CrossMark

Validating data acquired with experimental multimodal biometric system installed in bank branches

P. Szczuko[1] · A. Czyżewski[1] ⓘ · P. Hoffmann[1] ·
P. Bratoszewski[1] · M. Lech[1]

Abstract An experimental system was engineered and implemented in 100 copies inside a real banking environment comprising: dynamic handwritten signature verification, face recognition, bank client voice recognition and hand vein distribution verification. The main purpose of the presented research was to analyze questionnaire responses reflecting user opinions on: comfort, ergonomics, intuitiveness and other aspects of the biometric enrollment process. The analytical studies and experimental work conducted in the course of this work will lead towards methodologies and solutions of the multimodal biometric technology, which is planned for further development. Before this stage is achieved a study on the data usefulness acquired from a variety of biometric sensors and from survey questionnaires filled in by banking tellers and clients was done. The decision-related sets were approximated by the Rough Set method offering efficient algorithms and tools for finding hidden patterns in data. Prediction of evaluated biometric data quality, based on enrollment samples and on user subjective opinions was made employing the developed method. After an introduction to the principles of applied biometric identity verification methods, the knowledge modelling approach is presented together with achieved results and conclusions.

Keywords Biometry · Identity verification · Dynamic signature · Voice recognition · Face recognition · Rough set

1 Introduction

One of the significant trends in the world of current scientific research is rapid development of biometric verification methods. On the basis of the project IDENT ("Multimodal

✉ P. Szczuko
  ident@multimed.org

[1] Faculty of Electronics, Telecommunications and Informatics, Multimedia Systems Department, Gdańsk University of Technology, Gdańsk, Poland

biometric system for the verification of bank clients identity") authors of this paper working in an academic institution had an opportunity to build a scientific synergy with the specialists of the biggest Polish Bank (PKO BP), both in terms of technical cooperation, as well as in the domain of assessing the feasibility of biometric solutions, which provide the subject of joint research and development work. It is probable that the future identity document will include handwritten signature, face and voice features. Therefore, it was decided in the context of the project, to plan a broadened research scope in the domain of a multimodal experimental biometry in real banking environment. The research was planned to be carried-out in the following areas:

– the development of biometrics stations to be installed in bank branches,
– the further development of algorithms and methods for dynamic analysis of a handwritten signature,
– the implementation of methods for secure authentication by voice,
– the development of methods for the analysis of facial contour using the lidar imaging,
– the implementation of face recognition algorithms in RGB video,
– combining multiple biometric methods (modality fusion).

The engineered biometric stands were also completed by a hand vein scanner available on the market (Fujitsu Identity Management and PalmSecure 2017), serving for the purpose of reference to the methods implemented by the research team.

In the course of the research, the previous knowledge in the field of multimodal technology of identity verification was used and analyzed in the context of many research challenges. The research work was conducted in real banking environment, since PKO Bank Polski offered access to 100 teller stands located in 60 bank branches, which creates the opportunity to develop the experimental research at a large scale.

The development of mobile biometrics has accelerated sharply in recent years. The present verification solutions are based on fingerprint, voice or face scan. Unfortunately, none of the available methods of bank clients verification provides complete reliability for the whole range of possible cases. Therefore, the authors, on the basis of their experience, proposed the use of several biometric methods and their effective combination. A secure data communication system was also established for the purpose of biometric data transmission inside the bank communication network. The technology had been implemented in the form of a system for acquiring, storing, analyzing, and combining (fusing) biometric data. When using simultaneous biometric verification of many modalities, from the practical and research viewpoint, the significant issue is studying the level of acceptance of individual technology solution by both: banking clients and banking tellers.

Especially, combining the biometric methods should be beneficial for the person being verified, both in terms of improving the safety and convenience of using the services of (lowering the rate of false acceptance and false rejection errors). The central element of this process is to discern information on the convenience of individual methods and their acceptance.

The research on the development of the dynamic signature method (described in Section 2) is significant enough to lead to the development of the most acceptable method of identity verification in the society. A handwritten signature has been for many centuries the most prominent means of expressing one's approval, and to confirm the identity. Owing to the work carried out on the technology of biometric signature, it will be possible to reduce the risk of signature forgery by adding recorded extra dynamic measures of the signature. Therefore, an important step in our research, consequently leading to a large-scale implementation, is the further advancement of methods of acquisition, analysis

and interpretation of graphomotoric data. In particular, the leading role will be assigned to the methods of acquisition and playback of dynamic handwritten signature on the basis of multi-dimensional data obtained from accelerometric and gyroscopic sensors, pen-on-paper or pen-on-screen pressure sensors, as well as from the sensors of the pressure exerted on the handle of the pen. During the broadened research on the developed biometric pen technology (Lech and Czyżewski 2016; Lech et al. 2016; Shanker and Rajagopalan 2007), the action was taken to analyze the functioning of the individual components that make up the pen for the mobile applications. The engineered biometric pen and the whole biometric stand installed in one of bank branches are depicted in Fig. 1.

Besides the numerous practical applications of the face image analysis methods known from the literature (Braga 2017; Bhele and Mankar 2015; Borade et al. 2016; Papatheodorou and Rueckert 2007) and used for personal identification, the progress and spreading of the image acquisition technology resulted also in the market introduction of widely available lidar technologies (laser scanning to produce a spatial representation of the visual three-dimensional objects). A current example of a commercial incarnation of this type of technology is the Time-of-Flight (ToF) camera. It uses temporal relations of radiated and reflected light in order to visualize spatial objects. This way, an image extracted from the acquired data cloud can be used to restore the contour of the face, which is a strong distinctive feature of individuals. Hence, this modality has been also implemented our project to enable carrying out research in the field of applications of laser face contour as an innovative method of biometric identification (Bratoszewski and Czyżewski 2015). However, at the time of this paper preparation the results of processing of data acquired in this way were not available, yet.

We use also some typical biometric approaches, namely RGB image-based face recognition (as in Section 3) and voice biometry (Section 4). Moreover, a commercially available palm vein scanner was mounted on our experimental biometric stands for collecting additional biometric data.

The main purpose of the research scope presented in this paper is to verify correlation between objective characteristics of biometric samples and samples clusters and subjective assessment of ergonomics, user satisfaction and easiness of use of particular biometric traits.
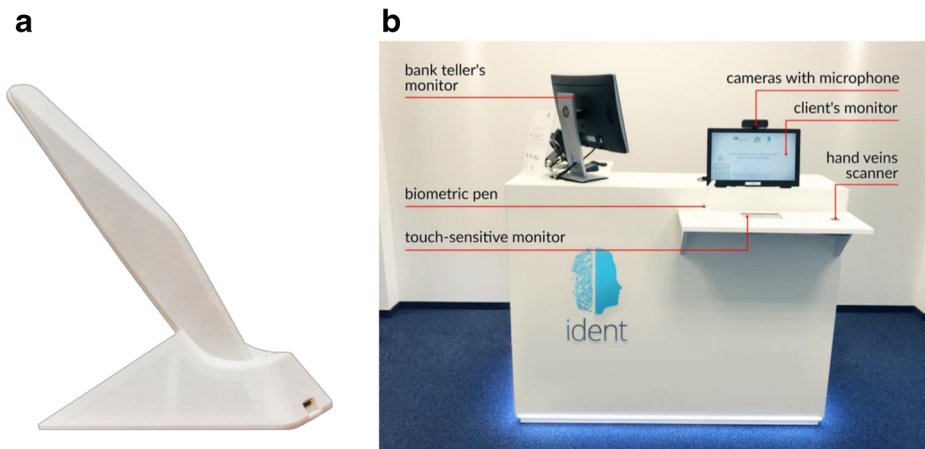

Fig. 1 Biometric pen developed during research (**a**) and biometric stand installed in one of bank branches (**b**)

By applying the Rough Set method, a decision-related sets can be approximated. The following key notions such as satisfied or dissatisfied, proficient or not proficient, reliable or unreliable are of interest. Data quality can be also expressed in terms of reliable or unreliable, stable or unstable data, etc. The rough representation is useful for classifying new cases during the enrolment phase, where the number of samples, stored in this phase, is relatively low. The prospect of repeatedly collecting new sample on each verification attempt should be considered. Therefore, based on initial, limited knowledge modeled from enrollment samples a recommendation should be automatically given, e.g. to reject most unstable trait, to repeat registration, or, in worst cases, to rely only on classical verification methods as an alternative procedure. Therefore, the results of soft computing-based processing of gathered data will serve a very important goal in terms of this project purpose.

2 Signature biometry

The signatures are acquired using the special wireless biometric pen, developed in the scope of the project. The biometric pen is equipped with a 3-axis accelerometer, 3-axis gyroscope, and a surface pressure sensor, plus a Bluetooth LE transmitter/receiver, and a rechargeable battery. The signature is being put down on the touch screen for the purpose of providing the visual feedback to the user. Whenever the user places the pen on the surface of the screen, triggering the pressure on the tip-mounted sensor and beginning to put down the signature, the data stream containing the acceleration and the angular positions is being transmitted via the Bluetooth LE channel. As both the acceleration and the angular position signals contain data for 3 axes, the signature verification has been based on 6 dynamic measures, which will be described later on in the paper, created employing the results of the dynamic time warping (DTW) algorithm. The method is based on the assumption that time-domain functions of two arbitrary authentic signatures entail less warping than the functions of the authentic and the forged signature. Therefore, the information derived from the DTW method and then used for the verification process represents the convergence of a "diagonal" and an optimal cost path in the accumulated cost matrix. The matrix contains the distances between every sample of functions F and G, representing values of a particular parameter, such as acceleration or angular position, of two arbitrary signatures of lengths m and n. The sequence of low numbers in cells close to the diagonal, indicates which samples of function F are closest in value to those of function G.The matrix is created in the manner given by (1), where $\gamma_{i,j}$ is the accumulated cost in cell (i, j), and d is the distance (2) between the elements of functions F and G (3 and 4):

$$\gamma_{ij} = d(f_i, g_j) + \min(\gamma_{i-1,j-1}; \gamma_{i-11,j}; \gamma_{i,j-11}) \tag{1}$$

$$d(f_i, g_j) = |f_i - g_j| \tag{2}$$

$$F = f_1, f_2, \ldots, f_i, \ldots, f_n \tag{3}$$

$$G = g_1, g_2, \ldots, g_j, \ldots, g_m \tag{4}$$

The distance metric given by (2) has been chosen empirically and it turned out that it outperformed other popular metrics, providing the best relation of FRR (false rejection rate) to FAR (false acceptance rate) measure during experiments. The standard back tracing

approach (Lech et al. 2016) of finding the optimal path **w**, given by (5) and (6), in the accumulated cost matrix has been used:

$$w = \{w_k, w_{k-11}, \ldots, w_0\} \qquad \max(m; n) \leq m + n - 1 \tag{5}$$

$$k_k = \begin{cases} (i-11, j-1) & \gamma_{i-1,j-1} = \min(\gamma_{i-1,j-1}; \gamma_{i-11,j}; \gamma_{i,j-11}) \\ (i-11, j) & \gamma_{i-1,j} = \min(\gamma_{i-1,j-1}; \gamma_{i-11,j}; \gamma_{i,j-11}) \\ (i, j-1) & \gamma_{i,j-1} = \min(\gamma_{i-1,j-1}; \gamma_{i-11,j}; \gamma_{i,j-11}) \end{cases} \tag{6}$$

