Face Liveness Detection by Learning Multispectral Reflectance Distributions

Zhiwei Zhang, Dong Yi, Zhen Lei, Stan Z. Li*
Center for Biometrics and Security Research & National Laboratory of Pattern Recognition
Institute of Automation, Chinese Academy of Sciences
{zwzhang,dyi,zlei,szli}@cbsr.ia.ac.cn

***Abstract*— Existing face liveness detection algorithms adopt behavioural challenge-response methods that require user cooperation. To be verified live, users are expected to obey some user unfriendly requirement. In this paper, we present a multispectral face liveness detection method, which is user cooperation free. Moreover, the system is adaptive to various user-system distances. Using the Lambertian model, we analyze multispectral properties of human skin versus non-skin, and the discriminative wavelengths are then chosen. Reflectance data of genuine and fake faces at multi-distances are selected to form a training set. An SVM classifier is trained to learn the multispectral distribution for a final Genuine-or-Fake classification. Compared with previous works, the proposed method has the following advantages: (a) The requirement on the users' cooperation is no longer needed, making the liveness detection user friendly and fast. (b) The system can work without restricted distance requirement from the target being analyzed. Experiments are conducted on genuine versus planar face data, and genuine versus mask face data. Furthermore a comparison with the visible challenge-response liveness detection method is also given. The experimental results clearly demonstrate the superiority of our method over previous systems.**

I. INTRODUCTION

Existing face recognition (FR) systems [1] are fragile to attacks of fake faces, since the detection of liveness of the captured faces has not been a built-in module. A typical FR system can be deceived by printed face pictures, video replays, or mimic masks. In general, fake faces have two main properties:

1) Large variations. Although the positive class, namely the genuine face, has limited variation (all genuine faces are human skins), the negative class, i.e., the fake faces, can range from photos, videos to masks and so on. When it comes to material level, the variety is even larger: take face mask for example, there are rubber mask, plastic mask, silica gel mask, etc. It's almost impossible to give a complete list. Some examples of fake faces are shown in Fig.1.
2) Indistinguishable under visible light. Fake is, by its definition, indistinguishable for human eyes. Therefore, without extra aid, only visual face images are insufficient and impossible for the detection of fake faces.

Previous works can be classified as visible liveness detection methods and multispectral liveness detection methods.

*Stan Z. Li is the corresponding author.

Fig. 1. Some fake face examples. Materials from left column to right are: silica gel, rubber, photo and video replay

The former is applied to face recognition systems which work under visible light. 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 kind of problem is that they can only detect planar faces such as photos. If the fake face is a mask, or even more simpler, some photos over a genuine face with eyes and mouth cut out as illustrated in [6], these methods will definitely fail. Therefore, the applications of such kind of methods are limited. It is also noticed that some researchers attempt to detect liveness from a single image [9], [10]. The clue they use is the illumination factor of the image. Apparently they can only be used for photo face detection and actually, at many times, these methods won't work well because photos can be made very vivid as genuine faces.

The other class is the multispectral methods, which detect the reflectance of object surface. To the best of our knowledge, there have been very few papers published in this field, among which two papers are most representative. In [3] Pavlidis and Symosekuses uses light at two wavelengths, and a simple threshold method to detect the genuine and fake faces. No experiments but only illustrations were reported in their paper. The second one [4] also selects light at two different wavelength and then LDA is used to make the final decision. However, this paper requires the distance between the user and the system is exactly 30cm, and they utilize users' forehead region to measure reflectance. Not only the forehead may be occluded, but also the exact distance is quite demanding and is impossible to execute in practice. Furthermore, the wavelengths they select are actually not as optimal as in [3], which we will show in section II-B. Thermal information is another choice, and we refer readers to a common facial thermal imagery database in [16]. However, it is clear that thermal radiation can pass through the wearing, such as the clothes, therefore it may not detect the case when attackers wear masks. Furthermore, the high cost also prevents its usage in real practice.

