A Holistic Automatic Method for Grouping Facial Features Based on their Appearance

Félix Fuentes-Hurtado, Jose Antonio Diego-Mas, Valery Naranjo, Mariano Alcañiz
Instituto de Investigación e Innovación en Bioingeniería, I3B
Universitat Politècnica de València
Camino de Vera s/n, 46022 Valencia, Spain.
Email: ffuentes@upv.es

***Abstract*—Classification or typology systems used to categorize different human body parts exist for many years. Nevertheless, there are very few taxonomies of facial features. A reason for this might be that classifying isolated facial features is difficult for human observers. Previous works reported low inter-observer and intra-observer agreement in the evaluation of facial features. Therefore, this work presents a computer-based procedure to classify facial features based on their global appearance automatically. First, facial features are located, extracted and aligned using a facial landmark detector. Then, images are characterized using the eigenpictures approach. We then perform a clustering of each type of facial feature using as input the weights extracted from the eigenpictures approach. Finally, we validate the obtained clusterings with humans. This procedure deals with the difficulties associated with classifying features using judgments from human observers and facilitates the development of taxonomies of facial features. Taxonomies obtained with this procedure are presented for eyes, noses and mouths.**

I. Introduction

For centuries, artists and researchers have tried to develop procedures to measure and classify human faces. Anthropometric facial analysis is used in different fields like surgery, forensic science, art, face recognition, emotion recognition and facial trait judgments [1]. In the last decades, new technologies have opened ways to automatically evaluate facial features and gestures, and computational methods for analysis of facial information are now applied to classify faces based on anthropometric or emotional criteria [2].

Classification or typology systems used to categorize different human body parts exist for many years. In 1940, William Sheldon developed somatotypes to describe the body build of an individual. Sheldon proposed a classification system in which all possible body types were characterized based on the degree to which they matched these somatotypes [3]. Taxonomies, as classification system, allow using a common terminology to define body part configurations providing a standardized way to describe them, and are widely used in many fields such as ergonomics and bio-mechanics, criminalistics, sports, medicine, design or apparel industry. In general, this kind of typology systems is intended for qualitative categorization based on the global appearance of body parts, although, in some cases, a quantitative analysis of some selected features is developed to obtain the classification.

In the case of facial features, taxonomies are useful, for example, in ergonomics, forensic, anthropology or crime prevention. New human-machine interaction systems and online activities like e-commerce, e-learning, games, dating or social networks, are fields in which facial features classifications are needed. In these activities it is common to use human digital representations that symbolize the users presence or that act as virtual interlocutor [4]. In this context, it is common to synthesize faces and facial expressions combining facial features [5], [6].

Several taxonomies of facial features can be found in the literature. For example, Vanezis's atlas [7] classifies 23 facial features, the Disaster Victim Identification Form (DVI) by Interpol categorizes 6, and the DVM database [8] 45 facial features. In Tamir [9], different shapes of the human nose are classified into 14 groups based on the analysis of 1793 pictures of noses. A similar approach was used for classifying human chin [10]. In these works, a big set of photographs were analyzed and classified based on the similarity of the features.

This approach, while intuitively logical, has several problems not only in the development of taxonomies, but also in its later use. The classification of facial features is obtained from the opinion of a limited group of human observers. Classic behavioral work has shown that humans brain integrates facial features into a gestalt whole when it processes face information (holistic face processing, [11]), decreasing our ability for processing individual features or parts of faces [12]. This part-whole effect makes difficult, for example, to recognize familiar faces from isolated features [13]. Moreover, there is a low inter-observer and intra-observer agreement in the evaluation of facial features [14]. Finally, creating this kind of taxonomies implies classifying a very big set of elements (the number of possible different features) in an undefined number of groups, and this kind of tasks easily overcomes our capacities for information processing [15], [16]. To deal with these problems, we propose a new procedure to develop facial trait taxonomies based on its appearance using computational methods for automatically classifying features.

