

Linguistic Descriptors in Face Recognition

Paweł Karczmarek¹ · Adam Kiersztyn¹ · Witold Pedrycz^{2,3,4} · Michał Dolecki¹

Received: 19 July 2017 / Revised: 20 February 2018 / Accepted: 16 June 2018 / Published online: 9 August 2018
© The Author(s) 2018

Abstract In this study, we propose linguistic descriptors-based approach to the problem of face identification realized by both humans and computers. This approach is motivated by an evident observation that linguistic descriptors offer an ability to formalize and exploit important pieces of knowledge describing human's face. These entities are used by people in face recognition and could be found of importance in building machine-oriented recognition schemes. Moreover, evident humans' abilities to recognize other individuals can be incorporated into computational face recognition problems as an invaluable

vehicle improving recognition rate of machine-oriented classifiers. Specifically, we propose an application of analytic hierarchy process to determine linguistic values of facial features. The experts' assessments of faces in terms of such attributes support coping with uncertainty captured through experts' decisions result in a set of useful assuring the desired property of inter-class similarities and between-class differences among faces. It is worth noting that the method presented in this study can be easily applied to any other classification problem with the presence of experts.

Keywords Linguistic descriptors · Membership functions · Information fusion · Analytic hierarchy process (AHP) · Face recognition

The authors are supported by National Science Centre, Poland (Grant No. 2014/13/D/ST6/03244). Support from the Canada Research Chair (CRC) program and Natural Sciences and Engineering Research Council is gratefully acknowledged (W. Pedrycz).

✉ Paweł Karczmarek
pawelk@kul.pl
Adam Kiersztyn
adam.kiersztyn@kul.pl
Witold Pedrycz
wpedrycz@ualberta.ca
Michał Dolecki
michal.dolecki@kul.pl

- ¹ Institute of Mathematics and Computer Science, The John Paul II Catholic University of Lublin, ul. Konstantynów 1H, 20-708 Lublin, Poland
- ² Department of Electrical and Computer Engineering, University of Alberta, Edmonton T6R 2V4 AB, Canada
- ³ Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia
- ⁴ Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

1 Introduction

Face recognition has been a challenging problem since more than two decades. The reason stems from the omnipresence of computers and a genuine need for identification of people with the use of biometric methods from which face recognition seems to be the most noninvasive. Moreover, the forensic identification methods are strongly related with biometric methods, particularly in case of face recognition. The literature of this field is rich and comprises various approaches. Most of the methods suffer from the lack of robustness to variance in pose, illumination, expression, age, occlusion, distance to the camera, and other factors. The existing methods could be augmented by bringing some crucial factors that are considered by humans when recognizing faces. With this regard, people are still better than machines, at least when recognizing familiar persons. No matter of the place of birth, race, education, and other factors coming from the living

environment, people describe the subjects in a similar manner. Of course, the factors like own race bias can skew this general opinion but for all the people any feature and its attribute such as “big nose” have the same obvious meaning. The area of computational face recognition methods is vast and covers many approaches being still developed and extended, e.g., well-known local descriptors, sparse representation, deep learning. The literature concerning the problem of human face recognition mainly focuses on two threads: biological foundations of recognition, namely detection of brain activity regions [1] or eye tracking [2], and the saliency of facial regions together with the ability to recognition of so-called familiar and unfamiliar faces [3, 4]. Analogical considerations, however from the computational point of view, were presented in [5–7] and others, where the relevance of particular facial features or facial regions (e.g., upper/lower or eye/nose/mouth area) was discussed. Those can result in determining the weights for the feature-based algorithms, where the aggregation of classifiers or the information fusion may be applied with a proper weight set. A brief survey of the papers can be found in [8]. It is worth noting that the task of face description and its parts has been discussed in numerous studies. Let us discuss some of them here. Government institutions (say, police) use the systems like Identikits, Evofit, and others [9] where the process of searching is carried out manually or automatically, and it is based on the manual (sketched) or automatic compositions of the images. It is important that finally, at the end of the process, the forensic expert has to confirm or reject the results of the search [10]. Very comprehensive and detailed guidelines on how to describe the criminal can be found in the standardizing documents [11], police websites with instructions for the witnesses of the crime like [12], and the textbooks for policemen [13]. An interesting approach to face identification and face retrieval was proposed in [14, 15]. A set of features considered in the studies was described in terms of linguistic descriptions including terms such as *small*, *medium*, *big*. Those descriptors are characterized (quantified) in terms of fuzzy sets. It is worth noting that the variants of this proposal included the emotions’ descriptions. An in-depth study of the way people describe human attributes was presented in [16]. Axiomatic fuzzy set was utilized to obtain a semantic facial descriptor in [17]. Interval value fuzzy TOPSIS technique was applied to 3D facial classification system in [18]. Recently, the results of AHP by experts were applied to a neural network classifier [19]. Moreover, the AHP was used to obtain the weights of facial features in [20]. Finally, the linguistic descriptors were measured by using experts’ votes in [21]. Other approaches based on linguistic descriptors expressed in terms of fuzzy sets, fuzzy geometries, granular computing, and others were described

in [22–28]. A comprehensive survey of methods utilizing the linguistic descriptors in face recognition can be found in [29]. Finally, it is noteworthy to add here that the AHP was used in object recognition tasks, however in a different manner than discussed in this study. The method was applied to image semantic representation in the image retrieval method [30, 31] and to face emotion recognition [32].

