Context Based Face Spoofing Detection Using Active Near-Infrared Images

Xudong Sun, Lei Huang and Changping Liu
Institute of Automation
Chinese Academy of Sciences
95 Zhongguancun East Road, 100190, Beijing, China
Email: {sunxudong2013, lei.huang, changping.liu}@ia.ac.cn

***Abstract*—In this paper, with the help of controllable active near-infrared (NIR) lights, we construct near-infrared differential (NIRD) images. Based on reflection model, NIRD image is believed to contain the lighting difference between images with and without active NIR lights. Two main characteristics based on NIRD images are exploited to conduct spoofing detection. Firstly, there exist obviously spoofing media around the faces in most conditions, which reflect incident lights in almost the same way as the face areas do. We analyze the pixel consistency between face and non-face areas and employ context clues to distinguish the spoofing images. Then, lighting feature, extracted only from face areas, is utilized to detect spoofing attacks of deliberately cropped medium. Merging the two features, we present a face spoofing detection system. In several experiments on self collected datasets with different spoofing media, we demonstrate the excellent results and robustness of proposed method.**

I. Introduction

Face recognition systems are being widely used in a variety of applications. One of the main problems is that the existing systems are vulnerable to face spoofing attacks. To defend against spoofing attacks conducted by photos, videos or 3D masks, face spoofing detection has received increasing interest in the recent years. Many anti-spoofing methods have been proposed.

Most face anti-spoofing techniques utilize visible light (VIS) images, and they can be roughly classified into three main categories depending on the employed features. The first category extracts texture information. Many methods have been proposed to analyze reflectance [1], [2], [3], power spectrum [4], image banding effects [5], multi-scale feature [6], [7], blurriness [2], [7] and other texture clues [8], [9]. They can achieve satisfying results, but high quality photos may fool this kind of feature. The second category uses motion information. Motion based methods mainly deals with human physiological response or human reaction movements, such as eye blinking [5], [10], human motion [5] and head rotation [11]. They use the relative features across frames, but are vulnerable to replay attacks. Approaches in the third category utilize other image clues such as 3D depth, multi-modal knowledge [12] and hybrid model of several features [5], [11], [13].

However, most existing methods are mainly concerned about facial texture and motion analysis of VIS spectrum, which is sensitive to the change of illumination. While face recognition of near-infrared (NIR) spectrum [14], [15] has been discussed a lot before and several datasets [15] are concerned about VIS and NIR images, there are only a few face spoofing detection methods dealing with infrared images. Kim *et al.* [1] exploits the reflectance disparity between VIR and VIS illuminations and separates masked fake faces. Sun *et al.* [12] analyzes correlation between thermal IR (TIR) and VIS spectrum faces to improve liveness detection ability. It is known that infrared components also exist in some light sources such as sunshine; sometimes when capturing NIR images, people have to make efforts to minimize environmental lighting under controlled illumination [14], which is not convenient. Even for indoor applications, various light conditions can drastically change appearances of NIR images with different spoofing media.

Moreover, context information has only drawn a little attention in the previous works. Pan *et al.* [10] compares the difference between reference scene images and input ones, and checks whether background scenes change suddenly. Nevertheless, reference scene images are sometimes difficult to obtain, and the method only works for a stationary camera. Yan *et al.* [5] shows that motion of face and background has high consistency for fake facial photos and low for genuine ones. Anjos *et al.* [13] measures motion correlation between face and background using movement direction. Both Yan and Anjos only measure overall motion between face and background, and the methods might be deceived by deliberately recorded videos. Besides, motion based methods may take a relatively long time to capture input image sequences. Komulainen *et al.* [8] detects the upper-body and the presence of display medium to conduct spoofing detection. The method is effective, but it could not work for large spoofing medium of which the boundary is outside the image, or spoofing medium with weird size that the descriptor fails to detect.

In this paper, we present a novel spoof detection solution using active near-infrared lights. We have three main contributions as follows:

(1) Active NIR lights are used to illuminate faces from the frontal direction. Using NIR instead of VIS images can avoid phenomena such as camera saturation, which overcomes uncontrolled illumination changes and provides suitable images for reflectance analysis. By turning on and off active near-infrared lights, we construct near-infrared differential (NIRD) images to represent face images with active NIR lighting.

