Face Liveness Detection using Variable Focusing

Sooyeon Kim[1], Sunjin Yu[2], Kwangtaek Kim[3], Yuseok Ban[1], and Sangyoun Lee[1]

[1]Department of Electrical and Electronic Engineering, Yonsei University, Republic of Korea
{sykim1221, van, syleee}@yonsei.ac.kr
[2]Department of Broadcasting and Film, Cheju Halla University, Republic of Korea
sjyu@chu.ac.kr
[3]Human Computer Interaction Group, Microsoft Research Asia, China
kwangk@microsoft.com

Abstract

As Face Recognition(FR) technology becomes more mature and commercially available in the market, many different anti-spoofing techniques have been recently developed to enhance the security, reliability, and effectiveness of FR systems. As a part of anti-spoofing techniques, face liveness detection plays an important role to make FR systems be more secured from various attacks. In this paper, we propose a novel method for face liveness detection by using focus, which is one of camera functions. In order to identify fake faces(e.g. 2D pictures), our approach utilizes the variation of pixel values by focusing between two images sequentially taken in different focuses. The experimental result shows that our focus-based approach is a new method that can significantly increase the level of difficulty of spoof attacks, which is a way to improve the security of FR systems. The performance is evaluated and the proposed method achieves 100% fake detection in a given DoF(Depth of Field).

1. Introduction

Surveillance systems have been developed briskly and the importance of those are also growing[17]. Instead of just using PIN code, industries have applied biometric authorization systems like face recognition in order to reinforce security[12]. Biometric traits are definitely strong factors to preserve one's personal information. However, attacks with spoofed faces are increasing and threatening many surveillance systems and privacy. Actually, there is a case that face recognition system embedded in cellular phones gives an approval to forged faces. For this reason, the need of defense technologies against spoof attacks is arising to protect users' private data and to provide secure mobile banking.

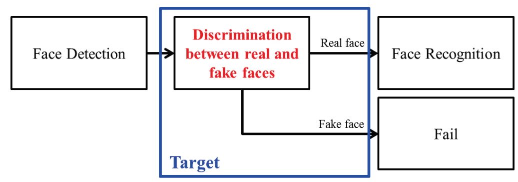


Figure 1. Target of this paper

There are several types of fake face samples from 2D photos to skin-like masks. Of those samples, making hard copies of faces by using a computer printer is the easiest and cheapest means to intruders. In previous works, a large number of researches have shown steady progress in developing anti-spoofing technologies over the last decade [11]. Most of those focus on applying features obtained from the analysis of textures, spectrums and motions to detect face liveness. However, they require extra devices such as infrared sensors or complex computation.

In this paper, we present a novel method to prevent from infringing face identification systems with forged 2D photos. Our proposed method utilizes variable focusing, one of camera features. For instance, any image taken from a high-end camera has both focused and out-focused(i.e. blurred) areas. The focus area can be controlled by users from foreground to background during a shot, though there is a limitation due to the unique effective focal length of each lens. With this function, we could find that there is an obvious difference in focus values between real and fake faces when two sequential images(in/out-focus) are collected from each subject, assuming that there is no big movement. Defocusing is used to estimate the depth information in practice[22]. It becomes a good feature to identify live faces. Real faces

978-1-4799-0310-8/13/$31.00 ©2013 IEEE

are solid, but images which are targeted as spoofed faces are flat. This feature makes our system be able to discriminate which face is real or not.

The remainder of this paper is organized as follows. In Section 2, we discuss related works on face liveness detection and a theoretical background about camera focusing. Our proposed system is described in Section 3. Database acquisition and experimental results are shown in Section 4. Finally, concluding remarks are provided in Section 5.

2. Related works

2.1. Face liveness detection

There are two main approaches to detect spoofed faces : intrusive and non-intrusive. Intrusive methods request users' cooperation such as turning ones head [1] and speaking some words, so those provide users with an inconvenient interface. For this reason, many researchers have put more efforts to develope non-intrusive methods using a visual camera. These are categorized as two groups, regarding the number of input images. The technologies in the first group are based on a single input image. Due to using a still-image, they exploit an analytic way. Texture and frequency components are typical and crucial characteristics for distinguishing fake faces. Kim *et al*.[8] applied Local Binary Patterns(LBP) for texture analysis and power spectrum for frequency analysis. Määttä *et al*.[13] and Bai *et al*.[3] also detected face liveness by examining micro texture with multiscale LBP. Peixoto *et al*.[16] proposed a method to detect and keep edges with difference of Gaussian under bad illumination. In [20], they extracted essential information for discrimination by using Lambertian model. Single image-based methods have low capacity and simplicity, but do not produce information about motion. Unlike single image-based approaches, multiple image-based methods exploit movements. Especially, signs of liveness, such as eye blink and head movements, are used intuitively[15, 7, 21, 2]. In addition, optical flow and various illumination are helpful to analyze the difference between real and fake faces [10, 4, 6]. To make the system robust, some methods use a combination of static and dynamic images[21, 18].