Recently, analysis of facial images has become a major research topic, and new computational methods for analysis of facial information have been developed. A comparison of these techniques shows two different approaches to deal with facial information [1]. The first one (structural approach) automatically encodes the geometry of faces using several significant points and relationships between them, doing a metric or morphological assessment of facial features. Examples of this kind of techniques are those based on SIFT feature descriptors

ISBN 978-90-827970-1-5 © EURASIP 2018

[17], point distribution models [18] or local binary patterns [19]. On the other hand, the holistic approach uses appearance-based representations, considering all available information and encompassing the global nature of the faces. Holistic techniques include, for example, Fisherfaces [20] or Eigenfaces [21]. Some work in facial features characterization has been done mixing structural and holistic techniques [22].

A. Our method

Classification methods of facial traits are needed in order to develop taxonomies. Research using computational methods is usually focused on the characterization of complete faces [23], [24]. However, less efforts have been done in facial trait classification based on its appearance. The objective of this work is to develop an appearance-based method to obtain a relatively low-dimensional vector of characteristics for facial traits, although its use for classifying any other kind of image by simmilarity is also possible. On this basis, large sets of three facial traits (eyes, noses and mouths) were characterized. Using this characterization, the traits were clustered obtaining new taxonomies for each facial feature independenttly. The procedure followed avoids the problems related to human limitations in classifying facial traits. On the one hand, the characterization and clustering of the traits were not based in human judgments. On the other hand, classifying new traits in one of the groups of the taxonomies becomes trivial and automatic. Finally, the procedure was tested comparing human opinions with automatically generated groups of traits.

II. Materials and Methods

This section details the methods followed for extracting the features from the faces and grouping them by appearance.

A. Facial feature database creation

The database employed in this work is the Chicago Face Database [25], which is formed by 290 male and 307 female faces of ages ranging from 17 to 65 and varying ethnicity (Asian, Black, Latino and White). Each target in the database is represented with a neutral expression photo that has been normalized by an independent rater sample. As perception of facial features are not ethnicity independent [26], the objective of this work was to create proof of concept working with male gender of White ethnicity (93 images). All photographs were normalized and have the same size, illumination conditions and position.

1) Facial feature extraction: In order to locate the facial features, we employed a facial landmark detector [27]. Then, each feature was extracted individually, centered within the image and crop so all images of a given type of feature have the same size, and aligned. Figure 1 shows the extraction process and the resulting extracted facial features. These comprise the eyes, nose and mouth, and are automatically clustered using the *eigenfaces* method, which employs a holistic approach based on the appearance of the images.

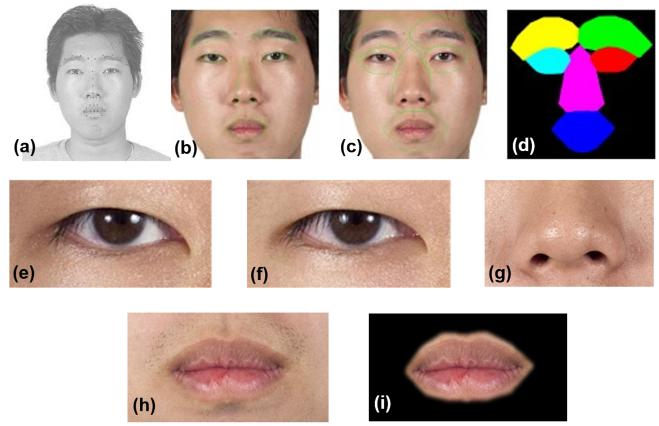

Fig. 1. Landmarks employed for facial feature extraction. (a) shows the detected landmarks, (b) the polygons formed by these landmarks, (c) the polygons thickened with $n = \inf$ and (d) the final masks used for extracting the features. (e), (f), (g) and (h) are the extracted facial features. In the case of the mouth, a black mask is applied to remove the possible moustache (i).

2) Facial feature clustering: In order to cluster the extracted features, the pixels of the images themselves could be used. In this manner, eyebrows, eyes, noses and mouths would be clustered in spaces of 169×96, 236×116, 187×118 and 191×104 dimensions respectively (the number of each image pixels). As can be noted, this approach yields very high dimensional spaces for clustering, which makes the method slow, very sensitive to noise and weak dealing with slight variations across images. Then, this work implements the *eigenfaces* [21] approach to obtain a relatively low-dimensional vector of characteristics which characterizes the features (the term eigenfaces is maintained although the method is now used over facial features). This method performs a PCA analysis over an ensemble of face images to form a set of basis features [28]. These basis images, known as eigenpictures, can be linearly combined to reconstruct images in the original set. This procedure allows for automatic, robust, fast and objective characterization of the facial features considering their global appearance while summarizing the central information.

