

Facial Behaviometrics: the Case of Facial Deformation in Spontaneous Smile/Laughter

Stefanos Zafeiriou[†]
[†]Dept. of Computing,
Imperial College London
180 Queen's Gate
London SW7 2AZ, U.K.
{s.zafeiriou,m.pantic}@icl.org

Maja Pantic^{†,*}
*EEMCS
University of Twente
Drienerlolaan 5
7522 NB Enschede
The Netherlands
PanticM@cs.utwente.nl

Abstract

In this paper we explore the use of dense facial deformation in spontaneous smile/laughter as a biometric signature. The facial deformation is calculated between a neutral image (as neutral we define the least expressive image of the smile/laughter episode) and the apex of spontaneous smile/laughter (as apex we define the frame of the maximum facial change/deformation) and its complex representation is regarded. Subsequently, supervised and unsupervised complex dimensionality reduction techniques, namely the complex Principal Component Analysis (PCA) and the complex Linear Discriminant Analysis (LDA), are applied at the complex vector fields for feature extraction. We demonstrate the efficacy of facial deformation as a mean for person verification in a database of spontaneous smiles/laughters.

1. Introduction

The role of rigid and non-rigid facial motion in the perception of faces is a rather popular research topic in experimental psychology. One of the first studies conducted demonstrated that humans can recognize the six universal expressions even only when facial motion is visible [1]. Since then, many researchers have shown that facial motion conveys useful information about gender [5, 11], age [4] and identity [14, 20, 11, 14, 6, 8, 17, 18, 16, 19, 21, 26, 13, 22].

In the first studies conducted [14, 20] the results indicated that humans can recognize moving faces (both familiar and unfamiliar) with a significantly better rate than still faces. In [11] it was shown that rigid head motion influences identity recognition process, while in [23] it was shown that visible speech motion increases the face recognition accuracy achieved by humans.

However, a series of studies published after [14] on how

non-rigid motion affects the perception of identity in human faces reported quite contradictory results. For example, in [6, 8, 17, 18, 16] the effectiveness of non-rigid motion was argued in the recognition of only highly familiar to the observers faces (e.g., famous faces). In [19] it was further suggested that the advantage of face motion was only significant when the displayed motion was distinctive. In other words some familiar faces do have quite characteristic facial motion patterns and these patterns constitute an additional cue to identity. Finally, in the experiments conducted in [19] rigid head movement did not provide identity recognition clues.

One of the first studies that demonstrated that there is a difference between dynamic and static stimuli in the perception of facial identity even when the stimuli were not degraded images of famous faces was published in [26]. In another study [21] further evidence was provided that non-rigid motion of a previously unfamiliar human face can affect identity decisions over extended periods of time. In [13] the authors studied the synergy and integration of facial form and facial motion in the perception of identity. Consistent evidence was also reported that non-rigid facial motion biases observers perception of identity. In [22] the authors investigated the effects of familiarity, facial motion, and direction of learn/test transfer on person recognition and verified once more that the presentation of moving faces resulted in better recognition rate than that of static ones. The authors reported that there was no advantage in learning moving versus static faces for the recognition of profile views of faces.

In order to facilitate the following discussion let us divide facial motion into that induced by speech, referred to as speech-related, (i.e., caused by speech production) and that not induced by speech, referred to as speech-unrelated, (i.e., motion defining facial behavior). Although there is

a significant amount of work on the role of facial motion (both speech related and speech-unrelated) during the processing of identity in humans, there is rather limited research on using speech-unrelated facial motion (e.g., from facial expressions) for building automatic person recognition/verification systems.

One of the first works which showed that the development of such a system is feasible was presented in [9]. In this work it was shown that changes in action units appearance can be used as a biometric trait. In detail, the different phases of facial action units, their sequence order and their duration were measured. The attained verification rates were directly comparable with one of the best commercial face recognition systems at that time. In [27] tracker displacement features between the neutral state and the apex of the expression were used for person verification where an equal error rate of about 40% was reported.

Even though there are not many studies about the individuality and the persistence of speech-unrelated facial motion, the problem of persistence and individuality of speech-related facial motion, and especially of lip motion, is relatively more studied. In [7] it was experimentally demonstrated that dense lip motion is useful for person verification. In [10] orientation maps of the lip motion fields were used for person verification. Finally, in [3] the repeatability of viseme production over time for any speaker and the distinctiveness of lip motions across speakers in videos of 3D faces was studied.

