A Model-Based Facial Expression Recognition Algorithm using Principal Components Analysis

N. Vretos, N. Nikolaidis and I.Pitas
Informatics and Telematics Institute
Centre for Research and Technology Hellas, Greece
Department of Informatics, Aristotle University of Thessaloniki
Thessaloniki 54124, Greece Tel,Fax: +30-2310996304
e-mail: vretos,nikolaid,pitas@aiia.csd.auth.gr

Abstract—**In this paper, we propose a new method for facial expression recognition. We utilize the Candide facial grid and apply Principal Components Analysis (PCA) to find the two eigenvectors of the model vertices. These eigenvectors along with the barycenter of the vertices are used to define a new coordinate system where vertices are mapped. Support Vector Machines (SVMs) are then used for the facial expression classification task. The method is invariant to in-plane translation and rotation as well as scaling of the face and achieves very satisfactory results.**

I. Introduction

Facial expression recognition in video sequences and still images is a very important research topic with applications in human-centered interfaces, ambient intelligence, behavior analysis etc. In [1], Ekman has established, based on an anthropological investigation, six main facial expressions (anger, surprise, happiness, disgust, fear and sadness) which are used in order to communicate human emotions. In many cases the neutral state is included along with the six expressions.

Although humans can easily recognize these facial expressions, this is not the case for algorithms that try to imitate this skill. During the last decade, many attempts involving a wide range of approaches have been undertaken to resolve this problem. Despite their diversity, most algorithms utilize information coming from the eyes, the mouth and the forehead region since these areas are considered the ones with the richest information for facial expressions recognition.

One consideration that has to be taken into account when designing facial expression recognition algorithms is the fact that a facial expression is a dynamic process that evolves over time and includes three stages [2]: an *onset* (attack), an *apex* (sustain) and an *offset* (relaxation) as described in [2]. Many facial expression recognition algorithms operate on the video frame (or still image) that corresponds to the expression *apex*. Based on the input data type used, facial expression recognition algorithms can be classified in two main categories: image (or image feature)-based and model-based ones. Each approach has its own merits. For instance, the image-based algorithms are faster, as no complex image pre-processing is usually involved. On the other hand, model-based approaches employ a 2-D or 3-D face model, whose fitting on the facial image [3] implies significant computational cost. Despite their computational effort, model-based approaches are popular because they can capture essential geometrical information for facial expression recognition [3]. An overview of the state of the art can be found in [2]

In this paper, we propose a novel model-based facial expression technique which utilizes the Candide facial grid (Figure 1), deformed so as to match the *apex* of an expression. So far, most model-based techniques use Candide vertex displacements, like in [3] or the Euclidian distance of pairs of Candide vertices [4]. In our case we use the location of the Candide vertices. Such information is sufficient to describe the geometry of the face and the facial features, and thus, can lead to facial expression recognition. However, this information is also extremely vulnerable to affine transformations (e.g. translation) of the face. To remedy this problem, we utilize principal components analysis in order to establish a new coordinate system that utilizes the evaluated eigenvectors and the vertices barycenter. As it will be proven later on, mapping the vertices to the new coordinate system ensures the method's robustness to translation, rotation and isotropic scaling of the face.

The paper is organized as follows: in Section II, we present the PCA-based feature selection procedure and prove the robustness of the method towards translation, scaling and rotation. In Section III, we describe the application of Support Vector Machines (SVM) upon the selected features for the recognition of 6 or $6+1$ different facial expressions and show experimental results on the Cohn-Kanade [5] database. Finally, conclusions are drawn in Section IV.

II. Principal Components Analysis and Feature Selection

The first step in the proposed approach is to track the Candide facial grid from the *onset* to the *apex* of the facial expression in a video sequence. In order to do so, we perform a manual localization of 7 vertices of the grid on the video frame corresponding to the onset of the facial expression as shown in Figure 2a. The rest of the Candide vertices are arranged in this frame through the application of a spring mesh-model. Then, we use the Kanade-Lucas-Tomasi (KLT) algorithm to track the grid vertices to the facial expression apex, as shown in Figure 2b. It has been proven that only 67 out of the