We chose 45 eigenvalues to characterize the facial features, since it allowed to recover the original image without apparent loss. The same value was chosen for all of them in order to facilitate the subsequent clustering process, bearing in mind that the explained variance was higher than 85% in all cases (86.22% for eyes, 91.36% for noses and 95.26% for mouths).

As an example of the information of the features that is captured using eigenfaces, Figure 2 shows a reduced set of original mouths (a), and the same set of mouths reconstructed using 45 eigenvalues before de-normalization (b).

Finally, and after reducing the dimensionality of the internal features data, the k-means algorithm is employed to cluster the features using their eigenvalues as characteristics. A drawback of using this method is that the number of clusters k must be predefined, and this is unknown *a priori*. The approach followed to face this problem was to perform several k-means executions with $k = 5, 6, 7 \ldots 25$, and to calculate the Dunn's Index for each clustering while monitoring the number of

ISBN 978-90-827970-1-5

Fig. 2. Original and reconstructed mouths before de-normalization using 45 eigenfaces. (a) Original mouths. (b) Reconstructed mouths before de-normalization.

existent *mono-clusters*, which are defined as clusters with only one (for mouths and noses) or two instances (for eyes). The Dunn's Index measures the compactness and separation of the clusters obtained for each k. A higher Dunn's Index points to a small intra-cluster variance and a high inter-cluster distance. Therefore, the number of clusters for each feature was selected as the k that maximized the Dunn's Index while keeping the number of mono-clusters equal or below 2. The mono-clusters clusters were removed in a post-processing step and their instances assigned to their closest cluster.

III. Results

A. Clustering facial features by appearance

Table I shows the cluster metrics obtained for each facial feature using the clustering procedure explained before. k refers to the original number of clusters and k_{final} to the final number of clusters after removing the mono-clusters (referred to as *mc* in the table). Furthermore, the Dunn's Index and the Silhouette Index are also given for each clustering.

TABLE I
Clusters obtained for each facial feature.

	k	k_{final}	# of mc	DI	SI
eyes	19	19	0	0.86	0.12
noses	12	12	0	0.62	0.17
mouths	11	9	2	0.55	0.21

Figure 3 shows the results of the clusterings. (a) presents the results of the eyes, (b) the results of the noses and (c) the ones of the mouths. The first column indicates the cluster name, the second one shows the representative of the cluster (i.e. the closest instance to the cluster's centroid) and the third column the distribution of instances within the clustering. To view all the images employed and see the obtained clustering, please refer to the author's website (https://gitlab.com/ffuentes/facial-feature-clust).

