Liveliness Detection in Face Recognition Systems Using rPPG

Jin Sung Kang
Sherif N. Abbas
Umang Yadav

4th April, 2017

Contents

1 Introduction 4

2 Literature Review 6

2.1 Hardware Approach . . . 6

2.2 Software Approach . . . 6

3 Proposed System 8

3.1 Face Detection . . . 8

3.2 Remote Photoplethysmogram Extraction . . . 8

3.3 Fourier Transform . . . 10

3.4 Classification . . . 10

4 Results 11

4.1 Database and Training . . . 11

4.2 Simulation . . . 11

4.3 Comparison . . . 12

5 Discussion 13

5.1 Security and Privacy Concerns . . . 14

6 Conclusion and Future Work 15

List of Figures

1.1 Face Spoofing Attack Techniques 5

3.1 System Diagram 9

List of Tables

4.1 IDIAP Database Information 11
4.2 Database distribution 12
4.3 Model Accuracy 12
4.4 Overall Results Comparison for Photo and Video Attacks on IDIAP Database 12

Chapter 1

Introduction

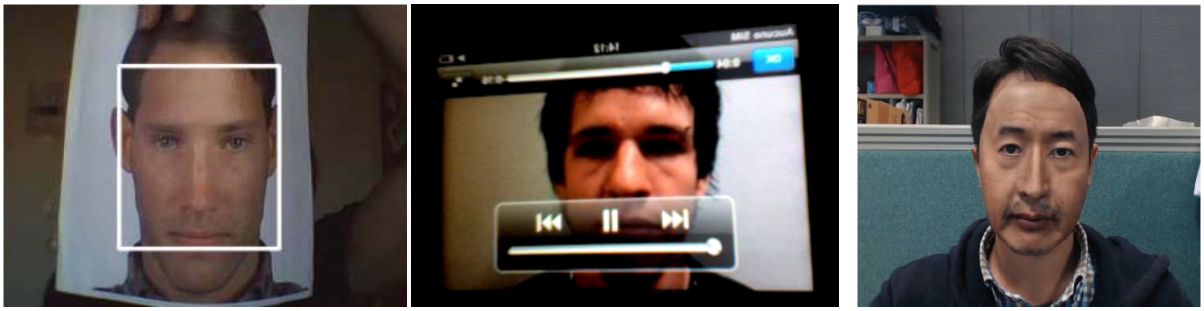
Nowadays, face has become the second largest deployed biometrics trait after fingerprint for human identification. This is because face recognition is socially accepted among almost all people as means of identification and because of the development of smart-phones with cameras and the spread of security cameras everywhere, face recognition became an emerging technology in many online access and security applications like mobile or Internet banking, accessing smart-phones or laptops, and surveillance in vital sites.

However, face is not hidden, it can be easily captured at a distant using a camera and hack a face recognition system by displaying an image of the captured face. Although it seems a very simple idea, it is very hard for a face recognition system to differentiate between a real face and a photo of a face. Nowadays, a high definition image of a face can be easily captured, or recorded using digital cameras installed everywhere (e.g. smart phones and security cameras) or can be easily downloaded from social networks like Facebook or Instagram. Moreover, due to the advances in 3D printers, a cheap 3D mask of face can be obtained [1].

A face recognition system can be easily spoofed by one of the following attacks [1]: (1) photo attacks: by displaying a printed face image or digital face photo (displayed from a mobile),(2) replay attacks: by displaying digital video of a face, (3) 3D attacks: by wearing a 3D mask. The three types of attack scenarios are shown in Figure **??**. A worse case is that a face recognition system can be hacked by a drawing of a face [2].

