

# Towards Generic Fitting Using Multiple Features Discriminative Active Appearance Models

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

- Solution for discriminative based Active Appearance Models (AAM).

## Abstract

- The model consists in a set of descriptors which are covariances of multiple features evaluated over the neighborhood of landmarks whose locations are governed by a Point Distribution Model (PDM). The covariances are a special set of tensors that lie into a Riemannian Manifold. It is possible to measure the dissimilarity and to update them, imposing temporal appearance consistency.
- The fitting method uses a combination of exhaustive local search, finding modes with mean-shift and clustering for each landmark independently. The global optimization then constrains each landmark location update by the PDM.

## Shape Model

$s = (x_1, x_2, \dots, x_{n-1}, x_n, y_1, y_2, \dots, y_{n-1}, y_n)^T$

$\Psi = \begin{bmatrix} x_1^2 & -y_1^2 & 1 & 0 \\ x_2^2 & -y_2^2 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_n^2 & -y_n^2 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \end{bmatrix}$

$s = s_0 + \sum_{i=1}^n p_i \phi_i + \sum_{j=1}^k q_j \psi_j$

$q = [s \cos(\theta) - 1 \quad s \sin(\theta) \quad t_x \quad t_y]^T$

Shape Covariance

$p$  - shape parameters  
 $q$  - pose parameters

## Patch-Based Descriptor

$f = \left[ x \quad y \quad I_x \quad I_y \quad \sqrt{I_x^2 + I_y^2} \quad \arctan\left(\frac{I_y}{I_x}\right) \quad I_{xx} + I_{yy} \right]^T$

Covariance of Multiple Features

$C_i = \frac{1}{T-1} \sum_{t=1}^T (F_{i,t} - \mu_{F_i})(F_{i,t} - \mu_{F_i})^T$

### Dissimilarity Between Covariances

$\rho(C_1, C_2) = \sqrt{\sum_{i=1}^m \lambda_i(C_1, C_2)}$

$\lambda_i(C_1, C_2)_{i=1, \dots, m}$   
 Generalized eigenvalues computed from  $\lambda_i C_1 x_i - C_2 x_i = 0$

### Mean over the Manifold

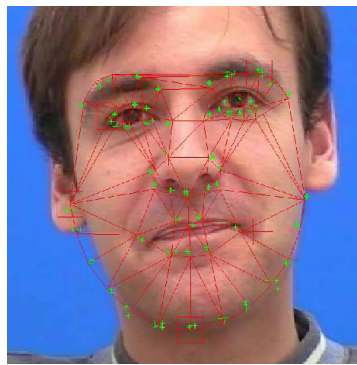
$\bar{C} = \arg \min \sum_{i=1}^n \rho(C, C_i)$

$\bar{C}^{(1)} = \exp_{\bar{C}} \left( \frac{1}{T} \sum_{t=1}^T \log_{\bar{C}}(C_t) \right)$

### Weighted Mean over the Manifold

$\bar{C}^{(1)} = \exp_{\bar{C}} \left( \frac{1}{\sum_{i=1}^n \rho(C_i, \bar{C})} \sum_{i=1}^n \rho(C_i, \bar{C})^{-1} \log_{\bar{C}}(C_i) \right)$

## Discriminative Active Appearance Models



### Shape + Similarity Warp

$W(x, p, q) = s_0 + \Phi p + \Psi q$

### Jacobians of the Warp

$\frac{\partial W(x, p, q)}{\partial p} = \Phi^T \quad \frac{\partial W(x, p, q)}{\partial q} = \Psi^T$

- Find local *optimal* displacements

$\Delta x_i^* = \arg \min_{\Delta x_i} \rho(C_i \{x_i + \Delta x_i\}, \bar{C}_i \{x_i\})$

- Constrain local updates to lie in the subspace spanned by the PDM

### Weighted LS Solution

$\Delta p = (\Phi W \Phi^T)^{-1} \Phi W \Delta x^*$

$\Delta q = (\Psi W \Psi^T)^{-1} \Psi W \Delta x^*$

- Diagonal Matrix of Weights

$W = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & w_n \end{bmatrix}$

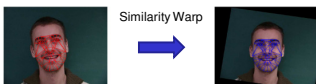
Goal

$\arg \min_{p, q} \sum_{i=1}^n \rho(C_i \{W(x_i, p, q)\}, \bar{C}_i \{x_i\})$

- Minimizing the covariance dissimilarity between the model and the covariance computed at a shifted location - constrained to be consistent with the PDM - for all the v patches

### Image Normalization

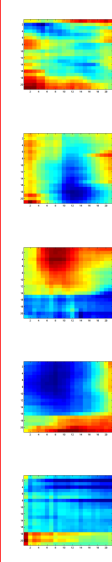
- The covariance computation is not invariant to scale and rotation effects



