UNIVERSIDADE DE COIMBRA
UNIVERSITY OF COIMBRA

ECCV 2012

Discriminative Bayesian Active Shape Models

Pedro Martins, Rui Caseiro, João F. Henriques, Jorge Batista

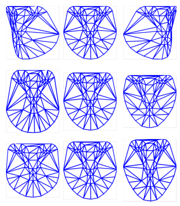
Institute of Systems and Robotics - University of Coimbra - Portugal

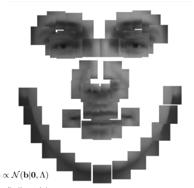
Overview:

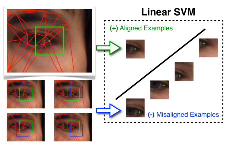
- **Goal:** Face alignment in unseen images.
- Closely related to Constrained Local Models (CLM) and Active Shape Models (ASM), where a set of local detectors is constrained to lie in the subspace spanned by a Point Distribution Model (PDM).
- Two step fitting approach:
  - (1) Local search using the local detectors (response maps for each landmark).
  - (2) Global optimization strategy that finds the PDM parameters that jointly maximize all the detection at once.
- New Bayesian global optimization strategy using second order statistics of the shape and pose parameters.

The Shape (PDM) and Appearance Models

$\mathbf{s} = (x_1, y_1, \dots, x_v, y_v)^T$ Patch-Based Appearance Ensemble of Local Detectors



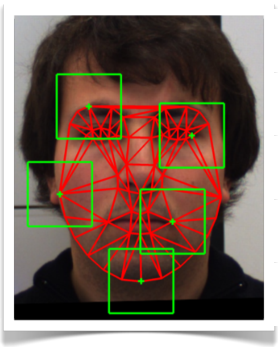


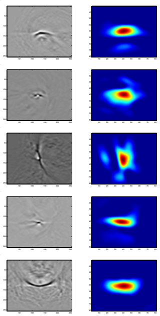


$\mathcal{D}_i^{\text{linear}}(\mathbf{I}(\mathbf{y}_i)) = \mathbf{w}_i^T \mathbf{I}(\mathbf{y}_i) + b_i$

$\mathbf{s} = \mathcal{S}(\mathbf{s}_0 + \Phi\mathbf{b}, \mathbf{q})$ → (**b**) shape and (**q**) pose parameters

Local Detectors (MOSSE Filters)

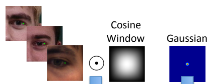




$\mathcal{F}^{-1}\{\mathbf{H}_i^*\}$ $p_i(\mathbf{z}_i)$

Correlation in Fourier Domain

$\mathbf{G} = \mathcal{F}\{\mathbf{I}\} \odot \mathbf{H}^*$



$$\min_{\mathbf{H}^*} \sum_{j=1}^{N} (\mathcal{F}\{\mathbf{I}_j\} \odot \mathbf{H}^* - \mathbf{G}_j)^2$$

MOSSE Filter

$$\mathbf{H}^* = \frac{\sum_{j=1}^{N} \mathbf{G}_j \odot \mathcal{F}\{\mathbf{I}_j\}^*}{\sum_{j=1}^{N} \mathcal{F}\{\mathbf{I}_j\} \odot \mathcal{F}\{\mathbf{I}_j\}^*}$$

$$\mathcal{D}_i^{\text{MOSSE}}(\mathbf{I}(\mathbf{y}_i)) = \mathcal{F}^{-1}\{\mathcal{F}\{\mathbf{I}(\mathbf{y}_i)\} \odot \mathbf{H}_i^*\}$$

The Alignment Goal

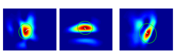
Given a shape observation (**y**), find the optimal set of shape (**b**) and pose parameters that maximize the posterior probability

$$\mathbf{b}^* = \arg\max_{\mathbf{b}} p(\mathbf{b}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{b})p(\mathbf{b})$$

Assuming:

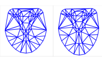
① Conditional independence between landmarks
② Close to a solution

$$p(\mathbf{b}|\mathbf{y}) \propto \left(\prod_{i=1}^{v} p(\mathbf{y}_i|\mathbf{b})\right) p(\mathbf{b}|\mathbf{b}_{k-1}^*)$$



