UDK 004.056
Preliminary communication

TOWARDS AN OPEN BIOMETRIC ONTOLOGY[1]

Miroslav Bača, Markus Schatten, Kornelije Rabuzin
Faculty of Organization and Informatics, Varaždin, Croatia
{miroslav.baca, markus.schatten, kornelije.rabuzin}@foi.hr

Abstract: *Over the last decade we faced a great number of publications in the field of biometrics. Many new biometric methods, techniques, models, metrics and characteristics were proposed. Due to this explosion of research, scientific and professional papers certain inconsistencies in terminology. What some authors call a biometric method, others call model, system or even characteristic. There wasn't enough effort in creating a unique systematization and categorization which would approach the stated issues and open new areas of research. We argue that it is possible to approach biometrics in a narrower and in a broader perspective. We observed biometrics in the narrower perspecive and created a unique framework for the systematization and categorization of biometric methods, models, characteristics and patterns based on a general biometric system. This systematization is a fundamental step forward towards the creation of an open biometrics ontology.*

Keywords: *systematization, categorization, biometrics, model, method, characteristic, open ontology.*

1. INTRODUCTION

Biometrics is not a new science, even if considered in the context of modern information technologies. In 1882 the Berillton system used photographs of people in addition to the measurement of height, arms, legs and finger length. In 1900 Scotland Yard accepted the Galton/Henry fingerprint classification system. In 1924 the Federal Bureau of Investigation organized a full division for fingerprint recognition, while in 1965 AFIS (Automated Fingerprint Recognition System) was introduced with more than 810 000 fingerprints in its database. Goldstein et al. published the first paper on face recognition in 1971. In 2000 the FBI introduced IAFIS (Integrated Automated Fingerprint Identification System) which currently contains more than 55 million fingerprints in its database, and an average of more than 2.2 million searches per month.

A first thing to be stated is that it is possible to approach biometrics in a broader and in a narrower perspective. In the broader perspective biometrics can be defined as the statistical research of biological phenomena; it is the use of mathematics and statistics in the understanding of living beings [7]. In the narrower perspective we can define biometrics

[1] Shown results came out from scientific project (Methodology of biometrics characteristics evaluation 016-0161199-1721), supported by Ministry of science, education and sport Republic of Croatia.

as the investigation of automated human recognition, based on physiological or/and psychological characteristics [2].

We approach biometrics in the narrower perspective in this paper thus we have to define terms used in this perspective. A biometric characteristic is a physiological or behavioral (psychological) characteristic of a person which is used for automated biometric recognition. Physiological characteristics are those people are born with. They are relatively consistent in time (for example the face structure, the iris of the human eye, the vascular pattern etc.). Behavioral or psychological characteristics are characteristics which are acquired or learned during time (for example handwritten signature, gait, keyboard typing dynamics or voice). We consider a biometric system to be unimodal if it uses only one biometric characteristic for the recognition process. If a system uses more than one biometric characteristic it is considered to be multimodal.

The term method comes from the ancient Greek word methodos which means a previously established or given path/way. A scientific method is a set of procedures a scientist uses during research in order to investigate and present its results [20, p. 29]. In such a context we can define a biometric method as a set of procedures which are used to process samples of a biometric characteristic in order to recognize a specified structure in the sample. The recognition of the owner of the sampled characteristic is a special case in such a definition, since biometric methods are not exclusively used for authentication and/or identification, but also for classification.

A biometric sample is any sampled data acquired from a person's biometric characteristic. Such samples include image, sound, video, time measurements and other data which can be used for biometric processing.

A model is a sample of a system (note that the term sample here is used in a different context than before). The acquisition of information about the original system is a model's primary usage [13, p. 241]. Thus a biometric model can be defined as a sample of a biometric human recognition system which provides information about a person's biometric characteristics. In this paper we will consider a biometric model to be any interrelated combination of biometric methods which allows for decision making based on the biometric characteristics of a person.