The main objective of this study is to propose a novel method based on linguistic descriptors which can be applied to the face recognition or face retrieval problem, particularly to the problem of criminal identification, with or without the usage of any numeric measures related with a particular face images. The process of identification becomes easy, intuitively appealing and can be conducted both by professional expert from the field of criminology or by the witness of the crime. We are interested in the use of a mechanism commonly encountered in multi-criteria decision-making theory, namely analytic hierarchy process, AHP [33–35]. The problem of facial recognition can be decomposed into two levels of hierarchy. At the higher level of the hierarchy, we form the weights related to the abstract facial features involved in the process of classification. At the lower level of the hierarchy, using again the AHP method, we transform the linguistic descriptions of the concrete facial features into the numeric variables specifying the importance of all the possible attributes related with a given feature. Additionally, our aim is to investigate how the AHP can improve the recognition rate when it is combined with some other methods based on the geometrical relationships present in the face. It is worth noting that the method presented in this study can be easily applied to any other expert-oriented task. Face recognition problem is treated here as a case study.

The paper is organized as follows. In Sect. 2, the role played by the AHP method is presented. In Sect. 3, we describe the methods of assignment of information coming from the numeric values of the features to their linguistic labels (descriptors) and the general scheme of processing. Section 4 covers the experimental results, while Sect. 5 offers conclusions and elaborates on the perspectives for the future work.

2 The Role of the Analytic Hierarchy Process

This section is devoted to a concise introduction to the AHP method as it was originally proposed in [33, 34]. Using this approach one can obtain the ranking and the priorities of any set of features or attributes under consideration. The algorithm is briefly outlined as follows. First, the hierarchy present in the problem is formed. The goal is positioned at the top, next the criteria are formulated, while

at the bottom of the hierarchy the set of alternatives is located. In our case, there are two objectives. First, we intend to produce the weights of the facial features, which can be utilized in the process of classification (whenever the weights can be applied to prioritize the classifiers). Second, we are interested in obtaining concrete degrees of membership of the facial features to the individual linguistic terms describing the set of attributes occurring in the population. For instance, we would like to know whether someone's nose is *short*, *middle*, or *long*, and to which extent it belongs to each of the classes, i.e., *short*, *middle*, or *long* noses.

At the next step of the algorithm, the expert (or a group of experts) quantifies the judgments between the elements (i.e., alternatives) of the hierarchy. These assessments are based on the pairwise comparisons of the elements. For n alternatives, the experts' responses are collected in the $n \times n$ matrix A , where n is a number of the options to be considered (in our case, facial features).

The experts generate the pairwise comparisons' results using the following scale [34, 35]: equal importance (1), weak importance (2), moderate importance (3), moderate plus (4), essential/strong importance (5), strong plus (6), very strong/demonstrated importance (7), very, very strong (8), extreme importance (9). A is called a reciprocal matrix, meaning that it satisfies the following requirements: For each element a_{ij} we have $a_{ij} = 1/a_{ji}$, $i, j = 1, \dots, n$, and $a_{ii} = 1$. Let us introduce the expression $v = (\lambda_{\max} - n)/(n - 1)$, where $\lambda_{\max} \geq n$ is a maximal eigenvalue of the reciprocal matrix A and the value $\mu = v/r$, where $r = 0, 0, 0.52, 0.89, 1.11, 1.25, 1.35, 1.40, 1.45, 1.49$ for $n = 1, \dots, 10$, respectively. These values concern the mean consistency indices of 500 randomly generated reciprocal matrices [36]. For the matrices of higher dimensionality, the methods generating the pertinent values are discussed in [37]. v , μ , and r are called inconsistency index, consistency ratio, and random inconsistency index, respectively. From a practical perspective, it is considered that the consistency ratio should not exceed the value 0.1 to assure the satisfactory level of consistency of results [35]. However, it can be difficult to obtain such level of consistency, especially when the non-numerical, intangible features are compared. The final ranking of the priorities is constructed using the values of the elements of the eigenvector of the matrix A associated with the maximal eigenvalue λ_{\max} .

As mentioned in the introduction, we use the AHP method to obtain the weights related to particular facial features and the degree of membership of specific features to the linguistic attributes. Therefore, if we can assume that we extract the most descriptive facial features, which can be relatively easily estimated by people when looking at the 2D facial image. The main idea of the process is that

the experts do the pairwise comparisons between the features answering the questions in the following form: To which extent the feature A is preferred over the feature B ? To put in a different way: To which extent one of attributes is preferred over other attributes of this feature? The algorithm results in the normalized vector $\mathbf{w} = [w_1, w_2, \dots, w_N]$ containing N weights associated with the considered features.