Likelihood from the local detectors

Prior from the PDM parameters transition



The Likelihood Term

Observed shape (**y**)

$$p(\mathbf{y}|\mathbf{b}) \propto \exp\left(-\frac{1}{2}(\underbrace{\mathbf{y} - (\mathbf{s}_0 + \Phi\mathbf{b})}_{\Delta\mathbf{y}})^T \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - (\mathbf{s}_0 + \Phi\mathbf{b}))\right)$$

Likelihood follow a Gaussian distribution

$p(\mathbf{y}|\mathbf{b}) \propto \mathcal{N}(\Delta\mathbf{y}|\Phi\mathbf{b}, \Sigma_{\mathbf{y}})$

Uncertainty covariance $\Sigma_{\mathbf{y}}$

$2v \times 2v$ Block diagonal

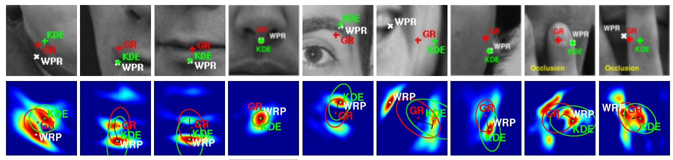
The Prior Term

$p(\mathbf{b}_k|\mathbf{b}_{k-1}) \propto \mathcal{N}(\mathbf{b}_k|\mu_{\mathbf{b}}, \Sigma_{\mathbf{b}})$

$\mu_{\mathbf{b}} = \mathbf{b}_{k-1}$

$\Sigma_{\mathbf{b}} = \Lambda + \Xi$ PCA eigenvalues + additive dynamic noise

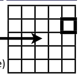
Local Optimization Strategies (Finding the Likelihood Parameters)



Patches under occlusion

$p_i(\mathbf{z}_i)$ Prob. $\mathbf{z}_i$ is aligned

$\mathbf{y}_i^c$ Center (current mesh estimate)



$\mathbf{z}_i = (x_i, y_i)$ Pixel candidate to the i[th] landmark location

$\Omega_{\mathbf{y}_i^c}$ Local search grid (L x L)

Weighted Peak Response (WPR)

$$\mathbf{y}_i^{\text{WPR}} = \max_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} (p_i(\mathbf{z}_i))$$

$$\Sigma_{\mathbf{y}_i}^{\text{WPR}} = diag(p_i(\mathbf{y}_i^{\text{WPR}})^{-1})$$

Gaussian Response (GR)

$$\mathbf{y}_i^{\text{GR}} = \frac{1}{d} \sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} p_i(\mathbf{z}_i)\mathbf{z}_i \qquad d = \sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} p_i(\mathbf{z}_i)$$

$$\Sigma_{\mathbf{y}_i}^{\text{GR}} = \frac{1}{d-1} \sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} p_i(\mathbf{z}_i)(\mathbf{z}_i - \mathbf{y}_i^{\text{GR}})(\mathbf{z}_i - \mathbf{y}_i^{\text{GR}})^T$$

Kernel Density Estimator (KDE)

$$\mathbf{y}_i^{\text{KDE}(\tau+1)} \leftarrow \frac{\sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} \mathbf{z}_i \, p_i(\mathbf{z}_i) \, \mathcal{N}(\mathbf{y}_i^{\text{KDE}(\tau)}|\mathbf{z}_i, \sigma_{h_j}^2 \mathbf{I}_2)}{\sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} p_i(\mathbf{z}_i) \, \mathcal{N}(\mathbf{y}_i^{\text{KDE}(\tau)}|\mathbf{z}_i, \sigma_{h_j}^2 \mathbf{I}_2)}$$

$$\Sigma_{\mathbf{y}_i}^{\text{KDE}} = \frac{1}{d-1} \sum_{\mathbf{z}_i \in \Omega_{\mathbf{y}_i^c}} p_i(\mathbf{z}_i)(\mathbf{z}_i - \mathbf{y}_i^{\text{KDE}})(\mathbf{z}_i - \mathbf{y}_i^{\text{KDE}})^T$$

2nd Order MAP Global Alignment (DBASM)

Likelihood

$$p(\mathbf{y}|\mathbf{b}) \propto \mathcal{N}(\Delta\mathbf{y}|\Phi\mathbf{b}, \Sigma_{\mathbf{y}})$$