In the previous work by Lech and Czyżewski (2016) another path tracing method had been proposed, which performed well with the measures representing in time domain the shape of the signature put down on the graphic tablet using an ordinary tablet stylus. In the work presented herein, for the 3-axis acceleration (ax, ay, az) and 3-axis angular position from a gyroscope (gx, gy, gz), 6 measures have been defined, denoted respectively by cax, cay, caz, cgx, cgy, cgz, representing a degree of convergence of DTW matrix "diagonal" and the optimal path. The convergence c for any of the six measures is defined as the sum of absolute differences of y positions of pixel belonging to the matrix "diagonal" d and pixel belonging to the optimal cost path w, for the same x position, according to (7).

$$c = \sum_{x=0}^{k} |y(d_x) - y(w_x)| \tag{7}$$

The process of verification involves a comparison of model signatures with those obtained in the current authentication session. The assessment of the degree of signature authenticity p, within a range [0 1], using the 6 measures involves a comparison of their values c' with threshold values cTHR, according to (8):

$$p = \begin{cases} 1 & c' < c_{THR} \\ \dfrac{c_{THR}}{c'} & c' > c_{THR} \end{cases} \tag{8}$$

where c' is a value of measure c obtained from DTW method after rescaling, according to (9), where n and m define the size of the accumulated cost matrix:

$$c' = 10000 \frac{c}{nm} \tag{9}$$

While the convergence c for any of the six measures is rescaled by the size of the accumulated cost matrix, the thresholds c_{THR} could be set to a fixed value. The value was equal to 300 and it was set empirically, providing the best FRR/FAR ratio. The global similarity ratio value is the average value of all the p values.

3 Face biometry

Face image processing chain starts with image acquired by a RGB camera with resolution of 1920×1080 pixels at 30 frames per second rate. Extraction of facial image parameters is preceded by face detection in the scene. The face detection methods involved in the processing were presented in more detail in the previous paper by Szczodrak and Czyżewski (2017). The parameterization procedure is described more precisely in the paper by Bratoszewski et al. (2017), nevertheless the main stages of the feature extraction process are recalled here. landmarks in total number of 77 are calculated on the detected face image. Based on the landmarks, distinctive face regions are found, i.e. eyes, eyebrows, nose, mouth. Each fragment of the image representing those regions and the whole face image are parameterized using Histogram of Oriented (HOG) Gradients and Local Binary Pattern (LBP) features. In

the last step, each parameter set calculated from separate image patches and the face image itself are concatenated. Subsequently the Linear Discriminant Analysis (LDA) is performed in order to create the feature vector of the size equal to 768 elements. Number of features was chosen upon an analysis of the previous research studies (Klontz et al. 2013; Jin and Zhang 2014). The samples of face landmarks (mouth, jaw, chin) found for 3 different persons together with corresponding first 6 values of feature vectors of the mouth region were presented in Fig. 2.

4 Voice biometry

The speaker identity verification is performed using Gaussian Mixtures (GMM) and Universal Background (UBM) Models (Chen et al. 2012). The Alize framework is used as the speaker recognition back-end (Alize 2017) and mel-frequency cepstral coefficients (MFCC) were employed for speech parametrization. In the first step the Universal Background Model (UBM) was trained based on the recordings prepared in real bank branches environment. Those recordings include speech data recorded by 84 participants in both quiet and noisy environments that were found in real banking outlets environment. Besides the performed recordings the speech material from the MOBIO dataset (McCool et al. 2012) was utilized for the processing in order to increase UBM inner variance.

All speech signals were recorded using a single microphone, with 44 kS/s sampling rate and 16-bit resolution. The MFCCs in 13 cepstral channels were extracted within 10 ms timebase, using window of the size equal to 25 ms. The number of MFCCs used in speaker modelling varies from 10 to 20 in literature (Gupta and Gupta 2013; Mermelstein 1980). This is mainly determined by the speech signal characteristics where most of the information is held in low frequency components of speech spectrum (i.e. formants) what coincides with the highest resolution of the mel filters. In this work the number of coefficients was chosen based on both literature studies and empirical approach and was set to 13. Furthermore, it has been long established that adding dynamic information (Furui 1982) increases the speaker recognition accuracy, therefore, the final acoustic feature vector was formed by combining zero-order MFCCs with delta and delta-delta features, resulting in 39 features in total

At the data acquisition step users were recording their utterances in four trials. First, three of them consisting of 17 seconds of speech were acquired in the users' enrolment process. The speech model training refers to the creating of the user-specific statistical Gausian Mixture Model (GMM) adapted from the UBM employing the maximum a posteriori criterion (Gauvain and Lee 1994). The fourth speech sample, 7 seconds-long was used for the purpose of users' identity verification.

4.87295, 2.71676, -5.28989,   -2.45700, -2.72461, -1.22167,   1.06519, 8.60224, -2.94332,
2.06025, -7.80630, 5.55962   4.60434, 4.27078, -6.34568   1.05944, 1.60684, -1.48528
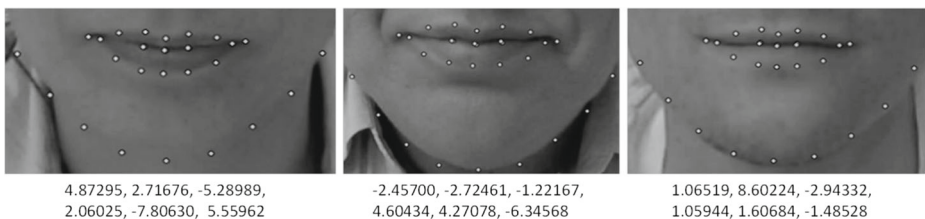
Fig. 2 Comparison of the face landmarks and feature vectors calculated for 3 different persons together with corresponding first 6 values of feature vectors of the mouth region (numbers below each photo)

In the experiments related to this work the gathered data were analyzed in the following manner: in order to determine the similarities between speech samples the log-likelihood of the utterance X over the model s (Larcher et al. 2013) was computed as in (10):

$$\log P(X \mid s) = \sum_{t=1}^{T} \log \sum_{c=1}^{C} \gamma_c N(x_t, \mu_c, \Sigma_c), \qquad (10)$$

where: X is represented as a sequence of acoustic features derived from speech recording sample (39 features per every signal frame) $X = \{x_t\}_{t \in T}$ of the length T, C is the number of distributions in s (set to 256 for our experiments) and γ_c, μ_c and Σ_c are the weight, mean and co-variance matrices of the cth distribution, respectively. The introduced log-likelihood equalization is used for the calculation of distances between the model s and the sample X in the distance-based analysis of biometric samples clusters as in Section 6.

In the experiments referring to the calculation of distances between each enrollment the samples were calculated using *log likelihood ratio* (LLR) defined by the Neyman-Pearson Lemma (Jiang 2005) as in (11):

$$\log \Lambda_M = \log p(H_M|x, C_v) - -log p(\neg H_M|x, C_v), \qquad (11)$$

where $p(\neg H_M|x, C_v)$ is the probability of H_M not being the right hypothesis given x (meaning that M is not the real speaker of x). Depending on the experiment assumptions different voice recordings of the speaker are used for forming the x vector.

5 Knowledge modeling method

Foundations of rough set theory were laid by Polish mathematician Pawlak (1982, 1991), and used with successfully in many domains, including biometric data feature selection (Mazumdar et al. 2010), behavioral patterns analysis (Bazan et al. 2005), discovery of medical knowledge (Tsumoto 2002), dimensionality reduction (Banerjee et al. 2007). The rough set prominent applications involve processing of real-life data, often imprecise, ambiguous and uncertain. Therefore, considering the subjective nature of data collected in the presented experiments, the rough set methodology is regarded adequate as a processing and a data mining tool. The theory is used for the approximation of a set by defining its upper and lower approximation: the first one including objects that may belong to the set, and the latter one including objects that certainly belong to the set. Both approximations are expressed as unions of the so-called atomic sets containing objects that are indiscernible, i.e. they have the same values of attributes (Fig. 3). Two objects x and y can be characterized by a defined set of attributes $\mathbf{P} \subseteq \mathbf{A}$ ($\mathbf{P}$ is a subset of all possible attributes in $\mathbf{A}$), and they are related by
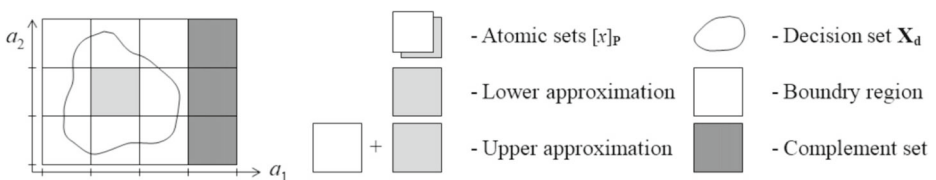

Fig. 3 Partition of the universe based on attributes a_1 and a_2 into atomic sets, and approximation of decision set $\mathbf{X}_d$

the indiscernibility relation if: $(x, y) \in \text{IND}(\mathbf{P})$, where $\text{IND}(\mathbf{P})$ is the equivalence relation defined as in (12):

$$\text{IND}(\mathbf{P}) = \{(x, y) \in \mathbf{U}^2 \mid \forall a \in \mathbf{P}, a(x) = a(y)\}, \tag{12}$$

where $a(x)$ is a value of attribute a of object x. In this work $\mathbf{P}$ is a set of features introduced in Section 7, and objects x are particular biometric identities.

All objects being in a relation with a given object x produce an equivalence class of this object $[x]_{\mathbf{P}}$. If $\mathbf{P}$ contains attributes sufficient for distinguishing between objects with different decisions, then the class $[x]_{\mathbf{P}}$ contains only objects with the same decision as x. It is assumed that a possible lack of discernibility between some objects inside the equivalence class is not diminishing the classification accuracy. Thus, the considered set of attributes $\mathbf{P}$ generates a partitioning of the universe of discourse $\mathbf{U}$ into atomic sets, which in turn form building blocks for representing decision classes expressed as rough sets. A set of all objects with a given decision d, is denoted as $\mathbf{X}_d$. Then $\mathbf{X}_d$, can be approximated by ist lower approximation $\underline{\mathbf{P}}\mathbf{X}_d$ (13):

$$\underline{\mathbf{P}}\mathbf{X}_d = \{x | [x]_{\mathbf{P}} \subseteq \mathbf{X}_d\}, \tag{13}$$

namely, a set of all objects x whose equivalence classes $[x]_{\mathbf{P}}$ are included within the decision class of interest $\mathbf{X}_d$. It can be interpreted as a set of objects whose attributes values allow for precise classification.