However, tools and skills for disguising have been gradually evolved. Masks and camouflages make it difficult to determine the truth. To tackle these problems, researchers have considered using extra sensors as well as a visual camera. Thermal and near infrared sensors are typical examples. Zhang *et al*.[23] proposed a method which measures reflectance of skin using near infrared sensor. Sun *et al*.[19] showed a thermal IR and visible light correlation system with a thermal infrared sensor. Although their suggestions can winnow out whether faces are real or not, those do not have merits in terms of cost and commercialization.

2.2. Backgrounds related to focusing

Our method discriminates fake faces from real ones, using the effect of out-focusing. To improve the liveness detection performance, we increase out-focusing effect. In order to make high out-focusing effect, depth of field(DoF) should be narrow. The DoF is the range between the nearest and farthest objects in a given focus. People perceive that entities in a DoF are sharp. There are three ways to have low DoF. The first is to open the aperture of a camera wider. This is related to having small F-stop. The second is to adjust the focal length longer. The last one is to decrease the distance between the camera and the subject. These options make it capable to obtain low DoF and a large variation effect in out-focusing[5].

For the purpose of computing the variations of focus from images taken in focus, a measure metric that can precisely estimate the level of focusing at pixels should be considered. We utilize Sum Modified Laplacian(SML)[14] as the focus value measurement in our method. The SML represents degrees of focusing in images and those values are represented as a transformed 2nd-order differential filter in Eq. (1). In Eq. (1), $SML(i,j)$ is the SML measurement at a pixel (i,j). N is the size of a window and T is an empirical threshold. The edges in focus look very sharp and have high values from the SML. On the contrary, a region of defocusing has severe blurring and losses on edge information that leads to low values of SML. This way, the SML measurement is used for interpreting the degree of focusing.

$$
\begin{gathered}
SML(i,j) = \sum_{x=i-N}^{i+N} \sum_{y=j-N}^{j+N} ML(x,y) \cdot t \\
ML(x,y) = \Delta_{ML}^2 I = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right| \\
t = \begin{cases} 1 & ML(x,y) \geq T \\ 0 & otherwise \end{cases}
\end{gathered} \tag{1}
$$

3. Proposed method

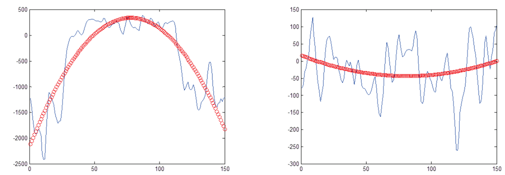
3.1. Basic constraint

We propose a novel method for detecting 2D fake faces. The key point of our method is using one of camera features, variable focusing. By controlling a camera, we can take pictures focused on facial components such as a nose and ears. In case of real faces, focused regions are clear and others are blurred due to depth information. In contrast, there is little difference between images taken in different focuses from a printed copy of a face, because they are not solid. We emphasize on this characteristic so as to discriminate fake faces from real ones.

Note that our method relies on the degree of DoF that determines the range of focus variations at pixels from the sequentially taken images. More specifically as mentioned

(a) (b) (c) (d)

Figure 2. Samples of input images (a) real face focused on nose, (b) real face focused on ears, (c) fake face focused on nose, and (d) fake face focused on ears



(a) (b)

Figure 4. Results of curve fitting (a) real face and (b) fake face (subject ID: 1)

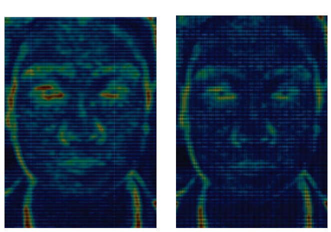
in chapter 2.2, if DoF is too wide to get a face focused in an ear or the nose, there could be no difference in SML measurements between fake and real faces. Therefore, input data for our method should be acquired by a camera that supports sufficiently narrow DoF.