In this project, a liveliness detection technique is employed to differentiate between real and fake faces for face recognition systems. The liveliness detection is based on magnifying and extracting the small color variations in the face that happen when the heart bumps blood in the whole body. The small color variations of the skin due to blood circulation in the human body is known as photo-plethysmo-gram PPG. Due to advances in signal processing this signal can be remotely detected (rPPG) and also be used to estimate the heart rate with good accuracy [6]. This signal is extracted from three different types of videos that represent three different spoofing or authentication scenarios: videos of real face (real

[1] That's my face: http://www.thatsmyface.com/



(a) Photo Attacks [3] (b) Video Attacks [4] (c) 3D mask attacks [5]

Figure 1.1: Face Spoofing Attack Techniques

scenario), videos of printed or digital photo (photo attack scenario) and video of a video (replay attack scenario). These signals were converted in the Fourier domain and classified using random forest classifier. Based on the achieved results, the proposed algorithm can differentiate between the the real and attack scenarios with accuracy 94.4%.

The remaining part of this report is organized as follows. Section 2 provides a literature review of face anti-spoofing techniques that have been implemented in previous works. Section 3 provides a detailed description of the proposed algorithm for face liveliness detection. Section 4 presents the achieved results for the proposed algorithm. Sections 5 and 6 discuss the results of the proposed algorithm and its application related to security in face recognition systems.

Chapter 2

Literature Review

Existing approaches for detecting different attack scenarios are discussed here briefly. These approaches can be divided into two categories as discussed hereafter.

2.1 Hardware Approach

In this technique an extra hardware component is added for the face recognition system to detect real faces. Previous works used different sensors to detect signals that indicate liveliness of the user like using thermal cameras to detect thermal waves coming out of a real face [7]. Using thermal imaging, Dhamecha et. al. were able to detect real and attack scenarios with an error rate of 13%. In [8], face biometrics were combined with voice for liveliness detection and an error rate of 2% was achieved on video attacks. Other approaches include using depth sensor to detect voluntary eye blinking and mouth movement following a request from the system as in [9] where an error rate of 3.5% was achieved. Based on previous hardware approaches, we can say that adding extra sensor can detect certain types of attack but not all of them. For example, using thermal imaging can detect video and digital photo attacks but may not be able to detect 3D or 2D masks as thermal waves can still be detected in this case. Another example, using depth sensor can detect photo (digital or print) and video attacks but may not be able to detect 3D mask attacks.

2.2 Software Approach

In this approach, an algorithm is developed that can extract extra features from the face that help in differentiating between real and fake faces. For example, face texture extraction techniques were employed in [3, 10, 11] using different algorithms like difference of Gaussians (DoG) filters, Gabor wavelets, and Local Binary Patterns (LBP). Error rates between 0.5% to 15% were achieved using this technique under different scenarios and databases. In [12, 13] face motion and eye blinking detection techniques were adopted using algorithms like Optical Flow lines (OFL) and Conditional Random Fields (CRF). This approach achieved error rates between 0.5% and 10% under different scenarios and databases. Finally, a new approach was adopted in [5] to extract remote photoplethysmography (rPPG) from face using correlation models for local rPPG signal. Using rPPG, an error rate between 9.9% and 16.2% was achieved on different databases.

In the next section, the proposed algorithm for face liveliness detection using rPPG is discussed in details.

Chapter 3

Proposed System

The proposed approach uses remote Photoplethysmogram(rPPG). This method is different from traditional PPG because the data can be collected remotely without having to install devices on the person. Compared to the traditional collection methods, it has more noise. The advantage of the method is the non-intrusive nature of the method and portability. This approach is also advantages because it does not need extra sensors to be installed to the original detection system. The images from the identification camera can be directly used to distinguish between attack and real authentications. The proposed approach involves four steps: face detection, rPPG extraction, Fourier transform, and classification as shown in figure 3.1.

3.1 Face Detection

Detecting the face and extracting its location is key to detecting attack methods. The face extraction should be robust to movements of the face while authenticating. Face authentication scheme may not have a platform for the user to sit, and this increases the chance of user movement. Most of the authentication image consists of background, and this is another why face detection can be effective to lower noise. Doing face detection cuts down processing time in the future steps because the system focuses on the important part of the image rather than the whole.