In this paper we conduct a preliminary study in order to assess the discriminative capability of facial deformation for person verification. That is, from a database of 22 persons with many episodes of spontaneous smiles and laughs we compute the facial deformation between the frames corresponding to the least expressive (close to neutral state) and to the apex (maximum intensity) states. For the dimensionality reduction of the motion vector fields we adopt a complex representation of the facial deformation and we then apply the complex Principal Component Analysis (PCA) and complex Linear Discriminant Analysis algorithms. To the best of our knowledge this the first study which demonstrates that facial deformation from a spontaneous expression can be used for automatic person verification.

The remainder of the paper is organized as follows. In Section 2 we briefly describe the method applied for computing facial deformation. In Section 3 we formulate the complex PCA and the complex LDA algorithms for dimensionality reduction of facial motion vector fields. Experimental results are described in Section 4. Finally, conclusions are drawn in Section 5.

2. Extracting Facial Deformation using Free Form Deformations

For the extraction of facial deformation (i.e., facial motion vector field) we adapted a method used for non-rigid registration of medical images [24]. This method uses an Free Form Deformations (FFD) model based on b-splines. The method was originally used to register breast Magnetic Resonance images, where the breast undergoes local shape changes as a result of breathing and patient motion and recently has been applied for the extraction of facial motion from 2D and 3D facial image sequences [25, 15].

Let Ω_t denote the gray-level image of the face region at frame t , where $\Omega_t(x, y)$ is the gray-level intensity at pixel (x, y) . Given a pixel (x, y) in frame t , let (\tilde{x}, \tilde{y}) be the unknown location of its corresponding pixel in frame s . Then, the nonrigid registration method is used to estimate a motion vector field \tilde{F}_s between frames t and s such that:

$$(\tilde{x}, \tilde{y}) = (x, y) + \tilde{F}_s(x, y) \quad (1)$$

To estimate \tilde{F}_t , we select a $U \times V$ lattice Φ_t of control points with coordinates $\phi_t(u, v)$ in Ω_t , evenly spaced with spacing d . Then, nonrigid registration is used to align Φ_t with Ω_s , resulting in a displaced lattice $\tilde{\Phi}_s = \Phi_t + \Phi_\delta$. \tilde{F}_t can be derived by b-spline interpolation from Φ_δ . To estimate $\tilde{\Phi}_s$, a cost function C is minimized. In [24] the normalized mutual information was used as the image alignment criterion. However, in the 2D low-resolution case considered here, not enough sample data are available to make a good estimate of the image probability density function from the joint histograms. Therefore, we use the sum of square distances as the image alignment criterion, i.e.,

$$C(\tilde{\Phi}_s) = \sum_{x,y} (\Omega_t(x, y) - \Omega_s(\tilde{x}, \tilde{y}))^2 \quad (2)$$

The full algorithm for estimating $\tilde{\Phi}_s$ (and, therefore, Φ_δ) can be found in [24]. We can calculate \tilde{F}_t using b-spline interpolation on Φ_δ . For a pixel at location (x, y) , let $\phi_t(u, v)$ be the control point with coordinate (x_0, y_0) that is the nearest control point lower and to the left of (x, y) , i.e., it satisfies:

$$x_0 \leq x < x_0 + d, \quad y_0 \leq y < y_0 + d. \quad (3)$$

In addition, let $\phi_\delta(u, v)$ denote the vector that displaces $\phi_t(u, v)$ to $\tilde{\phi}_s(u, v)$. Then, to derive the displacement for any pixel (x, y) , we use a b-spline interpolation between its 16 closest neighboring control points. This gives us the estimate of the displacement field \tilde{F}_t :

$$\tilde{F}_t(x, y) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(a) B_l(b) \phi_\delta(u+k-1, v+l-1), \quad (4)$$

where $a = x - x_0, b = y - y_0$ and B_n is the n th basis function of the uniform cubic b -spline, i.e.,

$$\begin{aligned} B_0(a) &= (-a^3 + 3a^2 - 3a + 1)/6 \\ B_1(a) &= (3a^3 + 6a^2 + 4)/6 \\ B_2(a) &= (-3a^3 + 3a^2 + 3a + 1)/6 \\ B_3(a) &= a^3/6 \end{aligned} \quad (5)$$