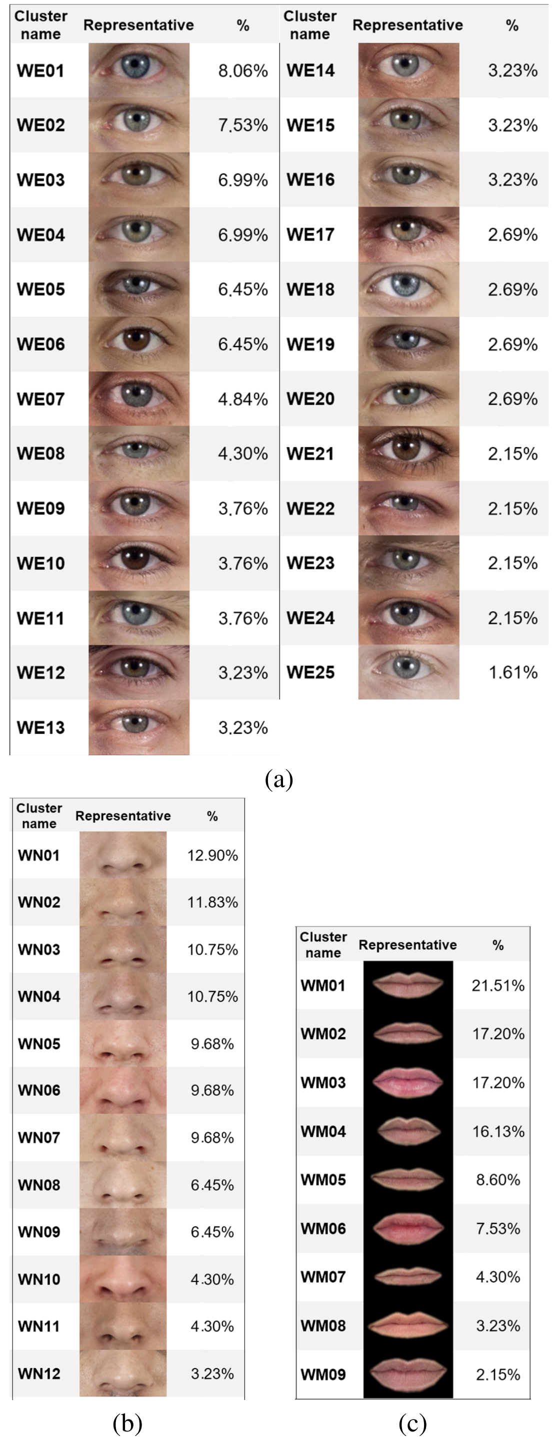

Fig. 3. Clustering results. (a) Shows the results for the eyes, (b) for the noses and (c) for the mouths.

IV. Validation of the procedure

The intuitively logical approach to validate the procedure is to compare the obtained taxonomies with those generated by human evaluators. However, as it has been aforementioned, classifying a big set of features in an undefined number of groups is a hard task considering human capabilities for information processing [15], [16]. Then, we measured the agreement of human evaluators with the proposed taxonomies instead. The main objectives were to reduce the number of features presented simultaneously and to simplify the decision that must be made. To do this, a survey composed of several stages was developed (Figure 4).

Initially, the image of one feature was selected from the entire dataset in a random way (target feature). Four different representative features were randomly selected (representative features are the features designated as representatives of their

ISBN 978-90-827970-1-5

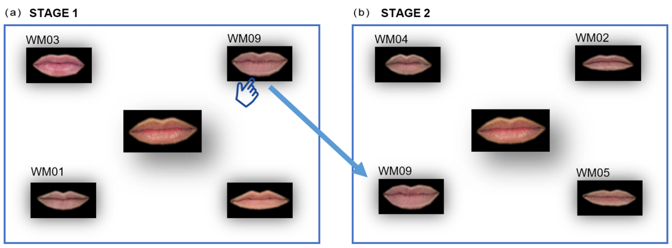

Fig. 4. Stages 1 and 2 of the survey procedure.

groups in the obtained taxonomy). In the first stage of the survey, the five features were presented to the evaluator in a web form (Figure 4 (a)). The target feature was in the center of the form, and the four representative features were at the corners. The evaluator was asked to select the representative feature most similar to the target feature. Then, the selected representative feature passed to the second stage in which a new form was composed like in Figure 4 (b). The target feature was in the center again, and the selected representative feature was at a corner of the form. Three new different representative features were randomly selected and situated in the remaining corners. This process was repeated until every representative feature was shown at least once. The cluster of the representative feature selected in the last stage was considered the result of the survey (i.e. the cluster to which the target feature belongs according to the opinion of the respondent). Using this procedure, the decision process was simplified because the number of simultaneous alternatives was reduced to four. As a drawback, the probability of one representative feature to be finally selected depends slightly on the stage in which it is shown.

21 white males and 11 white females aged between 25 and 46 years old participated in three surveys: eyes, noses and mouths. Table II show the results of the survey for eyes, noses and mouths. The first column of the table presents the finally selected cluster, where *Expected* means the cluster in which the target feature was grouped by the automatic procedure. 82 target mouths, 62 target eyes and 93 target noses were classified in the expected cluster. Since the distance between clusters can be measured through the eigenvalues of their representative features, it is possible to determine the distance from the expected cluster to each of the other clusters. The closer two clusters are, the more similar are the features they contain. In the aforementioned tables, 1^{st}, 2^{nd} and so on are closest clusters to the expected one. The number, the percentage and the cumulative percentage of features classified in each cluster are shown. The percentages of features classified in the expected cluster or in the three clusters closest to it were 73.0% for eyes, 81.0% for noses and 75.5% for mouths.