On the other hand the set of objects $\mathbf{P}\mathbf{X}_d$ is called upper approximation and it is defined as (14):

$$\mathbf{P}^- \mathbf{X}_d = \{x \mid ([x]_{\mathbf{P}} \cap \mathbf{X}_d) \neq \varnothing\}, \tag{14}$$

whereas it includes all objects whose equivalence classes have non-empty intersections with the decision class $\mathbf{X}_d$. It can be interpreted as a set of objects whose attribute values point to objects with the decision of $\mathbf{X}_d$, but some equivalent object(s) in $[x]_{\mathbf{P}}$ can have other decision as well (see Fig. 1).

The given subset of attributes $\mathbf{P}$ can be sufficient to generate such a partitioning of the universe of $x \in \mathbf{U}$ that decision classes are approximated correctly. The accuracy of rough set approximation of $\mathbf{X}_d$ is expressed as (15):

$$\alpha = |\underline{\mathbf{P}}\mathbf{X}_d|/|\mathbf{P}^-\mathbf{X}_d| \tag{15}$$

where $\alpha \in [0, 1]$ and $\alpha = 1$, is true for a precisely defined crisp (conventional) set.

In order to express the quality of a particular classification more conveniently for many decision classes a positive region is defined as (16):

$$\mathbf{POS}(\mathbf{P}) = \sum \underline{\mathbf{P}}\mathbf{X}_d \tag{16}$$

what is interpreted as a set of all objects with an accurate classification attributed to any of the modeled decision classes $\mathbf{X}_d$. If $\mathbf{POS} = \mathbf{U}$ then each object of universe is correctly classified, whereas $|\mathbf{POS}|/|\mathbf{U}|$ is a relative number of correctly classified objects related to all objects in the universe.

Application of the rough sets theory in decision systems often requires a minimal (the shortest) or the most convenient subset of attributes $\mathbf{RED} \subseteq \mathbf{P}$ resulting in the same quality of approximation as $\mathbf{P}$, called a reduct, therefore introducing the same indiscernibility relations: $\text{IND}(\mathbf{RED}) = \text{IND}(\mathbf{P})$. Numerous algorithms to calculate reducts are available. A greedy heuristic algorithm was applied for this work (Janusz and Stawicki 2012).

Usually, for attributes with continuous values, prior to the reduct calculation a discretization is performed. Maximum discernibility (MD) algorithm is applied, which analyses

attribute domain, sorts values present in the training set, takes all midpoints between values and finally returns the midpoint maximizing the number of correctly separated objects of different classes (Bazan et al. 2000; Nguyen 2001). Above procedure is repeated for every attribute.

Once the reduct is obtained, the attributes useful for expressing differences between objects of different classes are determined, then all cases in the decision table are analyzed, and finally decision rules are generated. Attributes $a_n \in$ **RED** of each object x_i are treated as an implication's antecedent, whereas the decision forms its consequent. Rules in the form of logic sentences are obtained as in (17):

$$IFa_1(x_i) = v_1 \ AND \ \ldots \ AND \ a_n(x_i) = v_n \ THEN \ d(x_i) = d_i. \tag{17}$$

At the classification phase these rules are applied for every object in the testing set, and then the decision is first predicted, and then compared to the actual one. More information on the rough set theory can be found in the literature (Pawlak 1982; Bazan et al. 2000; Nguyen 2001; Zhong et al. 2001).

6 Biometric samples characterization

Raw biometric signals, such as a voice recording, face image or acceleration and rotation data from the engineered biometric pen, were processed to extract features. From this point biometric samples are described by respective features and they are treated as points in the multidimensional feature space. These points will be characterized employing: 1) their relative distances, 2) average position, i.e. cluster center, and 3) distances of samples to the cluster center. Distances reflect stability of each biometric feature, as well as their changes over time. Appropriate metrics were introduced for each biometric modality: face image, signature, and voice, as is described below.

6.1 Signature samples distance metric

The process of calculating distances between two arbitrary signature samples or the sample and the cluster center involves a comparison of their p values for each measure (Section 3), forming a vector $\mathbf{v}_{sig}$ (18), according to (19):

$$\mathbf{v}_{sig} = \left\{ p(c_{ax}), p(c_{ay}), p(c_{az}), p(c_{gx}), p(c_{gy}), p(c_{gz}) \right\} \tag{18}$$

$$\| p', p'' \|_{signature} = \sqrt{\sum_{n=1}^{N} (p'_N - p''_N)^2} \tag{19}$$

where p', p'' are two compared p values, and N is a number of features.

6.2 Face samples distance metric

Face modality is represented by samples in 768-dimensional space of face features, described in Section 3. Therefore, for two samples the Euclidean distance metric can be

applied, serving as a similarity measure, which facilitates characterization of differences occurring inside the particular identity class (20):

$$\|A, B\|_{face} = \sqrt{\sum_{n=1}^{N}(A_n - B_n)^2} \qquad (20)$$

where A, B denote two compared samples, and N is a number of features.

6.3 Voice samples distance metric

Voice modality is represented by Gaussian Mixtures Models. For each speaker, regardless the number of collected recordings, 256 mixtures were created, each comprising 39 Gaussians as described in Section 4. A distance metric is based on log-likelihood between a model and a signal in question (19). Namely, to obtain the distance between two samples, a GMM model based on the first recording in the pair is created, and the probability that the second recording fits to the model is calculated (21). This metric is unsymmetrical, as the model based on sample i-th validates j-th sample at a different level than the model based on j-th sample validates i-th sample.

$$\|A, B\|_{voice} = \log P(B|A) \qquad (21)$$

Above metrics (18–20) will be denoted as $\| \cdot \|_{trait}$, where the *trait* may be signature, *face* or *voice*. The following methodology is compliant with each discussed modality. By applying the distance metric various features can be derived from all available samples.

6.4 Data cluster characteristics based on distances

For every modality, the following data characteristics were collected:

– minimal, mean, and maximal distances between all possible pairs of samples (22–24), reflecting existing similarities and dissimilarities between samples,
– mean value of all samples attribute values, representing the center of the samples cluster in n-dimensional space (25). In case of voice a model based on all samples ($\mathbf{C_v}$) was used instead (26).
– the distance of each sample from the cluster (27–28).

$$dist_{\min} = \min_{i,j=1...N}\{\|x_i, x_j\|, \ i \neq j\} \qquad (22)$$

$$dist_{mean} = 1/(N \cdot (N - 1)) \cdot \sum_{i,j=1...N, \ i \neq j}(\|x_i, x_j\|) \qquad (23)$$

$$dist_{\max} = \max_{i,j=1...N}\{\|x_i, x_j\|, \ i \neq j\} \qquad (24)$$

$$C = \left\{a_i = 1/N \cdot \sum_{j=1...N}(a_i(x_j)), \ i = 1...M\right\} \qquad (25)$$

$$C_v = \sum_{j=1...K}(w_i \, g(x_N|\mu_i, \ \Sigma_i)) \qquad (26)$$

$$Cdist_{\min} = \min_{i=1...N}\{\|C, x_i\|\} \qquad (27)$$

$$Cdist_{\max} = \max_{i=1..N}\{\|C, x_i\|\} \qquad (28)$$

where: $(x_i, x_j) \in \mathbf{U}^2$, N – number of objects, and M – number of features, w_i are normalized positive scalar weights of the Gaussian mixtures, K is the number of mixtures in the voice model, x_N is the vector of acoustic features for N objects (recordings), $g(x_N \mid \mu_i, \Sigma_i)$ are the Gaussian densities with mean vector μ_i and covariance matrix Σ_i (Jiang 2005).

The cluster of all enrolment samples was characterized by (22–28), and then a new cluster of all positive samples, i.e. union of sets from the enrolment and from verification procedure, was created and parameterized alike.

As it was previously stated, distances are calculated in the multidimensional feature space. To explain the purpose of introduced metrics a synthetic 2-dimensional samples and distances are visualized (Fig. 4). It can be observed, that some cluster changes occur as a result of adding new samples. In this case:

– cluster center shifted towards new samples,
– $dist_{min}$ decreased as a new sample appeared more similar to existing sample,
– $Cdist_{min}$ decreased as new cluster center is closer to existing sample,
– $Cdist_{max}$ decreased as new cluster center was shifted right, better balancing the distribution of samples, and cluster size decreased,
– $dist_{max}$ was not changed as two most dissimilar samples are still in the set.

Depending on the distribution of enrollment and new samples such changes can be more prominent or rather subtle. The higher the number of samples the less impact new samples have on cluster characteristics, as the cluster based on many samples better estimates real distribution. The important question related to this research is to predict changes of small clusters based on personal subjective characteristics of the user. It is assumed, that unreliable user or user feeling not comfortable with the particular biometric technology will register samples that are not representative during the enrollment, so verification samples may shift
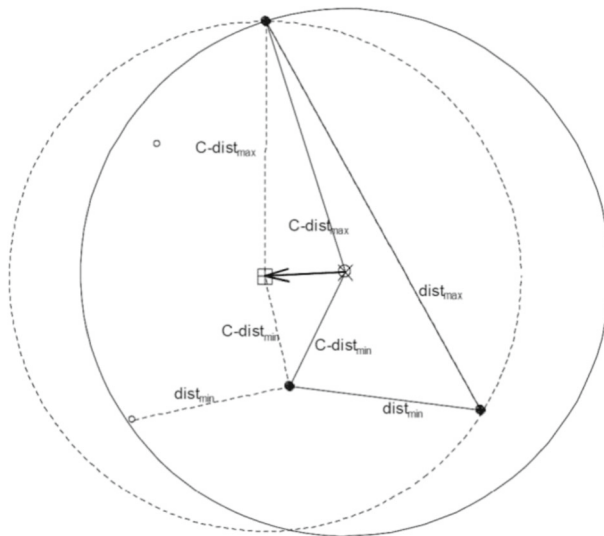

Fig. 4 Cluster characteristics changes as a result of adding new samples. Visualization of synthetic 2-dimensional samples. Legend: dots: 3 enrollment samples, crossed circle: cluster center, large circle: cluster size, open circles: 2 new samples, crossed square: new cluster center, arrow: cluster shift, dashed large circle: new cluster size

the cluster significantly. A reliable user will register representative samples, assuring their repeatability, especially in behavioral biometry traits.