To speed up the process and avoid local minima, we use a hierarchical approach in which the lattice density is being doubled at every level in the hierarchy. The coarsest lattice Φ_t^0 is placed around the point $c = (c_x, c_y)$ at the intersection of the horizontal line that connects the inner eye corners, and the vertical line passing through the tip of the nose and the center of the upper and bottom lip. Then,

$$\Phi_t^0 = \{(u, v) \mid \begin{array}{l} u \in [c_x - 2id, \dots, c_x + 2id] \\ v \in [c_y - 2id, \dots, c_y + 4id] \end{array} \} \quad (6)$$

where id is the distance between the eye pupils (i.e., Φ_t^0 consists of 35 control points). New control points are iteratively added in between until the spacing becomes $0.25id$ (approximately the size of a pupil), giving 1,617 control points. This has proven sufficient to capture most movements and gives a good balance between accuracy and calculation speed.

Having estimated \tilde{F}_t , we now have a motion vector field depicting the facial motion between frames s and t . Some examples of the facial motion vector field extracted can be found in Figure 1.

3. Dimensionality Reduction of Motion Fields

The dense motion vector fields computed from the previously described procedure are patterns of dimensionality $M \times N \times 2$ (where M and N are image rows and columns, respectively). In our experiments $M = N = 128$, hence the need of dimensionality reduction is evident. In the following we formulate complex Principal Component Analysis (PCA) and complex Linear Discriminant Analysis (LDA) for unsupervised and supervised dimensionality reduction of motion vector fields, respectively.

3.1. Complex PCA of motion fields

Let us denote by $\mathbf{x}_i \in \mathcal{C}^p$ the $p = NM$ -dimensional vector obtained by writing the vector field $\tilde{F}_t(x, y) = (u(x, y), v(x, y))$ in a complex form as $u(x, y) + jv(x, y)$ (where $j = \sqrt{-1}$) in lexicographic ordering. We assume that we are given a population of n samples $\mathbf{X} = [\mathbf{x}_1 | \dots | \mathbf{x}_n] \in \mathcal{C}^{p \times n}$, where p is the image resolution and n the number of samples. Without loss of generality, we assume zero-mean data. PCA finds a set of $k < n$ orthonormal bases $\mathbf{B}_k = [\mathbf{b}_1 | \dots | \mathbf{b}_k] \in \mathcal{C}^{p \times k}$ by minimizing the error function [12]

$$\epsilon(\mathbf{B}_k) = \|\mathbf{X} - \mathbf{B}_k \mathbf{B}_k^H \mathbf{X}\|_F^2. \quad (7)$$

where H denotes the complex conjugate operator.

The solution is given by the eigenvectors corresponding to the k largest eigenvalues obtained from the eigen-decomposition of the covariance matrix $\mathbf{X}\mathbf{X}^H$. Finally, the reconstruction of \mathbf{X} from the subspace spanned by the columns of \mathbf{B}_k is given by $\tilde{\mathbf{X}} = \mathbf{B}_k \mathbf{C}_k$, where $\mathbf{C}_k = \mathbf{B}_k^H \mathbf{X}$ is the matrix which gathers the set of projection coefficients.

For high dimensional data and Small Sample Size (SSS) problems (i.e. $n \ll p$), an efficient implementation of PCA in $O(n^3)$ can be used. Rather than computing the eigen-analysis of $\mathbf{X}\mathbf{X}^H$, we compute the eigen-analysis of $\mathbf{X}^H \mathbf{X}$.

Overall, **Algorithm 1** summarizes the steps of our PCA of the vector fields.

Algorithm 1. *Estimating the principal subspace of the vector fields*

Inputs: A set of n vector fields $\tilde{F}_i, i = 1, \dots, n$ of p pixels and the number k of principal components.

Step 1. Obtain \mathbf{x}_i by writing \tilde{F}_i in a complex form and in lexicographic ordering.

Step 2. Compute \mathbf{x}_i and form the matrix of $\mathbf{X} = [\mathbf{x}_1 | \dots | \mathbf{x}_n] \in \mathcal{C}^{p \times n}$ and compute the matrix $\mathbf{T} = \mathbf{X}^H \mathbf{X} \in \mathcal{C}^{n \times n}$.

Step 3. Compute the eigen-decomposition of $\mathbf{T} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ and denote by $\mathbf{U}_k \in \mathcal{C}^{p \times k}$ and $\mathbf{\Lambda}_k \in \mathcal{R}^{k \times k}$ the k -reduced set. Compute the principal subspace from $\mathbf{B}_k = \mathbf{X}\mathbf{U}_k\mathbf{\Lambda}_k^{-\frac{1}{2}} \in \mathcal{C}^{p \times k}$.