### Impose Temporal Appearance Consistency By Updating Model Covariances Across Time

$\bar{C}_i \{x_i\} \leftarrow \bar{C}_i^{-1} \{x_i\} \bar{C}_i^{-1} \{x_i\} \dots \bar{C}_i^{-1} \{x_i\}$

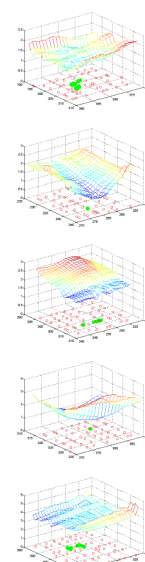
Mean over the Manifold



### Dissimilarity Response Maps

$\rho(C_i \{x_i + \Delta x_i\}, \bar{C}_i \{x_i\})$

↑ Sampled Cov.    ↑ Model Cov.

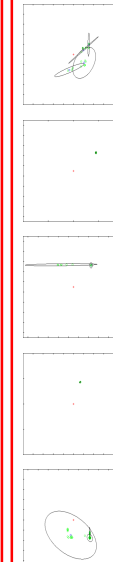


### Mean-Shift

$m_w(x) = \frac{\sum w_i(x) g\left(\frac{x-x_i}{h}\right)}{\sum w_i(x) g\left(\frac{x-x_i}{h}\right)}$

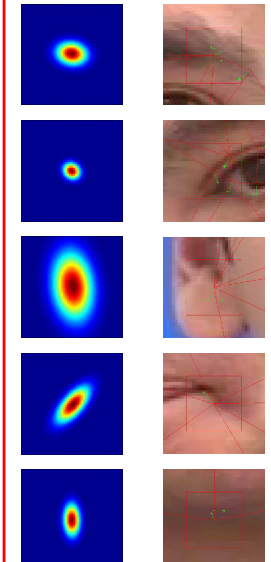
- Seed Weights

$w_i = \rho(C_i \{x_i + \Delta x_i\}, \bar{C}_i \{x_i\})^{-1}$



### Unsupervised Clustering

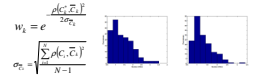
- Unsupervised learning of finite mixture models, IEEE TPAMI, M.Figueiredo and A.Jain



### Selecting the Best Cluster

$x_i^* = \arg \min \sqrt{(x_i^m - x_i)^T \Sigma_i^{-1} (x_i^m - x_i)}$

- Landmark Matching Score



## Algorithm:

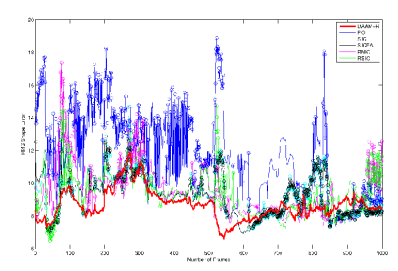
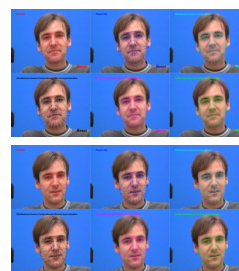
- Start with na estimate for the position of the face [Adaboost]
- For each landmark k

### Offline Computations:

- PDM  $(s_0, \Phi, \Psi)$
- Jacobians of the Warp  $(\frac{\partial W(x, p, q)}{\partial p}, \frac{\partial W(x, p, q)}{\partial q})$
- Shape Location Covariance  $(\Sigma_i)$
- Average Covariances  $(\bar{C}_i)$
- Statistics of Dissimilarity  $(\sigma_i)$

- Generate a PDM instance  $s = s_0 + \sum p_i \phi_i + \sum q_j \psi_j$
- Warp image into the base mesh  $I(x) \rightarrow I(W(x, p, q))$
- Response maps by exhaustive local search  $\rho(C_i \{x_i + \Delta x_i\}, \bar{C}_i \{x_i\})$
- Use mean-shift to find models (local minima)
- Unsupervised search for the clusters
- Select the best cluster  $x_i^*$
- Assign landmark matching weight  $w_i = e^{-\frac{\rho(C_i, \bar{C})}{2\sigma_i}}$
- Find weighted warp update  $\Delta p = (\Phi W \Phi^T)^{-1} \Phi W \Delta x^*$   
 $\Delta q = (\Psi W \Psi^T)^{-1} \Psi W \Delta x^*$
- Update shape and pose parameters  $p \leftarrow p + \Delta p \quad q \leftarrow q + \Delta q$

## Evaluation on the Talking Face Sequence



This work was supported by the Portuguese Science Foundation (FCT) by the project "Dinâmica Facial 4D para Reconhecimento de Identidades" with grant PTDC/EIA-CCO/108791/2008. Pedro Martins also acknowledges the FCT for support through the grant SFRH/BD/47178/2008.



Institute of Systems and Robotics  
 University of Coimbra