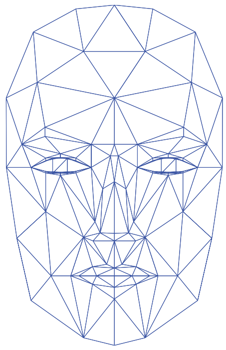
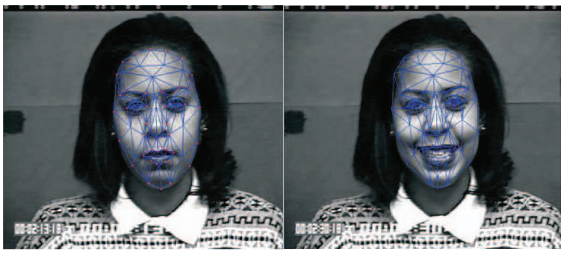


Fig. 1: The Candide face model



(a) (b)

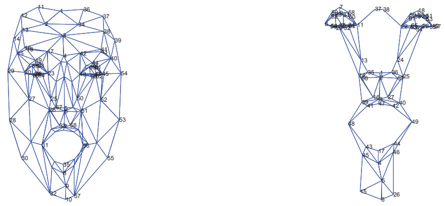
Fig. 2: a) Initialization of the Candide face model. b) Deformed Candide model.

104 Candide vertices carry important information for facial expression recognition [3]. Therefore only these vertices are retained. In Figure 3a the deformed facial grid corresponding to the surprise apex is shown. The 67 information-carrying vertices, are the ones at the mouth and eyes regions, as shown in Figure 3b.

The N=67 retained vertices form a point cloud in $\mathbf{R^2}$. We calculate the barycenter $P_m = (x_m, y_m)$ of these points by averaging the x and y coordinates for all points. Then, we subtract P_m from all points $P_i^{orig} = (x_i, y_i)$, $i = [1..67]$ obtaining $P_i = (x_i - x_m, y_i - y_m)$ that have zero mean coordinates. Finally, we calculate the 2×2 covariance matrix of these points:

$$\Sigma = \frac{1}{N-1} \left[\begin{array}{cc} \sum_{i=1}^{N}(x_i - x_m)(x_i - x_m) & \sum_{i=1}^{N}(x_i - x_m)(y_i - y_m) \\ \sum_{i=1}^{N}(y_i - y_m)(x_i - x_m) & \sum_{i=1}^{N}(y_i - y_m)(y_i - y_m) \end{array} \right] \tag{1}$$

To achieve invariance to in-plane rotation, translation and scaling, as will become apparent below, we adopt a new coordinate system whose origin is P_m and its axis are defined by the eigenvectors $\vec{v}_1$ and $\vec{v}_2$ of Σ and evaluate the coordinates of all points with respect to this system. In addition, we normalize the coordinates of each point with the inverse of the square root of the respective eigenvalue. The reason for this normalization will become apparent in the paragraphs below. Thus:



(a) (b)

Fig. 3: a) Deformed Candide face model corresponding to surprise. b) Retained deformed vertices from the model.

$$X_{new} = \frac{1}{\sqrt{\lambda_2}} \cdot P^T \cdot \vec{v}_2 \tag{2}$$

$$Y_{new} = \frac{1}{\sqrt{\lambda_1}} \cdot P^T \cdot \vec{v}_1 \tag{3}$$

where P is a 2×67 matrix containing all P_i, X_{new} and Y_{new} the 67×1 vectors that contains the normalized coordinates in the new coordinates system, λ_1 and λ_2 the eigenvalues of the covariance matrix Σ with $\lambda_1 > \lambda_2$ and $\vec{v}_1$, $\vec{v}_2$ the respective 2×1 eigenvectors. We have to notice that we assume that the variance in the y (vertical) axis of the original point cloud is bigger than the one in the x axis, an assumption that is valid due to facial proportions. The set of normalized points in the new coordinate system is used as the feature set for the expression recognition task.

In the following we prove the robustness of the proposed features (namely the Candide vertices expressed in the new coordinate system), with respect to translation, 2D-rotation (in-plane) and scaling of the face, and subsequently, of the Candide grid. First, since the origin of the new coordinate system is the barycenter of the vertices, a rigid translation of all vertices will lead to the translation of the origin by the same amount. Thus the vertices coordinates will remain unaltered.