Finally, notice that our framework also enables the direct embedding of new samples. Let a test vector field \tilde{F} of p pixels and the principal subspace \mathbf{B}_k of **Algorithm 1**. Obtain the test vector \mathbf{x} by writing \tilde{F} in a complex form and lexicographic ordering. Then, we can extract the low dimensional embedding \mathbf{y} by

$$\mathbf{y} = \mathbf{B}_k^H \mathbf{x}. \quad (8)$$

3.2. Complex Linear Discriminant Analysis of motion fields

Let us now assume that our training set vector fields consists of C client (person) classes $\mathcal{C}_1, \dots, \mathcal{C}_C$. LDA aims at finding discriminant projection bases by exploiting this class-label information [2]. Within our framework of complex motion field, we need to formulate LDA in the complex domain. Let us denote by $\mathbf{b} \in \mathcal{C}^{N(P)}$ a projection vector and by $y_i = \mathbf{u}^H \mathbf{x}_i$ the projection of \mathbf{x}_i onto \mathbf{b} . We start by minimizing the (sum of the) variances of the data assigned to a particular class

$$\begin{aligned} E_w(\mathbf{u}) &= \sum_{c=1}^C \sum_{\mathbf{y}_i \in \mathcal{C}_c} \|y_i - \tilde{m}^c\|^2 \\ &= \sum_{c=1}^C \sum_{\mathbf{y}_i \in \mathcal{C}_c} \|\mathbf{u}^H (\mathbf{x}_i - \mathbf{m}^c)\|^2 \\ &= \mathbf{u}^H \sum_{c=1}^C \sum_{\mathbf{x}_i \in \mathcal{C}_c} (\mathbf{x}_i - \mathbf{m}^c)(\mathbf{x}_i - \mathbf{m}^c)^H \mathbf{u} \\ &= \mathbf{u}^H \mathbf{S}_w^z \mathbf{u} \end{aligned} \quad (9)$$

where $\tilde{\mathbf{m}}^c = \mathbf{u}^H \mathbf{m}^c$, $\mathbf{m}^c = \frac{1}{N(\mathcal{C})} \sum_{\mathbf{x}_i \in \mathcal{C}_c} \mathbf{x}_i$ and \mathbf{S}_w is the complex within-class scatter matrix

$$\mathbf{S}_w \triangleq \sum_{c=1}^C \sum_{\mathbf{x}_i \in \mathcal{C}_c} (\mathbf{x}_i - \mathbf{m}^c)(\mathbf{x}_i - \mathbf{m}^c)^H. \quad (10)$$

We also wish to maximize the distances between the projected class-means

$$\begin{aligned} E_b(\mathbf{u}) &= \sum_{c=1}^C N(\mathcal{C}_c) \|\tilde{\mathbf{m}}^c - \tilde{\mathbf{m}}\|^2 \\ &= \sum_{c=1}^C N(\mathcal{C}_c) \|\mathbf{u}^H (\mathbf{m}^c - \mathbf{m})\|^2 \\ &= \mathbf{u}^H \sum_{c=1}^C N(\mathcal{C}_c) (\mathbf{m}^c - \mathbf{m})(\mathbf{m}^c - \mathbf{m})^H \mathbf{u} \\ &= \mathbf{u}^H \mathbf{S}_b \mathbf{u} \end{aligned} \quad (11)$$

where \mathbf{S}_b is the complex between-class scatter matrix

$$\mathbf{S}_b \triangleq \sum_{c=1}^C N(\mathcal{C}_c) (\mathbf{m}^c - \mathbf{m})(\mathbf{m}^c - \mathbf{m})^H \quad (12)$$

and \mathbf{m} is the total mean vector.

