The Eighth International Conference on Business Information Security (BISEC-2016), 15th October 2016, Belgrade, Serbia

MULTIMODAL BIOMETRIC AUTHENTICATION IN IOT: SINGLE CAMERA CASE STUDY

NEMANJA MAČEK
School of Electrical and Computer Engineering of Applied Studies, Belgrade and SECIT Security Consulting, nmacek@viser.edu.rs

IGOR FRANC
Belgrade Metropolitan University, Faculty of Information Technologies and SECIT Security Consulting, igor.franc@metropolitan.ac.rs

MITKO BOGDANOSKI
Military Academy General Mihailo Apostolski, Skoplje, Macedonia, mitko.bogdanoski@ugd.edu.mk

ALEKSANDAR MIRKOVIĆ
eSigurnost Association, Belgrade and SECIT Security Consulting, amirkovic@secitsecurity.com

Abstract: *This paper presents an approach to multimodal biometric authentication using face and iris biometric traits. Having in mind that variety of devices, such as laptops, smartphones and tablets have a high quality camera built in, it is possible to obtain images of iris and face simultaneously. By combining geometric and photometric techniques, such as fiducial point localization and Gabor filtering, features are extracted and biometric templates are generated. Generated templates are further used to identify and authenticate the user on any device which he is allowed to use, thus replacing the standard username – password authentication scheme with a single camera shot.*

Keywords: *Biometrics, Authentication, Iris, Face*

1. INTRODUCTION

According to ITU-T Y.2060 recommendation, Internet of things (IoT) is defined as a global infrastructure for the information society, enabling advanced services by interconnecting physical and virtual things based on existing and evolving interoperable information and communication technologies, while the thing is defined as an object of the physical world (physical things) or the information world (virtual things), which is capable of being identified and integrated into communication networks [1]. According to Gartner, Inc., 6.4 billion connected things will be in use in 2016, while approximately 20 billion devices on the IoT is expected in year 2020 [2].

Most of the criticism and controversies regarding IoT are related to privacy, autonomy, control and security of things. Variety of researchers have explored these areas, and many interesting results are reported in a literature. As an example, a report that states that the privacy of households using smart home devices could be compromised by analysing network traffic [3] is a bit spooky. Perera et al. have identified user consent, freedom of choice and anonymity as major privacy challenges in the IoT domain [4], which led to Privacy by Design principle being enforced in some applications, such as British Government smart metering program.

So where does the biometry fit in the IoT? It fits as a technology that can remove certain privacy and security issues off from some devices. Biometrics is defined as the science of establishing the identity of an individual based on physical, chemical or behavioural attributes of the person [5]. Due to distinctive nature of biometric traits and non-repudiation it offers, biometrics is frequently used to enhance the overall security of the system it is implemented in [6]. Biometric authentication offers the ease and convenience users want and the verification enterprises and manufacturers require for IoT because it is able to verify the true identity of the user. There are various industries in which biometrics can be integrated, ranging from smart homes, to the automotive industry, banking, and healthcare. According to Gartner, Inc., 30% of organizations will use biometric authentication for mobile devices by year 2016, while biometric sensors, such as premise security entry consoles, will total at least 500 million IoT connections in year 2018 [7]. Acuity Market Intelligence forecasts that within three years, biometrics will become a standard feature on smartphones as well as other mobile devices [8].

2. PROPOSED AUTHENTICATION SCHEME

Two most common authentication methods are single-factor passwords or PINs and multi-factor authentication, such as a card combined with a PIN. Both having their drawbacks, such as loosing cards or forgetting passwords made a clean path for biometrics to emerge as a new way of granting physical or logical access securely and conveniently. Biometrics can provide both single-factor and multi-factor authentication, and, having that said, one can differentiate two types of biometric systems: unimodal and multimodal. Unimodal systems employ single biometric sample, such as face or fingerprint. Multimodal

systems employ two or more modalities belonging to a same person, such as face and fingerprint. Many consider unimodal biometrics still not to be secure enough because of the limitations in biometric technology or using low cost sensors. Employing two or more modalities increases recognition accuracy, strengthens the proof [9], and reduces false rejection rates (FRR) and false acceptance rates (FAR). Multimodal biometric systems are based on information fusion that can be performed on several levels, typically at feature, match score or decision level.

