Identity and Expression Recognition on Low Dimensional Manifolds

Pedro Martins ♦ Jorge Batista
Institute of Systems and Robotics
Department of Electrical and Computer Engineering, University of Coimbra, Portugal

Abstract

- Face geometry extracted using Active Appearance Models (AAM).
- Low dimensional manifolds were then derived using Laplacian EigenMaps resulting in two types of manifolds, one for model identity and the other for expression. The recognition is composed by a two step cascade, where first the identity is predicted and then its associated expression model is used to predict the facial expression.
- The identity overall recognition rate was 96.8%.Facial expression results are identity dependent, the most expressive individual achieves 81.2% of overall recognition rate.

Active Appearance Models

- Generative nonlinear parametric models of shape and texture, commonly used to model faces.

Shape and Texture Models

$$s = (x_1, \ldots, x_v, y_1, \ldots, y_v)^T$$

$$s = s_0 + \sum_{i=1}^{n} p_i s_i$$

$$A(x) = A_0 + \sum_{i=1}^{m} \lambda_i A_i, x \in s_0$$

Fitting Goal

$$\sum_{x \in s_0} \left[A_0(x) + \sum_{i=1}^{m} \lambda_i A_i(x) - I(W(x;p)) \right]^2$$

Simultaneous Inverse Compositional Image Alignment

Warp **I** with **W(x,p)**  $I \rightarrow I(W(x;p))$ 

Error Image $E(x) = I(W(x;p)) - \left(A_0(x) + \sum_{i=1}^{m} \lambda_i A_i \right)$

Steepest Descent Images $SD_{SIC} = \left(\left(\nabla A_0 + \sum_{i=1}^{m} \lambda_i A_i \right) \frac{\partial W}{\partial p_1}, \ldots, \left(\nabla A_0 + \sum_{i=1}^{m} \lambda_i A_i \right) \frac{\partial W}{\partial p_n}, A_1(x), \ldots, A_m(x) \right)$

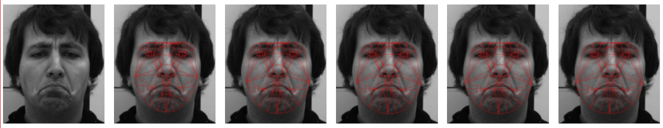
Hessian $H_{SIC} = \sum_{x \in s_0} SD_{SIC}^T(x) SD_{SIC}(x)$

Parameters SD Updates $\begin{pmatrix} \Delta p \\ \Delta \lambda \end{pmatrix} = -H_{SIC}^{-1} \sum_{x \in s_0} SD_{SIC}^T(x) E(x)$

Parameters Updates $W(x;p) \leftarrow W(x;p) \circ W(x;\Delta p)^{-1}$ $\quad \lambda \leftarrow \lambda + \Delta\lambda$

while $|\Delta p| < \varepsilon$ or max iterations reached

Fitting Example

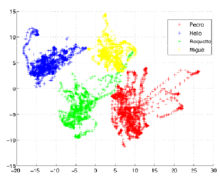


Input I 1st 2nd 5th 8th Final

Laplacian EigenMaps

- Nonlinear dimension reduction techniques that derive a low dimensional manifold lying in a higher dimensional more complex manifold.
- Given k feature points, $x_1, \ldots, x_k \in \Re^l$, a graph with k nodes is build. The embedding map is found by computing the eigenvectors of the graph Laplacian,
- Algorithm
  - Build the adjacency graph, each node is connected to the n nearest neighbors
  - Choose the weights for edges in the graph, $W_{ij} = 1$ is node i and j are connected by an edge
  - Eigen-decomposition of the graph laplacian, $(D-W)f = \lambda Df, D_{ij} = \sum_j W_{ji}$
  - Form the low-dimensional embedding $\Phi = (f_1 | \ldots | f_m)$

Identity Manifold

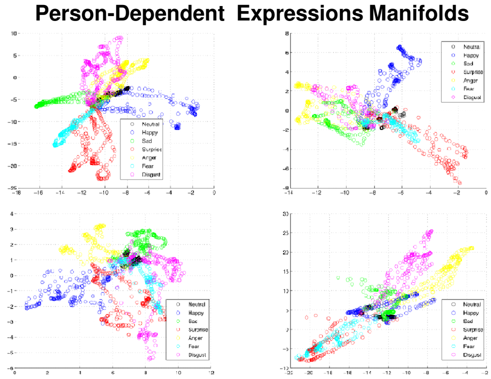


- Results for Identity Recognition

	Person 1	Person 2	Person 3	Person 4
Person 1	**98.11**	0.09	1.79	0
Person 2	1.32	**98.67**	0	0
Person 3	2.93	0.29	**94.50**	2.27
Person 4	1.29	0.13	2.32	**96.25**