To find K projections $\mathbf{U} = [\mathbf{u}_1 | \dots | \mathbf{u}_K] \in \mathcal{C}^{N(\mathcal{P}) \times K}$ we generalize $E_w(\mathbf{u})$ and $E_b(\mathbf{u})$ as

$$E_w(\mathbf{U}) = \text{tr}[\mathbf{U}^H \mathbf{S}_w \mathbf{U}] \quad (13)$$

$$E_b(\mathbf{U}) = \text{tr}[\mathbf{U}^H \mathbf{S}_b \mathbf{U}] \quad (14)$$

Then, the optimal projections are given by the solution to the following optimization problem

$$\begin{aligned} \mathbf{U}_o &= \arg \max_{\mathbf{U}} \text{tr}[\mathbf{U}^H \mathbf{S}_b \mathbf{U}] \\ \text{s.t. } &\mathbf{U}^H \mathbf{S}_w \mathbf{U} = \mathbf{I} \end{aligned} \quad (15)$$

The solution of the above noted problem is given by the K eigenvectors of $(\mathbf{S}_w)^{-1} \mathbf{S}_b$ corresponding to the K largest eigenvalues. For Small Sample Size (SSS) problems (i.e., $p \gg n$), the matrix \mathbf{S}_w is not invertible and the solution, in this case, is found by applying first the described complex PCA to preserve $n - C$ dimensions and then applying the complex LDA on the low-dimensional samples as

$$\mathbf{y} = \mathbf{U}^H \mathbf{B}_{n-C}^H \mathbf{x}. \quad (16)$$

where \mathbf{B}_{n-C} is the matrix with columns the $n - C$ projection bases obtained by the complex PCA and \mathbf{U} contains the $C - 1$ projection bases from the complex LDA. Similarity is measured by using the real part of the normalized correlation between the test sample \mathbf{y} and the training sample \mathbf{y}_i as

$$r(\mathbf{y}) = \text{Re} \left[\frac{\mathbf{y}^H \mathbf{y}_i}{\|\mathbf{y}\| \|\mathbf{y}_i\|} \right]. \quad (17)$$

4. Experimental Validation

The experiments were conducted in a newly collected audiovisual database. The audiovisual database has been collected for the purpose of studying laughter. It contains 22 subjects which were recorded while watching stimulus material, by two microphones, a video camera and a thermal camera. The primary goal was to elicit laughter, but also posed smiles, posed laughter and speech were also recorded. In total, 180 sessions are available with a total duration of 3h and 49min. There are 563 spontaneous smiles/laughter episodes, 849 speech utterances, 51 posed laughs, 67 speech-laughs and 167 other human noises annotated in the database. From the spontaneous smiles/laughter episodes we extracted the apex frames (the frames of the maximum facial change). We aligned the facial images using manually extracted eye coordinates. Then, we computed the non-rigid motion, using the method described in Section 2, between the apex frame and the frame that more closely resembles the neutral state from the particular episode. Examples of facial deformations fields used in our experiments can be found in Figure 1. From the color-coded vector field shown in this Figure it is evident that facial deformation is quite distinctive.

To fully validate our suggestion, we adopted a standard experimental protocol used in verification applications. To demonstrate that facial deformation conveys discriminative information that can be captured by a verification system, we measured the performance of a simple thresholding scheme. The protocol used was the following. Five experimental sessions were implemented by employing the leave-one-out (jackknife) and rotation estimates. In each session, 20% of the samples for every person were left out to be used as a test set (genuine claims). To implement test impostor claims, we rotated over the 22 person identities by considering the samples of each person in the test set as an impostor. By excluding any sample of the test impostor from the remaining four sessions, a training set consisted of 21 clients was built. The test impostor pretended to be one of the 21 clients and this attempt was repeated for all client identities. This way the impostor claims were produced. In a similar manner, test client claims were tested by employing the clients' samples from the session that was left out and those of the training set. Let $A_1, A_2, A_3, \dots, A_{22}$ be the identity codes of the persons included in the database. Figure 2 depicts the experimental protocol when person A_1 is considered to be an impostor and the samples of Session 5 is employed as test set. It can be seen that the training set is built of four out of the five available sessions each one consisting 21 out of the 22 available persons. The comparisons shown for person A_1 are repeated for all other persons in the database. Obviously, similar comparisons are made by rotating among the available sessions.