In this paper, we propose a novel liveness detection method using multi-spectral lighting. We start by analyzing how to distinguish fake faces multispectrally based on Lambertian model when the user-system distance is unlimited and variable. Then after measuring the albedo curves of different materials(skin and non-skin), two discriminative wavelengths are selected . A device is built to capture multi-spectral data of the face to be recognized and a classifier is trained on the multi-distance reflectance data set for the final liveness detection.

Compared with previous works, the advantages of our methods are obvious. Firstly, our method requires no user cooperation, and therefore is user-friendly and fast. Secondly, our multispectral method takes user-system distance factor into consideration, which is novel.

The rest of the paper is organized as follows: in section II, we analyze the problem based on the Lambertian reflectance model, followed by the multispectral light selection and classification process; in section III the system construction is given as well as three different experiments conducted under different cases, which show the superior of our method; in section IV, we conclude the paper.

II. MULTISPECTRAL LIVENESS DETECTION

Light beyond visual spectrum gives us hope to tackle the liveness detection problem. Indistinguishable fake faces may exhibit quite different properties under invisible light. In this section, we first give a reflectance analysis at multi-distances using Lambertian reflectance model in section II-A. Based on the analysis, the albedo curves of some materials and selection of the proper wavelengths are given in section II-B. In section II-C, the learning and classification process is presented.

A. Reflectance Analysis

According to the Lambertian reflectance model [14], the reflectance light intensity I at a location (x, y) is:

$$I(x, y) = A_0(x, y)r(x, y)\cos\theta(x, y) \tag{1}$$

in which $A_0(x, y)$ is the input light intensity at the facial location (x, y), $r(x, y)$ is the object albedo, and $\theta(x, y)$ is the angle between surface normal vector and the receiver's viewpoint.

According to the Beer-Lambert law [15], the attenuation of light through the air is

$$A_0 = Ae^{-cd} \tag{2}$$

in which A is the light source intensity, c is the attenuation coefficient in the air and d is the distance traveled. For simplicity, (2) is replaced by a function D, which is a monotone decreasing function of distance d, and by combining (1), we have:

$$I(x, y) = A(x, y)r(x, y)\cos\theta(x, y)D(d) \tag{3}$$

in which d is the distance between the object and the receiver.

The average value ave of I over an area Ω is,

$$\begin{aligned} ave &= \int\limits_{(x,y)\in\Omega} I(x, y)dxdy = \int\limits_{\Omega} (Arcos\theta D(d))\, dxdy \\ &= ArD(d) \int\limits_{(x,y)\in\Omega} cos\theta(x, y)dxdy \end{aligned} \tag{4}$$

assuming the uniform distribution of A, r and D in Ω. The average value ave is more robust and useful than I at a single point, and is used to represent reflectance intensity.

Given a genuine face $f1$ and a fake face $f2$, due to different material, $r_{f1} \neq r_{f2}$ exists under most of the wavelengths. Given a wavelength $w1$, from (4) we have

$$ave_{(f1,w1)} = A_{w1}r_{(f1,w1)}D(d1) \int\limits_{(x,y)\in\Omega} \cos\theta_{f1}dxdy \tag{5}$$

$$ave_{(f2,w1)} = A_{w1}r_{(f2,w1)}D(d2) \int\limits_{(x,y)\in\Omega} \cos\theta_{f2}dxdy \tag{6}$$

If there is no distance requirement, then $f1$ and $f2$ are undistinguishable because there always exist proper distance $d1$ and $d2$ ($d1$ is unnecessarily equal to $d2$) which makes

$$ave_{(f1,w1)} = ave_{(f2,w1)} \rightarrow \frac{r_{(f1,w1)}}{r_{(f2,w1)}} = \frac{D(d2)\int\cos\theta_{f2}}{D(d1)\int\cos\theta_{f1}} \tag{7}$$

