On Using Palm and Finger Movements as a Gesture-based Biometrics

Marcin Piekarczyk
Pedagogical University of Cracow
Institute of Computer Science and Computer Methods
2 Podchorazych Ave,
30-084 Krakow, Poland
e-mail: marp@up.krakow.pl

Marek R. Ogiela
AGH University of Science and Technology
Cryptography and Cognitive Informatics Research Group
30 Mickiewicza Ave.
30-059 Krakow, Poland
e-mail: mogiela@agh.edu.pl

***Abstract* — In this paper we consider possibility of using palm movements as effective behavioral biometric modality. We propose to exploit the hand motion characteristics gathered from 3D sensor device as input data for identification system. Currently for various reasons people are concerned about directly touching the biometric scanners, therefore touch-less and non-invasive approach seems to be useful in practical applications. In the described scheme palm and fingertip positions are tracked in real-time, while the defined gesture is performed. The proposed gesture is composed of the spatially arranged well-known signs like letters and numbers. The applied matching algorithm utilizes a combination of DTW and DCT techniques for comparing data series. The experiments show promising results using the proposed method.**

Keywords—biometrics; palm movements; finger tracking; gesture-based identification; behavioral biometrics

I. INTRODUCTION

In the contemporary world, where many activities are done very fast by using mobile devices or other types of computer-based machines or interfaces connected to global network, the need for reliable, cheap and user-friendly security systems is growing. In general the authorized access to different types of computer-based devices or common services as banking, ATMs, restricted zones can be performed with usage of password, smart cards or biometrics. Each of these technologies has both its own advantages and weak points. Although passwords means very well known, most common and easy to implement approach there is high risk and widespread tendency to choose too short, poor and finally insecure ones. Moreover, as people interact with many online systems and services they are forced to remember more and more passwords due to tendency to write them down, utilize one password for many purposes or other risky activities increases. On the other hand, smart cards and USB key sticks are very easy to use and are not significantly exposed on attacks based exploiting poor keys but can be easily misplaced or stolen. In case of biometric approach the physical or behavioral characteristics of the examined person are used for authentication purposes [21]. Very often people find biometrics intrusive and concerns about privacy issues especially when classical modalities and hard biometric data connected with fingerprint, iris scan, facial recognition, palm print or online handwritten signatures are collected and preprocessed [7].

To deal with such circumstances many promising solutions were proposed in last years. On the wave of popularity of mobile devices like smartphones and tablets, the concept of using gestures instead of text password were proposed [12]. There were developed different implementations like accelerometer-based recognition [20] or touch-based drawing pattern recognition [1][17][18][19]. In these both cases the human's hand dynamics is taken into consideration. Also, the exploitation of external controllers like commonly used mouse device was successfully tested with regard to dynamics of user's palm [8].

When we think about how modern user interfaces change, it is easily observe that in recent years touch-less spatial sensor devices have become more popular in different aspects of applicability. There are several examples that could be mentioned for example the well-known Kinect-type controllers applied in video game consoles, medical visualization in surgeries, virtual reality navigation or systems recognizing user's gestures/stances [2][3][4]. Considering the fast growing popularity and areas of application this type of devices we propose in this paper the utilization data acquired from palm tracking sensor device as input for biometric-based recognition system. In this concept we are especially focused on tracing a fingertip positions during gesture making because it gives the possibility to handle and exploit hidden traits of biomechanical characteristic of user's hand.

The rest of paper is outlined as follows. In Section II we analyze usefulness of the hand movements from biometric point of view. Section III covers the description of proposed method including steps like basic assumptions, preprocessing, encoding patterns and online verification. The following Section IV presents experimental results and in Section V we conclude the paper with some discussions on future works.