Let us now consider a specific face and its particular features. Each feature of this set can be described in terms of the available descriptors—this way we can produce vectors $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ which could be concatenated into a single vector presenting a description of a given face, say $\mathbf{f} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N]$. Let us assume that the whole set of such image descriptions is denoted by Ω . Our intent is to classify a new face as belonging/not belonging to one of the faces in Ω . The face is characterized by some vector \mathbf{g} . If we consider one of the possible applications, e.g., the classification process, it can be done, for instance, using the nearest neighbor rule by determining a minimal distance between \mathbf{g} and \mathbf{f} coming from Ω . To illustrate this, we can look, for example, at the feature eyebrow length. Let us assume that the set of faces were assessed by an expert and the expert's answers regarding the length of eyebrows were as follows: *short–middle*: 1–5, *short–long*: 1–9, *middle–long*: 1–7. Then, the pairwise comparison matrix A comes

in the form $A = \begin{bmatrix} 1 & 1/5 & 1/9 \\ 5 & 1 & 1/7 \\ 9 & 7 & 1 \end{bmatrix}$. Thus, its eigenvector

corresponding to the largest eigenvalue $\mathbf{f}_{\text{eyebrow length}} = [0.055, 0.173, 0.772]$. These values are related to the linguistic values *short*, *middle*, *long*. From this example, one can see that the eyebrow length is likely long rather than short or middle. Once all the features have been estimated, the face is described by the vector with entries in $[0, 1]$ being the result of the concatenation of all the N normalized eigenvectors built in the same way as for $\mathbf{f}_{\text{eyebrow length}}$. It is worth noting here, that the psychological studies suggest that people have difficulties with a numeric and linguistic estimation of the human physical attributes such as height and width [16]. Therefore, the method of pairwise comparisons can be a sound alternative here.

3 The Fusion of Information Coming from the Experts' Assessments and Numeric Measures

In the previous section, we have elaborated on a way on how to include the linguistic terms coming from the expert's opinion in the form of numerical values. The vectors obtained in this manner can be supplemented by the

numerical descriptors coming directly from the geometrical relations between the particular facial features. Starting from the linguistic description of the set of faces, we can get the degree of membership of the numeric values of measurable facial features (e.g., eyebrows or nose length) to the attributes such as short, medium, long. Intuitively appealing technique that is of interest here concerns the well-known *K*-means algorithm [38]. In this study, we present a model based on the above-mentioned *K*-means and membership functions, which can be easily extended by using other approaches.

3.1 *K*-Means and Features' Lengths Normalization

From all possible *N* facial features, some of them are measurable features, for which one can easily determine their numerical values such as a length of nose. In contrast to other non-measurable features like shape of the face, shape of the nose tip, there is a natural linear order of linguistic descriptors, say *short*, *medium*, *long*. This observation comes as the starting point for a selection of such measurable features whose specific values can serve as input data to be clustered by the *K*-means or the FCM method. Hence, the use of the classical *K*-means algorithm for a clustering of the investigated dataset into 3 groups, corresponding to the linguistic descriptions namely *small*, *average*, *big*, or *short*, *medium*, *long* depending upon each of these *M* measured features separately, arises as a sound alternative here. The clustering for each feature separately allows for a deeper analysis of the key differences between the considered faces and as a result we get *N* clusters, not

three general, multidimensional ones. It is based on data resulting from measuring the distances between the characteristic points on each of the faces in an image dataset. The location of the landmarks is visualized in Fig. 1a. For instance, forehead width can be obtained by the formula $0.5[d(P_1, P_2) + d(P_3, P_4)]$, where P_i denotes the coordinates of the *i*th point ($i = 1, \dots, 55$) shown in Fig. 1a. It should be emphasized that the measurable features discussed in this study serve as illustrative examples and do not exhaust the set of all measurable features. For example, for a feature “length of eyebrows” cluster centers corresponding to linguistic descriptors *short*, *medium*, *long* have been designed. Next, for each person, the degree of membership to each cluster is determined. These clusters are describing the possible values of feature “length of eyebrows”.

After determining the above-mentioned lengths of *M* measurable features $\alpha_i^k, i = 1, \dots, M, k = 1, \dots, m$ (*m* is a number of considered faces), the results are normalized. The distances are scaled by setting the same distance between the pupils for all the faces, namely $\alpha_i^{k*} = const \alpha_i^k$, where *const* is a scaling coefficient. In the sequel, they are clustered with the use of the *K*-means algorithm. Normalized in this manner, the length of a particular feature is the starting point for testing the degree of membership of every person to a cluster determined for each feature separately. After applying the *K*-means method, we receive a set of centers of clusters. More specifically, for each of the measurable facial parts, we receive three numerical values corresponding to the linguistic descriptors such as *small*, *average*, *big*. Denote such descriptions by $c_i, i = 1, \dots, M$. Taking into account these vectors, we can