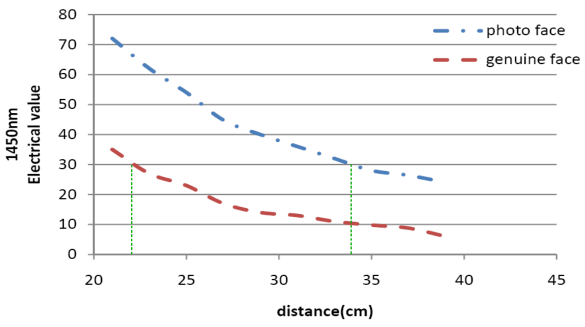
In order to handle the problem imposed by distance, we refer to a multispectral solution. If another wavelength $w2$ is added and received at the same area Ω, then we have

$$ave_{(f1,w2)} = A_{w2}r_{(f1,w2)}D(d1) \int\limits_{(x,y)\in\Omega} \cos\theta_{f1}dxdy \tag{8}$$

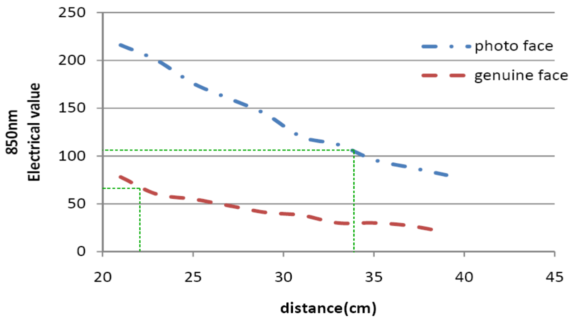
$$ave_{(f2,w2)} = A_{w2}r_{(f2,w2)}D(d2) \int\limits_{(x,y)\in\Omega} \cos\theta_{f2}dxdy \tag{9}$$

In order to distinguish $f1$ and $f2$ at wavelength $w2$, similar with (7) we get

$$ave_{(f1,w2)} \neq ave_{(f2,w2)} \rightarrow \frac{r_{(f1,w2)}}{r_{(f2,w2)}} \neq \frac{D(d2)\int\cos\theta_{f2}}{D(d1)\int\cos\theta_{f1}} \tag{10}$$



(a)



(b)

Fig. 2. An example for reflectance analysis at multi-distances. In (a), the reflectance-distance curves of a genuine face and a photo face at wavelength 850nm are given; then for a fixed reflectance value, two distances are obtained. In (b) the reflectance-distance curves at wavelength 1450nm are given; when the two distances of (a) are fixed, the reflectance values are different.

Combing (7) and (10), obviously if proper wavelengths $w1$ and $w2$ are selected which satisfy

$$\frac{r_{(f1,w1)}}{r_{(f2,w1)}} \neq \frac{r_{(f1,w2)}}{r_{(f2,w2)}} \tag{11}$$

then $f1$ and $f2$ can be distinguished without any distance requirement.

In Fig.2 an example is given for illustration. Fig.2(a) shows the reflectance[1] curves of a genuine face and a photo face at wavelength 1450nm as the distance increases (the choice of the wavelengths will be explained in next subsection). Given a reflectance value, 30 for example, the corresponding distance is about 22cm and 34.1cm for the genuine face and the photo face(the undistinguishable case). With the distance unchanged, now another light source of 850nm is added, and the reflectance curve is shown in Fig.2(b). Obviously, at the unchanged distance, the reflectance values at 850nm are different. Therefore, the genuine face can be classified from the fake photo face.

B. Select discriminative wavelengths

From section II-A, it is clear that proper wavelengths are crucial to correct classification of genuine-fake faces when there is no distance limitation. Also as stated in section I, there are great varieties for fake faces while the genuine faces are relatively constant, therefore it is better to analyze the reflectance properties of the skin rather than fake faces, based on which the proper wavelengths are to be selected.

From [11], a skin albedo curve can be found spanning a wide range, as shown in Fig.3. Wavelengths below 400nm is not considered because the ultraviolet rays are harmful to human beings. Visible wavelengths between 400nm and 700nm are not considered as well, because visible light source will make the users feel uncomfortable when lighting. One can easily notice that at wavelength 1450nm, the albedo of human skin is very low while at 800nm$\sim$900nm, the albedo is quite high. As current near-infrared face recognition systems also adopt 850nm as light source [2], we choose 1450nm and 850nm as our consideration. To further verify the choice, the albedo curves of some common materials, which are paper from a photo face, and two kinds of silica gel from two mask faces, are also tested and measured by us. The curves can be seen in Fig 4. Obviously the other materials do not possess such a high-low relationship at wavelength 850nm and 1450nm as human skin does, and these two wavelengths can satisfy (11). Therefore, our proper choice is 850nm and 1450nm. Obviously, the more light sources at different wavelengths are added, the more discriminative the reflectance value vector $\{ave_{(f,w1)}, ..., ave_{(f,wn)}\}$ is. But for the sake of simplicity, only two wavelengths are selected