Next, we describe the training procedure. It is applied

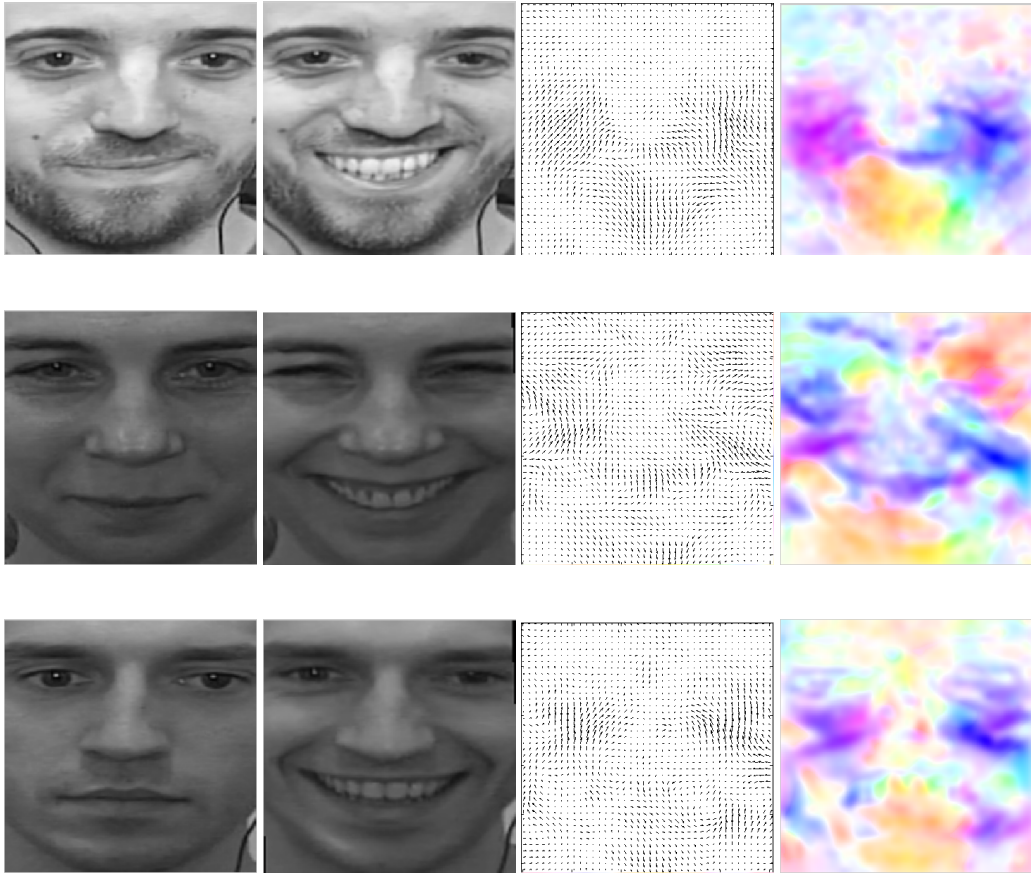


Figure 1. From left to right: The image that was considered as neutral. The apex of the episode. The vector field computed from the method described in section. Color-coded vector field.

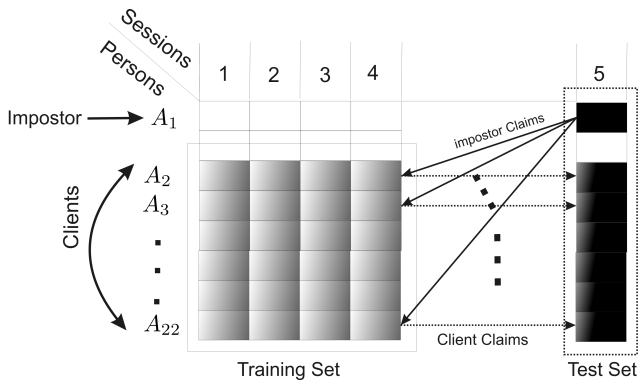


Figure 2. The verification protocol.

to the training set of the 21 clients. For each client, we had at least 8 samples at our disposal. Let us assume that person A_1 using one of his samples from session 5 pretends to be person A_2 during the test procedure. To test such a claim, we first implement the training stage, which includes:

- (i) the computation of PCA and/or LDA projections; (ii) the computation of the distances using (17) to be used as thresholds.

The defined thresholds should ideally enable the distinction between the distance measures that correspond to client claims and the distance measures that correspond to impostor claims. In the instantiation of Figure 2, the training procedure determines the projection matrices and 21 thresholds. Let us now explain how these thresholds are defined and incorporated in the final decision. For clarity purpose, we consider the case of person A_1 being an impostor and persons A_2, \dots, A_{21} being clients. We assume that person A_1 uses one of his samples to pretend to be person A_r . The measures (17) for every person calculated in the training set are used to form the distance vector $\mathbf{o}(r)$. The elements of the vector $\mathbf{o}(r)$ are sorted in ascending order and are used for the person specific thresholds on the distance measure. Let $T_Q(r)$ denoting the Q -th order statistic of the vector of distances, $\mathbf{o}(r)$. The threshold of the person r is chosen to be equal to $T_Q(r)$. Let r_1, r_2, \dots, r_s be the s instances of the person r in the training set. A claim of a person t is