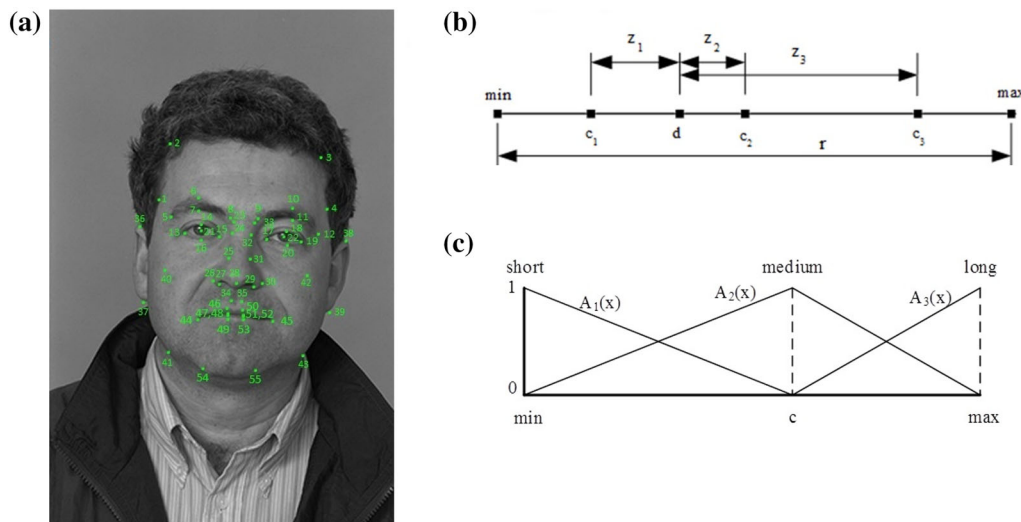


Fig. 1 a Location of selected face landmarks; the face coming from the FERET dataset [39]; b The relations between the cluster centers and a specific feature length. Note that the closer is the position of the numeric (measured in pixels) length *d* to a particular cluster center

then the greater is the value of z^* , see the description in the text; c Examples of membership functions of linguistic terms *short*, *medium*, and *long*

determine the degrees of membership to the respective centers in a following manner. Note that the degree of membership to the cluster should be determined by the distance between the numeric value of a concrete feature and the center of the cluster.

Assume that the vector $\mathbf{d}_k, k = 1, \dots, m$, contains the values of the measurable features for k th person. Based on the values of the K -means algorithm for each of the features, the distances to the centers \mathbf{c}_i are calculated. In other words, for each person $k = 1, \dots, m$ the elements of new vectors $\mathbf{z}_i^k, i = 1, \dots, M$, are given by $z_{i,j}^k = |d_{k,i} - c_{i,j}|$, where $j = 1, 2, 3$ is an index corresponding to *small*, *average*, and *long*, respectively. More precisely, if a specific eyebrows length is $d = 35$ and for this feature the centers vector \mathbf{c} has the form [25, 37, 47], then the vector containing distances between the numerical value and the centers of clusters is $\mathbf{z} = [10, 2, 12]$. In the next step, the vectors $\mathbf{z}_i^k, i = 1, \dots, M, k = 1, \dots, m$, are standardized as $z_{i,j}^{k*} = (r_i - z_{i,j}^k) / z_{i,j}^k$, where $r_i = \max_{1 \leq k \leq m} d_{k,i} - \min_{1 \leq k \leq m} d_{k,i}$ and $d_{k,i}$ denotes a dispersion of i th measurable feature. This procedure is illustrated in Fig. 1b. In our example, if this spread for the feature eyebrows length is 48, then $\mathbf{z}^* = [3.8, 23, 3]$. The final result is a set of vectors \mathbf{z}_i^{k*} which are normalized to the length of 1. These vectors are containing the degrees of memberships to particular clusters corresponding to linguistic values.

3.2 Formation of Membership Functions

We consider triangular membership functions as sound way of quantifying individual linguistic variables such as *short*, *medium*, and *long*, by introducing the following parameterized membership functions $A_1(x) =$

$$\begin{cases} 1, & x \leq x_{\min}, \\ \frac{c-x}{c-x_{\min}}, & x_{\min} < x \leq c, \\ 0, & c < x, \end{cases} \quad A_2(x) =$$

$$\begin{cases} \frac{x-x_{\min}}{c-x_{\min}}, & x_{\min} < x \leq c, \\ \frac{x_{\max}-x}{x_{\max}-c}, & c < x \leq x_{\max}, \\ 0, & \text{otherwise} \end{cases} \quad A_3(x) =$$

$$\begin{cases} 0, & x \leq c, \\ \frac{x-c}{x_{\max}-c}, & c < x \leq x_{\max}, \\ 1, & x_{\max} \leq x, \end{cases} \quad \text{where } x_{\min} = \min_{1 \leq k \leq m} x_k \text{ and }$$

$x_{\max} = \max_{1 \leq k \leq m} x_k$. As shown in Fig. 1c, the description of each feature is realized by means of three overlapping membership functions. The overlap is set to 1/2. These fuzzy sets exhibit a certain level of flexibility as the modal

value c of A_2 can be adjusted. Given the values of the vectors \mathbf{f}_j , these membership grades are used to adjust the value of c . For each of the M measurable features, we minimize the following sum $\sum_{j=1}^n \sum_{k=1}^m \sum_{l=1}^3 (A_l(x_k) - f_{k,l}^{(j)})^2$ where n is a number of experts (precisely, it is the number of independent AHP processes run for this feature), m is a number of examined face images, while $f_{k,l}^{(j)}$ are the elements of vectors assigned to the k th face, and x_k is a value of the feature's length. It is worth noting that other types of membership functions such as Gaussian ones could be considered here.