TABLE II
RESULTS OF THE VALIDATION SURVEY.

	Eyes			Noses			Mouths		
Selected	**N**	**%**	**Cum %**	**N**	**%**	**Cum %**	**N**	**%**	**Cum %**
Expected	62	31.0%	31.0%	93	46.5%	46.5%	82	41.0%	41.0%
1st	38	19.0%	50.0%	36	18.0%	64.5%	24	12.0%	53.0%
2nd	27	13.5%	63.5%	21	10.5%	75.0%	23	11.5%	64.5%
3rd	19	9.5%	73.0%	12	6.0%	81.0%	22	11.0%	75.5%
4th	12	6.0%	79.0%	8	4.0%	85.0%	17	8.5%	84.0%
5th	8	4.0%	83.0%	14	7.0%	92.0%	14	7.0%	91.0%
6th	5	2.5%	85.5%	3	1.5%	93.5%	6	3.0%	94.0%
7th	9	4.5%	90.0%	5	2.5%	96.0%	6	3.0%	97.0%
8th	4	2.0%	92.0%	6	3.0%	99.0%	6	3.0%	100.0%
9th	1	0.5%	92.5%	1	0.5%	99.5%			
10th	5	2.5%	95.0%	0	0.0%	99.5%			
11th	3	1.5%	96.5%	1	0.5%	100.0%			
12-24th	8	3.5%	100.0%						

V. Conclusions

There are very few classification systems or taxonomies for features, probably due to the complexity of this task, and to the human limited capacity for processing individual features compared to processing whole faces. Classifying the appearance of facial features requires a holistic approach considering all visible information. Therefore, we employed appearance-based representations of the features (*eigenfaces*) in order to classify them. The developed procedure groups the features considering all available information and encompassing their global nature. To validate the procedure, the agreement of human evaluators with the proposed taxonomies was measured. With the implemented system, more than 73.0% of the features were classified in the expected cluster or in the three clusters closest to it (75.5% of mouths, 73.0% of eyes and 81.0% of noses).

To the best of our knowledge, there are not similar studies to compare these results. Although more tests must be done, on the light of these results it can be concluded that the proposed automatic procedure is a good approach to classify facial features. Furthermore, the use of the proposed method is not restricted to facial features, and it should be possible to extend its use to automatically group any other kind of images by appearance.

Nevertheless, this study has some limitations. The experiment carried out employed 93 images of males with neutral expression from the Chicago Face Database. Therefore, the taxonomies obtained are only representative of the features of the faces belonging to this database. The representativeness of these taxonomies with respect to other populations must be carefully analyzed before using them. The objective of this work was not to achieve the taxonomies but to develop the automatic procedure to classify facial features based on their appearance. A more comprehensive face database can be used to obtain more representative taxonomies. On the other hand, future work must be done to extend this procedure to other facial features such as eyebrows, jawline, hair, and to obtain features taxonomies from faces of females. Moreover, the asymmetry of the face could be taken into account by introducing more horizontal distances to characterize the face.

References

[1] M. Rojas, D. Masip, A. Todorov, *et al.*, "Automatic prediction of facial trait judgments: Appearance vs. structural models," *PloS one*, vol. 6, no. 8, e23323, 2011.

[2] Y.-L. Tian, T. Kanade, *et al.*, "Handbook of face recognition," *Ch Facial Expression Analysis; Springer: Berlin/Heidelberg, Germany*, pp. 487–519, 2005.

[3] W. H. Sheldon, *Atlas of men: A guide for somatotyping the adult male at all ages*. Harper, 1954.

ISBN 978-90-827970-1-5 © EURASIP 2018

[4] A. Davis, J. Murphy, D. Owens, *et al.*, "Avatars, people, and virtual worlds: Foundations for research in metaverses," *Journal of the Association for Information Systems*, vol. 10, no. 2, p. 90, 2009.