Another term to consider is biometric pattern. A biometric pattern is a pattern which can be used in order to recognize a given biometric characteristic. Examples include the vein structure for retina, or the minutiae structure of a finger. A distinct term here is extracted biometric feature which is also a pattern but one acquired through biometric processing of a biometric sample. Examples of extracted biometric features include retinal signature templates for retina.. While patterns are actually a part of the characteristic, extracted features are acquired through mathematical or statistical analysis.

From this considerations we can that the fundamental concepts on which we will build our further systematization include (1) biometric methods, (2) biometric models, (3) biometric characteristics, (4) biometric samples and (5) extracted biometric features. In addition to this fundamental concepts we need to introduce the general biometric system developed by Wayman [9].

2. GENERAL BIOMETRIC SYSTEM

According to Wayman the general biometric system has five distinct subsystems (figure 1). These subsystems include (1) data collection, (2) transmission, (3) signal processing, (4) storage and (5) decision. This system is general because these components can be found in almost every biometric system, and can be applied either to unimodal or multimodal systems.

Every subsystem consists of elements which provide additional value to the quality of the system. The data collection subsystem consists of the biometric sample, the way in which the sample is presented and the sensor which samples the presented characteristic. The transmission subsystem consists of data compression, transmission and data expansion. The signal processing subsystem consists of feature extraction, quality control, and comparison (recognition). The decision subsystem consists of a decision making mechanism and the storage subsystem of a database and a sample (image) base.

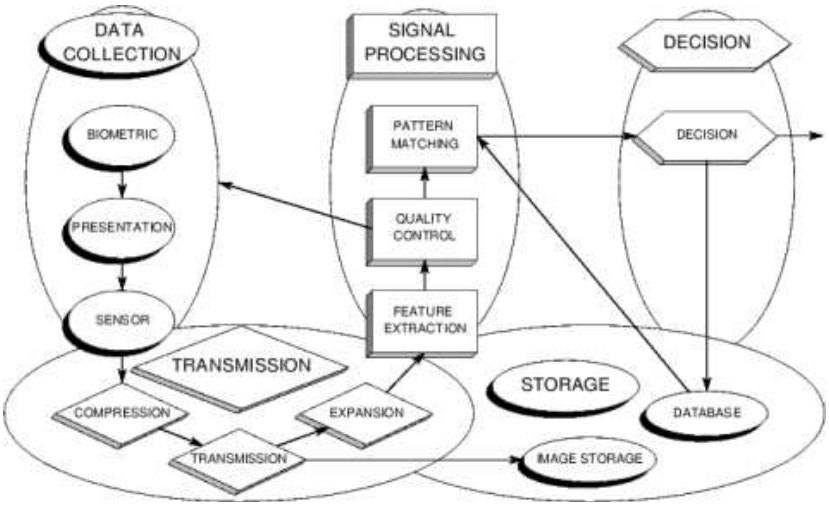

Figure 1. Waymans original general biometric system [21]

In order to use this system for our systematization we have to modify it using a general system theory approach. We need to define exact inputs and outputs for the system, and for each element in particular. We also want to define the system as a system of processes in order to extract biometric methods. The basic system diagram of the modified system is shown on figure 2.

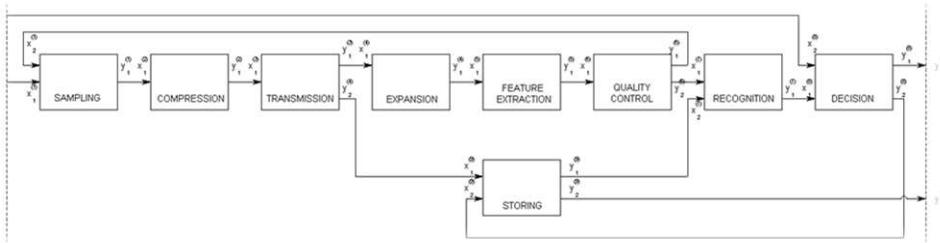

Figure 2. Basic system diagram of the modified system

The basic system diagram consists of nine elements. We left out some elements for a good reason. Biometric sample, presentation and sensor are actually part of a single process, the sampling process. If we would have taken them as independent elements the definition of input and output would have been disjointed. Also we joined the database and sample base elements into a singe storing process. The original diagram contradicts the general system theory rules since the element sample base does not have any outputs. Therefore, it wouldn't be considered a part of the system, or would be considered a parasite

