Rate) and FFR (False Fake Rate). The lower these two errors, the better the performance of the countermeasure.

To evaluate the influence of the countermeasure to the face recognition system, we usually fix at a evaluation point (e.g., FFR $= 1\,\%$). Once fixed, we can incorporate the countermeasure as a pre-processing step into the multi-spectral face recognition system oriented to reject fake samples, and generate the performance of the following three profiles.

- Baseline: The performance of the multi-spectral system;
- Baseline under attacks: The performance under spoofing attacks;
- Baseline + Countermeasure under attacks: The final performance of the multi-spectral system with countermeasure under spoofing attacks.

5.5.3 Results

After training the color and texture SVM classifiers, we apply them on the testing set. Because the differences between the genuine and fake samples are obvious, the two

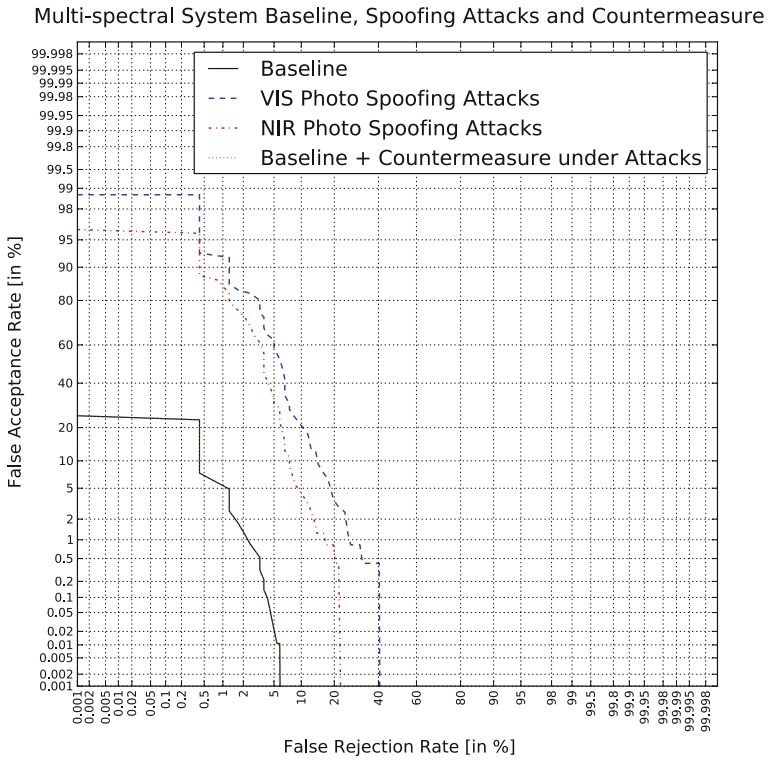

Fig. 5.13 DET curves for the multi-spectral system baseline, spoofing attacks, and countermeasure. Note that the plot of the "baseline + countermeasure under attacks" is invisible because it coincides with the baseline completely (perfect countermeasure)

SVM classifiers are easy to train and both achieve 100 % accuracy on the testing set. The color SVM classifier contains 20 support vectors and the texture SVM classifier contains 52 support vectors.

Due to the high performance of the two SVM classifiers, our countermeasure can reject all fake samples while allowing all genuine samples to pass. Therefore, the countermeasure can fully resist the photo attacks and do not produce any side effects to the multi-spectral system. The DET curve of the final system is shown in Fig. 5.13.

5.6 Conclusions

As shown in this chapter, mutli-spectral face recognition has high recognition rate and performs well to anti-spoof printed photo attacks. The success is mainly attributes to the complement of multi-spectral face images, as the VIS subsystem is robust to

NIR photo attack and the NIR subsystem is robust to VIS photo attack. Although the system works perfectly on the database, it is hard to say the anti-spoofing module in the system can appeal to the practical requirements. Limited by the scale and variations of the database, the introduced countermeasure, especially the texture classifier, may be over-fitting to the database. To really apply the anti-spoofing technologies in practice, we must collect more comprehensive spoofing databases, in larger scale, with more variations, such as pose, illumination, expression, etc., as similar as those in traditional face recognition.

Acknowledgments This work was supported by the Chinese National Natural Science Foundation Project #61070146, #61105023, #61103156, #61105037, National IoT R&D Project #2150510, European Union FP7 Project #257289 (TABULA RASA http://www.tabularasa-euproject.org), and AuthenMetric R&D Funds.