II. PALM MOVEMENTS AS BIOMETRICS

Biomechanical characteristics of user's palm and fingers movements are well recognized as sufficient data in biometrics [1]. The common techniques basing indirectly on this trait are online and offline handwritten signatures recognition as well as handwriting analysis [16]. In all these cases we observe the final effects of the drawing process which is strictly dependent on memorized patterns (e.g. signs, letters, signature) and also on mechanical (anatomical) abilities of one's hand. In this research we propose to use the motion sensing input device to gather the necessary data corresponding to hand-based gestures performed by the examined user. Utilization of the remote motion detector instead of touchable acquiring methods give us some advantages like non-invasiveness, mobility and fleetness. Remote motion sensors usually exploit non-invasive techniques based on optics, infrared or radio frequency (radar) so it provides very secure process of detecting the user's gestures. With reference to mobility it is worth to note that examined user doesn't need to use any additional equipment like, for example, electronic pen. It is similar situation to acquiring data from touchscreens. Probably the most important property, in this context, is fleetness which should be understood as lack of visual track of the performed shapes/drawings/gestures. In result it is significantly harder to make an attack based on imitating or counterfeiting movements/drawings performed by other person.

From technical point of view it is important to note that the motion sensor is able to trace how the position of palm and fingers changes during the time. Basically, the single fingertip position is represented as a vector in 3D Cartesian coordinate system. The data gathered from sensor device can be physically recorded as the separate time series where each series keeps information about variability of signal along chosen Cartesian axis as presented in Fig. 1. More advanced sensing devices are able to provide also additional data characterizing value of velocity or acceleration. This method of data representation enables to apply recognition strategies similar to these commonly used in the field of online handwritten signature verification. It usually includes techniques like frequency analysis, DTW, HMMs [16] and rarer syntactic-based approach [13][14][15].

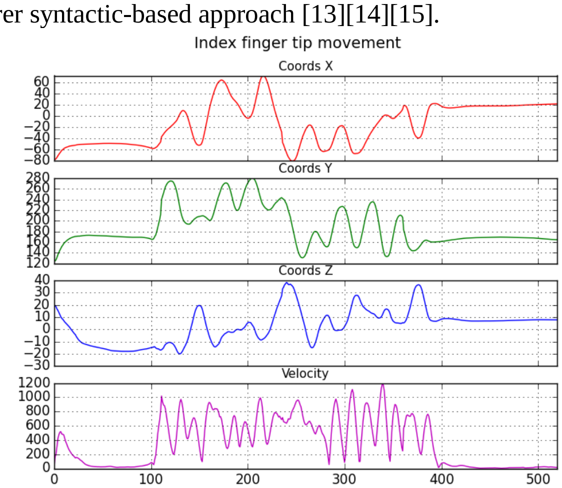

Figure 1. The sample of raw data acquired directly from remote motion sensor (3D coordinates and magnitude of the velocity for single fingertip).

Comprehensive analysis of the capability of this technology in the context of biometrics must include such types of hand-based gestures which have strong discriminative power. In other words the examined gesture should have the appropriate level of complication sufficient for obtaining the distinctive features of user's palm/fingers movements. To meet these requirements the following kinds of the gestures can be considered as input data for recognition system:

- fixed patterns,
- handwritten signatures,
- natural gestures,
- user-defined patterns.

In all mentioned cases the input data for the recognition system can be acquired for one or more fingertips or, if necessary, for other parts of analyzed hand like center of gravity, knuckles, etc. If more fingers are engaged in performing gesture it is very well from security point of view because the risk of imitating strongly decreases. From the other hand too complicated gesture can made up significant problem for the user who must perform the stable repeatable input patterns. Practical application should take into consideration both antagonistic factors to obtain the optimal solution.

III. Recognition of Fixed Pattern

In this paper the recognition problem based on fixed pattern analysis is presented and analyzed in details. Additionally, it is assumed that single fingertip is traced using information about spatial coordinates and velocity. Such formulated recognition task brings the main difficulty in the matter of small amount of input data, related only to single finger, which are available for the recognition algorithm.

A. *Preliminary Assumptions*

We assume the working environment is organized as presented in Fig. 2. The motion detector is placed vertically and is able to trace palm and fingers shifted above its active surface.