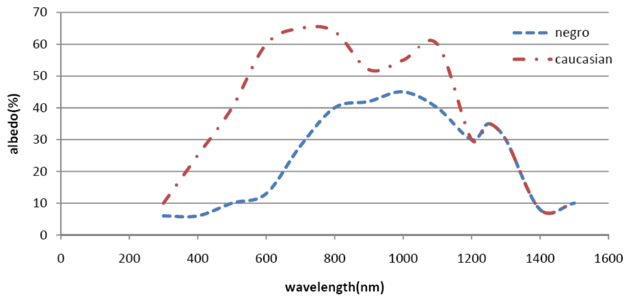


Fig. 3. The skin albedo curve of Caucasian and Negro, reproduced from [11]

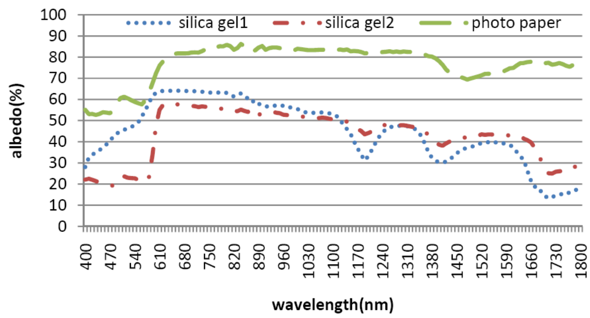


Fig. 4. We test the albedo curves of three materials, which are paper, and two kinds of silica gel.

[1]the reflectance is represented by the electrical value(voltage in this case) measured by a hardware system.

in this paper. Although there are only two wavelengths, the experiments show the discriminative power of our selection. In [4], however, the selected two wavelength is not the best. Besides 850nm, they select visible light at wavelength 560nm , which is not only uncomfortable to human eyes during experiment but also not as powerful as 1450nm when combined with 850nm.

C. Learning and classification at multi-distances

Based on the above discussion, with proper wavelengths, the fake faces can be detected very easily. However, as some simplifications and assumptions are added into the reflectance analysis, as well as the measure error in real practice, a learning and classification process has to be executed to achieve a high detection accuracy.

At the learning step, both positive and negative samples are measured at multi-distances to form a training database. It is believed that the multispectral facial reflectance distribution, namely the ave_{w1} and ave_{w2}, can be learned from the training database. Then an SVM classifier is trained as the Genuine-Or-Fake classifier. The reason for choosing SVM is that it is impossible to learn the multispectral reflectance distribution beforehand, and probably the distribution of genuine and fake faces are linearly inseparable. SVM is capable of calculating a nonlinear classifier, which is surly better than the mere threshold method in [3], and also better than a linear LDA method in [4]. Even if the distribution can be separated linearly, SVM is as good as LDA.

It is important to point out that in practice most of the face recognition systems require a proper face size, for example, with the eye distance more than 60 pixels. Similarly, we also expect that the distance between the face and the liveness detection device is about 20 $\sim$ 35cm. Such a broad requirement is clearly only a little deviation from the original intention of distance robust liveness detection.

III. Experiments

In this section, the system construction is firstly introduced in details in subsection A; then an experiment on detecting genuine faces and planar fake faces is conducted in B; in C we further test the effect of detecting genuine faces and many kinds of mask faces; in D a comparison between previous visible liveness detection method and our method is also given.