element which only takes input and doesn't give output. We also added a connection from the storing process to the environment of the system. It is a common situation (especially when analyzing rejections) that one wants to take a look at samples which caused the rejection (to find possible intruders or system failures for example). An incoming connection was also added to the decision making process, since decisions are often made not only from the information which is produced inside the biometric system, but also using other information not necessarily kept in the biometrics system database.

In the reminder of this paper we will present a complete analysis of the system in order to get a better understanding of methods used, and in order to create a precise systematization on behalf of the presented system.

2.1. SAMPLING

An input to the sampling process (x_1^1) is the biometric characteristic placed so that it enables acquisition by the sensor. Another input (x_2^1) is the feedback information from quality control which indicates if a sample has to be acquired again due to insufficient quality. The output (y_1^1) is the digitized sample of the presented characteristic.

In this process we can observe different methods used in order to acquire a digitized sample. Which method will be used depends on the given biometric characteristic (if the characteristic is the human face, 2D or 3D photography might be used, if the characteristic is the human gait a video camcorder might be the solution, if the characteristic is the human voice a microphone and a digitizer would probably be the right choice, if the characteristic is the human body odor, sensors which will analyze the odor's chemical structure might be used, if we are trying to sample heartbeats an electrocardiograph would be the right solution etc.). All these methods are in the technical domain of biometrics thus we won't consider them to be biometric methods.

2.2. COMPRESSION

The compression process lowers the amount of sampled data in order to ease transmission and storage. The raw digitized sample is the input (x_1^2) to this process, while the output (y_1^2) is the compressed digital sample.

Two types of methods are used for compression of digital data: (1) lossless compression and (2) lossy compression. Both types of methods and their specific instances are in the domain of information theory and thus won't be considered as biometric methods.

2.3. TRANSMISSION

The transmission process takes the compressed sample as its input (x_1^3), transmits it through an eventually lossy communication channel and outputs (y_1^3), the transmitted sample. Note that the transmitted sample is not necessarily the same as the input sample due to possible data loss in transmission.

Methods used in this process are also in the domain of information theory and/or data and network communication and therefore will not be considered to be biometric methods in the following.

2.4. EXPANSION

The expansion is the process of reconstructing compressed data. The input (x_1^4) to this process is the digitized, compressed and transmitted biometric sample. The output (y_1^4) is a reconstructed sample, which is suitable for further processing.

Methods used in this process depend on the previously used data compression methods. Signal restoration methods are included in this process. These types of methods are in the domain of information theory and especially in the domain of digital signal processing , and thus won't be considered to be biometric methods.

2.5. FEATURE EXTRACTION

The feature extraction process takes the expanded and eventually restored sample as its input (x_1^5) and extracts predefined features (y_1^5) using various transformations. These features will be used for biometric recognition. Methods used in this process are in the domain of biometrics and will be given additional attention in a later section.

We need to mention here that the selection of methods to be used for feature extraction depends on the biometric characteristic. We can classify methods of feature extraction into special and general methods since some methods can only be used with a special type of biometric characteristic while others can be used with different ones.

2.6. QUALITY CONTROL

Based on the extracted features (x_1^6), the quality control process must decide whether a biometric sample is of sufficient quality (y_1^6). If not, it has to be sampled again (y_2^6).

2.7. RECOGNITION

The recognition process takes the extracted features (x_1^7)and referent features (x_2^7) from the database and tries to compare them in order to recognize a given characteristic (x_1^7). Depending on the type and count of referent features used from the database we can apply identification, authentication or classification methods.

Methods used here also include biometric ones especially pattern recognition methods. Which methods will be used depends heavily on the feature extraction methods used before.