3.2 Remote Photoplethysmogram Extraction

Remote photoplethysmogram extraction uses the feed from the face extraction as an input. The principal behind the extraction is to detect the changes in the colour of the face. Colour of the face changes minutely every time the heart pumps blood to the face and it is difficult to see with naked eyes. Image processing technique can specifically target the frequency range of human heart beat to improve visibility. Amplifying the change to 100 to 150 times works best in practice. The nine second segment of the authentication scheme is amplified using methods proposed by [6]. The amplified video is processed by averaging the intensity of the face for each channel. The mean of each channel is subtracted from the signal to get a cleaner signal.

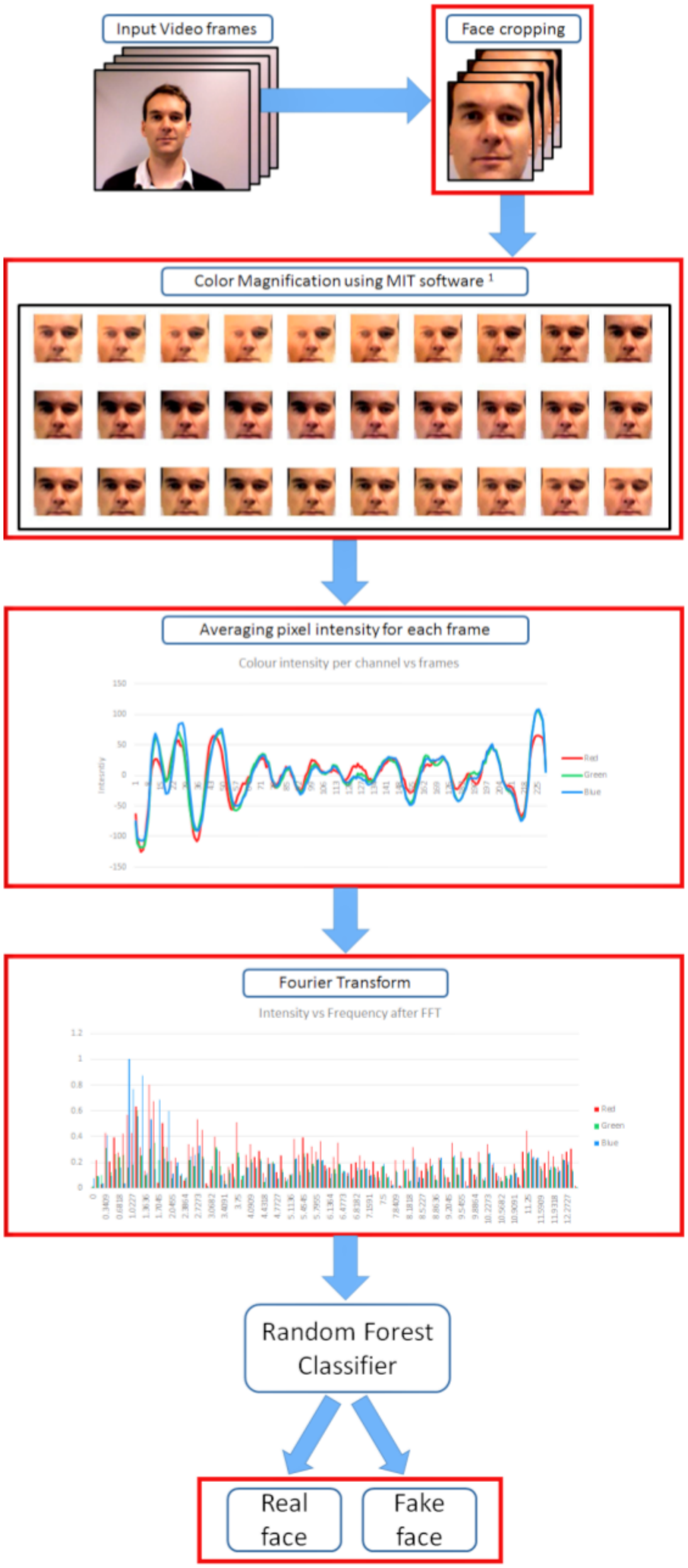


Figure 3.1: System Diagram

3.3 Fourier Transform

The signal that has been processed is then transformed into frequency domain using Fourier transform. PPG and hear beat are repetitive signal and understanding its frequency components is key. Fast Fourier transform offers a method to extract frequency components of the signal. In this method frequency spectrum outside of the heartbeat (0.5Hz to 1Hz) were also used. This is because this region may hold valuable information that may determine if the authentication is an attack or real. Fast Fourier transform is performed on each channel. This system uses a nine second of video at 25 frames per second. The total number of frames that are used for FFT is 220 frames. This produces in 111 bins of frequencies including zero per channel.

3.4 Classification

This step determines if the provided video segment is an attack or real authentication. Fourier transform produces features that are used for classification. There are 333 length feature vector. The classification uses random forest classifier. This classifier is an ensemble learning classification method using decision trees. The random forest classifier was chosen based on experiments. This is a binary classifier between real and attack authentication, and it does not differentiate between various other types of attack.

Chapter 4

Results

4.1 Database and Training

The database is from IDIAP Replay Attack Dataset [4]. This database consists of video recordings of 50 people. There are two major lighting conditions, which are from adverse and controlled lightings. It also has two ways that the attack is presented. One attack is performed by holding the media in front of the camera with the hands, and the other was fixed on a platform. The medium that were used to show the print photo attacks were iPhone 3GS, iPad, and a printed image. The medium to show video attacks were iPhone 3GS and iPad.