A. System Construction

The system is constructed on the near infrared face recognition system [2] with modifications. The system includes two groups of LED lights to provide active light sources. The two groups of LED are interweaved and evenly located as a rectangle. Because it is impossible to purchase any cameras corresponding to 1450nm light, two photodiodes are used instead to receive the reflectance light at both wavelengths. For photodiodes, the electrical values(voltage in this case), which is measured by a circuit, represent the sum of light intensity over certain areas, which equals ave explained in section III.A. If the receiving area of the photodiodes are

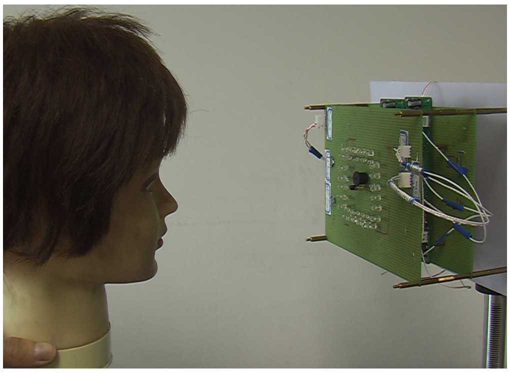


Fig. 5. System Overview

restricted to a certain size by limiting the receiving angle of the photodiode, then the two photodiodes have the same size of receiving area, making the electrical value at two wavelengths could represent the average facial reflectance properly. A view of our system is shown in Fig.5.

B. Genuine faces vs Planar faces

In this section we mainly concentrate on the planar faces, the most common fake face type. Planar faces are named because the faces are in 2 dimensional plane rather than in 3 dimensional space. Common planar faces are photos, and video replays. 40 face images from FRGC database are selected and are printed as photos, among which 20 are printed on high-quality paper and the other 20 are printed on normal paper. 20 faces are also shown on a laptop screen as the video replays. Each photo and video replay is measured 5 times, each time at different distances ranging from 20cm to 35cm. We also selected 40 persons to form the genuine face database. Each person is measured 3 times, each time at different distances as well. Therefore a database containing 300 negative samples (200 for photos and 100 for video replays) and 120 positive samples is built. A 5-fold cross validation is taken to give the final testing results, which is shown in Table 1. LibSVM 2.89 [12] is used as the classifier with default parameters.

The multispectral reflectance distribution of the photo faces and genuine faces are shown in Fig 6. In experiment it is found that the electrical values of the video replays are much higher than these of the photos and genuine faces (approximately 3 times more than the photos). If the electrical values of video replays as well as those of the genuine and photo faces are plotted in the same image, the latter will become unclear to see. Therefore the distributions of the video replays are simply omitted in Fig 6.

C. Genuine faces vs Mask faces

We further test the genuine faces versus mask faces. Compared with planar faces, mask faces are less common and more difficult to detect. Firstly, masks are 3D objects, whose surface terrain are more like a genuine face than a 2D planar face; secondly, the materials of masks, such as silica

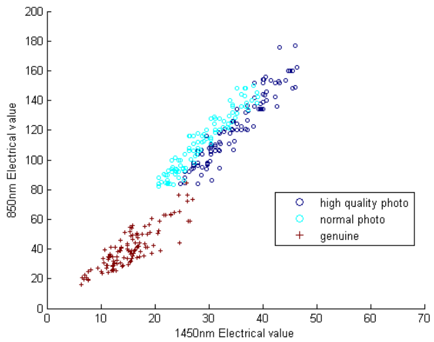


Fig. 6. The reflectance distribution of genuine faces and photos

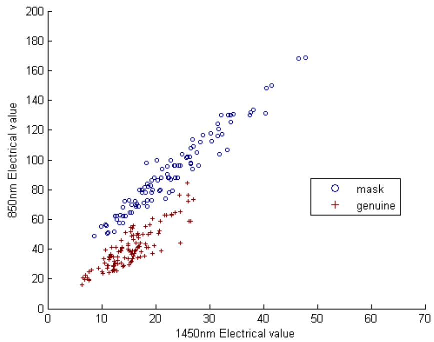


Fig. 7. The reflectance distribution of genuine faces and mask faces

gel and rubber, is more closer to human skins in reflectance than paper-based photos and glass-based video screens.