3.3 Classification Process

The main flow of the proposed classification process is outlined in Fig. 2. A group of experts describes a certain face by quantifying its features with the use of the AHP method. The results of assessments of the individual features are concatenated into the vectors representing activation levels of face descriptors forming a certain linguistic space. The vectors can be easily averaged using, for instance, arithmetic mean. In parallel, the original measures of facial features are hold in the form of the vectors containing the memberships to the linguistic values *short*, *medium*, *long*. The vectors formed on a basis of the linguistic terms and the numeric values of the measurable facial features are used to carry out classification. An intuitively appealing alternative is the nearest neighbor (NN) classification algorithm in which a weighted Euclidean distance is considered: $d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sqrt{w_i} (x_i - y_i)^2$, where $\mathbf{x} = [x_1, \dots, x_n], \mathbf{y} = [y_1, \dots, y_n]$, and $\mathbf{w} = [w_1, \dots, w_n]$. Considering that p experts took part in the process of pairwise comparisons of N features, one obtains p vectors of weights, i.e., $\mathbf{w}_1, \dots, \mathbf{w}_p$ describing the importance of facial cues, namely $\mathbf{w}_i = [w_{i,1}, \dots, w_{i,N}], i = 1, \dots, p$. Similarly, as in the case of concatenated vectors corresponding to particular facial features, we can reform the weight vectors and build the element vectors $\mathbf{v}_i = [v_{i,1}, \dots, v_{i,Q}], i = 1, \dots, p$. Here, Q is the number being the sum of all the linguistic values corresponding to the N facial features. By stretching we mean that the weight corresponding to the feature from above example, i.e., the eyebrow length, is associated with elements of the vector \mathbf{v}_i (corresponding to the three potential linguistic values *short*, *medium*, *long*). Finally, the weights are averaged, i.e., $v = (v_1 + \dots + v_p) / p$.

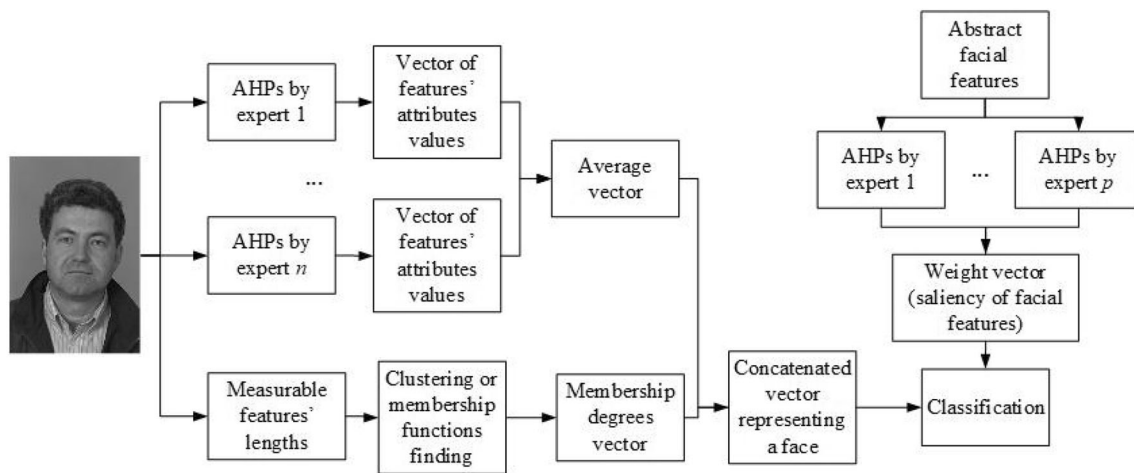


Fig. 2 An overall flow of classification

4 Experimental Studies

The experimental study is completed for the well-known FERET dataset [39]. We consider the first 50 images, called *ba* and the first 50 images coming from the subset referred to as *bk*. The first group of images stands for set *A* (we treat it as the training set), while the second one forms set *B* (testing set). Three experts (our laboratory members or friends) were asked to describe the images belonging to the set *A*, and 3 experts were asked to describe the images belonging to the set *B*. Two of them filled the questionnaires regarding both sets. More precisely, each of them played a role of a crime witness describing a facial image using the pairwise comparisons described above. The experts completed the questionnaires using some specially prepared forms to make their task easier. In this manner, we got 3 questionnaires for the set *A* and 3 questionnaires for the set *B*. For the needs of experiments, the questionnaires were formed as special tables prepared in spreadsheet program where the questions were given in the form presented in Sect. 2. Moreover, we asked an experienced expert from the field of face recognition (being our laboratory member) to describe $N = 52$ of the most descriptive, in our opinion, facial features. These facial features come with linguistic descriptors. The cues selected in this way along with their descriptors are intuitive, easy to identify and compare by experts in the fields of forensics, cognitive psychology, etc., or, what is probably the most desired in the context of application of the proposed method, for potential witnesses of the crime which have to describe a criminal to be identified. In the process of running the AHP, expert survey results concerning the estimation of interrelations between 52 facial features, were utilized. The features were chosen with use of the standards described in [11, 12, 40]. The non-measurable features were: gender, origin, shape of the face, hair length, hair texture, hairstyle,

forehead shape, forehead skin, forehead profiling, eyebrows direction, eyebrows shape, shape of the lower eyelid, direction of the fissures, placement of eyes shallow, shape of the nasal bridge, shape of the nasal tip, ears protrusion, ears shape, size of the earlobe, position of the earlobe, earlobe shape, cheeks fullness, shape of the opening between lips, mouth fullness, depth of the philtrum, chin shape, chin prominence, chin details. The measurable features were: forehead width, forehead height, eyebrows length, distance between the eyebrows, eyebrows position, eyebrows thickness, distance between eyelids, fissures length, inter-eye distance, nose length, nose width, width of the nasal bridge, height of the nasal bridge, nostrils, size of the nose holes, ears length, length of the cheeks, width of the cheeks bones, mouth width, upper lip height, lower lip height, width of the philtrum, chin size.