3.2. Algorithm

In the first step, we take two sequential pictures by focusing the camera on facial components. One is focused on a nose(I_N) and the other is on ears(I_E). The nose is the closest to the camera lens, while the ears are the farthest. The depth gap between them is sufficient to express a 3D effect. Figure 2 shows samples of input images. From these images, faces are detected. In this paper, we assume that frontal faces are already found before our liveness detection algorithm runs.

In order to judge the degree of focusing, SMLs are calculated. In Figure 3, (a) is SML of a real face focused on the nose and (b) is that on an ear. Visually bright pixels on the image mean large variations which are also represented as high values computed from the SML measurement and vice versa. We denote both SMLs by S_N and S_E.

After obtaining SMLs, we subtract S_E from S_N (S_N - S_E). This aims to maximize the SML gap between regions of nose and ears. For one-dimensional analysis, we sum differences of SMLs (DoS) in each of columns. If DoS is accumulated in every row, changes of SMLs will be blended and this makes it difficult to collate characteristics. Sums of DoS of real faces show similar patterns consistently, whereas those of fake faces do not. The differences in the patterns between real and fake faces are used as features to detect face aliveness.



(a) (b)

Figure 3. SMLs of images (a) focused on nose(S_N) and (b) focused on ears(S_E)(real face)

However, unrefined distributions are not proper to be employed for distinguishing faces, because there are a lot of noises. Therefore, curve fitting is applied to extract meaningful features. The sum of DoS is analogous to a curvature of a quadratic equation $y = ax^2 + bx + c$. This has three coefficients, $\mathbf{A} = \begin{bmatrix} a & b & c \end{bmatrix}^T$. To find them, we perform error minimization as Eq. (2).

$$
\begin{array}{ccccc}
\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_W \end{bmatrix} & = & \begin{bmatrix} x_1^2 & x_2^2 & \dots & x_W^2 \\ x_1 & x_2 & \dots & x_W \\ 1 & 1 & \dots & 1 \end{bmatrix}^T & \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\
\mathbf{Y} & = & \mathbf{X}^T & \mathbf{A}
\end{array}
\tag{2}
$$

$$\mathbf{e} = \mathbf{Y} - \mathbf{X}^T\mathbf{A}$$

$$\mathbf{A} = (\mathbf{X}\mathbf{X}^T)^{-1}\mathbf{X}\mathbf{Y}$$

$\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & \dots & y_W \end{bmatrix}^T$ is sum of DoS and $\begin{bmatrix} x_1 & x_2 & \dots & x_W \end{bmatrix}^T$ is $\begin{bmatrix} 1 & 2 & \dots & W \end{bmatrix}^T$. W is the width of I_N and I_E.

Figure 4 represents sums of DoS(blue line) and results of curve fitting(red circles). The sums of DoS of real faces fit to a convex curvature like (a) in Figure 4, but those of fake faces do not. Also, due to cubic effect, the difference between maximum and minimum values of a real face is much larger than that of a fake one.

In Figure 5, blue circles stand for coefficients of real

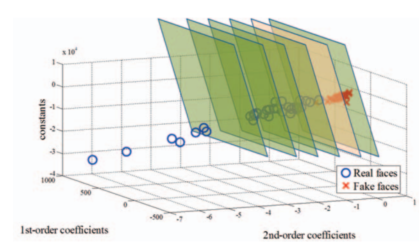


Figure 5. Distribution of features and Classifier

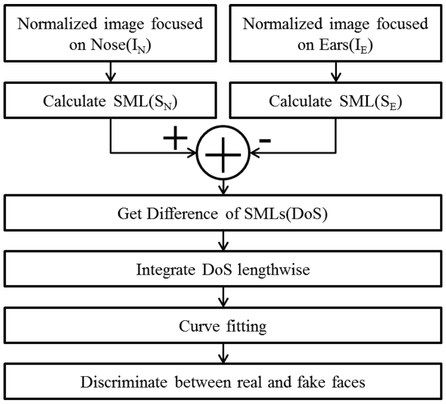
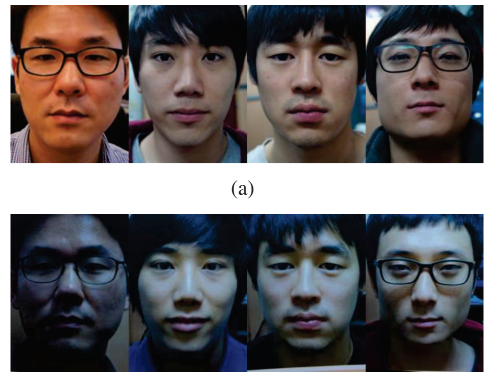