The database consisted of 500 videos with the distribution that is shown on Table 4.2. The database was split into training and test datasets. The 80% of the videos were training and 20% were test. For equal representation of the various types of attacks, each of the attack were first divided into training and test dataset, and combined to created the global training and test dataset.

4.2 Simulation

The classification accuracy of the model was measured with 5-fold cross validation. The results are show in Table 4.3. The model performed at 94.4% accuracy on the test set. When the test set was sub-divided into different types of attacks, there was a slight difference in accuracy. Digital photo attacks and video attacks were classified at 100%, but the printed photo attack did poorly at 87.5%. Real authentication method score 97.5%.

Table 4.1: IDIAP Database Information

Resolution	320×240
Length	9 seconds
Frame Rate	25 fps

Table 4.2: Database distribution

Authentication Type	Number of videos
Digital Photo Attack	160
Video Replay Attack	160
Printed Photo Attack	40
Real Authentication	140

Table 4.3: Model Accuracy

Authentication Type	Classification Accuracy (%)
Test Set	94.4
Digital Photo Attack	100
Video Replay Attack	100
Printed Photo Attack	87.5
Real Authentication	97.5

4.3 Comparison

Since, not everyone use the same database to evaluate their system performance and also, most of the people use subset of available datasets to report their system performance. So, it is not fair to compare results straight away. But, to get an idea of our system performance, here in Table 4.4, we have compared our results with some existing methods which were evaluated on photo and video attacks using same database. It is also worth mentioning that we are comparing only final results here. Our final results suffers from the poor classification of Print Photo attacks.

Table 4.4: Overall Results Comparison for Photo and Video Attacks on IDIAP Database

Paper	Technique	Classification Error (%)
Freitas et.al. [14]	Dynamics of facial texture using Local Binary Patterns	7.5%
Bhardwaj et.al. [15]	Motion Detection using Histogram of Oriented Optical Flows after Eulerian magnification	1.25%
Chingovska et.al. [4]	Face Texture using LBPs	15%
Maatta et.al. [11]	Texture + Shape combining + LBPs + Gabor Waveles + HOG	0.5%
Our Proposed Method	Skin color change magnification using Eulerian Technique	5.6%

Chapter 5

Discussion

In this project, we have shown the feasibility of Liveliness detection as Anti spoofing measure in face recognition systems. This approach is simpler than many of the other existing sophisticated techniques presented in section 2. Although, this simplicity comes with its own advantages and disadvantages. Here, in this section, our proposed system is discussed from the technical and performance point of view.