The results of our experiments are collected in Table 1. The application of AHP has been found a useful tool. Particularly, when more than one expert takes part in the estimation process, we may anticipate reaching a good recognition rate level (more than 90%). The results show that the application of the methods assessing the degrees of memberships to the corresponding linguistic values represented by membership functions or clusters can strongly supply a description process realized by the experts. Information fusion coming from the experts' linguistic descriptions and the measurements of the features' lengths in the form of concatenated vector improves the accuracy of the classification algorithm. For instance, if the face images are described by a single expert, in 94% of situations the subject has been correctly identified. The efficiency of the algorithm improves when more experts are involved. The participation of two experts in face evaluation arises as a good and relatively inexpensive alternative. When the *K*-means is used to build an augmented feature vector, good results are reported when only two experts are

Table 1 Average recognition rates

No. of experts in sets A–B	1–1	2–1	2–2	3–1	3–2	3–3
AHP + weights	74	90	96.67	96	100	100
	72.89	91.33	96.44	96	100	100
AHP + <i>K</i> -means + weights	94.22	97.33	99.33	98.67	100	100
	94.67	97.11	99.11	99.33	100	100
AHP + PSO + weights	75.33	90	96.44	96.67	100	100
	74.44	90.44	95.78	96	100	100
AHP + PSO + <i>K</i> -means + weights	94.44	97.56	99.33	99.33	100	100
	94.67	97.78	99.11	99.33	100	100
AHP + triang. mf. + weights	87.56	93.11	96.22	95.33	98	100
	87.56	93.56	96.67	97.33	99.33	100
AHP + <i>K</i> -means + triang. mf. + weights	94	97.78	99.33	98.67	100	98
	94.22	96.67	98.44	98.67	100	100
AHP + normalized lengths + weights	77.33	91.78	96.89	96.67	100	100
	78.44	92.44	96.67	96.67	100	100
Other ling. descr.-based methods						
AHP + neural network [19]	96		95.3			98
Voting [20]	38.94	48.13	62.75	53.79	69.41	78.18
Fuzzy sets [21]	20.8	29.3	43.6	35.3	53.5	67.2
Methods with no experts						
	Rec. rate		+weights			
Lengths	66		68			
<i>K</i> -means	56		54			
Triang. mf.	58		64			
LBP(6 × 6), MBLBP (7,6 × 6)	88		–			
MBLBP (3,6 × 6), MBLBP (5,6 × 6)	84		–			

The even lines of the results present the results obtained with weights associated to the saliency of each of the facial features

involved in the description of the training set and one expert is answering the questions regarding the testing images (recognition rate over 97%). The use of the weights produced by the AHP process completed in the linguistic facial space results in the improvement in the performance at the level of 6%. Finally, it is worth to note that the optimization of the experts' answers regarding both concrete descriptions of specific facial features and the weights assigned to the considered cues by well-known particle swarm optimization algorithm (PSO) [41] with the termination criterion that the inconsistency index should not exceed the 0.1 value leads to slight improvement in the results up to 1.5% recognition rate. These special results are denoted in Table 1 by AHP + PSO and AHP + PSO + *K*-means. In our case, we set the number *n* of particles in the PSO algorithm to be 30 and the number of generations as 200. To compare our proposal based on linguistic descriptors with other algorithms, we have

chosen the method based on so-called local descriptors. The last four rows shown in Table 1 contain the recognition rates obtained for well-known local descriptors, namely local binary patterns (LBP) [42] and multi-scale block local binary pattern (MBLBP) [43]. They are obtained using the same FERET subset when considering the following setup: Each image was divided into $n \times n$ rectangles of equal size. The best results were obtained for $n = 6$. In the case of MBLBP, the blocks of pixels were built from 3, 5, and 7 pixels, respectively. Note that in most cases our method outperforms the machine-based feature extraction approaches such as LBP and MBLBP. Moreover, we present 3 methods related to linguistic descriptors, namely AHP with no information about distances [19], voting on the chosen feature lengths [20], and fuzzy sets based on the weights obtained directly from the users [21].

5 Conclusions and Future Studies

In this study, we have presented a novel application of the analytic hierarchy process regarded here as a useful vehicle to develop linguistic descriptors of facial features obtained from the experts' evaluations of the faces. This approach produced very good results, particularly when it was fused with a standard classifier. The average recognition rates varying from 94 to 100% show the efficiency of the method in applications, where the presence of experts becomes necessary and very important, e.g., in forensics. Moreover, we have introduced the method of determining the weights of facial cues, which significantly improved the accuracy of classification process. Furthermore, the method can be easily improved by invoking optimization methods. Here, as an example, the application of a well-known PSO has led to improved results. Future work may focus on automation of the process, applying other weights and various aggregation techniques, i.e., the aggregation of the corresponding elements of reciprocal matrices or the aggregation of the AHP results, assessing the weights for specific elements of matrices, deepened investigation of the manner of calculating such weights, and an application of the method to other fields of decision making, where the experts play a significant role. Moreover, it is worth to examine other forms of classifiers like SVM or fuzzy Sugeno classifier, see, e.g., [44].