Taking into account uncertainty of subjective data as well as imprecision of signals features the rough set methodology is regarded as an efficient data mining tool, by definition aimed at modeling and processing of decision sets approximations. Finally, the presented work concludes with selection of relevant features, generation of rules for classification of new cases, thus acting as a framework encompassing all phases of data mining, through modeling up to exploitation defined for the application of biometric identity verification.

Other algorithms such as decision trees, and neural networks are not entirely suitable for processing of imprecise and approximate data. Thus in this research the subjective responses were collected by means of questionnaires with discrete answers. Therefore, the rough set theory was chosen as the most appropriate approach.

The goal of the work presented further on is to verify relations between objective characteristics of biometric samples sets and subjective features related to users' performance, their experience and subjective opinions.

7 Subjective characteristics of biometric traits

After each biometric identity registration and validation attempt users and the bank tellers were asked questions related to a predefined set of answers and 2 open questions of technical nature (Tables 1 and 2).

The distribution of answers is presented in the form of statistics below (Fig. 5). 126 persons took part in the pilot exploitation tests, providing their opinions.

It should be stressed that "0" interpreted as refusal to answer was frequent for almost every question, and it seems that users were not always motivated enough to provide replies. In turn, some answers, although different than "0", were not reliable, as e.g. user gender expected to be "1" or "2" has also other numerical values. It was expected that biometric pen ergonomics will be critically commented because it differs in bulk from a typical pen, but only one person of 126 reported that it is inconvenient.

8 Objective characteristics of collected biometric samples

Distances for all collected samples are shown in Figs. 6, 7 and 8. The radial plots show distances in the clusters only, but not absolute clusters position. All clusters are centered on the plot, thus in each case $Cdist = 0$ is situated in the middle of the plot. At each angle around the center a different biometric identity representation is plotted, sorted by an increasing value of $Cdist_{max}$. Distances of enrollment samples are shown by black dots, mean of all distances for all identities is shown as a black circle. Distances of new samples are shown by gray dots. It can be observed that in numerous cases new samples are placed significantly further away from the center, thus falling outside the region defined by mean distance, so that generally, thresholds for the similarity measure must not be defined on the basis of enrollment samples only.

The analysis of distances in the clusters of biometric samples reveals that for face and for voice an identity initially described by a compact cluster tends to remain relatively compact after adding new samples. The face recognition is a standard biometric trait, easy to obtain in controlled lighting conditions, thus repeatability of the samples features is high. The voice utterance is one of behavioral biometric traits, requiring a moderate level of rigor and

Table 1 User questions

Question	Short names	Answers*
How fast was the biometric samples registration process in your opinion?	uFast	1 – very slow, 2 – slow, 3 – average, 4 – fast, 5 – very fast
Was (signature/voice/face) registration easy and intuitive?	uEasyPen, uEasyVoi, uEasyFac	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes (separate answer for each trait)
Was the (signature/voice/face) biometry a reliable and convenient way of verification?	uReliPen, uReliVoi, uReliFac,	1 – too complex, 2 – complex, 3 – average, 4 – rather straightforward, 5 – very convenient, 6 – hard to judge (separate answer for each trait)
Which one of biometric modalities was the most inconvenient?	uHard, hardPen, hardVoi, hardFac	1 – signature, 2 – voice, 3 – face
Why the modality was inconvenient?	uWhy	1 – was not starting, 2 – usage was hard, 3 – verification failed, 4 – other reasons
Are you willing to use your biometric traits during each banking activity?	uWill	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes
Was the biometry registration environment private enough?	uPriv	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes
Will biometry increase safety of banking operations?	uSecu	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes
Were you afraid about registering your biometric traits in the bank biometric database?	uAfra	1 – definitely not, 2 – no, 3 – yes, 4 - definitely yes
State your gender.	uSex	1 – male, 2 – female
State your age.	uAge	1 – < 18, 2 – 18–24, 3 – 25–28, 4 – 29–36, 5 – 37–48, 6 – 49–60, 7 – 71–75, 8 – 75 <
State your education.	uEdu	1 – primary, 2 – secondary, 3 – above
State your residence type:	uAdd	1 – village (< 1000), 2 – small town (1000–5000), 3 – town (5000–20000), 4 – city (20000–30000), 5 – large city (> 300000)

* each question could also be replied with 0 = no reply

Table 2 Consultant (bank teller) questions

Question	Short names	Answers*
How long the registration process lasted in minutes?	cFast	Number of minutes
Was your assistance required during the biometric registration?	cHelp	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes
Was (signature/voice/face) registration easy for the client?	cHardPen cHardVoi cHardFac	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes (separate answer for each trait)
Was (signature/voice/face) registration by the client hard and cumbersome for you or requiring assistance?	cCumberPen cCumberVoi cCumberFac	1 – definitely not, 2 – no, 3 – yes, 4 – definitely yes (separate answer for each trait)
Free opinions on traits registration.	–	Descriptive opinion and comments on registration process
Free opinions on registration hardware.	–	Descriptive opinion and comments on hardware operation
State your gender.	–	1 – male, 2 – female
State your age.	–	$1 - < 18$, $2 - 18–24$, $3 - 25–28$, $4 - 29–36$, $5 - 37–48$, $6 - 49–60$, $7 - 71–75$, $8 - 75 <$
State your education.	–	3 – primary, 4 – vocational, 5 – secondary, 6 – above
State your residence type.	–	1 – village (< 1000), 2 – small town (1000–5000), 3 – town (5000–20000), 4 – city (20000–30000), 5 – large city (> 300000)

* each question could also be replied with 0 = no reply
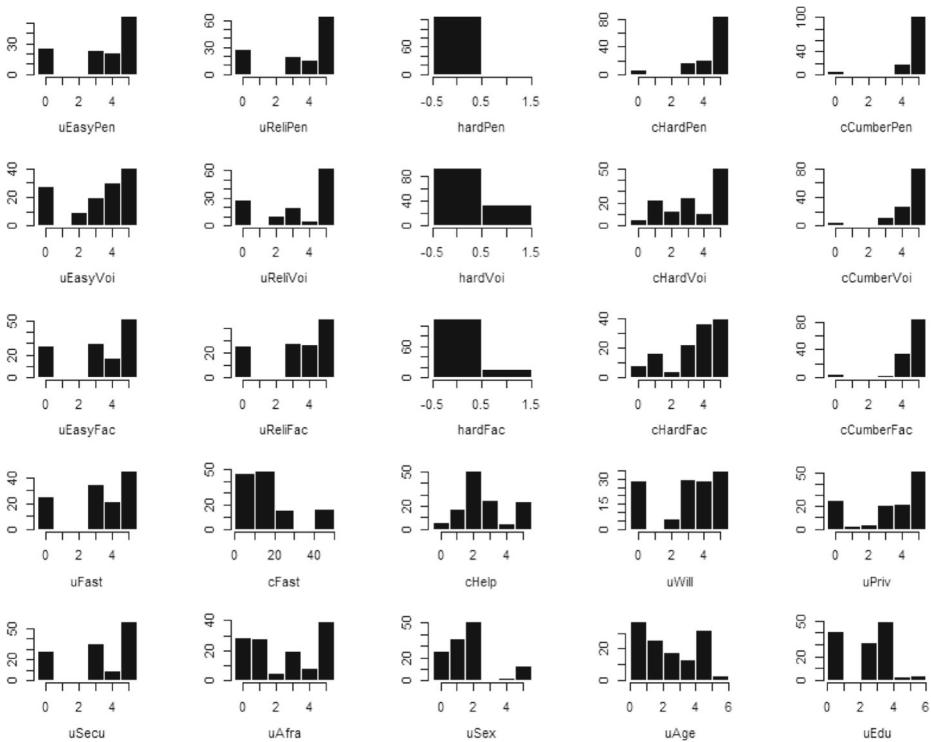
Fig. 5 Histograms of answers in user and consultant questionnaires

discipline, e.g. maintaining neutral tone not impacted by hoarseness or any speakers' health issues, hence in the tested group the stability of obtained clusters is high. The signature belongs also to behavioral biometric traits, but it requires the highest level of cooperation,
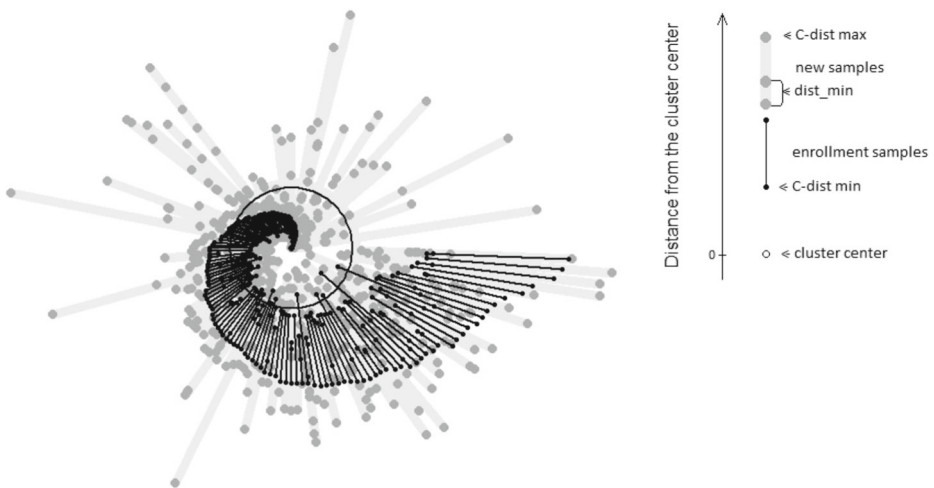
Fig. 6 Signature samples characteristics (description of the convention used on the plot is in the text on the right side of the plot)
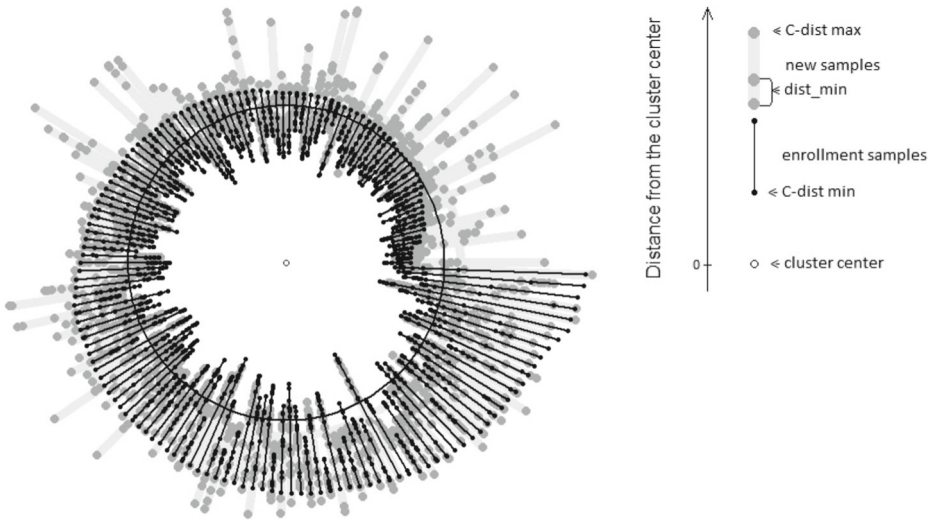
Fig. 7 Face samples characteristics (description of the convention used on the plot is in the text on the right side of the plot)

as each aspect of signing is analyzed, including pauses, rotation and acceleration changes and other. It turned out in our tests that almost 20% of identities have new signature samples further away from the clusters centers than enrollment samples distances. Moreover, for each trait of given biometric ID following 9 metrics expressing data quality were calculated:

- $trait_{dp, \min, mean, \max}$ = min, mean, and max distance among all pairs of positive samples;
- $trait_{p, \min, mean, \max}$ = min, mean, and max distance between positive samples and cluster center;
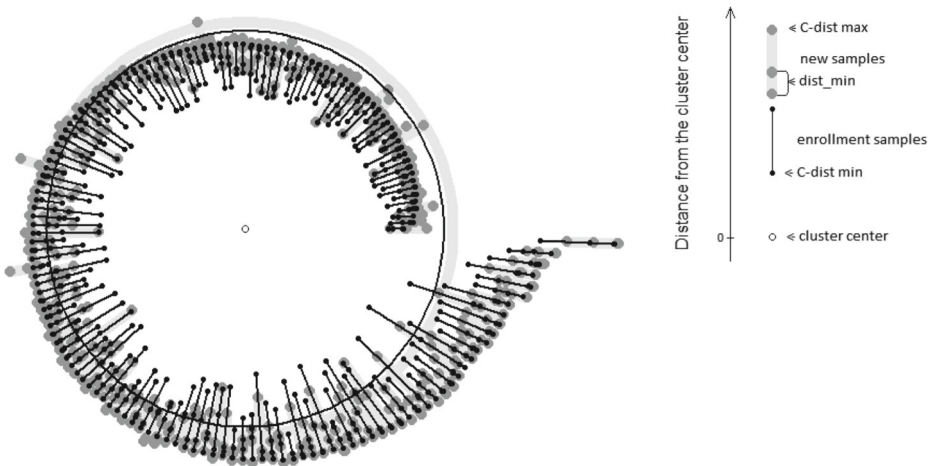
Fig. 8 Voice samples characteristics (description of the convention used on the plot is in the text above)

- $trait_{width}$ = range between traitp max and traitp min;
- $trait_{CLdrift}$ = drift of the clusters, i.e. distance between clusters centers before, and after adding evaluation samples;
- $trait_{dif\ max}$ = difference between maxima of distances in enrollment cluster and evaluation cluster.

Above values allow for an objective characterization of collected data, and for a comparison between enrollment data and validation data (Section 10).

9 Modeling knowledge on objective and subjective characteristics of biometric samples

Data processing was performed in *R* programming environment (Gardener and Beginning 2016) with RoughSets package (Riza et al. 2015). It is a mathematical calculation environment offering data importing, scripted processing, and visualization, extensible by numerous additional libraries and packages.

With regards to the rough set methodology presented above, an individual classifier is trained for each biometric trait and for each respective subjective feature related to this trait, and then applied in the following steps:

1. *Decision table* is constructed by selecting only biometric identities x_i with non-empty and non-zero answers in the questionnaire. Attributes of a particular trait $a_n(x_i)$ are extracted, and the respective subjective feature is taken as a decision d_i;
2. *Attributes $a_n(x_i)$ are discretized* by the local algorithm: the best cuts on every attribute are determined separately (discretization limits the number of possible values – for attributes in this study there are 1 to 6 cuts, splitting the values into 2 to 7 discrete ranges, accordingly);
3. *Deriving a reduct* **RED** $\subseteq$ **P** based on attributes a_n of objects x_i from the decision table;
4. *Calculating rules* by using the reduct **RED** and decisions d_i;
5. *Classifying decision table* by applying rules from previous step;
6. *The process repeats* for other number of discretization cuts.

This process was performed for each biometric trait and for each subjective opinion of clients and bank tellers on this trait to be modeled by the rough set. Results of accuracy determination of subjective opinions on each biometric trait are presented in Fig. 9, whereas the number of reducts used at each step is reported in Table 3 and some selected relations between objective features of voice biometry and subjective user experience metrics for given number of discretization cuts are illustrated in Figs. 10 and 11.

In numerous cases reducts with only 1 or 2 attributes are derived, because other attributes would not introduce further improvement in classes separation. Each classifier maintains up to $(c + 1) \cdot |\textbf{RED}|$ rules (c is the number of discretization cuts, and $|\textbf{RED}|$ is the number of attributes in the reduct). If the number of rules is high (e.g. hardFac feature for 3 cuts has 4 attributes, thus 16 rules are derived), then the modeled knowledge is sufficient to express differences between objects, and the relative size of positive region is close to 1.

The most important outcome for the practical implementation of the described procedure is the high accuracy of classification of cases such as:

- signature feature "hardPen" strongly correlates with SigLe-dif.mean and sigTotle, and it is based on these values for the user experience can be rated;
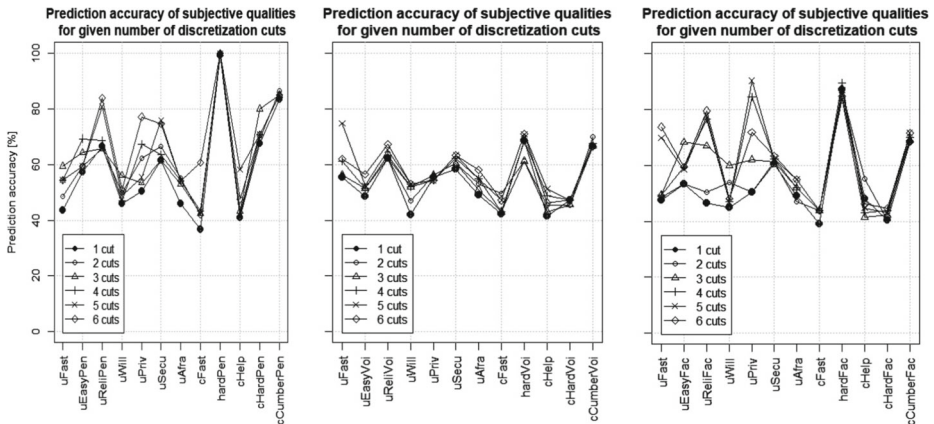
Fig. 9 Accuracy of prediction of face, signature, and voice biometry subjective features according to client and consultant opinions

- signature feature "cCumberPen" correlates with repeatability of signals, particularly with distances (differences) falling between stored sig_{dpmin}, sig_{dpmean} values;
- voice feature "hardVoi" correlates with repeatability of signals expressed in the metric voi_{dpmax}.

It should be pointed out that a new biometric identity can be handled twofold:

- it can be added to the database as a new reference. This would require an extension of the decision table with this new case, the calculation of new discretization cuts, reducts selection and rule generation over the decision set of all available identities. Such a process could result in invalidation of current models and it may entail emergence of new ones, potentially more accurate ones.
- it can be classified with regards to current models. This requires a discretization of the new identity features by the current discretization cuts, and then applying of decision rules, to determine the output features. This attempt will not change current models, but it may only utilize the knowledge extracted from previous cases.

10 Prediction of evaluation samples quality, based on enrollment samples and user subjective opinions

Following the same methodology of modeling the knowledge by creating rough sets, now the relation between other type of input data and output data is modeled, namely:

- set of real enrollment samples and user subjective opinions collected during this phase,
- objective metrics of evaluation samples quality.

Here, key quality metrics are of interest, especially observed drift between cluster of enrollment features and new cluster composed of enrollment and evaluation samples. If the large drift can be predicted beforehand, at the enrollment phase, then additional user training can be planned, to assure the best possible quality (e.g. repeatability of biometric traits), or particular trait can be ignored in the future.

Table 3 Number of reducts, accuracy and relative size of positive region for given number of used discretization cuts and for defined subjective feature

	1 cut			2 cuts			3 cuts			4 cuts			5 cuts			6 cuts		
	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.
Signature biometry subjective features modelling																		
uFast	1	43	0	1	48	0	2	59	0.14	1	54	0	1	54	0	1	54	0
uEasyPen	2	57	0	1	59	0	2	64	0.17	2	69	0.23	1	59	0	1	59	0
uReliPen	3	66	0	1	65	0	1	65	0	1	69	0	2	81	0.40	2	84	0.60
uWill	2	46	0	2	50	0.03	2	56	0.01	1	47	0	1	49	0	1	51	0
uPriv	1	50	0	3	62	0.11	1	53	0	2	67	0.17	1	55	0	2	77	0.46
uSecu	2	61	0	2	66	0.04	1	61	0	1	63	0	2	76	0.34	2	74	0.37
uAfra	1	46	0	2	55	0	1	53	0	1	54	0	1	54	0	1	54	0
cFast	1	36	0	1	42	0	1	42	0	1	43	0	1	43	0	2	61	0.29
hardPen	1	99	0.77	1	100	1	1	100	1	1	100	1	1	100	1	1	100	1
cHelp	1	41	0	1	42	0	1	43	0	1	42	0	2	58	0.17	1	47	0
cHardPen	2	67	0	1	70	0	3	80	0.56	1	71	0	1	71	0	1	71	0
cCumberPen	2	83	0.05	3	87	0.28	1	85	0	1	85	0	1	85	0	1	85	0

Table 3 (continued)

Most frequent signature features

Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts
sigPosMean	33	sigPosMax	7	sigdmin	4	sigLemax	2	SigLemean	1		
SigLedif.mean	17	sigTotle	7	sigCLdri	3	sigdpmean	2	sigCmean	1		
sigLemin	15	sigdpmin	5	sigdifmax	2	sigdmax	1				
Voice biometry subjective features modelling											
uFast	2	55	0	1	56	0	1	56	0	74	0.19
uEasyVoi	1	48	0	1	51	0	1	51	0	52	0
uReliVoi	1	62	0	1	62	0	1	62	0	65	0
uWill	1	42	0	1	47	0	1	52	0	53	0
uPriv	2	55	0	1	56	0	1	56	0	54	0
uSecu	1	58	0	1	60	0	1	60	0	63	0
uAfra	2	49	0	2	53	0.03	1	51	0	55	0
cFast	2	42	0	2	49	0.02	1	43	0	47	0
hardVoi	2	68	0	1	61	0	1	61	0	71	0.33
cHelp	2	41	0	1	45	0	1	43	0	51	0
cHardVoi	3	47	0.02	1	45	0	1	45	0	47	0
cCumberVoi	1	66	0	2	69	0.13	1	67	0	67	0