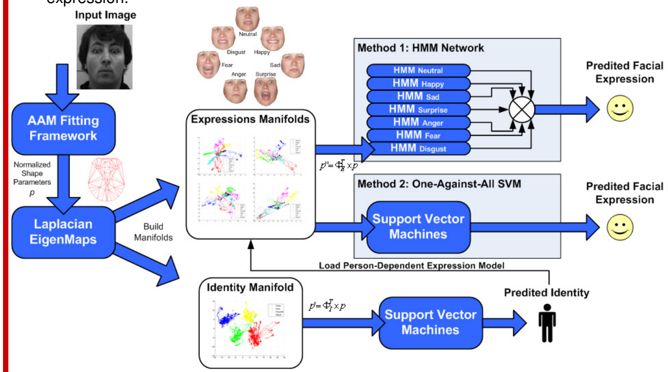
- Overall Recognition Rate SVM=96.88%

Person-Dependent Expressions Manifolds



Simultaneuos Identity and Facial Expression Recognition

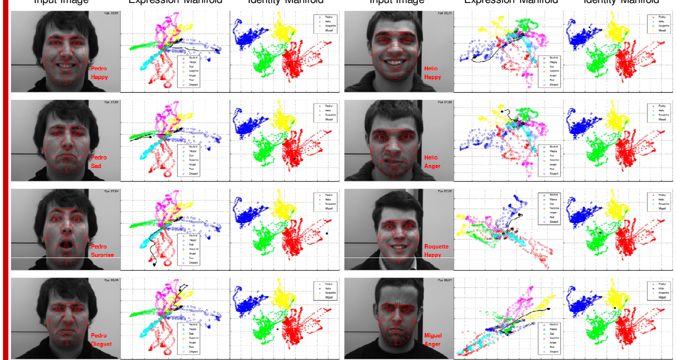
- Face geometry extracted using Active Appearance Models (AAM).
- Low dimensional manifolds were then derived using Laplacian EigenMaps resulting in two types of manifolds, one for model identity and the other for expression.
- The recognition is composed by a two step cascade, where first the identity is predicted and then its associated expression model is used to predict the facial expression.



Experimental Results

- For evaluation proposes a Facial Dynamics Database was build.
4 Individuals / 7 Different Facial Expressions / 4 Folds / 640 x 480 / Total of 6670 Frames

Input Image Expression Manifold Identity Manifold Input Image Expression Manifold Identity Manifold



Pers1	Neut.	Happ.	Sad	Surp.	Anger	Fear	Disg.
Neut.	**69.9** **58.4 (3)**	9.2 0 (0)	2.3 3 (0)	0 0 (0)	1 16 (1)	1 0 (0)	16.8 22.5 (0)
Happ.	0 1.2 (0)	**84.6** **95 (4)**	3.3 0 (0)	10.4 0 (0)	1.7 3.8 (0)	0 0 (0)	0 0 (0)
Sad	0 0.6 (0)	0 0 (0)	**100** **97.9** (4)	0 1.5 (0)	0 0 (0)	0 0 (0)	0 0 (0)
Surp.	0.7 0 (0)	0 0 (0)	0 0.7 (0)	**99.3** **99.3** (4)	0 0 (0)	0 0 (0)	0 0 (0)
Anger	2.4 0 (0)	0 2.7 (0)	0.8 5.1 (0)	0.6 0.6 (0)	**84.4** **87.7** (4)	0.3 1.2 (0)	11.4 2.7 (0)
Fear	0 0 (0)	0.8 0 (0)	0 0 (0)	38.7 29.4 (1)	0 2.2 (0)	**60.6** **68.4** (3)	0 0 (0)
Disg.	2.5 0 (0)	0 0 (0)	0 4.1 (0)	37.6 0.9 (0)	20.7 32.8 (1)	0 0 (0)	**39.2** **62.1** (3)

- Overall Recognition Rate: SVM=76.85% HMM=81.27% (full sequence)