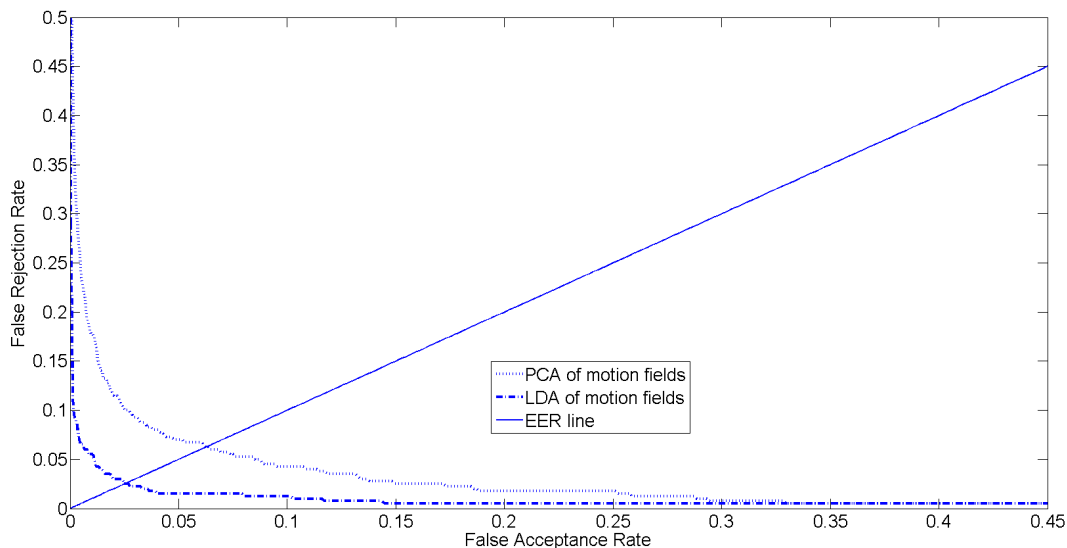


Figure 3. The ROC curves for the facial motion vector fields.

considered valid if $\max_j \{D_t(r_j)\} > T_Q(r)$ where $D_t(r_j)$ is the distance between the sample of test person t and the reference sample r_j .

The performance of verification systems is measured in terms of the false rejection rate (FRR) achieved at a fixed false acceptance rate (FAR). There is a trade-off between FAR and FRR. That is, it is possible to reduce either of them with the risk of increasing the other one. This trade-off between the FAR and FRR can create a curve where FRR is plotted as a function of FAR (while altering the threshold value). This curve is called receiver operating characteristic (ROC) curve [28, 30, 29]. The performance of a verification system is often quoted by a particular operating point of the ROC curve where FAR = FRR. This operating point is called equal error rate (EER).

The ROC curves using PCA and LDA of facial motion vector fields are depicted in Figure 3. The EER achieved by the applied PCA method was 6.3% while for the applied LDA 2.5%. The attained results provide a strong indication that person specific facial motion, in our case for spontaneous smiles, is useful for automatic person verification.

5. Conclusion

We presented a preliminary evaluation of the discriminative power of facial deformation from spontaneous smiles/laughters. We employed a complex representation of facial motion and applied complex PCA and LDA algorithms for dimensionality reduction. Further research includes the incorporation of dynamic information of dense facial deformation for person verification.

6. Acknowledgement

We thank the MEng. Student Tobias Gierk for annotating the database. This work has been funded by the European Research Council under the ERC Starting Grant agreement no. ERC-2007-StG-203143 (MAHNOB)."

References

- [1] J. Bassili. Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 4(3):373–379, 1978. 13
- [2] 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*, 19(7):711–720, July 1997. 15
- [3] L. Benedikt, D. Cosker, P. Rosin, and D. Marshall. Assessing the uniqueness and permanence of facial actions for use in biometric applications. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 40(3):449–460, 2010. 14
- [4] D. Berry. What can a moving face tell us? *Journal of Personality and Social Psychology*, 58(6):1004–1014, 1990. 13
- [5] D. Berry. Child and adult sensitivity to gender information in patterns of facial motion. *Ecological Psychology*, 3(4):348–366, 1991. 13
- [6] V. Bruce, Z. Henderson, C. Newman, and A. Burton. Matching Identities of Familiar and Unfamiliar Faces Caught on CCTV Images. *Journal of Experimental Psychology: Applied*, 7(3):207–218, 2001. 13
- [7] H. Cetingul, Y. Yemez, E. Erzin, and A. Tekalp. Discriminative analysis of lip motion features for speaker identification