References

1. Duc N, Minh B (2009) Your face is not your password. Black Hat Conference
2. Li J, Wang Y, Tan T, Jain AK (2004) Live face detection based on the analysis of fourier spectra. In: Jain AK, Ratha NK (eds) Society of photo-optical instrumentation engineers (SPIE) conference series, vol 5404, pp 296–303
3. Tan X, Li Y, Liu J, Jiang L (2010) Face liveness detection from a single image with sparse low rank bilinear discriminative model. Proceedings of the 11th European conference on computer vision: Part VI., ECCV'10. Springer, Berlin, pp 504–517
4. Maatta J, Hadid A, Pietikainen M (2011) Face spoofing detection from single images using micro-texture analysis. In: Biometrics (IJCB), 2011 international joint conference on, pp 1–7
5. Pan G, Sun L, Wu Z, Lao S (2007) Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In: Computer vision, 2007. ICCV 2007. IEEE 11th international conference on, pp 1–8
6. Kollreider K, Fronthaler H, Bigun J (2008) Verifying liveness by multiple experts in face biometrics. In: Computer vision and pattern recognition workshops, 2008. CVPRW '08. IEEE computer society conference on, pp 1–6
7. Kollreider K, Fronthaler H, Bigun J (2005) Evaluating liveness by face images and the structure tensor. In: Automatic identification advanced technologies, 2005. Fourth IEEE workshop on, pp 75–80
8. Chetty G, Wagner M (2004) Liveness verification in audio-video speaker authentication. In: In proceeding of the 10th ASSTA conference, pp 358–363. Macquarie University Press
9. Anjos A, Marcel S (2011) Counter-measures to photo attacks in face recognition: a public database and a baseline. In: Biometrics (IJCB), 2011 international joint conference on, pp 1–7
10. Bai J, Ng TT, Gao X, Shi YQ (2010) Is physics-based liveness detection truly possible with a single image? In: Circuits and systems (ISCAS), Proceedings of 2010 IEEE international symposium on, pp 3425–3428
11. Zhang Z, Yan J, Liu S, Lei Z, Yi D, Li S (2012) A face antispoofing database with diverse attacks. In: Biometrics (ICB), 2012 5th IAPR international conference on, pp 26–31
12. Pavlidis I, Symosek P (2000) The imaging issue in an automatic face/disguise detection system. In: Computer vision beyond the visible spectrum: methods and applications, 2000. Proceedings of the IEEE workshop on, pp 15–24
13. Kim Y, Na J, Yoon S, Yi J (2009) Masked fake face detection using radiance measurements. J Opt Soc Amer A 26(4)
14. OSU thermal imagery database. http://www.cse.ohiostate.edu/otcbvs-bench/

15. Yi D, Lei Z, Hu Y, Li SZ (2013) Fast matching by 2 lines of code for large scale face recognition systems. CoRR abs/1302.7180
16. Li SZ, Chu R, Liao S, Zhang L (2007) Illumination invariant face recognition using near-infrared images. IEEE Transactions on pattern analysis and machine intelligence 26(Special issue on biometrics: progress and directions)
17. Viola P, Jones M (2011) Robust real time object detection. IEEE ICCV workshop on statistical and computational theories of vision. Vancouver, Canada
18. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models: their training and application. CVGIP: image underst 61:38–59
19. Fukunaga K (1990) Introduction to statistical pattern recognition, 2nd edn. Academic Press, Boston
20. Ahonen T, Hadid A, Pietikainen M (2004) Face recognition with local binary patterns. In: Proceedings of the European conference on computer vision. Prague, Czech, pp 469–481
21. Yi D, Lei Z, Li SZ (2011) A robust eye localization method for low quality face images. In: International joint conference on biometrics (IJCB). Washington, pp 15–21
22. Saragih J, Lucey S, Cohn J (2011) Deformable model fitting by regularized landmark mean-shift. Int J Comput Vis 91:200–215
23. Milborrow S, Morkel J, Nicolls F (2010) The MUCT landmarked face database. Pattern recognition association of South Africa (2010). http://www.milbo.org/
24. Rother C, Kolmogorov V, Blake A (2004) Grabcut: interactive foreground extraction using iterated graph cuts. ACM Trans Graph 23(3):309–314
25. Phillips PJ, Moon H, Rizvi SA, Rauss PJ (2000) The FERET evaluation methodology for face recognition algorithms 22(10):1090–1104
26. Okada K, Steffens J, Maurer T, Hong H, Elagin E, Neven H, von der Malsburg C (1998) The Bochum/USC face recognition system and how it fared in the FERET phase III test
27. Belhumeur P, Hespanha J, Kriegman D (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans PAMI 19(7):711–720