Figure 6. Flowchart of our proposed method



(a)

(b)

Figure 7. Normalized images of (a) real and (b) fake faces

faces and red crosses mean those of fake faces. As shown, features of fake faces are gathered and are not overlapped with those of real ones. This distribution helps us recognize which one is forged. In order to detect fake faces, we set a plane equation $px + qy + rz + s = 0$ for a classifier. $\begin{bmatrix} p & q & r \end{bmatrix}^T$ is a normal vector of the plane and is equal to a principal axis of data. $\begin{bmatrix} x & y & z \end{bmatrix}^T$ is a set of coefficients of a quadratic equation extracted in the previous step. When we substitute the feature into the plane equation, if the result is higher than 0, input face is real. But, if not, it is fake. The flowchart of our algorithm is represented in Figure 6.

4. Experiments

4.1. Database

There is no open facial database which has various focal areas and existing data are not satisfying our demand. So, we obtained frontal face images from 12 subjects by using a SONY-NEX5 camera without regard to lighting conditions. For fake faces, we printed out photos with a Fuji Xerox ApeosPort-II C5400 printer.

Table 1. Experimental parameters

Distance between the camera and the subject(cm) (D)	Focal length(mm) (L)	F-stop (F)	Depth of Field(cm) (DoF)
40	16	2.8	6.84
		4.0	9.74
		7.1	17.9
	28	4.0	3.04
		7.1	5.44
50	16	2.8	10.8
		4.0	15.5
	28	4.0	4.83
		7.1	8.65

We emphasized that DoF is important to decide the performance of our system. To examine characteristics of features in accordance to DoF, we adjusted the distance between the camera and the subject, focal length and F-stop. Table 1 presents nine cases and each has different DoF from 3.04cm to 17.9cm. According to DoF, we categorized those as three groups such as within 5cm, 10cm and 20cm. Table 2 shows the number of faces.

Before making a decision which face is real or not, we assumed that faces were detected and were normalized based on locations of eyes. The distance between eyes is fixed as 60 pixels and the size of images is 200 by 150. Figure 7 shows samples of normalized faces.

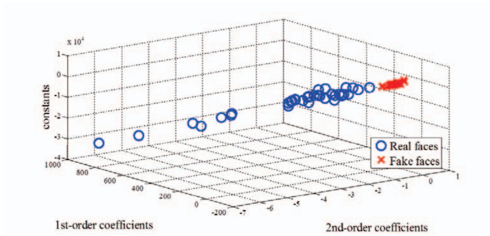
4.2. Experimental results

To evaluate the classification rate, we had 5-fold experiments in each case for cross validation[9]. In Table 2, 80% of images were used for training and residuals were for testing. False Acceptance Rate(FAR) and False Rejection Rate(FRR) were calculated for quantitative analysis. FAR is a rate of the numbers of fake images misclassified as real and FRR is a rate of the numbers of real images misclassified as fake. As stated in Table 3, when the range of DoF is within 5cm, the average of FAR is 2.86% and that of FRR is 0.00%. On the other hand, in case of DoF within 20cm, the average of FAR is 9.77% and that of FRR is 9.67%. These results show how crucial to make DoF small.

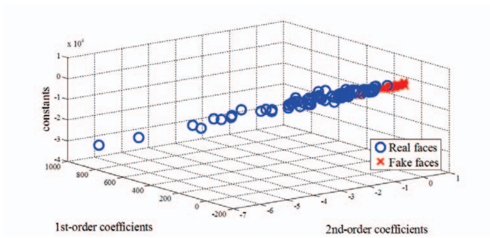
Using the total database, we could obtain the classifier which has 0.00% Equal Error Rate(EER). Figure 8 depicts distributions of features of real(blue circles) and fake(red

Table 2. The number of images

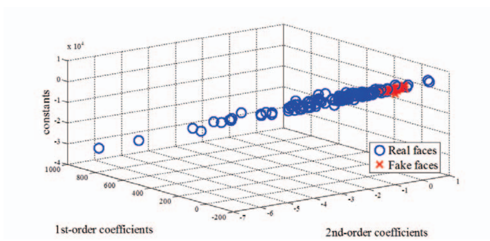
DoF	The number of Real faces	The number of Fake faces
within 5cm	35	35
within 10cm	84	84
within 20cm	123	123