- and speech-reading. *Image Processing, IEEE Transactions on*, 15(10):2879–2891, 2006. 14
- [8] F. Christie and V. Bruce. The role of dynamic information in the recognition of unfamiliar faces. *Memory & cognition*, 26(4):780–790, 1998. 13
- [9] J. Cohn, K. Schmidt, R. Gross, and P. Ekman. Individual differences in facial expression: Stability over time, relation to self-reported emotion, and ability to inform person identification. In *Multimodal Interfaces, 2002. Proceedings. Fourth IEEE International Conference on*, pages 491–496. IEEE, 2002. 14
- [10] M. Faraj and J. Bigun. Audio-visual person authentication using lip-motion from orientation maps. *Pattern recognition letters*, 28(11):1368–1382, 2007. 14
- [11] H. Hill and A. Johnston. Categorizing sex and identity from the biological motion of faces. *Current Biology*, 11(11):880–885, 2001. 13
- [12] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 12(1):103–108, Jan. 1990. 15
- [13] B. Knappmeyer, I. Thornton, and H. Bulthoff. The use of facial motion and facial form during the processing of identity. *Vision Research*, 43(18):1921–1936, 2003. 13
- [14] B. Knight and A. Johnston. The role of movement in face recognition. *Visual Cognition*, 4(3):265–273, 1997. 13
- [15] S. Koelstra, M. Pantic, and I. Patras. A dynamic texture based approach to recognition of facial actions and their temporal models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(11):1940–1954, november 2010. 14
- [16] K. Lander and V. Bruce. Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, 2000. 13
- [17] K. Lander, V. Bruce, and H. Hill. Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, 15(1):101–116, 2001. 13
- [18] K. Lander, F. Christie, and V. Bruce. The role of movement in the recognition of famous faces. *Memory & Cognition*, 27(6):974–985, 1999. 13
- [19] K. Lander and L. Chuang. Why are moving faces easier to recognize? *Visual Cognition*, 12(3):429–442, 2005. 13
- [20] G. Pike, R. Kemp, N. Towell, and K. Phillips. Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition*, 1997. 13
- [21] D. Roark, S. Barrett, M. Spence, H. Abdi, and A. O’Toole. Psychological and neural perspectives on the role of motion in face recognition. *Behavioral and cognitive neuroscience reviews*, 2(1):15, 2003. 13
- [22] D. Roark, A. O’Toole, H. Abdi, and S. Barrett. Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception*, 35(6):761–773, 2006. 13
- [23] L. Rosenblum, D. Yakel, N. Baseer, A. Panchal, B. Nodarse, and R. Niehus. Visual speech information for face recognition. *Attention, Perception, & Psychophysics*, 64(2):220–229, 2002. 13
- [24] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *Medical Imaging, IEEE Transactions on*, 18(8):712–721, 1999. 14
- [25] G. Sandbach, S. Zafeiriou, M. Pantic, and D. Rueckert. A dynamic approach to the recognition of 3d facial expressions and their temporal models. In *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition (FG’11), Special Session: 3D Facial Behavior Analysis and Understanding*, Santa Barbara, CA, USA, March 2011. 14
- [26] I. Thornton and Z. Kourtzi. A matching advantage for dynamic human faces. *Perception*, 31(1):113–132, 2002. 13
- [27] S. Tulyakov, T. Slowe, Z. Zhang, and V. Govindaraju. Facial expression biometrics using tracker displacement features. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–5. IEEE, 2007. 14
- [28] S. Zafeiriou, A. Tefas, and I. Pitas. Exploiting discriminant information in elastic graph matching. In *IEEE International Conference on Image Processing, 2005 (ICIP 2005)*, volume 3, pages III–768, 2005. 18
- [29] S. Zafeiriou, A. Tefas, and I. Pitas. The discriminant elastic graph matching algorithm applied to frontal face verification. *Pattern Recognition*, 40(10):2798–2810, 2007. 18
- [30] S. Zafeiriou, A. Tefas, and I. Pitas. Learning discriminant person specific facial models using expandable graphs. *IEEE Transactions on Information Forensics and Security*, 2(1):50–55, Mar. 2007. 18