In addition, 2D rotation invariance, is provided by the fact that the two axes of maximum variance (eigenvectors) evaluated throw PCA rotate rigidly along with the points. To prove this suppose we multiply our set of points with an arbitrary rotation matrix R. Thus the matrix containing the rotated points will be $P' = R \cdot P$. The covariance matrix of the rotated points will be:

$$\Sigma' = \frac{1}{N-1} RP \cdot (RP)^T = R \frac{1}{N-1} P \cdot P^T R^T = R \Sigma R^T \tag{4}$$

The eigenanalysis of Σ' will provide:

$$R \Sigma R^T \cdot \vec{v}_i' = \lambda' \cdot \vec{v}_i' \Longrightarrow \tag{5}$$

$$R^T R \Sigma R^T \cdot \vec{v}_i' = R^T \lambda' \cdot \vec{v}_i' \Longrightarrow \tag{6}$$

$$\Sigma \cdot R^T \vec{v}_i' = \lambda' \cdot R^T \vec{v}_i' \tag{7}$$

where Σ' is the covariance of the rotated points, λ_i', $\vec{v}_i'$ its eigenvalues and eigenvectors, and ^{T} the matrix transpose. Equation (7) implies that $R^T \vec{v}_i'$ is an eigenvector of Σ thus if $\vec{v}_i$ is an eigenvector of Σ then it holds that $\vec{v}_i' = R\vec{v}_i$. In other words, the eigenvectors of the rotated points are those of the original points rotated by the same rotation matrix. Furthermore, the eigenvalues of the set of the rotated points are the same as the ones of the original matrix, i.e. $\lambda_i = \lambda'_i$. For the vector X'_{new} contining the X coordinates (in the new coordinate system) of the rotated points, one can easily see that:

$$X'_{new} = \frac{1}{\sqrt{\lambda'_2}} \cdot P'^T \cdot \vec{v}'_2 \Longrightarrow \quad (8)$$

$$X'_{new} = \frac{1}{\sqrt{\lambda_2}} \cdot (RP)^T \cdot R\vec{v}_2 \Longrightarrow \quad (9)$$

$$X'_{new} = \frac{1}{\sqrt{\lambda_2}} \cdot P^T \cdot R^T \cdot R\vec{v}_2 \Longrightarrow \quad (10)$$

$$X'_{new} = \frac{1}{\sqrt{\lambda_2}} \cdot P^T \cdot \vec{v}_2 \Longrightarrow \quad (11)$$

$$X'_{new} = X_{new} \quad (12)$$

The same holds obviously for Y'_{new} and Y_{new}. Thus the points coordinates in the defined coordinates system do not change with rotation and thus our feature are invariant to in-plane rotation of the face.

Finally, invariance with respect to isotropic scaling can be proven as follows: by scaling in both dimensions with the same factor s it is easy to see that the new set of points will be $P'' = s \cdot P$ with covariance matrix $\Sigma'' = s^2 \cdot \Sigma$. The eigenanalysis of Σ provides that:

$$\Sigma \cdot \vec{v}_i = \lambda_i \cdot \vec{v}_i \Longrightarrow \quad (13)$$

$$s^2 \cdot \Sigma \cdot \vec{v}_i = (s^2 \lambda_i) \cdot \vec{v}_i \Longrightarrow \quad (14)$$

$$\Sigma'' \cdot \vec{v}_i = \lambda''_i \cdot \vec{v}_i \quad (15)$$

where $\lambda''_i = s^2 \cdot \lambda_i$ the eigenvalues of the scaled points. The above equation also implies that Σ'' has (as expected) the same eigenvectors $\vec{v}_1$, $\vec{v}_2$ with Σ. From (2) and (3) we have that:

$$X''_{new} = \frac{1}{\sqrt{\lambda''_2}} \cdot (s \cdot P)^T \cdot \vec{v}_2 \quad (16)$$

$$= \sqrt{\frac{1}{s^2 \cdot \lambda_2}} \cdot (s \cdot P)^T \cdot \vec{v}_2 \quad (17)$$

$$= \frac{1}{s} \cdot \sqrt{\frac{1}{\lambda_2}} \cdot (s \cdot P)^T \cdot \vec{v}_2 \quad (18)$$

$$= \frac{1}{\sqrt{\lambda_2}} \cdot P^T \cdot \vec{v}_2 \quad (19)$$

$$= X_{new} \quad (20)$$

Thus the x coordinates of the scaled points in the new coordinate system will remain unaltered. The same holds obviously for Y_{new}.