Pers2	Neut.	Happ.	Sad	Surp.	Anger	Fear	Disg.
Neut.	**67.8** **47.3 (2)**	0 1.7 (0)	6.4 0 (0)	0 0 (0)	0 0 (1)	25.8 41.9 (2)	0 9 (0)
Happ.	1 0 (0)	**78.7** **70.3 (3)**	0 0 (0)	17.1 25.9 (1)	0 0 (0)	3 3.8 (0)	0 0 (0)
Sad	1.7 0 (0)	0 0.4 (0)	**86.9** **93.4** (4)	5.5 0.7 (0)	0 0 (0)	0 1.4 (0)	5.9 4.1 (0)
Surp.	0.7 0 (0)	25.9 0 (0)	0.8 0.4 (0)	**41.6** **97.3** (4)	0 0.8 (0)	26.7 1.5 (0)	4.2 0 (0)
Anger	1.4 0 (0)	0 0 (0)	1.8 1.8 (0)	0 0 (0)	**79.3** **91.7** (4)	0.5 0.9 (0)	17 5.5 (0)
Fear	2.9 0 (0)	0 1.2 (0)	2 2 (0)	57.4 31.1 (2)	0 0 (0)	**33.6** **61.5** (2)	4 4 (0)
Disg.	1.6 15.3 (1)	17.2 0.7 (0)	4 0 (0)	22.8 9.1 (0)	1.6 1.9 (0)	2.9 2.9 (0)	**50.2** **70** (3)

- Overall Recognition Rate: SVM=62.56% HMM=75.95% (full sequence)

Pers3	Neut.	Happ.	Sad	Surp.	Anger	Fear	Disg.
Neut.	**43.7** **61.3 (2)**	0 20.1 (1)	20.1 0 (0)	25.6 0 (0)	0 18.6 (1)	10.5 0 (0)	0 0 (0)
Happ.	3.9 0.8 (0)	**80.5** **96.5 (4)**	0.4 0 (0)	6.5 2.6 (0)	0 0 (0)	3.9 0 (0)	4.8 0 (0)
Sad	8.3 0.4 (0)	0 0 (0)	**72.5** **94.8** (4)	0 0.9 (0)	10.5 2.2 (0)	2.6 1.3 (0)	6.1 0.4 (0)
Surp.	5.3 23.4 (1)	6.9 0 (0)	0 0 (0)	**66** **76** (3)	0 0 (0)	21.8 0.5 (0)	0 0 (0)
Anger	4.3 0 (0)	0.5 2.4 (0)	25.7 2.4 (0)	0 1.9 (0)	**61.9** **92.9** (4)	1 0.5 (0)	6.7 0 (0)
Fear	21.3 21.3 (1)	23.1 0 (0)	0 6.9 (0)	18.8 16.9 (1)	0 1.3 (0)	**23.1** **42.3** (2)	13.8 7.5 (0)
Disg.	10.1 27 (1)	2 1.4 (0)	37.1 16.2 (1)	10.8 0 (0)	5.4 9.5 (0)	2 1.3 (0)	**32.4** **44.3** (2)

- Overall Recognition Rate: SVM=54.30% HMM=73.20% (full sequence)

Pers4	Neut.	Happ.	Sad	Surp.	Anger	Fear	Disg.
Neut.	**52.5** **25 (1)**	17.5 8 (0)	0 0 (0)	18 39 (2)	0 28 (1)	0 0 (0)	12 0 (0)
Happ.	4.7 0 (0)	**90.2** **96.8 (4)**	0 0 (0)	3.7 0.5 (0)	0 0 (0)	1.1 0 (0)	0 3.7 (0)
Sad	2 1.5 (0)	12.5 30.7 (1)	**42.7** **53.3** (3)	0 6.5 (0)	0 8 (0)	0 0 (0)	42.7 0 (0)
Surp.	1.9 1.4 (0)	2.8 2.8 (0)	0 0 (0)	**56.5** **66.4** (3)	0 2.8 (0)	32.7 26.6 (1)	6 0 (0)
Anger	2.2 0 (0)	0 0.8 (0)	0 2.2 (0)	0 4.8 (0)	**55.7** **87.7** (4)	0 0.4 (0)	42.1 4 (0)
Fear	1.4 0 (0)	3.3 0 (0)	0 0 (0)	16.2 22 (1)	0 0 (0)	**75.6** **80** (3)	3.3 0 (0)
Disg.	0.5 0 (0)	6.5 0 (0)	0 0 (0)	0 6.5 (0)	1 0 (0)	0 0.5 (0)	**92** **93** (4)

- Overall Recognition Rate: SVM=66.47% HMM=71.30% (full sequence)

This work was funded by the FCT grant SFRH/BD/45178/2008

Institute of Systems and Robotics
University of Coimbra