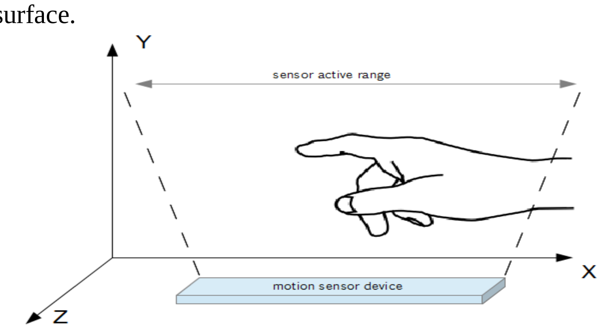

Figure 2. The spatial arrangements of a working space prepared for data acquisition.

As a input gesture previously defined pattern composed of spatially arranged signs is designed. The proposed pattern is presented in Fig. 3.

A X

0 H

Figure 3. Prearranged gesture: fixed pattern composed of spatially organized signs.

From practical point of view the proposed pattern has not too short length and is non-trivial. It is rather easily to remember for human, because consists of well-known signs and is geometrically complicated in sufficient extent. In real-world implementations any other pattern fulfilling similar criteria can be applied.

During the data acquisition phase, where examined user performs the gesture, only fingertip associated with index finger is observed. It is most trained finger of human hand so it should be relatively easily for user to emulate drawing the pattern using it. Moreover, for this reason, it does no matter whether left or right hand finger will be used. Because the proposed pattern is designed as plain, two-dimensional structure it is assumed that the gesture will be performed along XY-plane (see Fig. 2).

Last assumption is strictly related to continual nature of observed signals. The recognition system must be able to evaluate precisely when user begins and ends his activity while the gesture is performed. Due to this we assume that three phases are made during proper signal acquisition:

- relocating the hand inside the sensor range,
- performing the gesture (pattern),
- relocating the hand outside the sensor range.

Between these three phases user should keep his hand motionless throughout approximately 2 seconds. It gives the possibility to isolate the real signal effectively.

B. Preprocessing

According the assumptions described in subsection A of this chapter the input data in fixed pattern approach can be expressed in a formal notation as following time series:

$$X = [x_1, \dots, x_n] \quad (1)$$

$$Y = [y_1, \dots, y_n] \quad (2)$$

$$V = [v_1, \dots, v_n] \quad (3)$$

where X and Y represents fingertip coordinates along XY-plane and V means a magnitude of its velocity. A comprehensive signal gathered by motion detector during single experiment can be expressed as (4).

$$S = \begin{bmatrix} X \\ Y \\ V \end{bmatrix} \quad (4)$$

The exemplary view of a raw signal corresponding to fixed-shape gesture (Fig. 3) is available in Fig. 4 (graphical trajectory) and Fig. 5 (time series). The trajectory is obviously not visible for the user while the gesture is performing. Due to this it usually differs in some extent from original pattern (Fig. 3) since it represents personal realization of the initial gesture enriched by some artifacts associated with transitions between spatial elements.

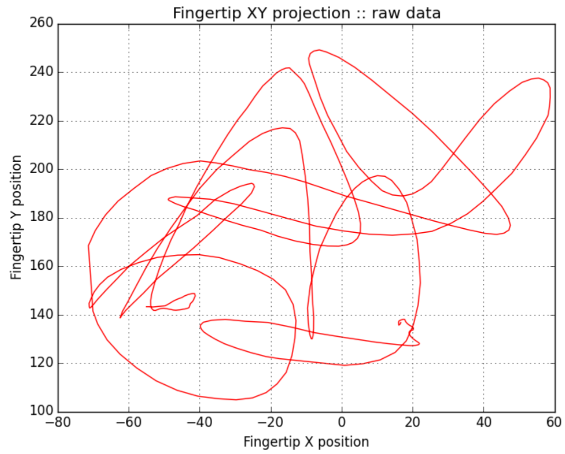

Figure 4. Virtual trajectory associated with the performed gesture.