(2) Given that most spoofing attacks need a display medium

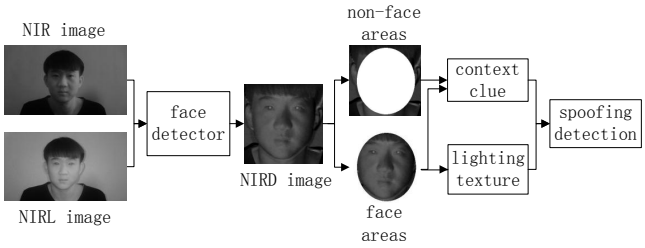

Figure 1. Proposed face spoof detection algorithms based on NIRD images.

and some of the non-face areas are shown by the spoofing medium, we exploit consistency of pixel intensities from face and non-face areas on the basis of NIRD images. Compared to other methods based on background consistency, our proposed methods do not rely on motion analyses or reference scene images; moreover, human need not to follow additional instructions.

(3) For deliberately cropped media, we employ lighting texture analysis to tell this kind of spoofing faces apart. The additional lighting feature adds robustness to our spoofing detection system.

II. SYSTEM OVERVIEW

We use a near-infrared spectrum camera with resolution of 1280×720 to capture images. On the camera, we mount one active NIR light-emitting diode (LED) with the central wavelength of 850 nm to provide frontal lighting. The LED lighting source has limited power, but for indoor use when the camera-face distance is from 30 to 80 cm, it is strong enough to provide clear NIR face images. General framework of our proposed method is illustrated in Fig.1.

For input NIR images, first of all, we use Viola-Jones face detector from OpenCV library to locate face regions. In most conditions, the original NIR images are usually too dark while NIR images with additional light are too bright. The abnormal conditions may do harm to the accuracy of existing face detector, so we adopt a simplified automatic exposure correction [16] to preprocess the input images, in which we adopt gamma compression on dark images and gamma expansion on bright images. After face regions are detected, images are normalized according to eye locations.

Employing our methods proposed in section III, we can obtain NIRD images from normalized ones. Utilizing face regions defined by face detector and a proper proportional size, we employ an elliptical mask to obtain face areas. The non-face areas are chosen outside the elliptical mask.

Next, we employ context clues and lighting texture from NIRD images. For the former, we try to exploit pixel intensity correlation between face and non-face areas and propose a consistency measurement. The feature is utilized to detect whether non-face areas share the same lighting conditions with face areas. For the latter, to make sure that our proposed method is robust enough to detect medium cropping or other spoofing attacks, we analyze the texture information only from face regions in NIRD images.

Finally, we construct a fusion descriptor based on these two features and we evaluate Support Vector Machine (SVM) algorithms on fusion model to make the final decision.

III. MODELING OF NIRD IMAGES

A. Reflectance Analysis

In our case, we make several assumptions about light sources in our experiments:

(a) There is one common fluorescent lamp as the main light source near the human, denoted as I_1

(b) There is an active near-infrared light to the front of the human, the light may decay sharply at a long distance, denoted as I_2

(c) The lighting component from other remote light sources, including computer screen and natural light, is denoted as ambient lighting I_a

According to the Lambertian reflectance model[17], the light reflectance can be decomposed in to diffuse reflection and specular reflection. In other words, the intensity at position x of a face image I is described theoretically as $I(x) = I_a + \sum_i f_i(d) \cdot (I_{i,d}(x) + I_{i,s}(x))$, where diffuse and specular reflection of the ith light are denoted as $I_{i,d}$ and $I_{i,s}$ respectively. $f_i(d)$ is light attenuation function at distance d. According to assumptions, attenuation function of NIR light $f_2(d)$ falls off quickly.

We capture two kinds of images sequentially: original NIR images without NIR lighting (denoted as NIR images) and images with active NIR lighting (denoted as NIRL images). For the former, the image intensities consist of I_a, $I_{1,d}$ and $I_{1,s}$, and for the latter, they also have $I_{2,d}$ and $I_{2,s}$ components, as shown in formula (1):

$$\begin{aligned} I &= I_a + f(d_1) \cdot (I_{1,d} + I_{1,s}) \,, \\ I^{(L)} &= I_a + f(d_1) \cdot (I_{1,d} + I_{1,s}) + f(d_2) \cdot (I_{2,d} + I_{2,s}) \,. \end{aligned} \tag{1}$$

When preparing experiment data, we are trying to shorten the time between turning on and off the active NIR lights. It can be assumed that human head poses do not change much in the process, which makes the diffuse and specular reflection of fluorescent lamp I_1 stay almost the same. The unchanged head poses also lead to the unchanged light attenuation values. Moreover, the ambient lighting I_a is unrelated to active NIR lights and can be regarded as a constant number.