Alternative approach to multimodal biometrics is presented in this paper: it is the replacement of typical username – password authentication method with two different biometric samples belonging to the same individual. For example, a user's face can be captured by a camera and by using Principal Component Analysis his identity can be determined. This equals providing a username to a system. Once identified, a user provides a fingerprint to a sensor and a system verifies the generated template with one stored in the database belonging to that user. This equals verifying a password that user have provided to a system. If user is successfully identified and verified, access is granted. With high quality cameras available on many devices, two biometric samples can be captured simultaneously: face and iris. The system presented in this paper is based on aforementioned authentication scheme and employs face and iris samples captured by a high quality camera. The efficiency of proposed approach is experimentally evaluated using CASIA databases, collected by the Chinese Academy of Sciences' Institute of Automation [10].

3. IMPLEMENTATION: FACE AND IRIS

The proposed authentication scheme employs two images obtained by a single high quality camera. The image of the face is used to identify the user and the image of the iris is further used to verify the identity.

Face recognition is convenient, non-intrusive authentication method. There are various feature extraction methods reported in the literature and, roughly, they can be classified either as geometric or photometric approaches. Geometric approaches are based on developing the model based on geometric distances between fiducial points, while the photometric approaches are based on extracted statistical values [11]. Before the face features are extracted, input image is pre-processed. Pre-processing steps include image size normalization, background removal (region of interest selection), translation and rotational normalizations and illumination normalization. Normalization increases system robustness against posture, facial expression and illumination. Pre-processing is crucial as the robustness of a face recognition system greatly depends on it. The photometric normalization techniques used in this research are described in [12].

Gabor wavelets based feature extraction technique is reported to provide good results [13] and according to that is used in this research. Let (x, y) specify the position of a light impulse in the visual field, let θ denote the orientation of the filter and λ, σ, γ, and φ denote the parameters of the wavelet (wavelength, Gaussian radius, aspect ratio and phase, respectively). A family of family of two-dimensional Gabor kernels [14] is used:

$$W(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma'^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \varphi\right) \quad (1)$$

$$x' = x\cos(\theta) + y\sin(\theta) \quad (2)$$

$$y' = -x\sin(\theta) + y\cos(\theta) \quad (3)$$

The same values that have been reported in [15] are used in this research: orientation θ ranging from 0 to $7\pi/8$ with a step of $\pi/8$, wavelength $\lambda=\{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$, phase $\varphi=\{0, \pi/2\}$, Gaussian radius σ equal to wavelength and aspect ratio $\gamma=1$.

A set of Gabor filters is used with 5 spatial frequencies and 8 distinct orientations, resulting in 40 different Gabor filters. Filter responses are obtained by convolving these filters with a simple face image, and these representations display desirable locality and orientation performance. When Gabor filters are applied to each pixel of the image, the filtered vector is high dimensional, which further leads to very large computational and storage costs. This problem can be solved without degrading overall robustness by obtaining Gabor features only at ten extracted fiducial points: three on each eye, two on the lips and two on the nose, as shown on Image 2. Fiducial points are extracted by analysis of the chrominance components in the YCbCr colour space: as an example, eyes present high values of Cb and small values of Cr component [16], while the geometric points of nose and mouth are extracted using Sobel filter [17].

Image 1: Fiducial points that Gabor features are obtained on

Each fiducial point will be represented by a Jet vector of n components, where n denotes the number of filters. Having that said, face is represented by a feature vector containing $10n$ real coefficients.