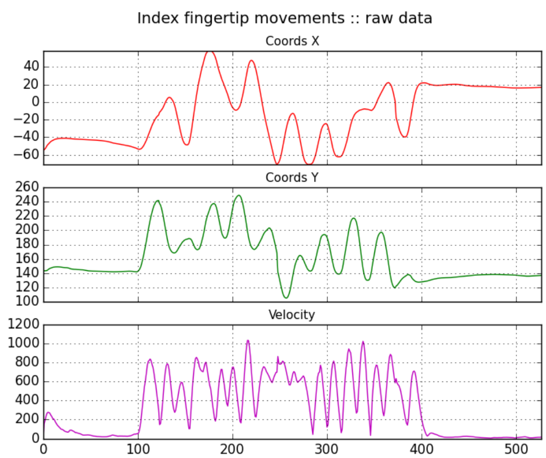

Figure 5. The time series equivalent to trajectory presented at Fig. 4.

The raw time series required some standardization to enable future recognition. Our approach assumes that effective range of signal is estimated on the basis of velocity analysis and then imposed on others series. In the first step it is necessary to cut off these parts of signal when no real hand activity took place. First, we need to remove about 500 ms from the beginning and the end for the sake of the instable phases connected to the activity of relocating the hand inside an outside the sensor range. Then, we have to cut off the parts of signals when no real hand activity took place. These periods can be estimated by searching the places where the velocity signal suddenly increasing over predetermined threshold. Additionally, to refine the signal more precisely the signal is limited to first and last maxima. The standardized range obtained by these operations is finally applied to all input signals.

The last operation required in the preprocessing step is normalization. It is calculated for all series according to (5) to adjust its amplitudes into the common range. It our approach the unity-based normalization to restrict values into the range [0,1] is used.

$$X^{norm} = (X - X_{min})/(X_{max} - X_{min}) \quad (5)$$

In the result we obtain clear and undisturbed data (Fig. 6) which are essential for the further analysis.

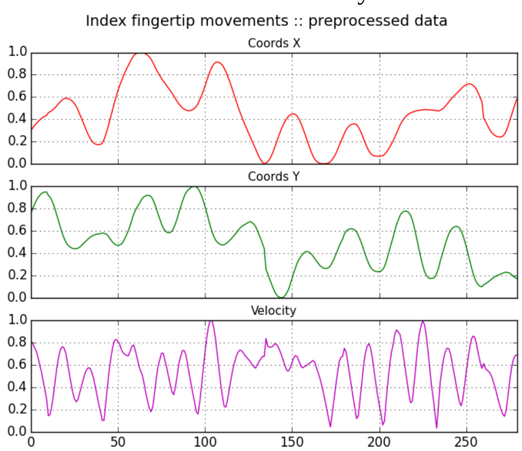

Figure 6. The result of preprocessing for the data series presented at Fig. 5.

C. *Feature-based Pattern Representation*

The three-pass signal (Fig. 6) obtained after preprocessing contains all crucial information required for the efficient recognition. From this point of view the vector $S = [X, Y, V]^T$ (4) can be simply interpreted as a comprehensive pattern representation. We propose to make a transformation such pattern into the frequency-based space. The global features obtained in this way should provide the better projection of the pattern properties. It is especially important when we realize that the same user-originated series are often characterized with high variability and unevenness. In our approach the series are transformed using discrete cosine transform (DCT) in accordance with DCT-II variant defined in (6) [10].

$$y[k] = 2 \cdot \sum_{n=0}^{N-1} x[n] \cos\left[\frac{\pi}{N}\left(\frac{2n+1}{2}\right)k\right] \quad (6)$$

$$k = 0, \ldots, N-1$$

Additionally, each calculated y[k] factor is multiplied by a scaling factor f determined as (7) or (8) to achieve the orthogonal version of DCT-II.