As shown in results, our system outperforms many software and hardware based methods with overall classification accuracy of 94.4% and 100% for digital and video attacks. Classification power of the system mainly comes from the magnification of the subtle color change and intrinsic nature of Digital Device LED Display screen. We are using Eulerian color magnification technique described in [6] which has its own limitation and which in turn limits the range of operation for our system. Eulerian color magnification technique can only be used for very small frequency changes. Although, pulse rate or heart rate satisfies this requirement but, for better classification in adverse environment, filters used in [6] need to be personalized. In spoofing with digital device as medium, better classification of the real vs fake can also be attributed to the fact that, LED display screens has more blue light intensity than red and green to reduce the power consumption. For now, this fact helps us for anti spoofing in video and digital photo attacks and we achieved 100% classification accuracy for both. If, in future some other technology for display comes into market than, our system needs to be upgraded. Faces in video frames, have minute natural motions always. We are cropping the face details with predefined points on face. Even then, motion incorporates as noise in frames. This is not a prominent noise factor but, influences the classification accuracy slightly.

For the print photo attacks, we only received, 87.5% classification accuracy, which isn't acceptable for real life deployment of the system. Poor accuracy of print photo attacks is justified in the sense that, we only had 40 training samples for this case as opposed to 100+ for other cases. Ideally, system should give near 100% accuracy for print photo attacks as well if it is also giving similar results for digital photo and replay attacks. More experimentation is needed for this particular case.

For the purpose of feasibility test and for proof of concept, we did experiments in Fourier

Domain with random forest classifier. Random forest classifier was selected based on the empirical results. We also had done experiments with Adaboost, Gradient Boosted Trees and found random forest tree superior among all.

Here in our system, we haven't much bothered about the false negatives, since it can be tolerated to certain extent. Also there are very few cases, where we expect to get false negatives. For example, if a person is wearing make up such that, it changes the light spectrum distribution then, we might see a false negative.

If we consider hardware cost, system performance, runtime and algorithm complexity then, we can argue that, our system is better than many other existing techniques. Our system doesn't require any separate sensor or hardware. System performance is acceptable with one or two exceptions.

5.1 Security and Privacy Concerns

For our system to work, it requires certain minimum length of video frames. So, it is also important priority that, system stores this information safely. Moreover, since our system uses rPPG for liveliness detection, it additionally contains medical data. Although, this data is not used as anti-spoofing measure, but inference can be made about person's health using this data. Any system using such techniques, should clearly follow HIPA (Health Information Protection Act) in Ontario.

It might seem innocent to share someone's heart rate, but there has been cases where these informations were used to infer health status. There was a case where the husband reported an increased heartbeat over a long period of his wife taken on her fitbit, he shared the information on-line to a forum. People were correctly able guess that the wife was pregnant [16]. In this case the information was made public by the husband and it revealed a happy surprise, but this could have taken a completely different direction. These information should be encrypted so that we can ensure the security and privacy of the user. This is even more important because it is a medical information. It would be difficult to justify the system, and its safety if information collected are not safety stored.

Chapter 6

Conclusion and Future Work

In conclusion, we can say that, our system is simpler than other sophisticated techniques with better classification accuracy. Discrimination power of the system lies in color change magnification and intrinsic characteristic of display screens.

Although, much can be done to improve the system. Our future work consists of trying out different techniques to form feature vector, instead of elementary Fourier transform. More sophisticated methods can be used to extract rPPG with different brands of Phones/Tablets to compensate for different color saturation in every models. Experimentation is needed with different combinations of classifier and feature vectors. Also, to assess and compare system performance with other techniques, we also need to evaluate our method with different publicly available datasets.