2.8. DECISION MAKING

The decision making process takes the data generated in the recognition process (x_1^8) and data from the environment (x_2^8) of the system (e. g. user information, an organizations database, expert systemsetc.) and decides about future actions. These action can include system access approval or denial, success of classification, and other action. From a basic perspective there are two possible results: (1) the system matched the sample or (2) the system did not match the sample.

We should mention here that with our modification the purpose of the system (e. g. authentication, identification, classification, enrollment etc.) is transparent and thus the model is on a higher level of abstraction [18].

Methods used in this process are in the domain of decision theory and knowledge based systems and thus won't be considered to be biometric methods.

2.9. STORAGE

The storage process takes digitized samples (x_1^9), extracted features (x_2^9) and other data (x_3^9) from the environment (e. g. user information) and stores them in the database for retrieval. Thus the outputs of the process equal its inputs depending on the given situation. Database design and implementation are in the domain of database theory.

3. SYSTEMATIZATION

We will build our systematization of biometric methods based on the types of biometric characteristics. There are different approaches in classifying biometric characteristics: (1) hard and soft, (2) traditional and non-traditional, or (3) physiological and behavioral or psychological. We will use the latter in our further presentation.

We use multiple criteria for our systematization: (1) the type of the characteristic to which a biometric method is applicable, (2) the general biometric system or more precisely its element feature extraction, quality control and recognition (the type of method) and (3) the type of biometric pattern which is used by the method. In this way we are able to construct the following table (figure 3). Completing Table 1 with relevant references would require additional 200-300 references, which is not feasible.

Table 1. Systematization of biometric methods, characteristics and patterns

		Pattern	Structure	Extraction	Quality control	Recognition
Biometric characteristics	**Physical**	Body odor [6]	Chemical structure			Template matching
		DNA [12]	Micro satellites	Templates or feature extraction, Restriction fragment length polymorphism (RFLP), Polymerase chain reaction (PCR), Amplified fragment length polymorphism (AmpFLP), Short tandem repeats (STR), Hidden Markov model (HMM)	Hidden Markov model	Comparison of actual samples, Pattern analysis, Hidden Markov model
		Ear	Cartilage structure, Geometrical characteristics	Principal component analysis (PCA)		Template matching
		ECG	Fiducial lines			Stable feature matching
		Face [4]	Geometrical structures, Vein structure	Eigenvector, Local Feature Analysis (LFA), Principle Component Analysis (PCA),	Eigenvector, Automatic face processing	Eigenvector, Automatic face processing, Neural

		Automatic face processing, Elastic nets, Shape from shading (3-D template), Holographic Quantum Neural Networks, Feature mapping, Face Monitoring, Wavelet/elastic matching		networks, Elastic nets, Template matching
Finger [6]	Minutiae, Core, Ridges Cores, Ridges, Geometrical description (2D and 3D)	Receiver Operating Characteristic (ROC) curves		Global pattern matching, Minutia matching, Template matching
Hair	Color			Template matching
Hand	Geometrics characteristics			Template matching
Head [3]	Three dimensional geometric characteristic	Eigenvector	Eigenvector	Eigenvector
Hyper spectral images	Types of tissues	Spectral comparison of combinations of tissue types		Spectral comparison of combinations of tissue types
Iris [9]	Pigment, Epical cell structure, Pupils, Iris parts	Wavelet analysis, 2D Gabor wavelets		Template matching
Lips	Geometrics characteristics	Mathematical morphology analysis		
Nail	Keratin micro waves, Skin wrinkles, Fingerprint lines			Template matching
Otto acoustic emissions	Types of emission			Template matching
Palm [15]	Minutiae, Core, Ridges	Palm imaging		Template matching
Retina [14]	Vein structure	Retinal Signature Template Extraction		Template matching
Skin spectroscopy	Skin layer thickness, Interfaces between skin layers, Waves swing, Cell dimension in layers, Cell density in layers, Layers chemical structure, Skin absorption spectrum			Characteristic optical pattern matching
Sole	Minutiae, Core, Ridges Cores, Ridges, Geometrics characteristics			Template matching
Sweat pore	Types of pores, Relatively distances between pores	Sweat pores analysis		