By making use of normalized face images generated by methods in section II, we can ensure as accurate pixel correspondence as we can. On the other hand, near-infrared lighting also makes our methods more convincing, which will be inspected in detail in section III-B.

By subtracting I from $I^{(L)}$ in formula (1), the impact by active NIR light can be properly calculated. Thus, we can calculate the near-infrared differential (NIRD) images by:

$$I^{(NIRD)} = I^{(L)} - I = f(d_2) \cdot (I_{2,d} + I_{2,s}) \,. \tag{2}$$

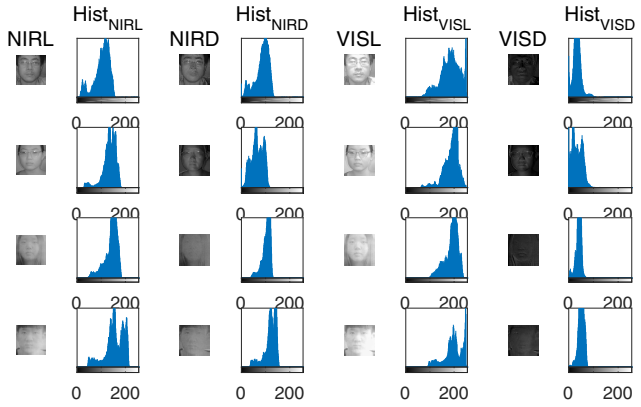

Figure 2. The figure illustrates images and histograms. The 1st, 3rd, 5th, 7th columns show NIRL, NIRD, VISL, VISD images respectively, and columns with even number show the corresponding histograms. The first two rows are genuine faces and the last two rows are spoofing faces.

B. Active NIR versus VIS Images

It has already been reported in [14] that the impact caused by visible ambient lighting can be reduced a lot by using NIR imaging systems, even if there are visible light lamps on left or right side. NIR spectrum can make images into good illumination condition, in which images have suitable pixel intensities, i.e., good contrast and not saturated. In this paper, we use active NIR light from the frontal direction to overcome possible uncontrolled illumination changes, since unsaturated image pixels are critical to our algorithm and NIRD images, which will be shown later.

According to our assumption, the light source I_1 is a typical "cool white" fluorescent lamp, which contains plenty of visible light ray components with only a few near-infrared components [18]. For VIS images, the intensities of all pixels can be higher than those in NIR images. In many situations, when additional light (I_2 in our assumption) is turned on, the VIS pixel intensities are often saturated. In other words, the input intensity signal strength exceeds the limit of camera sensors, which causes camera saturation and blooming [19]. So, the formula (2) might no longer be correct for VIS images and the lighting difference made by I_2 could not be separated properly. On the other hand, given that infrared component of I_1 is relative lower, NIR images are more likely to be unsaturated and have good pixel contrast.

To verify these analyses, besides NIR and NIRL images, we capture visible light images (denoted as VIS images) and VIS images with additional light (denoted as VISL images). The lighting images and differential images of both spectra, together with their corresponding histograms are illustrated in Fig.2.

It can be clearly noted that some of the pixel intensities approach maximum value in VISL images; in contrast, NIRL images are usually unsaturated. Moreover, most of the pixel intensities in VIS differential images are apparently lower than those in NIR differential images. This phenomenon just validates our analysis before. NIR images are indispensable for our proposed algorithms. The detailed results of using VIS spectrum images are shown in section V-C.

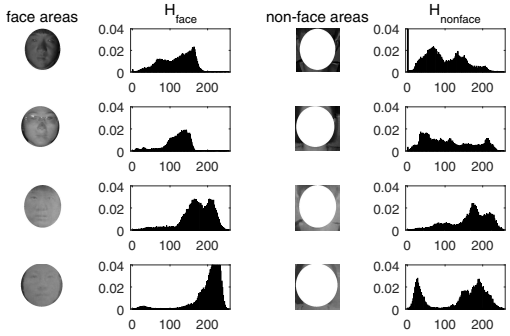

Figure 3. Examples of face and non-face areas, and their corresponding histograms. The first two rows demonstrate genuine faces and the last two rows indicate spoofing faces. The second and the fourth columns are histograms of column one and column three.