III. Facial Expression Classification Experiments

We use Support Vector Machines (SVMs) classifiers for recognizing facial expressions classes. SVMs were chosen due to their good performance in various practical pattern recognition applications [6]-[9], and their solid theoretical foundations.

SVMs minimize an objective function under certain constraints in the training phase, so as to find the support vectors, and subsequently use them to assign labels to the test set. Many SVMs variants exist. These include both linear and non linear forms, with different kernels being used in the latter. Six and seven class multiclass SVMs were used in our case.

We use the Cohn-Kanade [5] facial expressions database in order to evaluate our method. Firstly, we have initialized the Candide grid onto the onset frame of the database videos and tracked it till the facial expression apex frame. In our experiments, we used only the *apex* phase of the expression. The method was applied to 440 video frames (resulting from an equal number of videos): 35 for anger 35 for disgust 55 for fear 90 for happiness 65 for saddness 70 for surprise and finally 90 for the neutral state. We have conducted experiments for the recognition of either 6 or 6+1(neutral) facial expressions.

We shall first present the experiments for 6 facial expressions. In order to establish a test and a training set, we used a modified version of the leave-one-out cross-validation procedure, where, in each run, we exclude 20% of the grids for of each facial expression from the training set and use them to form the test set. Thus, in order to process all data, five runs were conducted and the average classification accuracy was calculated. In Table I, results are drawn from

Kernel	Degree	Recognition Rate
RBF	3	**88.69%**
RBF	5	88.18%
RBF	4	88.15%
RBF	7	87.79%
RBF	8	87.79%
RBF	6	87.59%
RBF	2	87.26%
Polynomial	3	**88.71%**
Polynomial	2	88.17%
Polynomial	4	87.70%

TABLE I: Results for radial basis function (RBF) and polynomial kernels with different degrees.

classifiers involving different parameters and kernels. It can be seen that the results are practically the same for all tested SVM configurations. The confusion matrices for the RBF and polynomial kernel parameters that achieved the best performance are depicted in Tables II and III. The fact that for different kernels and different parameters we obtain the same results is obviously an advantage for the proposed algorithm, since it potentially indicates good generalization properties. This desirable behavior can be attributed to the utilized features and is particularly important since many classification algorithms, using SVMs, suffer in generalization due to the known sensitivity of the SVMs with respect to their parameters [10].

We performed experiments for the 6+1 expressions as well (Tables IV, V and VI). For most algorithms, the classification accuracy when recognizing 6+1 expressions is worse than the one in the 6 classes case. In our case though, there is a slight performance improvement on the overall accuracy which can be perhaps attributed to the nature of the feature space. Most probably in this space the neutral class is far from all other classes and, thus, does not alter the 6 class results significantly. This fact can be noticed in the confusion

	Ang	Dis	Fea	Hap	Sad	Sur
Ang	84.8	0	0	0	8.9	0
Dis	6.1	93.3	2.2	2.2	2.5	1.4
Fea	0	0	89.1	11.0	3.8	0
Hap	3.0	6.7	6.5	86.8	3.8	1.4
Sad	6.1	0	0	0	81.0	0
Sur	0	0	2.2	0	0	97.1

TABLE II: Confusion matrix for polynomial kernel of 3^{rd} degree

	Ang	Dis	Fea	Hap	Sad	Sur
Ang	81.1	0	0	0	6.7	0
Dis	8.1	93.3	0	2.2	2.7	1.5
Fea	0	0	89.4	10.9	2.7	0
Hap	5.4	6.7	6.4	85.9	4.0	0
Sad	5.4	0	0	1.1	84.0	0
Sur	0	0	4.3	0	0	98.5

TABLE III: Confusion matrix for radial basis function (RBF) kernel of 3^{rd} degree.

matrices where the rates for the 6 classes are practically the same as in the previous experiments. On the other hand, as the neutral class exhibits 100% recognition rate the overall accuracy of the algorithm increases. The overall classification accuracy for the cases of Tables IV-VI is 90.22%, 90.02% and 89.49% respectively.