As it is quite difficult for us to acquire face-like masks, only 20 masks are selected, each sampled 5 times at different distances to form a 100 negative database. The materials of the 20 masks are: 4 plastic masks, 6 silica gel masks, 4 paper pulp masks, 4 plaster masks and 2 sponge masks. Once again 5-fold cross validation is taken to give a final testing results. Because each material class only contains a few masks, it is meaningless to calculate a detection accuracy for each class. Instead a detection accuracy treating all masks as a whole is given in Table 2 and the multispectral distribution is shown in Fig.7. From Fig.7 it can be seen that the distribution difference between genuine faces and mask faces are not as much as that of the planar faces. This is mainly because there are more material types for mask faces than mere plane faces, and the 3D surface also makes them more similar to genuine faces. Our result is not as high as that in [4], but it cannot be asserted that our method is worse. Reasons for different detection accuracy between ours and [4] are: 1) the mask faces used in the experiments are different; 2) they fix the distance as 30cm, and use the forehead for test. On the other hand, our method has no such fixed distance requirement and is based on the whole face. Actually, our method is more convenient in real practice, yet their requirement is too heavy for the users.

TABLE I

DETECTION ACCURACY OF GENUINE FACE VS PLANAR FACE

5 cross	1	2	3	4	5
genuine vs photo	93.6%	92.9%	89.6%	93.9%	91.0%
G vs P average			92.2%		
genuine vs video	100%	100%	100%	100%	100%
G vs V average			100%		

TABLE II

DETECTION ACCURACY OF GENUINE FACE VS MASK FACE

5 cross	1	2	3	4	5
accuracy	90.5%	87.8%	90.4%	90.2%	87%
average			89.18%		

D. Comparison with visible liveness detection

To present a complete comparison of current face liveness detection methods, a comparison experiment is also conducted between the visible liveness detection method and our multispectral one. As stated above, visible liveness detection method is essentially to find the biological facial motions. The motion detected here is the most frequently used motion type – eye blinking. During experiment, a period of 10 seconds is set for the users to blink in front of the camera. 10 photos, 10 photos with eyes cut out and 10 genuine faces are used during the experiment, as shown in Fig.8. With eyes cut out on the photos, people can hide behind the photos to show blinking, as does in [6]. For visible liveness detection, a Haar + GentleBoost [13] classifier is trained to detect the blinking. The classifier, also known as eye state classifier, is trained on an eye state database as does in [5], and is used to give a blink score to each eye state image: positive scores mean open state and negative scores mean close state. A training database containing 2000 open eye images and 1000 close eye images is collected; a testing database of the same size is also collected, on which the classifier achieves an accuracy of 98.7% in classifying open eyes and an accuracy of 98.3% in classifying close eyes. A complete blink process shown in Fig.9; for the multispectral liveness detection, our method is employed and the training data is from the above experiments. The liveness detection results are shown in Table 3.

From Table 3 it can be seen that with eyes cut out from the photos, traditional visible liveness detection method can still detect motion, namely the blinking. In this case, visible liveness detection method cannot even detect photos properly, let alone other types of fake faces which are capable of exhibiting motions as well. On the other side, our multispectral method only depends on the material covering the faces, and therefore can properly detect fake faces. Furthermore, waiting for the users' interaction is quite time-consuming, which is set to be 10 seconds in this experiment; yet our method only takes 1 second to measure reflectance. This comparison experiment proves the superior of our method than the visible liveness detection method.

However, an important note should be declared: challenge-

Fig. 8. Fake faces attacking the blinking-based visible liveness detection method. From left to right are photos, photos with eyes cut out,and the genuine faces.

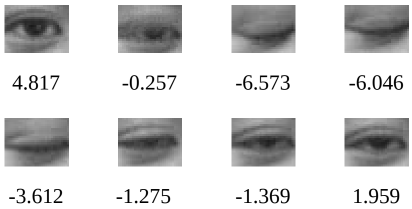


Fig. 9. Every eye state is given a Blink Score by the Haar+GentleBoost classifier in a complete blink process..

response based visible liveness detection method is not the competitor of the multispectral livenss detection method but a useful complement. The combination of the two methods will produce more precise results: when any of them fails the detection, the other one may work well to prevent any wrong decision, because the attackers have to pass both of the two methods to be verified alive.