IV. Fusion Feature

A. Context Clue

Since that spoofing attacks need a display medium, there exist some "background" pixels around faces which are actually displayed in spoofing media. They do not belong to the true background, and we call them non-face areas in spoofing images in contrast with face areas. In genuine images, non-face pixels belong to backgrounds.

For pixels in the background, the distances are much longer than those in face areas; light attenuation function may reduce rapidly so we believe pixel intensities are much smaller than those in face areas. In other words, the pixels in the background will not suffer from great changes due to active NIR light. In spoofing images, some non-face areas have similar reflecting appearance to face areas, especially when they are close to each other. But in genuine images, non-face pixels are quite distinct from pixels in face areas.

Above all, we exploit different distribution of pixel intensities between face and non-face areas. The correlation between pixels in different areas is quite robust and even if background scene is combined with photos or photos are placed close to the camera, the consistency of pixel intensity property keeps still in spoofing images. We use entropy consistency to obtain spoofing clues. If pixels in face and non-face have high consistency, we consider the input face as fake one; otherwise, if the consistency is low, maybe there is no spoofing medium around faces, indicating that the input face may be a genuine one.

The histograms of both face and non-face areas are calculated. For the numbers of pixels may not be equal, the two histograms are normalized to make sure ℓ_1-normalization of each histogram is equal to 1. They are illustrated in Fig.3, denoted as H_{face} and $H_{nonface}$ respectively. Fig.3 demonstrates that shapes of real faces are not ideal ellipse and we use a relative larger ellipse mask. Thus our ellipse face areas may contain some pixels which belong to non-face areas. It is not a good signal for our consistency calculation. To handle this problem, we

construct a third histogram H_{diff} by subtracting $H_{nonface}$ from H_{face}:

$$H_{diff}(i) = \max(H_{face}(i) - H_{nonface}(i), 0)\,, \qquad (3)$$

where i is an element of all intensity values and $H(i)$ is histogram value of intensity i.

Now we define a three-dimension context consistency feature as the concatenation of three entropies of the three histograms: H_{diff}, H_{face} and $H_{nonface}$.

The entropy of histogram H is defined as $entropy(H) = -\sum(H(i). * \log H(i))$, and entropy of zero value is defined as 0.

Sometimes, the coat collar, long hair, hat and other decoration may affect the histogram of non-face areas severely. They have approximate the same distance as genuine faces, and they can reflect the active light just as the spoofing medium dose, which may confuse our proposed methods. Luckily, this interference around the faces is not as uniform as the spoofing medium and it usually does not surround human faces in all directions.

To deal with problems like this, we use part based model and bisect the whole image equally along horizontal and vertical directions. Thus the image is decomposed into four regions. Then, we extract context consistency feature individually in each part. One major drawback of the decomposition is that some discriminative information carried by the whole image may be scattered and lost in the process. Consequently, we also employ the whole face image as one big part and obtain the same feature from it.

In this way, there are five parts in total, including four splitted regions and one whole image, in which the context consistency feature is extracted individually. Finally, the features are concatenated together as context clues to the input NIRD face image.

B. Lighting Texture

Proposed context clues are robust where the spoofing medium is hidden and the boundary could not be detected directly. But if the medium is deliberately cropped and only face areas remain, or there is something around genuine faces in all directions, context clues may not work as desired. So, another feature is needed to assist, which should be extracted only from face areas to handle troublesome situations. The lighting texture is examined in this section.

Given that distance between face and camera is relative longer than the depth of genuine faces and warped images, the attenuation values for all pixels in face areas stay approximately the same or they change slowly and smoothly. The attenuation values are considered to be a constant number and the formula (2) can be simplified as:

$$I^{(NIRD)}(x) = k_d \cdot \omega_d(x) \cdot E + k_s \cdot \omega_s(x) \cdot E\,. \qquad (4)$$

In formula (4), E is the light intensity and it is believed to be constant at all pixels. $\omega_d(x)$ and $\omega_s(x)$ are geometric factors. k_d and k_s are weighting factors for diffuse and specular components respectively, which contain constant parameters such as attenuation values and material parameters. Material parameters are related to materials, pixel position and angle of incident light, but for opaque material and frontal lighting, they are usually regraded as constant, too [3], [17]. Based on NIRD images, we provide a brief analysis of reflectance differences in genuine faces and three types of spoofing medium studied in this paper:

1) Genuine faces: Normally, specular highlights are located around several specific locations, such as nose tips, cheeks, forehead and glasses. The geometric factors are quite sophisticated, so diffuse reflection varies from areas to areas, and shades might exist on NIRD images.
2) Photos printed on common A4 paper: Specular reflectance is quite weak and there are few specular highlights. It has been reported that printing process probably reduce image contrast [19]. Even if they are warped, the surface normals are smooth, so they may obtain gradually changing diffuse reflection.
3) Photos printed on resin-coated (RC) paper: They have smooth and highly polished surfaces, and there might be more specular reflectance components. Moreover, the glossy ink layer on RC papers makes the surface scatters a lot in NIR spectrum and enhance diffuse reflection.
4) Photos displayed on tablet screen: The screen could not be bent and has no geometric distortion, so the surface normals are similar for all pixels. Specular reflection usually appears because of the screen. It emits lights itself and diffuse components may be affected.

As analyzed above, differences exist in both diffuse and specular components between genuine and spoofing faces. Specular reflectance can contribute the large pixel intensities only in a small area and raise the deviation of the whole image, while diffuse reflectance may enlarge pixel intensities on a large scale and reduce some edge information. So the differences will lead to various pixel intensities dramatically in NIRD face areas, and the spoofing images are likely to show quite a different intensity distribution compared to that of genuine ones.

Many statistics figures [2] have been explored to conduct spoofing detection and they have shown excellent results. In this paper, after pixels in face areas are collected, three statistical figures are employed to represent the intensity distribution: i) the mean intensities of face area pixels μ, ii) the standard deviation σ, and iii) the skewness γ. The three figures contain statistical distribution of lighting textures, forming texture clues of three dimensions for face areas.

V. EXPERIMENT

A. Dataset Overview

Due to the characteristic of proposed method, it is quite hard to employ it to current popular datasets. In this paper, we utilize several self-collected anti-spoofing datasets on different spoofing media to verify our proposed method. NIR images and

Figure 4. Some samples in our NIRD image datasets. Images from top row to down are genuine faces, A4 paper photos, RC paper photos and tablet screen photos respectively. The first and third columns are NIR images, and the second and fourth columns are NIRL images.

Table I
THE NUMBER OF IMAGE PAIRS IN OUR DATASET.

	Genuine	A4 Spoof	RC Spoof	Tablet Spoof
NIR pairs	1099	1092	375	796
VIS pairs	1106	1140	426	791

NIRL images are sequentially taken with active NIR light off and on. For genuine faces, all people in datasets are allowed to change their head poses freely. If some subjects wear glasses, we use their photos with glasses as attacking images.

We use three methods to conduct spoofing attacks: high-quality color photos printed on resin-coated (RC) papers, photos printed on A4 papers and photos displayed on the tablet screen. In all printed spoofing attacks, photos are allowed to be warped freely. To demonstrate the benefit of active near-infrared images in our proposed algorithm, besides NIR and NIRL images, we also capture VIS and VISL images using the same spoofing media.

There are 30 subjects taking part and there are 13650 images (6825 pairs) with detected face regions in total, as listed in Table I. The images in which faces cannot be located are discarded, and the situation often appears for NIR images and RC spoofing images. Due to the space limitations, normal visible light images are not illustrated and some gray-scale images of near-infrared spectrum are shown in Fig.4.

B. Baseline Methods

As far as we know, there are no algorithms concerned on processing differential images for spoofing detection. We conduct some texture based methods on collected images as comparisons.

According to recent research [20], although LBP based methods are proposed relatively a long time before, they are still powerful local texture descriptors in terms of face spoofing detection such as CoA-LBP, and most of them have public access code [20], which makes it much easier and more convincing to make comparisons. Besides, multiscale local binary pattern (MsLBP) and related methods draw much attention in previous works [6] [7], we combine MsLBP [6] and multi color space [9] as one baseline method, which works better than merely CoA-LBP in our dataset. Besides, Lee *et al.* [4] also presents a method based on entropy using power spectra, and it is chosen as another baseline method.

C. Results and Analysis

In this section, a fifteen-fold cross-validation SVM classifier, with images from twenty six subjects as training sets and images from the other four subjects as testing sets, is applied for evaluation for both proposed methods and baseline methods. Mean half total error rate (HTER) and mean equal error rate (EER) are used to evaluate the performance.