	Teeth [8]	Geometrics characteristics, Dental insertions characteristics	Feature indexing, Teeth segmentation, Dental film classification, Anisotropic diffusion, Active contour models, Adaptive threshold, Pixel classification, Root Shape Extraction, Radiograph segmentation, Gumline detection, Crown Shape Extraction		Pattern matching
	Thermogram	Bones density, Fat density, Vein density			Template matching
	Toe	Minutiae, Core, Ridges			Template matching
	Vascular structure	Vein structure	Random line tracking		Template matching
Behavioral	Gait [1]	Templates of running or waking	Hidden Markov Model, Static, Activity-Specific Parameters, Motion-history method, Frequency domain, Data distribution statistics, Image Self-Similarity – EigenGait	Hidden Markov Model, Motion-history method	Hidden Markov Model, Walker identification method, Correlation, Motion-history method
	Hand grip [17]	Hand grip strength, Hand grip structure, Skin characteristic, Subcutaneous structure			Pattern matching
	Keystroke dynamics [16]	Keystroke rhythm, Typing structure, Time interval between strokes, Hold time in stroke, Ways of input password	Neural networks		Template matching, Neural networks
	Mouse movement dynamics	Specifically mouse movement and using of keys, Specifically keyboard use	Neural networks		Template matching, Neural networks
	Signature [5]	Writing angle, Signature time, Speed and acceleration while writing, Number of pencil lifts, Pressure strength			Template matching
	Smile	Ways of skin moving while smiling			Template matching
	Structure of brain waves	Structure of brain waves			Template matching

Voice [10]	Ways of pronunciation of words, Voice altitude, Voice intonation	Hidden Markov Model, Recurrent Neural networks, Wavelet transform, Linear Pedictive Coding	Hidden Markov Model	Hidden Markov Model, Language model subsetting, Recurrent Neural networks, Template matching - dynamic time warping, Grammar and language models

Such a systematization has the following advantages. (1) it is open which means that new biometric methods, characteristics and patterns can be added at will without changing the structure of the systematization, (2) every method can be uniquely classified according to the given criteria (type of characteristic, type of method, type of pattern). Because of (2) we can introduce a new classification of biometric methods: (i) special methods (used only for one characteristic), (ii) general methods (used for more than one characteristic). Other classifications similar to this (e. g. with regard to method type or pattern type) can be added at will. Other advantages include (3) empty fields which should be subject to future research, (4) clear distinction between biometric methods and biometric models. While a biometric method covers only one field in the table, a biometric model would cover at least two. For example PCA (Principal Component Analysis) would be a method for feature extraction, while HMM (Hidden Markov Model) and neural networks would be samples of a biometric system according to our previous definition of biometric model.

4. OPEN ONTOLOGY OF BIOMETRICS

An ontology (in the perspective of information and computer sciences) models concepts, instances, relations and attributes which can exist for an agent or a group of agents inside a domain. Ontologies are used to reason about the objects within the domain.

During this systematization of biometrics we defined some fundamental concepts an their interrelationships and thus this systematization can be used as a fundament in building an open ontology of biometrics. Such an open ontology would let us formalize biometrics even more than this systematization did, and would us give the possibility to reason about specific objects inside the domain of biometrics. We call such an ontology open since the amount of knowledge grows and thus any ontology should be updated. Open in this context could be understood in the meaning of open in the open source paradigm where applications change over time with the goal to fit the users needs better.

From our perspective the concepts of biometric system, method, model, characteristic, pattern and extracted feature are fundamental concepts in the domain of biometrics. Other concepts such as user, enrollment, authentication, identification, classification, security etc. should be added to make such an ontology more precise.

5. CONCLUSION

In this paper we drew attention to the problem of therminological inconsistencies in the field of biometrics. In order to solve this problem we developed a systematization of biometrics on behalf of a modified general biometric system and a classification of biometric characteristics. This systematization has the following advantages: (1) openness, (2) unambiguity, (3) identification of fields for future research, and (4) clear distinction between biometric method and biometric model.