[5] J. A. Diego-Mas and J. Alcaide-Marzal, "A computer based system to design expressive avatars," *Computers in Human Behavior*, vol. 44, pp. 1–11, 2015.

[6] P. Sukhija, S. Behal, and P. Singh, "Face recognition system using genetic algorithm," *Procedia Computer Science*, vol. 85, pp. 410–417, 2016.

[7] P. Vanezis, D Lu, J Cockburn, *et al.*, "Morphological classification of facial features in adult caucasian males based on an assessment of photographs of 50 subjects," *J. of Forensic Science*, vol. 41, no. 5, pp. 786–791, 1996.

[8] S. Ohlrogge, *Anthropological Atlas of Female Facial Features*, ser. Schriftenreihe angewandte forensische Anthropologie. Verlag für Polizeiwissenschaft Prof. Dr. Clemens Lorei, 2009, ISBN: 9783866760721.

[9] A. Tamir, "Numerical survey of the different shapes of the human nose," *Journal of Craniofacial Surgery*, vol. 22, no. 3, pp. 1104–1107, 2011.

[10] ——, "Numerical survey of the different shapes of human chin," *Journal of Craniofacial Surgery*, vol. 24, no. 5, pp. 1657–1659, 2013.

[11] J. J. Richler, O. S. Cheung, and I. Gauthier, "Holistic processing predicts face recognition," *Psychological science*, vol. 22, no. 4, pp. 464–471, 2011.

[12] J. Taubert, D. Apthorp, D. Aagten-Murphy, and D. Alais, "The role of holistic processing in face perception: Evidence from the face inversion effect," *Vision research*, vol. 51, no. 11, pp. 1273–1278, 2011.

[13] J. Davidoff and N. Donnelly, "Object superiority: A comparison of complete and part probes," *Acta psychologica*, vol. 73, no. 3, pp. 225–243, 1990.

[14] S. Ritz-Timme, P. Gabriel, Z. Obertovà, *et al.*, "A new atlas for the evaluation of facial features: Advantages, limits, and applicability," *International journal of legal medicine*, vol. 125, no. 2, pp. 301–306, 2011.

[15] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information.," *Psychological review*, vol. 63, no. 2, p. 81, 1956.

[16] A. Scharff, J. Palmer, *et al.*, "Evidence of fixed capacity in visual object categorization," *Psychonomic bulletin & review*, vol. 18, no. 4, pp. 713–721, 2011.

[17] E. Meyers and L. Wolf, "Using biologically inspired features for face processing," *International Journal of Computer Vision*, vol. 76, no. 1, pp. 93–104, 2008.

[18] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 6, pp. 681–685, 2001.

[19] T. Ahonen, A. Hadid, *et al.*, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[20] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[21] M. Turk *et al.*, "Eigenfaces for recognition," *J. of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[22] B. Klare and A. K. Jain, "On a taxonomy of facial features," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, IEEE, 2010, pp. 1–8.

[23] D. Cazzato, M. Leo, and C. Distante, "A statistical approach to automatically detect how many persons appear in a video," *Open Journal of Information Security and Applications*, vol. 1, no. 1, pp. 11–23, 2014.

[24] M. Leo, D. Cazzato, T. De Marco, and C. Distante, "Unsupervised approach for the accurate localization of the pupils in near-frontal facial images," *Journal of Electronic Imaging*, vol. 22, no. 3, p. 033 033, 2013.

[25] D. S. Ma, J. Correll, and B. Wittenbrink, "The chicago face database: A free stimulus set of faces and norming data," *Behavior research methods*, vol. 47, no. 4, pp. 1122–1135, 2015.

[26] M. Walker, F. Jiang, T. Vetter, *et al.*, "Universals and cultural differences in forming personality trait judgments from faces," *Social Psychological and Personality Science*, vol. 2, no. 6, pp. 609–617, 2011.

[27] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental face alignment in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1859–1866.

[28] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *Josa a*, vol. 4, no. 3, pp. 519–524, 1987.