The face recognition process employs multi-layer perceptrons (MLPs). The system employs as many neural networks as much persons we want to identify (see Image

2). Inputs to the each network are Gabor coefficients and distances between fiducial points: the distance between the centres of the eyes, the distance between two eyes, width and height of nose, width of mouth and the distance between nose and mouth. Each network is trained by different samples of the same person obtained by rotation, translation and variation of the lighting and sufficient number of information about different persons. During the identification phase, only one network is allowed to identify the person – output neuron of only one network is allowed to have the positive value. If two or more networks provide positive outputs, the person is considered as not identified and the system repeats the recognition operation.

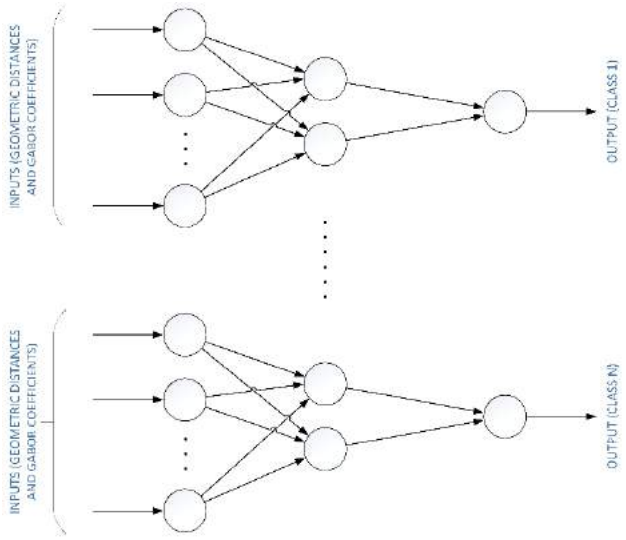

Image 2: Set of neural networks for face recognition

Once the user is identified via neural network face recognition, the system verifies the user by matching the generated iris template with the one stored in the database belonging to that user.

Iris is as first roughly localized from the obtained image in the YCbCr colour space [16] and further pre-processed. Once converted to grayscale, the outer radius of iris patterns and pupils are localized with Hough transform that involves a canny edge detector to generate an edge map. A poorly localized iris will result in unsuccessful segmentation and incorrectly generated biometric template – iris code. Hugh transform identifies positions of circles and ellipses [18]. It locates contours in an *n*-dimensional space by examining whether they lie on curves of a specified shape. Localization process is presented by Image 3.

Once an iris image is localized, regions of interests are defined and it is transformed into fixed-size rectangular image. The normalization process employs Daugman's rubber sheet model that remaps the iris image $I(x, y)$ from Cartesian (x, y) to polar coordinates (r, θ) [19]:

$$I\left(x\left(r,\theta\right), y\left(r,\theta\right)\right) \rightarrow I\left(r,\theta\right) \tag{4}$$

Parameter r is on the interval [0, 1] and θ is the angle [0, 2π]. If iris and pupil boundary points along θ are denoted by (x_i, y_i) and (x_p, y_p), respectively, the transformation is performed according to equations (2) and (3):

$$x\left(r,\theta\right) = \left(1-r\right) x_p\left(\theta\right) + x_i\left(\theta\right) \tag{5}$$

$$y\left(r,\theta\right) = \left(1-r\right) y_p\left(\theta\right) + y_i\left(\theta\right) \tag{6}$$

The rubber sheet model does not compensate rotational inconsistencies. However it produces a normalized representation with constant dimensions (see Image 4) set by angular and radial resolution by taking pupil dilation size inconsistencies into the account [20].