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Haxby, J.V., Ungerleider, L.G., Horwitz, B., Maisog, J.M., Rapoport, S.I., Grady, C.L.: Face encoding and recognition in the human brain. *Proc. NAS USA* **93**, 922–927 (1996)
- Duchowski, A.J.: *Eye Tracking Methodology. Theory and Practice*. Springer, London (2007)
- Johnston, R.A., Edmonds, A.J.: Familiar and unfamiliar face recognition: a review. *Memory* **17**, 577–596 (2009)
- Karczmarek, P., Pedrycz, W., Kiersztyn, A., Rutka, P.: A study in facial features saliency in face recognition: an analytic hierarchy process approach. *Soft. Comput.* **21**, 7503–7517 (2017)
- Ekenel, H.K., Stiefelhagen, R.: Generic versus salient region-based partitioning for local appearance face recognition. *Adv. Biom. LNCS* **5558**, 367–375 (2009)
- Heisele, B., Blanz, V.: Morphable models for training a component-based face recognition system. In: Zhao, W., Chelappa, R. (eds.) *Face Processing, Advanced Modeling and Methods*, pp. 439–462. Elsevier (2005)
- Kwak, K.-C., Pedrycz, W.: Face recognition: a study in information fusion using fuzzy integral. *Pattern Recognit. Lett.* **26**, 719–733 (2005)
- Karczmarek, P., Pedrycz, W., Reformat, M., Akhouni, E.: A study in facial regions saliency: a fuzzy measure approach. *Soft. Comput.* **18**, 379–391 (2014)
- Frowd, C.D., Hancock, P.J.B., Carson, D.: EvoFIT: a holistic, evolutionary facial imaging technique for creating composites. *ACM Trans. Appl. Percept.* **1**, 19–39 (2004)
- Spaun, N.A.: Face recognition in forensic science. In: Li, S.Z., Jain, A.K. (eds.) *Handbook of Face Recognition*, pp. 655–670. Springer, London (2001)
- FISWG Facial Identification Scientific Working Group: Facial image comparison feature list for morphological analysis, Version 1.0. https://fiswg.org/FISWG_1to1_Checklist_v1.0_2013_11_22.pdf (2014) Accessed 15 July 2017
- Chicago Police Department: How to describe a suspect. <https://portal.chicagopolice.org/portal/page/portal/ClearPath/Get%20Involved/Hotlines%20and%20CPD%20Contacts/How%20to%20Describe%20a%20Suspect> (2013) Accessed 15 July 2017
- Lindsay, R.C.L., Ross, D.F., Read, J.D., Togli, M.P.: *The Handbook of Eyewitness Psychology: Volume II: Memory for People*. Psychology Press, Mahwah (2007)
- Fukushima, S., Ralescu, A.L.: Improved retrieval in a fuzzy database from adjusted user input. *J. Intell. Inf. Syst.* **5**, 249–274 (1995)
- Nakayama M., Miyajima K., Iwamoto H., Norita T.: Interactive human face retrieval system based on linguistic expression. In: *Proceedings of IIZUKA'92*, vol. 2, pp. 683–686 (1992)
- LaVergne, D., Tiferes, J., Jenkins, M., Gross, G., Bisantz, A.: Linguistic estimations of human attributes. In: *Proceedings of HFESAM'16*, vol. 60, pp. 318–322 (2016)
- Ren, Y., Li, Q., Liu, W., Li, L.: Semantic facial descriptor extraction via axiomatic fuzzy set. *Neurocomputing* **171**, 1462–1474 (2016)
- Ramalingam, S., Maheswari, U.: A fuzzy interval valued fusion technique for multi-modal 3D face recognition. In: *2016 IEEE International Carnahan Conference on Security Technology (ICCST)*, pp. 1–8 (2016)
- Dolecki, M., Karczmarek, P., Kiersztyn, A., Pedrycz, W.: Face recognition by humans performed on basis of linguistic descriptors and neural networks. In: *Proceedings of 2016 International Joint Conference on Neural Networks (IJCNN 2016)*, pp. 5135–5140 (2016)
- Karczmarek, P., Kiersztyn, A., Pedrycz, W., Dolecki, M.: Linguistic descriptors and analytic hierarchy process in face recognition realized by humans. In: Rutkowski L. et al. (eds.), *Artificial Intelligence and Soft Computing, 15th International Conference, ICAISC, Part I, LNAI 9692*, pp. 584–596 (2016)
- Kiersztyn A., Karczmarek P., Dolecki M., Pedrycz W.: Linguistic descriptors and fuzzy sets in face recognition realized by humans. In: *Proc. 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1120–1126 (2016)
- Al-Hmouz, R., Pedrycz, W., Daqrouq, K., Morfeq, A.: Development of multimodal biometric systems with three-way and fuzzy set-based decision mechanisms. *Int. J. Fuzz. Syst.* <https://doi.org/10.1007/s40815-017-0299-9> (2017)
- Benhidour, H., Onisawa, T.: Interactive face generation from verbal description using conceptual fuzzy sets. *J. Multimed.* **3**, 52–59 (2008)
- Kumar, N., Berg, A.C., Belhumeur, P.N., Nayar, S.K.: Describable visual attributes for face verification and image search. *IEEE Trans. PAMI* **33**, 1962–1977 (2011)
- Kurach, D., Rutkowska, D., Rakus-Andersson, E.: Face classification based on linguistic description of facial features. In:

- Rutkowski, L. et al. (eds.) *Artificial Intelligence and Soft Computing*, Part II, LNAI 8468, pp. 155–166 (2014)
26. Lee, S.Y., Ham, Y.K., Park, R.H.: Recognition of human front faces using knowledge-based feature extraction and neuro-fuzzy algorithm. *Pattern Recognit.* **29**, 1863–1876 (1998)
 27. Martínez, G.E., Mendoza, O., Castro, J.R., Rodríguez-Díaz, A., Melin, P., Castillo, O.: Comparison between Choquet and Sugeno integrals as aggregation operators for pattern recognition. *NAFIPS* **2016**, 1–6 (2016)
 28. Rahman, A., Sufyan Beg, M.M.: Face sketch recognition using sketching with words. *Int. J. Mach. Learn. Cybern.* **6**, 597–605 (2015)
 29. Karczmarek, P., Kiersztyn, A., Rutka, P., Pedrycz, W.: Linguistic descriptors in face recognition: a literature survey and the perspectives of future development. *Proc. SPA* **2015**, 98–103 (2015)
 30. Cheng, S.C., Chou, T.C., Yang, C.L., Chang, H.Y.: A semantic learning for content-based image retrieval using analytical hierarchy process. *Expert Syst. Appl.* **28**, 495–505 (2005)
 31. Chou, T.C., Cheng, S.C.: Design and implementation of a semantic image classification and retrieval of organizational memory information systems using analytical hierarchy process. *Omega* **34**, 125–134 (2006)
 32. Cheng, S.C., Chen, M.Y., Chang, H.Y., Chou, T.C.: Semantic-based facial expression recognition using analytical hierarchy process. *Expert Syst. Appl.* **33**, 86–95 (2007)
 33. Saaty, T.L.: *The Analytic Hierarchy Process*. McGraw-Hill, New York (1980)
 34. Saaty T.L.: What is the analytic hierarchy process?. In: Mitra, G. (ed.) *Mathematical Models for Decision Support*. NATO ASI Series, vol. F48, pp. 109–121 (1988)
 35. Saaty, T.L., Vargas, L.G.: *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*. Springer, New York (2012)
 36. Saaty, T.L., Mariano, R.S.: Rationing energy to industries: priorities and inputoutput dependence. *Energy Syst. Policy* **3**, 85–111 (1979)
 37. Saaty, T.L.: *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*. AHP Series, vol. 6. RWS Publications, Pittsburgh (2000)
 38. Hartigan, J.A., Wong, M.A.: A k -means clustering algorithm. *J. R. Stat. Soc. Ser. C* **28**, 100–108 (1979)
 39. Phillips, P.J., Wechsler, J., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. *Image Vis. Comput.* **16**, 295–306 (1998)
 40. Czerw, Z.: Human identification using appearance. In: Kedzierski, W. (ed.) *Forensic Technique*, vol. II, pp. 141–171. WSPol, Szcztyno (1995)
 41. Kennedy, J.F., Eberhart, R.C., Shi, Y.: *Swarm Intelligence*. Academic Press, San Diego (2001)
 42. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: *Proceedings of the 8th European Conference on Computer Vision*, LNCS, 3021, pp. 469–481 (2004)
 43. Chan, C.-H., Kittler, J., Messer, K.: Multi-scale local binary pattern histograms for face recognition. In: *ICB 2007*, LNCS 4642, pp. 809–818 (2007)
 44. Acharya, U.R., Molinari, F., Sree, S.V., Chattopadhyay, S., Nge, K.-H., Suri, J.S.: Automated diagnosis of epileptic EEG using entropies. *Biomed. Signal Process. Control* **7**, 401–408 (2012)

Paweł Karczmarek received his Ph.D. degree in mathematics from the University of Gdańsk, Poland (2010). He is currently an assistant professor at Institute of Mathematics and Computer Science of The John Paul II Catholic University of Lublin. He was a postdoctoral fellow in the Department of Electrical and Computer Engineering of the University of Alberta (2011–2012). His current research areas are pattern recognition, computer vision, granular computing, computational intelligence, multi-criteria decision-making theory, and computational mathematics.

Adam Kiersztyn received Ph.D. degree in mathematics from Faculty of Mathematics, Physics and Computer Science, Maria Curie-Skłodowska University, Lublin, Poland. His current research interests include fuzzy measures, face recognition, quantitative methods, time-series analysis, construction of predictive models, and decision-making theory. Moreover, he is interested in the application of mathematics in other areas of science like biology or economy.

Witold Pedrycz is Professor and Canada Research Chair in the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He is also with the Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland. His main research directions involve computational intelligence, fuzzy modeling and granular computing, knowledge discovery and data mining, pattern recognition, knowledge-based neural networks, relational computing, and software engineering.

Michał Dolecki is an assistant professor at Institute of Mathematics and Computer Science of The John Paul II Catholic University of Lublin. He received Ph.D. degree in Computer Science from Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, Poland. His research interest can be described as biologically inspired calculations especially artificial neural networks and genetic algorithms. Additionally, he is interested in cryptography and data protection.