Each method is examined on three kinds of spoofing media (A4, RC and tablet spoofing), and we also test each method on images of all three media. Numerical results for NIR spectrum are shown in Table II. As mentioned before, to verify the meirt of using NIR images, the same experiments are performed on VIS images and the results are shown in Table III.

Table II
COMPARISONS FOR EER (%) AND HTER (%) ON NIR DATASET.

Method	A4 Spoof		RC Spoof		Tablet Spoof		ALL	
	EER	HTER	EER	HTER	EER	HTER	EER	HTER
Proposed	**1.08**	**3.31**	**0.25**	5.49	**0.11**	**0.25**	**1.20**	**2.96**
MsLBP[6]	1.36	3.72	0.33	**4.84**	0.29	0.95	3.04	6.87
Entropy[4]	6.13	11.93	4.81	8.92	2.62	5.96	4.76	10.23

In Table II, we notice that when spoofing images of all spoofing media are mixed up together, various texture features might confuse the final classifier, leading to relative poor performances. It occurs to both two baseline methods, but our proposed method suffers less from it.

One reason might be that we adopt several ways for hiding the display medium and fooling the system mentioned in [8] when collecting images, such as combining background scenes into spoofing photos and making photos close to the sensor. With proper camera focus, images captured from spoofing faces may visually look similar to those from genuine ones, which fools existing methods such as MsLBP and entropy of power spectra to some extent. Fortunately, even if the media are hidden or cropped, the spoofing faces still rely on spoofing media, and reflectance information is different from genuine faces, especially in NIRD images. We employ reflectance consistency and lighting clues to conduct spoofing detection and it performs excellently on NIR images of all spoofing media.

Table III
COMPARISONS FOR EER (%) AND HTER (%) ON VIS DATASET.

Method	A4 Spoof		RC Spoof		Tablet Spoof		ALL	
	EER	HTER	EER	HTER	EER	HTER	EER	HTER
Proposed	4.81	13.23	**0.29**	17.70	**0.08**	**0.52**	4.01	10.07
MsLBP[6]	**0.82**	**3.59**	6.16	**9.87**	0.10	0.61	**1.99**	**5.04**
Entropy[4]	6.96	18.13	22.28	16.63	5.03	11.45	8.52	16.03

For all three spoofing media of both spectra, some results seem to be regular. Taking tablet spoofing images for examples, for a given spectrum, ERR and HTER values are always the smallest for every method. It is related to the property of

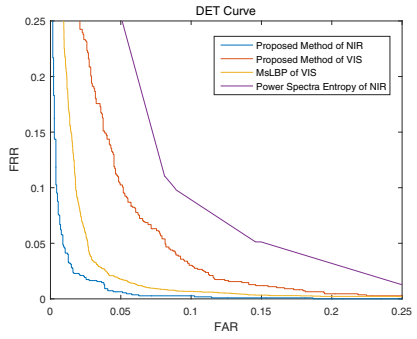

Figure 5. Results obtained on four experiments.

tablet itself. The tablet screen emits VIS lighting rather than NIR lighting. The images shown by tablet screen may not be influenced a lot by the addition active light. Except the obvious highlights caused by screens, the other parts have similar appearances because of flat and unwarped screens. These objective reasons make it easier to distinguish tablet spoofing faces apart.

Comparing the results of both VIS and NIR spectra from Table II and Table III, proposed method behaves quite different. As mentioned earlier, the results of NIR spectrum are much better than those of VIS spectrum. As for power spectra entropy [4], the same thing happens and most of the numerical results deteriorate quickly with VIS images. Entropy based methods are just too vulnerable when dealing with color images with unconstrained VIS lighting. Interestingly, the results show MsLBP [6] does better than the other two methods with VIS images, and even better than MsLBP on NIR images. The VIS images do not worsen MsLBP, but benefit the method instead. This may be explained that MsLBP is only concerned about the texture of input images on different scales, which relies on image intensity. It does not care about the reflectance component. While NIR images are gray-scale and quite dark for the majority of time, NIR cannot provide enough intensity information such as color texture as VIS spectrum.

As discussed before, it seems unfair to compare our method with MsLBP of NIR spectrum, so we use MsLBP of VIS spectrum instead. Fig.5 shows the results of four experiment settings: proposed method of both spectra, MsLBP of VIS spectrum and entropy of NIR spectrum. All spoofing images of different spoofing media are used in these experiments. It should be noted that proposed method achieve perfect performance at EER of 1.2% with one channel gray-scale NIR images, which beats baseline method MsLBP with color VIS images and outperforms baseline methods using entropy on power spectra. In addition, the usage of NIR differential images benefits our proposed method a lot.