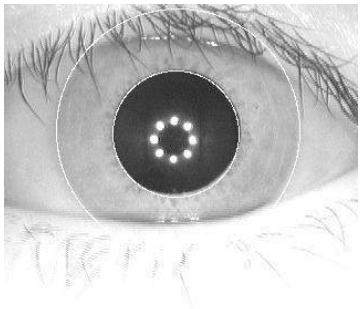

Image 3: Localized iris

Image 4: Normalized iris

There is a variety of iris feature extraction methods reported in the literature, such as Gabor filtering, log-Gabor filtering, zero-crossings of 1-D wavelets and Haar encoding (wavelet method). 1-D log-Gabor filtering is validated as suitable iris feature extraction method in various researches by other authors. A normalized image is broken into a number of 1-D signals that are convolved with 1-D Gabor wavelets. Let f_0 denote centre frequency, and σ the bandwidth of the filter. The frequency response of 1-D log-Gabor filter [21] is given by:

$$G\left(f\right) = \exp\left(-\left(\log \frac{f}{f_0} \right)^2 \Big/ 2\left(\log \frac{\sigma}{f_0} \right)^2 \right) \tag{7}$$

Phase quantization is applied to four levels on filtering outputs (each filter produces two bits of data for each phasor) and the quantized phase data is used to encode an iris pattern into a bit-wise biometric template. The number of bits in the biometric template depends on angular and radial resolution and the number of used filters. Biometric template size used in this research is 9600 bits.

Iris biometric template is generated for each user in the enrolment phase, after training the neural network for face recognition. During the authentication, after successful used identification via face recognition, captured iris image is used to create iris code and measure Hamming distance with the one stored in the database belonging to that user. Let n denote the number of bits in the iris codes x and y of equal length and $n_d(x,y)$ denote number of positions at which the corresponding bits are different. Hamming distance is given by:

$$d(x,y) = \frac{n_d(x,y)}{n} \tag{7}$$

A match is considered to be perfect if d=0, while random strings are expected to provide a distance d=0.5. For iris codes, identical irises are expected to provide Hamming distance d=0.08, while verification is considered to be successful for values d<0.32.

4. EXPERIMENTAL EVALUATION

Performance of the proposed solution is experimentally evaluated using MATLAB R2016a (feature extraction and neural networks) and Python 2.7 (iris matching and scripting). As this research does not deal with image capturing hardware, CASIA-IrisV4 and CASIA-FaceV5 databases are used to evaluate the performance of the proposed authentication scheme. 50 randomly chosen subjects were used from CASIA-FaceV5 and each subject was accompanied with a randomly chosen subject from CASIA-IrisV4 Interval database.

The first step of the experiment is training the neural networks for face recognition. 50 MLP neural networks were trained with the 60% of the database subset used as a training set and remaining 40% as a testing set. Experimental results for different number of filters and accompanying orientations are given in the Table 1.

Table 1: Average face recognition rates

Wavelets	Orientations	Accuracy
5	$\theta = \{0\}$	98.2%
10	$\theta = \{0, \pi/8, 7\pi/8\}$	98.7%
15	$\theta = \{0, \pi/4, \pi/2\}$	99.1%
20	$\theta = \{0, \pi/4, \pi/2, 3\pi/4\}$	99.4%
25	$\theta = \{0, \pi/8, \pi/4, \pi/2, 3\pi/4\}$	99.6%

As Gabor wavelets represent feature points in special frequency at different orientations, it was expected that the MLP recognition rates will increase with the number of Gabor wavelets, which is proven with the results presented in Table 1. With 15 or more Gabor wavelets, recognition accuracy is over 99% and is considered to be satisfactory. One should note that if MLP is trained only with geometric distances as inputs, it provides significantly lower recognition accuracy. As an example, recognition accuracy ranging between 79.7% and 84.2% is reported in [22] if geometric distances only are used with neural networks.

The next step in the experiment is verification of the iris. For each successful face recognition attempt a set of 5 irises belonging to that person and a set of 5 irises belonging to a randomly selected imposter was provided to iris matcher. This allowed us to measure overall accuracy as well as FRR rates (the percentage of valid inputs which are incorrectly rejected). Experimental results are given in Table 2 and graphically presented on Images 5 and 6.