$$f = \frac{1}{\sqrt{4N}} \; if \; k = 0 \quad (7)$$

$$f = \frac{1}{\sqrt{2N}} \; if \; k \in [1, N-1] \quad (8)$$

As the output we receive a finite number of coefficients in descending order with reference to value. In result the output $DCT(S)$ matrix (9) is obtained. Moreover, DCT as opposed to Fourier Transform has a unique property that most energy (high values) is accumulated in only several first coefficients (Fig. 7). Such collection of high energy factors can be treated as global features describing the examined signal.

$$DCT(S) = \begin{bmatrix} DCT(X) \\ DCT(Y) \\ DCT(V) \end{bmatrix} \quad (9)$$

The important challenge is how many coefficients should be taken into consideration? In our approach at least 10 values have been considered during the experiments. Of course, the greater number of factors will be resulted in more accurate representation of the original signal. As Fig. 7 illustrates the 50 values should be fully sufficient to obtain a good precision.

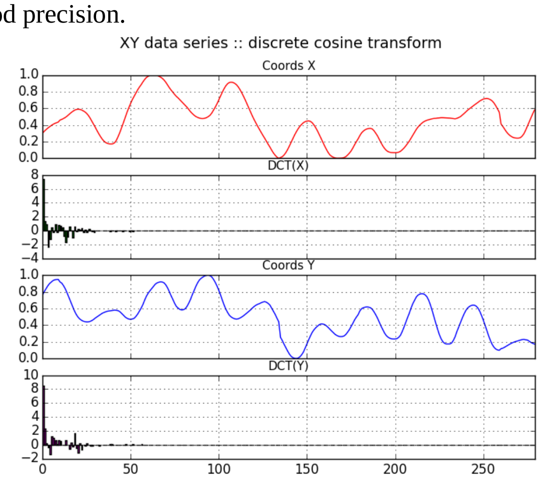

Figure 7. The data series XY transformed by DCT.

From the other hand, in real-world applications we should be interested in to store in memory and to make the calculations with as limited features as possible. Additionally, the representation of the signal not as full-length original series but in simplified and shortened form can be treated as an strong advantage in the context of security. In contemporary strategies there is a tendency to store the sensitive data using indirect and uncomprehensive method of representation to protect from open access to personal-oriented information.

D. *Verification Scheme*

The proposed recognition system (Fig. 8) operates in two phases indicated as offline and online. The offline step is responsible for gathering (motion sensor) and preparing (calculations) the collection of patterns (learning set). Each pattern should be stored in preprocessed form as DCT-calculated matrix (9). During the online phase the examined gesture performed in real-time is transformed also into the DCT-matrix and considered as candidate signal CS (10).

$$DCT(CS) = \begin{bmatrix} DCT(CX) \\ DCT(CY) \\ DCT(CV) \end{bmatrix} \quad (10)$$

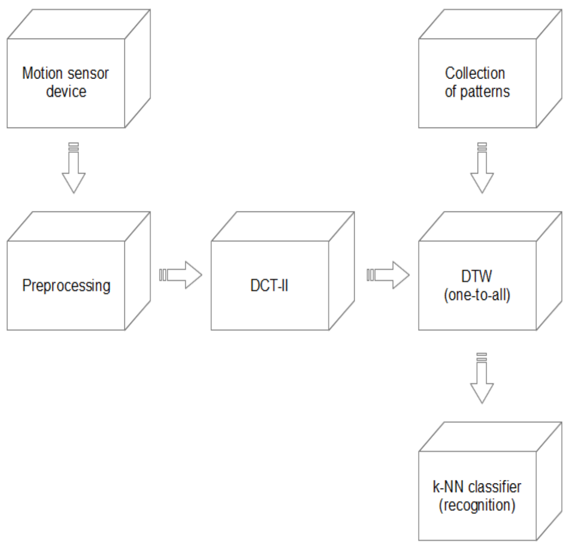

Figure 8. Proposed recognition system based on DCT | DTW | k-NN chain.