Table 3 (continued)

Feature	1 cut reduct size	accu-racy	pos. reg.	2 cuts reduct size	accu-racy	pos. reg.	3 cuts reduct size	accu-racy	pos. reg.	4 cuts reduct size	accu-racy	pos. reg.	5 cuts reduct size	accu-racy	pos. reg.	6 cuts reduct size	accu-racy	pos. reg.
Most frequent voice features	Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts	
voidpmin		33		voiPCdmin	13		voidmean	5		voidpmax	3			3		voiCmean	1	
voiCmin		17		voidmin	7		voidmax	4		voiCmax	2			2				
Face biometry subjective features modelling																		
uEasyFac	1	53	0	2	53	0	2	68	0.03	1	59	0	1	58	0	1	59	0
uReliFac	1	46	0	1	50	0	2	67	0.05	2	76	0.28	2	78	0.34	2	80	0.48
uWill	1	45	0	2	54	0	2	60	0.01	1	47	0	1	47	0	1	48	0
uPriv	1	50	0	1	50	0	2	62	0.04	3	84	0.67	3	90	0.78	2	72	0.39
uSecu	2	60	0	1	61	0	1	61	0	1	62	0	1	62	0	1	63	0
uAfra	2	49	0	1	47	0	1	52	0	1	52	0	1	55	0	1	55	0
cFast	1	39	0	1	44	0	1	44	0	1	44	0	1	44	0	1	44	0
hardFac	1	87	0	3	86	0.69	4	84	0.97	3	90	0.87	3	87	0.95	2	85	0.78
cHelp	3	48	0.03	3	55	0.13	1	41	0	1	43	0	1	45	0	1	46	0
cHardFac	2	40	0	1	40	0	1	42	0	1	44	0	1	43	0	1	46	0
cCumberFac	1	68	0	1	68	0	1	69	0	1	70	0	1	72	0	1	72	0
Most frequent face features	Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts	
facdifmax		13		facCLdri	12		facpdmean	9		facpdmin	9		facpdmin	3		facCmax	1	
facPCdmean		15		facCmin	10		facCmean	7		facpdmax	7		facpdmax	3				
facdmax		15		facdmin	9		facPCdmax	4		facPCdmin	4		facPCdmin	3				

Fig. 10 Relations between objective features of signature biometry and subjective user experience metrics for given number of discretization cuts. Circular plots left sides denote contribution of objective features, numbers indicate number of reducts including given feature. On the right side classification accuracy [%] is given
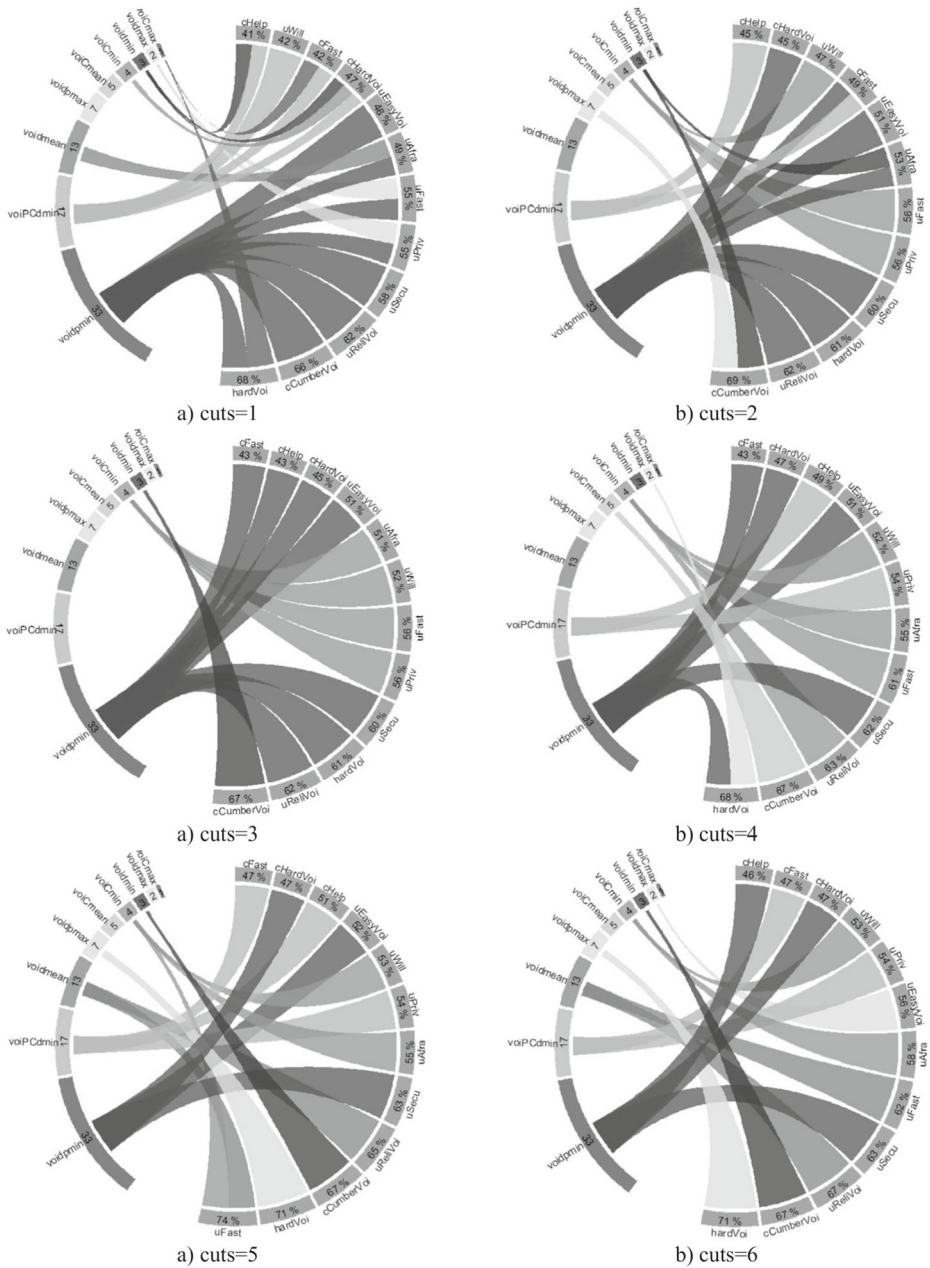
Fig. 11 Relations between objective features of voice biometry and subjective user experience metrics for given number of discretization cuts. Circular plots left sides denote contribution of objective features, numbers indicate number of reducts including given feature. On the right side classification accuracy [%] is given

Table 4 Number of reducts, accuracy and relative size of positive region for given number of used discretization cuts and for defined objective quality metric classified into 4 classes based on quartile ranges

Signature biometry objective features modelling

Feature	1 cut			2 cuts			3 cuts			4 cuts			5 cuts			6 cuts		
	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.
sigdpmin	1	100	1.0	1	100	1.0	1	100	1.0	1	100	1.0	1	100	1.0	1	100	1.0
sigdpmean	3	57	0.0	1	60	0.0	2	68	0.12	1	66	0.0	1	66	0.0	1	66	0.0
sigdpmax	4	61	0.15	2	62	0.0	3	79	0.47	2	73	0.24	1	71	0.0	1	73	0.0
sigCLdri	1	42	0.0	2	56	0.03	1	50	0.0	1	52	0.0	1	53	0.0	1	53	0.0
sigdifmax	3	79	0.23	1	75	0.0	3	82	0.48	3	89	0.60	1	78	0.0	1	78	0.20
sigPCwidth	4	72	0.07	2	74	0.0	2	74	0.01	2	75	0.24	1	74	0.0	2	79	0.36
sigPosMin	4	56	0.04	3	65	0.15	2	67	0.0	2	66	0.04	2	70	0.13	1	62	0.0
sigPosMean	3	62	0.0	3	69	0.16	3	80	0.47	3	87	0.65	2	72	0.14	1	67	0.0
sigPosMax	3	58	0.0	1	56	0.0	1	61	0.0	1	64	0.0	1	67	0.0	1	67	0.0

Most frequent signature features

Feature	1 cut		2 cuts		3 cuts		4 cuts		5 cuts	
	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature
sigdmin	11	sigLe3xmean	7	sigLedif.mean	4	sigLemax	2	uWill	1	
sigCmax	10	uReliPen	6	sigLemax	3	sigLemean	2	uSecu	1	
sigdmax	9	sigCmean	6	uAfra	3	uFast	1			
sigCmean	8	sigCmax	6	uEasyPen	2	sigLemean	1			
sigdmean	7	uPriv	4	sigCmin	2	sigTotle	1			

Voice biometry objective features modelling

Feature	1 cut			2 cuts			3 cuts			4 cuts			5 cuts			6 cuts		
	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.
voiPCdmean	5	75	0.23	2	63	0.0	2	72	0.01	2	74	0.16	2	76	0.20	2	78	0.21
voiPCdmax	5	75	0.19	4	90	0.57	2	88	0.59	2	92	0.70	2	93	0.76	2	93	0.76
voidpmin	3	61	0.04	3	89	0.35	3	94	0.79	2	93	0.56	2	94	0.70	2	95	0.75
voidpmean	4	75	0.19	1	100	1.0	1	100	1.0	1	100	1.0	1	100	1.0	1	100	1.0
voidpmax	3	65	0.10	2	69	0.04	2	72	0.10	2	67	0.29	2	71	0.30	1	58	0.21
voidifmax	1	100	1.0	2	82	0.32	2	81	0.27	3	90	0.67	2	83	0.35	3	95	0.86
voiPCwidth	4	67	0.03	3	99	0.94	3	99	0.97	3	100	1.0	1	97	0.69	1	97	0.92
voiPCdmin	2	56	0.0	3	88	0.30	2	92	0.43	2	93	0.46	1	92	0.23	1	92	0.23

Table 4 (continued)

Most frequent voice features

1 cut		2 cuts		3 cuts		4 cuts		5 cuts		6 cuts	
Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts
voiCmin	33	voidmax	13	cHardVoi	4	uEasyVoi	1	uEasyVoi	1	uAfra	1
voidmin	19	voiCmax	11	cCumberVoi	2	uSecu	1	uSecu	1		
voiCmean	14	voidmean	8	hardVoi	2	uPriv	1	uPriv	1		