(a)



(b)



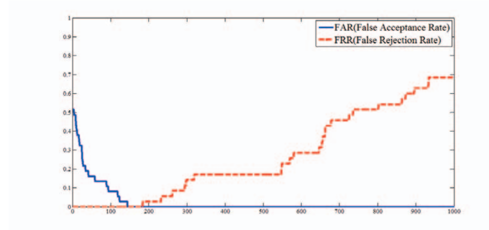
(c)

Figure 8. Feature distributions (a) DoF within 5cm, (b) DoF within 10cm, (c) DoF within 20cm

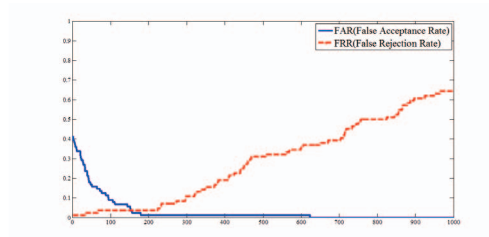
crosses) faces. When the DoF is within 20cm, features are overlapped as Figure 8(c). This dispersion causes high EER(9.57%) in Figure 9(c). However, when the DoF is within 5cm, features are divided into two groups clearly and the EER is 0.00%. Figure 10 is the Receiver Operating Characteristics(ROC) curves and shows the performances of three cases. As shown in Figure 10 and Table 4, the case

Table 3. Results(%) (# of misclassified images / # of test images)

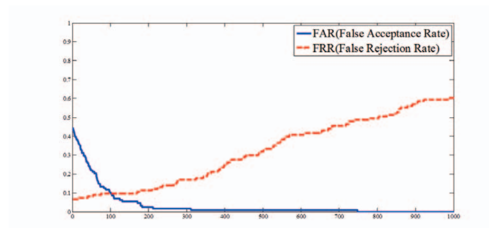
		1	2	3	4	5	Avg
within 5cm	FAR	0.00 (0/7)	0.00 (0/7)	0.00 (0/7)	14.3 (1/7)	0.00 (0/7)	2.86
	FRR	0.00 (0/7)	0.00 (0/7)	0.00 (0/7)	0.00 (0/7)	0.00 (0/7)	0.00
within 10cm	FAR	11.8 (2/17)	0.00 (0/17)	5.88 (1/17)	0.00 (0/17)	0.00 (0/16)	3.54
	FRR	11.8 (2/17)	0.00 (0/17)	0.00 (0/17)	5.88 (1/17)	0.00 (0/16)	3.54
within 20cm	FAR	8.00 (2/25)	4.00 (1/25)	16.0 (4/25)	12.5 (3/24)	8.33 (2/24)	9.77
	FRR	16.0 (4/25)	16.00 (4/25)	8.00 (2/25)	0.00 (0/24)	8.33 (2/24)	9.67



(a)



(b)



(c)

Figure 9. FAR and FRR curves of (a) DoF within 5cm, (b) DoF within 10cm, (c) DoF within 20cm

of within 5cm-DoF has the best performance. Although we acquired the small number of database, we could construct the classifier which has the high accuracy, if DoF is sufficiently narrow.

Table 4. Equal Error Rate(EER)(%)

	within 5cm	within 10cm	within 20cm
EER	0.00	3.47	9.57

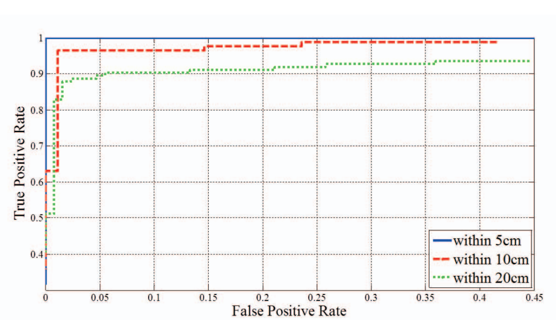


Figure 10. ROC curves of three cases

5. Conclusion

We proposed a face liveness detection method using characteristics of variable focusing. If DoF is adjusted as small as possible, our system can have better performance in discriminating whether faces are real or not. In the experiment, with a small DoF, the method can detect all fake faces. In addition, we think that data acquisition will be possible by exploiting compact cameras or camera-embedded cellular phones. Moreover, the proposed method can be easily adapted for the devices without any other sensors. For the future work, we will increase the number of differently focused data which may gaurantee more reliable performance. Also we will consider counter measures keeping from diverse spoofed faces including 3D attack by using effective texture analysis.