Subsequently, it is calculated the DTW-based distance (12) between candidate signal and matrices related to in-memory patterns.

$$DTW(CS, S_i) = \begin{bmatrix} DTW(DCT(CX), DCT(S_i)) \\ DTW(DCT(CX), DCT(S_i)) \\ DTW(DCT(CX), DCT(S_i)) \end{bmatrix} \quad (11)$$

In calculations standard DTW [11] using the Manhattan distance as a local cost measure has been used. As signals are treated data series consisted of DCT-calculated coefficients. In result the set of points in 3D space is achieved where k-NN classifier based on Minkowski distance (12) can be applied.

$$d = (\sum_i |x_i - y_i|^p)^{\frac{1}{p}} \quad (12)$$

IV. EXPERIMETAL RESULTS

To test the initial accuracy the small database was used. This database contains four people aged from 30 to 40 years. Every person performed the proposed gesture (Fig. 3) for the ten times. Three randomly chosen samples belonged to each person have been selected as representative learning set. The remaining ones were exploited as input data series. In case of k-NN classification different values of parameter k (3 and 5) were considered. As a motion sensor during the data acquisition phase the Leap Motion controller [9] has been exploited. The results received during the tests are presented in Tab. I-II. With reference to DCT parametrization the number of coefficients in range from 10 to 100 has been considered.

TABLE I. SYSTEM PERFORMANCE FOR 3-NN CLASSIFIER

Error type	**3-NN classification**				
	DCT-10	***DCT-20***	***DCT-30***	***DCT-50***	***DCT-100***
FRR	7.1	3.6	0.0	3.6	3.6
FAR	2.4	1.2	0.0	1.2	1.2

TABLE II. SYSTEM PERFORMANCE FOR 5-NN CLASSIFIER

Error type	**5-NN classification**				
	DCT-10	***DCT-20***	***DCT-30***	***DCT-50***	***DCT-100***
FRR	0.0	3.6	3.6	0.0	0.0
FAR	0.0	1.2	1.2	0.0	0.0

For better view the achieved accuracy can be compared with recognition where DCT encoding is omitted (Tab. III).

TABLE III. SYSTEM PERFORMANCE FOR DIRECT DTW MATCHING

Error type	**DTW**	
	3-NN	***5-NN***
FRR	28.6	32.1
FAR	9.5	10.7

The most promising results were achieved with system configuration where DCT subsequence with 50 coefficients and 5-NN classifier were applied.

V. CONCLUSIONS

In the paper the concept of utilization of hand-based motion characteristics for the verification purposes has been discussed. We have also proposed the automatic recognition scheme when finger-based motion and fixed-shape gesture can be exploited. Preliminary experiments reveal the feasibility of such system type to effective work based on palm and finger motion dynamics. Therefore work is currently being done in enhancing pre-alignment and deep testing against the considerably larger dataset.

In the future we plan to study the applicability of similar models in the context of more complicated user-defined and multi-fingers gestures. Besides, we consider as crucial to investigate a usefulness motion-based dynamic palm characteristics in field of cryptography especially in field of key generation, key exchange protocols and secret sharing.

ACKNOWLEDGMENT

We kindly acknowledge the support of this study by a Pedagogical University of Cracow Statutory Research Grant.

REFERENCES

[1] M. Frank, et al., "*Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication*," Information Forensics and Security, IEEE Transactions on,Vol. 8, Iss. 1, 2013, pp. 136-148.

[2] T. Hachaj, M. R. Ogiela, M. Piekarczyk, "*Dependence of Kinect sensors number and position on gestures recognition with Gesture Description Language semantic classifier*," Proc. of the Federated Conference on Computer Science and Information Systems, Series: Annals of Computer Science and Information Systems, Vol. 1, 2013, pp. 571-575.

[3] T. Hachaj, M. R. Ogiela, M. Piekarczyk, "*Real-time recognition of selected karate techniques using GDL approach*," Image Processing and Communications: challenges 5 (ed. R. Choras), Series: Advances in Intelligent Systems and Computing, Vol. 233, Heidelberg, Springer, 2013, pp. 99-106.