Face biometry objective features modelling

Feature	1 cut			2 cuts			3 cuts			4 cuts			5 cuts			6 cuts		
	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.	reduct size	accu-racy	pos. reg.
facpdmin	1	47	0.0	1	59	0.0	1	76	0.0	1	76	0.0	1	79	0.0	1	79	0.0
facpdmean	3	59	0.05	2	67	0.01	1	68	0.0	1	69	0.0	1	69	0.0	1	69	0.0
facpdmax	2	52	0.0	2	60	0.01	1	54	0.0	1	61	0.0	1	61	0.0	1	65	0.0
facdifmax	3	62	0.02	1	63	0.22	1	63	0.22	1	69	0.22	1	69	0.22	1	70	0.22
facCLdri	3	44	0.0	1	34	0.0	2	49	0.0	1	34	0.0	1	37	0.0	1	38	0.0
facPCwidth	4	60	0.26	3	64	0.23	2	65	0.21	2	68	0.21	2	70	0.26	1	62	0.0
facPCdmin	2	58	0.0	1	63	0.0	2	71	0.12	1	68	0.0	1	68	0.0	1	69	0.0
facPCdmean	3	59	0.05	1	62	0.0	1	64	0.0	1	65	0.0	1	65	0.0	1	67	0.0
facPCdmax	3	51	0.0	2	56	0.0	2	59	0.04	1	59	0.0	2	68	0.16	1	61	0.0

Most frequent face features

1 cut		2 cuts		3 cuts		4 cuts		5 cuts		6 cuts	
Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts	Feature	No. of reducts
facCmean	20	facdmax	14	cFast	1	uFast	1	uFast	1	facCmin	1
facdmin	15	cHardFac	6	cHelp	1	uWill	1	uWill	1		
facCmax	15	uPriv	4	uSecu	1	uEasyFac	1	uEasyFac	1		

Table 5 Number of reducts, accuracy and relative size of positive region for given number of used discretization cuts and for defined objective quality metric classified into 2 classes based on mean value

Signature biometry objective features modelling

Feature	1 cut reduct size	1 cut accuracy	1 cut pos. reg.	2 cuts reduct size	2 cuts accuracy	2 cuts pos. reg.	3 cuts reduct size	3 cuts accuracy	3 cuts pos. reg.	4 cuts reduct size	4 cuts accuracy	4 cuts pos. reg.	5 cuts reduct size	5 cuts accuracy	5 cuts pos. reg.	6 cuts reduct size	6 cuts accuracy	6 cuts pos. reg.
sigdpmin	2	92	0.13	1	90	0.0	2	93	0.71	1	92	0.62	2	96	0.81	1	92	0.62
sigdpmean	3	85	0.03	2	86	0.0	1	85	0.25	2	89	0.49	3	97	0.93	2	91	0.70
sigdpmax	1	87	0.35	1	87	0.35	3	94	0.78	3	96	0.92	2	92	0.68	1	87	0.35
sigCLdri	2	69	0.0	2	72	0.0	1	73	0.0	2	77	0.22	2	81	0.32	1	73	0.0
sigdifmax	3	78	0.22	4	85	0.40	1	74	0.19	2	82	0.07	1	77	0.0	1	77	0.20
sigPCwidth	3	83	0.05	2	82	0.05	2	81	0.08	2	83	0.34	2	85	0.43	2	86	0.54
sigPosMin	1	83	0.0	1	83	0.0	2	87	0.16	1	83	0.0	2	87	0.40	2	88	0.53
sigPosMean	2	87	0.0	3	89	0.19	2	90	0.12	2	90	0.40	1	86	0.25	1	86	0.25
sigPosMax	1	84	0.0	1	84	0.0	2	85	0.24	1	84	0.0	1	84	0.0	1	86	0.0

Most frequent signature features

Feature (1 cut)	No. of reducts	Feature (2 cuts)	No. of reducts	Feature (3 cuts)	No. of reducts	Feature (4 cuts)	No. of reducts	Feature (5 cuts)	No. of reducts	Feature (6 cuts)	No. of reducts
sigdmean	24	sigdmin	8	sigCmean	3	uReliPen	3	sigmax	2	sigmax	1
sigdmax	10	SigLe3xmean	7	sigLemin	3	uWill	3	cHardPen	2	cHardPen	1
sigCmax	9	sigLemean	5	sigTotle	3	sigLemax	3				
uAfra	9	SigLedif.mean	3	uPriv	2	uEasyPen	2				

Voice biometry objective features modelling

Feature	1 cut reduct size	1 cut accuracy	1 cut pos. reg.	2 cuts reduct size	2 cuts accuracy	2 cuts pos. reg.	3 cuts reduct size	3 cuts accuracy	3 cuts pos. reg.	4 cuts reduct size	4 cuts accuracy	4 cuts pos. reg.	5 cuts reduct size	5 cuts accuracy	5 cuts pos. reg.	6 cuts reduct size	6 cuts accuracy	6 cuts pos. reg.
voidpmin	2	88	0.0	1	88	0.0	3	94	0.79	1	87	0.25	2	93	0.68	2	95	0.75
voidpmean	2	95	0.50	3	100	1.0	1	95	0.86	2	99	0.97	1	97	0.90	2	100	1.0
voidpmax	2	95	0.41	2	96	0.52	2	97	0.61	2	97	0.70	1	94	0.51	2	99	0.97
voidifmax	5	89	0.23	3	90	0.44	3	91	0.65	1	83	0.0	3	96	0.79	3	97	0.93
voiPCwidth	2	77	0.02	2	79	0.0	2	81	0.04	2	82	0.15	3	91	0.54	2	85	0.31
voiPCdmin	2	94	0.02	3	99	0.84	2	97	0.86	2	97	0.90	2	97	0.94	2	99	0.97
voiPCdmean	2	97	0.92	2	99	0.81	3	100	1.0	2	100	1.0	1	98	0.95	1	99	0.98
voiPCdmax	3	98	0.74	2	98	0.88	2	100	1.0	1	98	0.88	1	98	0.95	1	99	0.97

Table 5 (continued)

| Feature	1 cut reduct size	1 cut accu-racy	1 cut pos. reg.	2 cuts reduct size	2 cuts accu-racy	2 cuts pos. reg.	3 cuts reduct size	3 cuts accu-racy	3 cuts pos. reg.	4 cuts reduct size	4 cuts accu-racy	4 cuts pos. reg.	5 cuts reduct size	5 cuts accu-racy	5 cuts pos. reg.	6 cuts reduct size	6 cuts accu-racy	6 cuts pos. reg.
Most frequent voice features																		
Feature	No. of reducts			Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts	
voiCmin	25			voidmax	9		cFast	3		uFast	3			1				
voiCmean	21			voidmean	8		uWill	3		uPriv	3			1				
voiCmax	16			voidmin	8		uEasyVoi	2		uAfra	2			1				
Face biometry objective features modelling																		
facpdmin	2	88	0.0	2	89	0.01	1	88	0.42	1	88	0.42	2	92	0.77	3	100	1.0
facpdmean	2	90	0.0	2	92	0.40	2	92	0.67	2	93	0.78	2	94	0.71	1	90	0.27
facpdmax	2	84	0.34	2	83	0.31	2	86	0.36	3	95	0.81	2	89	0.68	2	91	0.71
facdifmax	5	81	0.14	2	80	0.23	1	80	0.0	1	80	0.0	1	81	0.0	1	81	0.0
facCLdri	4	65	0.05	2	64	0.0	1	58	0.0	1	62	0.0	2	73	0.15	1	69	0.0
facPCwidth	2	75	0.0	2	75	0.30	2	78	0.02	2	80	0.32	2	80	0.44	2	83	0.49
facPCdmin	2	88	0.03	2	88	0.06	1	88	0.0	2	91	0.53	1	88	0.0	2	92	0.71
facPCdmean	3	92	0.18	2	92	0.40	2	92	0.67	2	93	0.78	1	90	0.27	3	100	1.0
facPCdmax	2	79	0.0	3	86	0.45	2	86	0.31	2	88	0.39	2	89	0.53	2	90	0.62
Most frequent face features																		
Feature	No. of reducts			Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts		Feature	No. of reducts				
facCmean	29			facdmax	14		uAfra	3		cFast	3		cFast	1				
facCmax	24			facCmin	6		cHelp	3		uPriv	3		uPriv	1				
facdmin	20			cHardFac	3		hardFac	1										
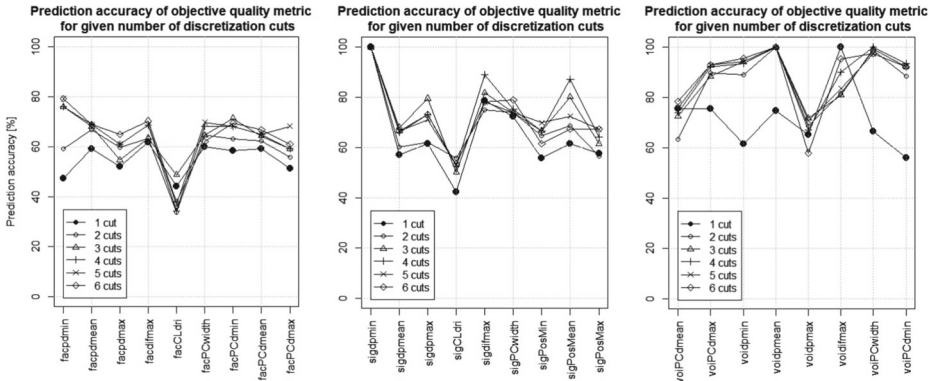
Fig. 12 Accuracy of prediction of quality metrics for face, signature, and voice biometry. Classification was performed into 4 classes based on quartile ranges

10.1 Decision discretization

At this stage, the goal is to provide the prediction for all objective metrics expressing quality of biometric samples introduced above. Those metrics are expressed as real numbers, but the rough sets are capable to perform classification of objects into discrete categories only, therefore the continuous domains of metrics values have to be discretized in advance.

For every predicted metric, a discretization is performed twofold: into four quartile ranges and into two ranges based on the mean value, defined employing the whole dataset of 126 biometric IDs (collections of samples representing all users). Other approaches are possible, e.g. based on interquartile range rule for outliers detection, but they were not used here.