IV. Conclusion

In this paper we propose a distance robust face liveness detection method, which performs better than previous works. Firstly an analysis on facial reflectance at multi-distances is given and after measuring the albedo curves of some materials, two discriminative wavelengths are selected to build our multispectral system. Samples are measured at multi-distances, and then a SVM classifier is trained to learn the distribution. Experiments are conducted on genuine faces vs plane faces, genuine faces vs mask faces, showing the effectiveness of our method. Furthermore we make a comparison with visible liveness detection method, proving the power of our method. Compared with previous works, our method does not adopt time-consuming and user-unfriendly interactions; and we consider the distance factor, which is crucial in real practice.Our future work will lie on improving the detection accuracy on mask faces.

V. Acknowledgement

This work was supported by the Chinese National Natural Science Foundation Project #61070146, National Science and Technology Support Program Project #2009BAK43B26, and AuthenMetric R&D Funds.

References

[1] Stan Z. Li and Anil K Jain, Handbook of Face Recognition , Springer, New York, 2004.

TABLE III

VISIBLE VS MULTISPECTRAL LIVENESS DETECTION

genuine detected	photo(10)	cutted photo(10)	genuine(10)
visible	0%(0/10)	90%(9/10)	100%(10/10)
multipsepctral	0%(0/10)	0%(0/10)	100%(10/10)

[2] Stan Z. Li, RuFeng Chu, Shengcai Liao and Lun Zhang, "Illumination Invariant Face Recognition Using Near-Infrared Images", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, April 2007.

[3] Ioannis Pavlidis and Peter Symosek, "The imaging issue in an automatic face/disguise detection system", *proceedings of IEEE workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, 2000.

[4] Youngshin Kim, Jaekeun Na, Seongbeak Yoon, and Juneho Yi, "Masked fake face detection using radiance measurements", *Journal of the Optical Society of America A*, vol. 26, no. 4, 2009.

[5] Gang Pan, Lin Sun, Zhaohui Wu and Shihong Lao, "Eyeblink-based Anti-Spoofing in Face Recognition from a GenericWebcamera", *the 11th IEEE International Conference on Computer Vision*, Rio de Janeiro, October, 2007.

[6] K. Kollreider, H. Fronthaler and J. Bigun, "Verifying Liveness by Multiple Experts in Face Biometrics", *IEEE Computer Vision and Pattern Recognition Workshops*, Anchorage, 2008.

[7] K. Kollreider, H. Fronthaler and J. Bigun, "Evaluating Liveness by Face Images and the Structure Tensor", *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, October, 2005.

[8] G. Chetty and M. Wagner, "Liveness Verification in Audio-Video Speaker Authentication", *In 10th Australian Int. Conference on Speech Science and Technology*, December, 2004.

[9] J. W. Li, Y. H. Wang, T. N. Tan, and A. K. Jain, "Live Face Detection Based on the Analysis of Fourier Spectra", *In Proc. SPIE Biometric Technology for Human Identification*, vol. 5404, pp. 296–303, January 1999.

[10] Xiaoyang Tan, Yi Li, Jun Liu and Lin Jiang, "Face Liveness Detection from A Single Image with Sparse Low Rank Bilinear Discriminative Model", *In Proceedings of the European Conference on Computer Vision*, 2010.

[11] R. Rox Anderson, B.S. and John A. Parrish M.D., "The Optics of Human Skin", *The Journal of Investigative Dermatology*, vol.77, no.1, 1981.

[12] Chih-Chung Chang and Chih-Jen Lin , LIBSVM, http://www.csie.ntu.edu.tw/ cjlin/libsvm/

[13] Jerome Friedman, Trevor Hastie and Robert Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting", *In Annals of Statistics*, vol. 28, pp.2000-2043, 1998.

[14] Ronen Basri and David W. Jacobs, "Lambertian Reflectance and Linear Subspaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.25, no.2, 2003.

[15] J. D. J. Ingle and S. R. Crouch, "Spectrochemical Analysis", Prentice Hall, New Jersey, 1988.

[16] OSU Thermal Imagery Database. http://www.cse.ohio-state.edu/otcbvs-bench/.