Additionally to the systematization classifications of biometric methods with regards to

type of method, type of pattern and type of characteristic has emerged. Methods for quality control are identified as the field were most future research is possible.

We argued that this systematization defines some of the fundamental concepts of biometrics and thus can be used as a basis for the development of an open ontology of biometrics. The implementation of such an ontology is part of the further research of the authors.

REFERENCES

[1] J. Ashbourn. Practical Biometrics, Springer-Verlag, NJ, 2003.

[2] M. Bača, M. Uvod u računalnu sigurnost, Narodne novine, Zagreb, 2004. In Croatian

[3] M.Bača. M. Čubrilo. K. Rabuzin. Biometric in ITS security, IIS'05 Conference proceedings of 16th International Conference "Information and Intelligent Systems", Varaždin, 2005.

[4] M. Bača. Ž. Hutinski. K. Rabuzin. Using Face Recognition System in Ship Protection Process, Traffic & Transportation, Scientific Journal on Traffic and Transportation Research, Vol. 18, No.2, Zagreb, 2005.

[5] R. Bolle. J. Connel. S. Pankanti. N. Ratha. A. Senior. Guide to Biometrics, Springer-Verlag, NJ, 2003.

[6] J. Chirillo. S. Blaut. Implementing Biometric Security. Wiley Publishing, Inc., USA, 2003.

[7] R. H. Giles. Lasting Forests Glossary, on-line <http://fwie.fw.vt.edu/rhgiles/appendicies/glossb.html> (28.02.2005.)

[8] K. Jain. H. Chen. S. Minut. Dental biometrics: Human identification using dental radiographs, In. Proc. Of 4th AVBPA, Guildford, UK, 2003.

[9] K. Jain. S. Pankanti. S. Prabhakar. L. Hong. A. Ross. J. L. Wayman: Biometrics Tutorial. Michigan State University,

[10] on-line <http://biometrics.cse.msu.edu/icprareareviewtalk.pdf>, (06.11.2005.)

[11] S. Makthal, A. Ross. Synthesis of iris images using Markov random fields. In Proceedings of 13th European Signal Processing Conference (EUSPICO), Turkey, 2005.

[12] M. Munich. Q. Lin. Explicit modeling of common acoustic features for character recognition. In Proc. Of the 12th European Signal Processing Conf. (EUSPICO 2004), Vienna, 2004.

[13] S. Nanavati. M. Thieme. R. Nanavati. Biometrics Identity Verification in a Networked World. Wiley Publishing, Inc., USA, 2002.

[14] D. Radošević. Osnove teorije sustava. Nakladni zavod Matice Hrvatske, Zagreb, 2001. In Croatian

[15] P. Reid. Biometrics for Network Security. Prentice Hall, Upper Saddle River, NJ, 2004.

[16] K. Saeed. P. Charkiewicz. A New Approach for Hand-Palm Recognition. Kluwer

Academic Publishers, USA, 2005.

[17] D. X. Song. D. Wagner. X. Tian. Timing analysis of keystroke and timing attacks on SSH. 10th USENIX Security Symposium, Washington, USA, 2001.

[18] [J. L. Wayman. A. K. Jain. D. Maltoni. D. Maio. Biometric Systems, Technology, Design and Performance Evaluation. Springer-Verlag London, Berlin, Heidelberg, 2005.

[19] J. L. Wayman. Generalized Biometric Identification System Model. National Biometric Test Center, Collected Works 1997. - 2000., San Jose State University, 2000., pp. 25 - 31.

[20] X. Zhu. I. Waibel. Segmenting Hands of Arbitrary Color. In. Proc. IEEE Int. Conference in Automatic Face and Gesture Recognition, 2000.

[21] M. Žugaj. K. Dumičić. V. Dušak. Temelji znanstvenoistraživačkog rada. Metodologija i metodika, Fakultet organizacije i informatike, Varaždin, 1999.

[22] J. Wayman, A. Jain, D. Maltoni, D. Maio: Biometric Systems", Springer, 2005.

Received: 16 March 2007

Accepted: 2 November 2007