10.2 Prediction results

Tables 4 and 5 contain detailed results of modeling rough set relation between input data and each quality metric. The tables show the number of features included in the reduct, type
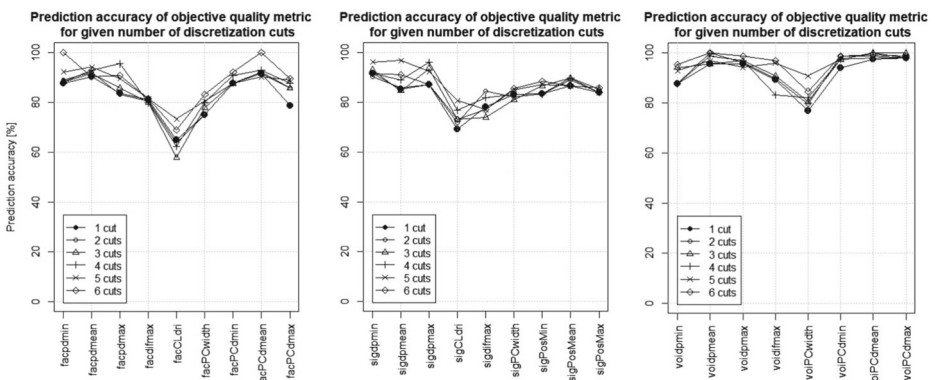
Fig. 13 Accuracy of prediction of quality metrics for face, signature, and voice biometry. Classification was performed into 2 classes based on mean value

of features, prediction accuracy and relative size of positive region. Two attempts are documented: classification into four classes based on quartiles (Table 4), and into two classes based on the mean value (Table 5).

It can be observed, that classification of metric value into two ranges (above or below the mean value) is generally more accurate than classification into respective quartile (Figs. 12 and 13). The least accurate predictions are obtained for the drift of cluster for face images (fac_{CLdri}), as it can be noticed that future images of face rely not on subjective features, and not on face features at the enrollment time only, but on other factors, such as: lighting conditions, change in make-up, and facial hair. On the other hand, the cluster drift can be predicted with relatively high accuracy (80%) only for two classes problem (higher or lower than the mean value) for models with 5 cuts.

11 Conclusions

As is seen from the results of the study, each enrolment cluster reflects quality and stability of samples during the limited period of time. Wide clusters are expected to be related to inability of the user to provide biometric sample in a repeated manner, due to e.g. biometric pen ergonomics differing from a typical pen, voice issues related to stress or noisy background conditions, or face capture issues due to unstable position in front of a camera etc.

In turn, all positive samples cluster includes enrollment samples as well as all other samples collected after arbitrary time intervals. The cluster width is related to differences between old samples and the newer ones. The cluster center is expected to drift over time to account personal characteristics changes, such as voice harshness due to aging, signature improvements being a result of repeated usage of the biometric pen, face image changes due to aging or facial hair, make-up, etc.

The cluster drift would be more prominent in case of rejection of the oldest samples. The rejection time limit definition should be determined on the basis of more prolonged studies employing the group of users, in order to enable assessing the problem of aging and any time-related changes of biometric traits.

The presented method can be also applied for all biometric features fused together and analyzed in unison. This can potentially reveal cross-modalities relations, such as both behavioral traits (speech and signature) features revealing the dependency on user emotions, or feeling of convenience, or familiarity with the technology.

References

Alize (2017). Open source recognition, University of Avignon, http://mistral.univ-avignon.fr. Accessed: 01 Oct 2017.

Banerjee, M., Mitra, S., Banka, H. (2007). Evolutionary rough feature selection in gene expression data. *IEEE Transactions on Systems, Man, and Cybernetics Part C: Applications and Reviews*, *37*(4), 622–632.

Bazan, J.G., Nguyen, H.S., Nguyen, S.H., Synak, P., Wroblewski, J. (2000). Rough set algorithms in classification problem, chapter 2. In Polkowski, L., Tsumoto, S., Lin, T.Y. (Eds.) *49–88*. Heidelberg: Physica-Verlag, https://doi.org/10.1007/978-3-7908-1840-6_3.

Bazan, J.G., Peters, J.F., Skowron, A. (2005). Behavioral pattern identification through rough set modelling. In Ślęzak, D., Yao, J., Peters, J.F., Ziarko, W., Hu, X. (Eds.) *Rough sets, fuzzy sets, data mining, and granular computing. RSFDGrC 2005. Lecture notes in computer science*, Vol. 3642. Berlin: Springer.

Bhele, S.G., & Mankar, V.H. (2015). Recognition of faces using discriminative features of LBP and HOG descriptor in varying environment. In *2015 International conference on computational intelligence and communication networks (CICN)* (pp. 426–432). Jabalpur.

Borade, S.N., Deshmukh, R.R., Ramu, S. (2016). Face recognition using fusion of PCA and LDA: Borda count approach. In *2016 24th Mediterranean conference on control and automation (MED)* (pp. 426–432). Athens. https://doi.org/10.1142/S0219467806002239.

Braga, M. (2017). Facial recognition technology is coming to Canadian airports this spring, CBC News, http://www.cbc.ca/news/technology/cbsa-canada-airports-facial-recognition-kiosk-biometrics-1.4007344. Accessed 01 Oct 2017.

Bratoszewski, P., & Czyżewski, A. (2015). Face profile view retrieval using time of flight camera image analysis. In Kryszkiewicz, M., Bandyopadhyay, S., Rybinski, H., Pal, S. (Eds.) *Pattern recognition and machine intelligence. PReMI 2015. Lecture notes in computer science*, Vol. 9124: Springer, https://doi.org/10.1007/978-3-319-19941-2_16.

Bratoszewski, P., Czyżewski, A., Hoffmann, P., Lech, M., Szczodrak, M. (2017). Pilot testing of developed multimodal biometric identity verification system. In *Proc.* signal processing, algorithms, architectures, arrangements, and applications (pp. 184 –189). Poznań, 20.9.2017–22.9.2017.

Chen, W., Hong, Q., Li, X. (2012). GMM-UBM for text-dependent speaker recognition. In *International conference on audio, language and image processing* (pp. 432–435). Shanghai.

Fujitsu Identity Management and PalmSecure (2017). https://www.fujitsu.com/au/Images/PalmSecure_Global_Solution_Catalogue.pdf. Accessed 01 Oct 2017.

Furui, S. (1982). Comparison of speaker recognition methods using statistical features and dynamic features. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *29*, 342–350.

Gardener, M., & Beginning, R. (2016). The statistical programming language. See also: https://cran.r-project.org/manuals.html. Accessed 01 Oct 2016.

Gauvain, L., & Lee, C.-H. (1994). Maximum a posteriori estimation for multivariate gaussian mixture observations of Markov chains. In *IEEE International conference on acoustics, speech, and signal processing, ICASSP* (Vol 2, pp. 291–298).

Gupta, A., & Gupta, H. (2013). Applications of MFCC and vector quantization in speaker recognition. In *2013 International conference on intelligent systems and signal processing (ISSP)* (pp. 170–173). Gujarat.

Janusz, A., & Stawicki, S. (2012). Applications of approximate reducts to the feature selection problem. *Proceedings of International Conference on Rough Sets and Knowledge Technology (RSKT)*, *6954*, 45–50.

Jiang, H. (2005). Confidence measures for speech recognition a survey. *Speech Communication*, *45*(4), 455–470. https://doi.org/10.1016/j.specom.2004.12.004.

Klontz, J.C., Klare, B.F., Klum, S., Jain, A.K., Burge, M.J. (2013). Open source biometric recognition. In *2013 IEEE Sixth international conference on biometrics: theory, applications and systems (BTAS)* (pp. 1-8). IEEE.

Larcher, A., Bonastre, J.-F., Fauve, B.G.B., Lee, K.-A., Levy, H., Li, H., Mason, J.D.D., Parfait, J.-Y. (2013). ALIZE 3.0 - open source toolkit for state-of-the-art speaker recognition. In *Proceedings of the annual conference of the international speech communication association, INTERSPEECH* (pp. 2768–2772).

Lech, M., & Czyżewski, A. (2016). A handwritten signature verification method employing a tablet. *Signal Processing, Algorithms, Architectures, Arrangements, and Applications, Poznań*, 21.9.2016–23.9.2016. https://doi.org/10.1109/SPA.2016.7763585.

Lech, M., Bratoszewski, P., Czyżewski, A. (2016). A handwriten signature verification system XXXII Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Gliwice Przeglåd Telekomunikacyjny + Wiadomości Telekomunikacyjne. https://doi.org/10.15199/59.2016.8-9.77.

Jin, J., & Zhang, L. (2014). Celebrity face image retrieval using multiple features. In *International conference on neural information processing* (pp. 119–126). Cham: Springer.

Mazumdar, D., Mitra, S., Mitra, S. (2010). Evolutionary-rough feature selection for face recognition. In Peters, J.F., Skowron, A., Słowiński, R., Lingras, P., Miao, D., Tsumoto, S. (Eds.) *Transactions on rough sets XII. Lecture Notes In Computer Science*, Vol. 6190. Berlin: Springer.

McCool, C., Marcel, S., Hadid, A., Pietikinen, M., Matjka, P. (2012). Bi-modal person recognition on a mobile phone: using mobile phone data. In *IEEE ICME Workshop on hot topics in mobile mutlimedia*. Melbourne.

Mermelstein, D. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *28*(4), 357–366.

Nguyen, S.H. (2001). On efficient handling of continuous attributes in large data bases. *Fundamenta Informatics*, *48*(1), 61–81.

Papatheodorou, T., & Rueckert, D. (2007). 3D face recognition, face recognition. In Kresimir D., and Mislav G., (eds.), InTech. https://doi.org/10.5772/4848.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences*, *11*, 341. https://doi.org/10.1007/BF01001956.

Pawlak, Z. (1991). *Rough sets theoretical aspects of reasoning about data*. Kluwer.

Riza, S.L., Janusz, A., Ślęzak, D., Cornelis, C., Herrera, F., Benitez, J.M., Bergmeir, C., Stawicki, S. (2015). RoughSets: data analysis using rough set and fuzzy rough set theories. https://github.com/janusza/RoughSets. Accessed 01 Oct 2016, https://cran.r-project.org/web/packages/RoughSets/index.html, Accessed 01 Oct 2016.

Shanker, P., & Rajagopalan, A. (2007). A.N.: off-line signature verification using DTW. *Pattern Recognition Letters*, *28*, 1407–1414.

Szczodrak, M., & Czyżewski, A. (2017). Evaluation of face detection algorithms for the bank client identity verification. *Foundations of Computing and Decision Sciences*, *42*(2), 137–148. https://doi.org/10.1515/fcds-2017-0006.

Tsumoto, S. (2002). Discovery of approximate knowledge in medical databases based on rough set model. In Lin, T.Y., Yao, Y.Y., Zadeh, L. (Eds.) *Data mining, rough sets and granular computing. Studies in fuzziness and soft computing*, Vol. 95. Heidelberg: Physica.

Zhong, N., Dong, J., Ohsuga (2001). Using rough sets with heuristics for feature selection S. *Journal of Intelligent Information Systems*, *16*, 199. https://doi.org/10.1023/A:1011219601502.