[4] T. Hachaj, M. R. Ogiela, "*Full-body gestures and movements recognition: user descriptive and unsupervised learning approaches in GDL classifier*," Applications of Digital Image Processing XXXVII, edited by Andrew G. Tescher, Proc. of SPIE Vol. 9217, 921704, doi: 10.1117/12.2061171, 2014.

[5] S. Hashia, C. Pollett, and M. Stamp, "*On using mouse movements as a biometric*," Proceeding in the International Conference on Computer Science and its Applications, Volume 1, 2005.

[6] D. Impedovo, G. Pirlo, R. Plamondon, "*Handwritten Signature Verification: New Advancements and Open Issues*," *International Conference on Frontiers in Handwriting Recognition (ICFHR), 2012*, IEEE, pp. 367-372.

[7] A. Jain, R. Bolle, S. Pankanti, eds., "*Biometrics: personal identification in networked society*", Vol. 479, Springer Science & Business Media, 2006.

[8] Z. Jorgensen, T. Yu, "*On mouse dynamics as a behavioral biometricsfor authentication*", Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, ACM New York, USA, 2011, pp. 476-482.

[9] Leap Motion, https://www.leapmotion.com/

[10] J. Makhoul, "*A Fast Cosine Transform in One and Two Dimensions*," IEEE Transactions on acoustics, speech and signal processing, Vol. 28(1), 1980, pp. 27-34.

[11] M. Muller, "*Information Retrieval for Music and Motion*," Springer-Verlag, 2007, doi: 10.1007/978-3-540-74048-3.

[12] G. Niezen, G. P. Hancke, "*Gesture recognition as ubiquitous input for mobile phones*," International Workshop on Devices that Alter Perception (DAP 2008), in conjunction with Ubicomp, 2008, pp. 17-21.

[13] M. R. Ogiela, M. Piekarczyk, "*Random graph languages for distorted and ambiguous patterns: single layer model*," Proceedings of the Sixth International Conference on Innovative Mobile and Internet Services in ubiquitous computing (IMIS2012), pp. 108-113, 4-6 July 2012, Palermo, Italy, doi: 10.1109/IMIS.2012.147.

[14] M. Piekarczyk, M. R. Ogiela, "*Hierarchical Graph-Grammar Model for Secure and Efficient Handwritten Signatures Classification*," Journal of Universal Computer Science, Vol. 17, Iss. 6, 2011, pp. 926 – 943, doi: 10.3217/jucs-017-06-0926.

[15] M. Piekarczyk, M. R. Ogiela, "*Matrix-based hierarchical graph matching in off-line handwritten signatures recognition*," Proceedings of 2nd IAPR Asian Conference on Pattern Recognition, IEEE, 2013, pp. 897-901, DOI 10.1109/ACPR.2013.164.

[16] R. Plamondon, S. N. Srihari, "*Online and off-line handwriting recognition: a comprehensive survey*," Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 22(1), 2000, pp. 63-84.

[17] N. Sae-bae, et al., "*Multitouch gesture-based authentication*," Information Forensics and Security, IEEE Transactions on,Vol. 9, Issue 4, 2014, pp. 568-582.

[18] M. Shahzad, A. X. Liu, A. Samuel, "*Secure unlocking of mobile touch screen devices by simple gestures: you can see it but you can do it*", Proceedings of the 19th annual international conference on Mobile computing & networking, ACM New York, USA, pp. 39-50.

[19] N. Zeng, et al., "*You are how you touch: user verification on smartphones via tapping behaviors*", Proceedings of IEEE 22nd International Conference on Network Protocols, IEEE, 2014, pp. 221-232.

[20] J. Wu, et al., "*Gesture recognition with a 3-d accelerometer*", Proc. of Ubiquitous intelligence and computing, Springer Berlin Heidelberg, 2009, pp. 25-38

[21] R. V. Yampolskiy, V. Govindaraju, "*Behavioural biometrics: a survey and classification*", International Journal of Biometrics Vol. 1(1), 2008, pp. 81-113.