	Ang	Dis	Fea	Hap	Sad	Sur	Neu
Ang	82.4	3.3	0	0	9.3	0	0
Dis	5.9	93.3	2.3	1.0	3.5	1.4	0
Fea	0	0	90.9	10.7	3.5	0	0
Hap	5.9	3.3	4.5	87.4	3.5	0	0
Sad	5.9	0	0	1.0	79.1	0	0
Sur	0	0	2.3	0	0	98.6	0
Neu	0	0	0	0	1.2	0	100.0

TABLE IV: Confusion matrix for radial basis function (RBF) kernel of 5^{th} degree (6+1 expressions).

	Ang	Dis	Fea	Hap	Sad	Sur	Neu
Ang	81.8	3.3	0	0	10.3	0	0
Dis	6.1	93.3	2.3	1.0	3.4	1.4	0
Fea	0	0	90.9	10.7	3.4	0	0
Hap	6.1	3.3	4.5	87.4	3.4	0	0
Sad	6.1	0	0	1.0	78.2	0	0
Sur	0	0	2.3	0	0	98.6	0
Neu	0	0	0	0	1.1	0	100.0

TABLE V: Confusion matrix for radial basis function (RBF) kernel of 2^{nd} degree (6+1 expressions).

IV. Conclusions and Future Work

In this paper, we have introduced a method for automatic facial expressions recognition using principal components analysis on the vertices of the Candide model. SVMs are used for the classification task. Results show clearly that the proposed method is a good framework towards model-based facial expression recognition. The achieved facial expression

	Ang	Dis	Fea	Hap	Sad	Sur	Neu
Ang	81.2	3.3	0	0	11.2	0	0
Dis	6.2	93.3	2.2	1.0	3.4	1.4	0
Fea	0	0	88.9	11.0	3.4	0	0
Hap	6.2	3.3	6.7	88.0	3.4	1.4	0
Sad	6.2	0	0	0	77.5	0	0
Sur	0	0	2.2	0	0	97.1	0
Neu	0	0	0	0	1.1	0	100.0

TABLE VI: Confusion for polynomial kernel of 1^{st} degree (6+1 expressions).

classification accuracy is approximately 90%. The main advantage of the employed feature space is its robustness towards scaling, in-plane rotation and translation of the face.

It is also worth noting that the method is not sensitive to the SVM configuration with respect to kernel and kernel parameters, which is an indication that the method has good generalization properties. Future work will aim in the experimental evaluation of this claim.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 211471 (i3DPost).

References

[1] P. Ekman, "Facial expression and emotion," *Personality: Critical Concepts in Psychology*, vol. 48, pp. 384–92, 1998.

[2] M. Pantie and LJM Rothkrantz, "Automatic analysis of facial expressions: the state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1424–1445, 2000.

[3] I. Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 172–187, 2007.

[4] K. Kahler, J. Haber, and H.P. Seidel, "Geometry-based muscle modeling for facial animation," *Proc. Graphics Interface 2001*, pp. 37–46, 2001.

[5] T. Kanade, J.F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," *Proceedings of the fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG00)*, pp. 46–53, 2000.

[6] A. Tefas, C. Kotropoulos, and I. Pitas, "Using support vector machines to enhance the performance of elasticgraph matching for frontal face authentication," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 7, pp. 735–746, 2001.

[7] H. Drucker, D. Wu, and V.N. Vapnik, "Support Vector Machines for Spam Categorization," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, 1999.

[8] A. Ganapathiraju, J. Hamaker, and J. Picone, "Support Vector Machines for Speech Recognition," *Fifth International Conference on Spoken Language Processing*, 1998.

[9] M. Pontil and A. Verri, "Support vector machines for 3 D object recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 637–646, 1998.

[10] A. Zien, G. Ratsch, S. Mika, B. Scholkopf, T. Lengauer, and K.R. Muller, "Engineering support vector machine kernels that recognize translation initiation sites," *Bioinformatics*, vol. 16, no. 9, pp. 799–807, 2000.