Acknowledgements

This work was supported by the Industrial Strategic technology development program(10040018, Development of 3D Montage Creation and Age-specific Facial Prediction System) funded by the Ministry of Knowledge Economy(MKE, Korea). Also, this was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST)(No. 2011-0016302).

References

[1] A. Ali, F. Deravi, and S. Hoque. Liveness Detection Using Gaze Collinearity. In *2012 Third International Conference on Emerging Security Technologies(EST)*, 2012.

[2] A. Anjos and S. Marcel. Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline. In *IJCB '11 Proceedings of the 2011 International Joint Conference on Biometrics*, 2011.

[3] J. Bai, T.-T. Ng, X. Gao, and Y.-Q. Shi. Is Physics-based Liveness Detection Truly Possible with a Single Image? In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 3425–3428, 2010.

[4] W. Bao, H. Li, N. Li, and W. Jiang. A liveness detection method for face recognition based on optical flow field. In *International Conference on Image Analysis and Signal Processing*, pages 233–236, 2009.

[5] G. M. Cope. Depth of Field:The Misunderstood Element in Image Design, 2010.

[6] C. H. Huang and J. F. Wang. SVM-based One-Against-Many Algorithm for Liveness Face Authentication. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, pages 744–748, 2008.

[7] H. Jee, S. Jung, and J. Yoo. Liveness Detection for Embedded Face Recognition System. *International Journal of Biological and Life Sciences*, 1:235–238, 2005.

[8] G. Kim, S. Eum, J. K. Suhr, D. I. Kim, K. R. Park, and J. Kim. Face Liveness Detection Based on Texture and Frequency Analyses. In *2012 5th IAPR International Conference on Biometrics (ICB)*, pages 67–72, 2012.

[9] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2*, IJCAI'95, pages 1137–1143. Morgan Kaufmann Publishers Inc., 1995.

[10] K. Kollreider, H. Fronthaler, and J. Bigun. Non-intrusive liveness detection by face images. *Image and Vision Computing*, 27:233–244, 2009.

[11] O. Khm and N. Damer. 2D Face Liveness Detection: an Overview. In *2012 BIOSIG - Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2012.

[12] S. Z. Li and A. K. Jain. *Handbook of Face Recognition*. Springer, 2011.

[13] J. Määttä, A. Hadid, and M. Pietikäinen. Face Spoofing Detection From Single Images Using Micro-Texture Analysis. In *IJCB '11 Proceedings of the 2011 International Joint Conference on Biometrics*, 2011.

[14] S. K. Nayar and Y. Nakagawa. Shape from Focus. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16:824–831, 1994.

[15] G. Pan, L. Sun, W. Zhaohui, and Y. Wang. Monocular camera-based face liveness detection by combining eyeblink and scene context. *Telecommunication Systems*, 47:215–225, Aug. 2011.

[16] B. Peixoto, C. Michelassi, and A. Rocha. Face liveness detection under bad illumination conditions. In *IEEE 2011 International Conference on Image Processing*, pages 3557–3560, 2011.

[17] T. Raty. Survey on Contemporary Remote Surveillance Systems for Public Safety. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 2010.

[18] W. R. Schwartz, A. Rocha, and H. P. Edrini. Face Spoofing Detection through Partial Least Squares and Low-Level Descriptors. In *IJCB '11 Proceedings of the 2011 International Joint Conference on Biometrics*, 2011.

[19] L. Sun, W. Huang, and M. Wu. TIR/VIS Correlation for Liveness Detection in Face Recognition. In *Computer Analysis of Images and Patterns Lecture Notes in Computer Science*, volume 6855, pages 114–121, 2011.

[20] X. Tan, Y. Li, J. Liu, and L. Jiang. Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model. In *ECCV 2010 Lecture Notes in Computer Science*, volume 6316, pages 504–517, 2010.

[21] R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena, G. Murgia, M. Ristori, and F. Roli. Fusion of multiple clues for photo-attack detection in face recognition systems. In *IJCB '11 Proceedings of the 2011 International Joint Conference on Biometrics*, 2011.

[22] R. Veerender, K. Acharya, J. Srinivas, and D. Mohan. Depth Estimation Using Blur Estimation in Video. *International Journal of Electronics and Computer Science Engineering*, 2012.

[23] Z. Zhang, D. Yi, Z. Lei, and S. Z. Li. Face liveness detection by learning multispectral reflectance distributions. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 2011.