VI. CONCLUSION

A novel solution based on active near-infrared images is proposed for face spoofing detection. Unlike most existing spoofing detection techniques, NIR spectrum is adopted and differential image is analyzed. The context consistency is calculated between face and non-face areas to detect the existence of spoofing media. Lighting texture is employed to deal with extreme situations such as cropped spoofing media. Experiment results demonstrate the promising results and robustness for different spoofing media. Future works will be aimed at using image spatial information to enhance the performance.

REFERENCES

[1] Y. Kim, J. Na, S. Yoon, and J. Yi, "Masked fake face detection using radiance measurements," *JOSA A*, vol. 26, no. 4, pp. 760–766, 2009.

[2] D. Wen, H. Han, and A. K. Jain, "Face spoof detection with image distortion analysis," *Information Forensics and Security, IEEE Transactions on*, vol. 10, no. 4, pp. 746–761, 2015.

[3] X. Tan, Y. Li, J. Liu, and L. Jiang, "Face liveness detection from a single image with sparse low rank bilinear discriminative model," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 504–517.

[4] T.-W. Lee, G.-H. Ju, H.-S. Liu, and Y.-S. Wu, "Liveness detection using frequency entropy of image sequences," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2367–2370.

[5] J. Yan, Z. Zhang, Z. Lei, D. Yi, and S. Z. Li, "Face liveness detection by exploring multiple scenic clues," in *Control Automation Robotics & Vision (ICARCV), 2012 12th International Conference on*. IEEE, 2012, pp. 188–193.

[6] J. Yang, Z. Lei, S. Liao, and S. Z. Li, "Face liveness detection with component dependent descriptor," in *Biometrics (ICB), 2013 International Conference on*. IEEE, 2013, pp. 1–6.

[7] S. R. Arashloo, J. Kittler, and W. Christmas, "Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features," *Information Forensics and Security, IEEE Transactions on*, vol. 10, no. 11, pp. 2396–2407, 2015.

[8] J. Komulainen, A. Hadid, and M. Pietikainen, "Context based face anti-spoofing," in *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013, pp. 1–8.

[9] Z. Boulkenafet, J. Komulainen, and A. Hadid, "face anti-spoofing based on color texture analysis," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2636–2640.

[10] G. Pan, L. Sun, Z. Wu, and Y. Wang, "Monocular camera-based face liveness detection by combining eyeblink and scene context," *Telecommunication Systems*, vol. 47, no. 3-4, pp. 215–225, 2011.

[11] J. Komulainen, A. Hadid, M. Pietikainen, A. Anjos, and S. Marcel, "Complementary countermeasures for detecting scenic face spoofing attacks," in *Biometrics (ICB), 2013 International Conference on*. IEEE, 2013, pp. 1–7.

[12] L. Sun, W. Huang, and M. Wu, "Tir/vis correlation for liveness detection in face recognition," in *Computer Analysis of Images and Patterns*, 2011, pp. 114–121.

[13] A. Anjos, M. M. Chakka, and S. Marcel, "Motion-based counter-measures to photo attacks in face recognition," *Biometrics, IET*, vol. 3, no. 3, pp. 147–158, 2014.

[14] S. Z. Li, S. R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 627–639, 2007.

[15] R. S. Ghiass, O. Arandjelović, A. Bendada, and X. Maldague, "Infrared face recognition: A comprehensive review of methodologies and databases," *Pattern Recognition*, vol. 47, no. 9, pp. 2807–2824, 2014.

[16] L. Yuan and J. Sun, "Automatic exposure correction of consumer photographs," in *European Conference on Computer Vision*. Springer, 2012, pp. 771–785.

[17] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 2, pp. 218–233, 2003.

[18] C. A. Parker and W. Rees, "Correction of fluorescence spectra and measurement of fluorescence quantum efficiency," *Analyst*, vol. 85, no. 1013, pp. 587–600, 1960.

[19] O. Bimber and D. Iwai, "Superimposing dynamic range," in *ACM Transactions on Graphics (TOG)*, vol. 27, no. 5. ACM, 2008, p. 150.

[20] D. Gragnaniello, G. Poggi, C. Sansone, and L. Verdoliva, "An investigation of local descriptors for biometric spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 849–863, 2015.