Table 2: Overall system accuracy and FRR

Wavelets	Accuracy	FRR
5	97.5%	2.4%
10	97.9%	~2%
15	98.3%	1.6%
20	98.7%	~1%
25	99.1%	0.8%

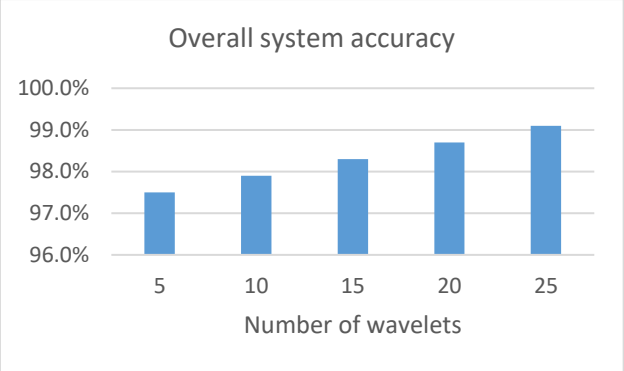

Image 5: Overall system accuracy

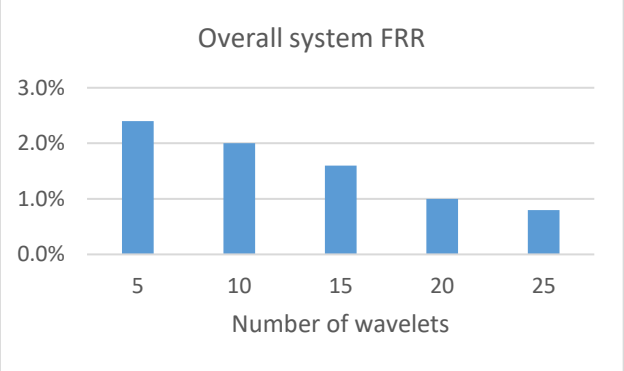

Image 6: Overall system false rejection rates

According to the experimental results we can conclude that the system which operates with face recognition that employs 25 wavelets and iris verification module based on 9600 bit iris code and d<0.32 Hamming distance provides 99.1% accuracy and less than 1% false rejection rate. To put it in the simple words, one in a hundred genuine authentication attempts is rejected. Regarding false acceptance rate, it is virtually set to zero as system employs face recognition system that is very hard for imposter to

trick as it employs both geometric and photometric features. Even if an imposter somehow manages to bypass face recognition (for example, an identical twin may somehow manage to do so), a single shot that captures the image will also capture the iris. And iris verification is very hard for an imposter to trick, as this biological trait is considered as one that distinguishes individuals with highest precision. For example, even a twin could not do so, as irises of identical twins differ as much as irises of unrelated persons due to a chaotic colour pattern in the eye and high degree of randomness iris possesses.

5. CONCLUSION

Various commercial products regarding biometrics and IoT exist on the market. For example, HYPR-2 enables one to secure any device with fingerprint, voice, face and eye recognition. Some solutions extend decentralized biometric authentication down to the firmware level, thus securely transforming smart things into biometric things. However, to the best of our knowledge, no commercial products employ multimodal biometrics in a way as it is described in this paper. The proposed solution is not computationally expensive, and does not require tons of storage space. Only one high quality camera is required on the device, and nowdays it can be found on most of smartphones and mobile devices (such as Samsung Galaxy S7 Edge) and users are further allowed to identify themselves and verify the identity using one shot. The only drawback of the proposed solution is the acceptability of iris biometrics and privacy concerns on stored templates. Further work will be focused on additional security countermeasures, such as implementing cancellable biometrics in this authentication scheme, which will preserve the privacy of stored biometric templates. Additionally, authors are about to explore different face recognition methods and see if any of them can improve overall accuracy and reduce processing and storage costs.