In this project, we haven't evaluated our system against 3D-Masks attack. Although, we are certain about good classification accuracy even in that scenario. Future work includes evaluation of system against 3D masks attack to affirm this hypothesis.

One of the drawbacks of this system is, it requires video frames for certain amount of time. Videos used in this project were of length 9 seconds. Our future work includes, testing our system performance against shorter length videos. We are optimistic for good results with videos of length up to 5 seconds.

Potential applications of this method is not just limited to anti spoofing. But it can also be used for non-intrusive, remote physiological measurement such as Heart Rate, Blood pressure and Respiration Rate. Also psychological measurements such as mental stress, emotional valence and individual emotions [17]. Our future work includes investigating these domains of application using our system.

Our system can also be incorporated into multi-modal system for robustness. Recently, many people have shown the possibility of PPG alone as identity measure [18]. There is no doubt, PPG has discriminative power. We also think, we can exploit this discriminative power for spoof-proof identity system in future.

Bibliography

[1] J. Galbally, S. Marcel, and J. Fierrez, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access*, vol. 2, pp. 1530–1552, 2014.

[2] M. Lewis and P. Statham, "Cesg biometric security capabilities programme: method, results and research challenges," in *Biometrics Consortium Conference. Arlington, Virginia*, vol. 200, 2004.

[3] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A face antispoofing database with diverse attacks," in *Biometrics (ICB), 2012 5th IAPR international conference on.* IEEE, 2012, pp. 26–31.

[4] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," 2012.

[5] S. Liu, P. C. Yuen, S. Zhang, and G. Zhao, "3d mask face anti-spoofing with remote photoplethysmography," in *European Conference on Computer Vision.* Springer, 2016, pp. 85–100.

[6] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph. (Proceedings SIGGRAPH 2012)*, vol. 31, no. 4, 2012.

[7] T. I. Dhamecha, A. Nigam, R. Singh, and M. Vatsa, "Disguise detection and face recognition in visible and thermal spectrums," in *Biometrics (ICB), 2013 International Conference on.* IEEE, 2013, pp. 1–8.

[8] G. Chetty and M. Wagner, "Liveness detection using cross-modal correlations in face-voice person authentication." in *INTERSPEECH*, 2005, pp. 2181–2184.

[9] K. Kollreider, H. Fronthaler, and J. Bigun, "Verifying liveness by multiple experts in face biometrics," in *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on.* Ieee, 2008, pp. 1–6.

[10] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "Lbp- top based countermeasure against face spoofing attacks," in *Asian Conference on Computer Vision.* Springer, 2012, pp. 121–132.

[11] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using texture and local shape analysis," *IET biometrics*, vol. 1, no. 1, pp. 3–10, 2012.

[12] K. Kollreider, H. Fronthaler, and J. Bigun, “Non-intrusive liveness detection by face images,” *Image and Vision Computing*, vol. 27, no. 3, pp. 233–244, 2009.

[13] G. Pan, L. Sun, Z. Wu, and S. Lao, “Eyeblink-based anti-spoofing in face recognition from a generic webcamera,” in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.

[14] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, “Lbp-top based countermeasure against face spoofing attacks,” in *International Workshop on Computer Vision With Local Binary Pattern Variants - ACCV*, Nov. 2012, p. 12.

[15] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, and R. Singh, “Computationally efficient face spoofing detection with motion magnification,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, June 2013, pp. 105–110.

[16] A. Jackson, “Husband and wife never expected their fitbit would tell them this.” [Online]. Available: http://www.cnn.com/2016/02/10/health/fitbit-reddit-pregnancy-irpt/

[17] “Transdermal optical imaging (toi),” http://www.nuralogix.com/transdermal-optical-imaging.html, accessed: 2017-04-04.

[18] A. R. Kavsaolu, K. Polat, and M. R. Bozkurt, “A novel feature ranking algorithm for biometric recognition with {PPG} signals,” *Computers in Biology and Medicine*, vol. 49, pp. 1 – 14, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010482514000687