REFERENCES

[1] International Telecommunication Union, "Overview of the Internet of things," Recommendation ITU-T Y.2060, June 15, 2012.

[2] Gartner, Inc., "Gartner Says 6.4 Billion Connected Things Will Be in Use in 2016, Up 30 Percent From 2015", Nov. 10, 2015. Last time visited: August 18, 2016.

[3] J. Singh, T. Pasquier, J. Bacon, H. Ko, D. Eyers, David "Twenty Cloud Security Considerations for Supporting the Internet of Things" IEEE Internet of Things Journal. 3 (3): 1–1, 2015.

[4] C. Perera, R. Ranjan, L. Wang, Lizhe, S. Khan, A. Zomaya, "Privacy of Big Data in the Internet of Things Era", EEE IT Special Issue Internet of Anything, 6., 2015.

[5] A. K. Jain, A. Ross, "Introduction to Biometrics", In "Handbook of Biometrics", A. Jain et al. (Eds), Springer, 2008.

[6] P. Balakumar, R. Venkatesan, "A Survey on Biometrics based Cryptographic Key Generation Schemes", International Journal of Computer Science and Information Technology & Security, Vol. 2, No. 1, pp. 80-85, 2012.

[7] Entertech, "Biometrics to Secure the Internet of Things", Dec. 9, 2015. Last time visited: August 20, 2016.

[8] Acuity Market Intelligence, "Biometric Smartphones Are Officially Mainstream", Published by PR Newswire, Feb 11, 2016. Last time visited: August 20, 2016.

[9] L. Hong, A. K. Jain, S. Pankanti, "Can multibiometrics improve performance?", In Proceedings of IEEE Workshop on Automatic Identification Advanced Technologies, pp. 59-64, NJ, USA, 1999.

[10] Biometrics Ideal Test, http://biometrics.idealtest.org

[11] M. H Yang, D. Kriegman, N. Ahuja, "Detecting faces in images: A survey", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, pp. 34-58, 2002.

[12] R. Rouhi, M. Amiri, B. Irannejad, "A Review on Feature Extraction Techniques in Face Recognition", Signal & Image Processing: An International Journal (SIPIJ) Vol. 3, No. 6, pp 1-14, 2012.

[13] R. Chellappa, C. Wilson, S. Sorihey, "Human and machine recognition of faces: a survey", Proceedings of the IEEE, Vol. 83, pp. 705-740, May 1995.

[14] N. Petkov, P. Kruizinga, "Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells", Biological Cybernetics, pp 83-96, 1997.

[15] L. Wiskott, J. M. Fellous, N. Kruger, C. Malsburg, "Face Recognition by Elastic Bunch Graph Matching", Intelligent Biometric Techniques in Fingerprint and Face Recognition, Ch. 11, pp. 355- 396, 1999.

[16] R. L. Hsu, M. A. Mottaleb, A. K. Jain, "Face detection in color images", IEEE Trans on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 696-706, May 2002.

[17] H. Rowley, S. Baluja, T. Kanade, "Neural Network-based face detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, No. 1, pp. 23-38, Jan. 1998.

[18] D. J. Kerbyson, T. J. Atherton, "Circle Detection using Hough Transform Filters", Fifth International Conference on Image Processing and its Applications, Edinburgh, UK, 04 – 06 July 1995, pp. 370-374.

[19] J. Daugman, "How iris recognition works", Circuits and Systems for Video Technology, IEEE Transactions on, 14(1) pp. 21-30, 2004.

[20] G. Amoli, N. Thapliyal, N. Sethi, "Iris Preprocessing", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, No. 6, pp. 301-304, 2012.

[21] D. J. Field, "Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells", Journal of the Optical Society of America, Vol. 4, No. 12, 1987

[22] Y. B. Jemaa, S. Khanfir, "Automatic local Gabor features extraction for face recognition", International Journal of Computer Science and Information Security, Vol. 3, No. 1, pp. 116-122, 2009.