Mean is function of **b,** covariance independent of **b**

Bayes's Theorem for Gaussian variables

Prior

$$p(\mathbf{b}_k|\mathbf{b}_{k-1}) \propto \mathcal{N}(\mathbf{b}_k|\mu_{\mathbf{b}}, \Sigma_{\mathbf{b}})$$

Posterior

$$p(\mathbf{b}_k|\mathbf{y}) \propto \mathcal{N}(\mathbf{b}_k|\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma} = (\Sigma_{\mathbf{b}}^{-1} + \Phi^T \Sigma_{\mathbf{y}}^{-1} \Phi)^{-1}$$

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}(\Phi^T \Sigma_{\mathbf{y}}^{-1} \mathbf{y} + \Sigma_{\mathbf{b}}^{-1} \mu_{\mathbf{b}})$$

2nd order estimate

Model the Covariance of **b**, by a Linear Dynamic System (LDS)

$$\mathbf{b}_k = \mathbf{A}\mathbf{b}_{k-1} + q, \quad q \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{b}})$$

$$\Delta\mathbf{y} = \Phi\mathbf{b}_k + r, \quad r \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{y}})$$

Posterior

$$p(\mathbf{b}_k|\mathbf{y}_k, \dots, \mathbf{y}_0) \propto \mathcal{N}(\mathbf{b}_k|\boldsymbol{\mu}_k^{\mathbf{F}}, \boldsymbol{\Sigma}_k^{\mathbf{F}})$$

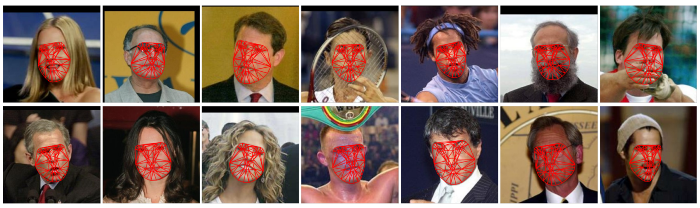
LDS update equations:

$$\mathbf{K} = \mathbf{P}_{k-1}\Phi^T(\Phi\mathbf{P}_{k-1}\Phi^T + \Sigma_{\mathbf{y}})^{-1}$$

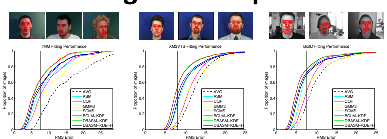
$$\boldsymbol{\mu}_k^{\mathbf{F}} = \mathbf{A}\boldsymbol{\mu}_{k-1}^{\mathbf{F}} + \mathbf{K}(\mathbf{y} - \Phi\mathbf{A}\boldsymbol{\mu}_{k-1}^{\mathbf{F}})$$

$$\boldsymbol{\Sigma}_k^{\mathbf{F}} = (\mathbf{I}_n - \mathbf{K}\Phi)\mathbf{P}_{k-1}$$

Qualitative Results - Labeled Faces in the Wild

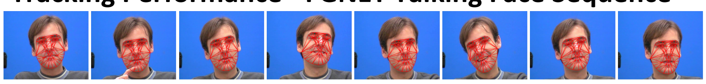


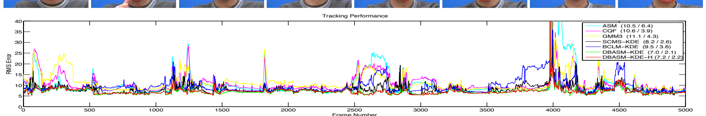
Evaluating Global Optimization Strategies



Reference 7.5 RMS	IMM (240 images)		XM2VTS (2360 images)		BioID (1521 images)	
ASM	50.0		30.7		70.0	
DBASM-WPR (our method)	**56.7**	(+6.7)	**45.1**	(+14.4)	**75.4**	(+5.4)
CQF	45.4		10.9		47.0	
GMM3	40.8	(-4.6)	10.4	(-0.5)	51.7	(+4.7)
BCLM-GR	48.3	(+2.9)	15.9	(+5.0)	54.2	(+7.2)
DBASM-GR (our method)	**50.4**	(+5.0)	**18.0**	(+7.1)	**62.2**	(+15.2)
SCMS-KDE	54.6		35.7		69.0	
BCLM-KDE	57.1	(+2.5)	43.4	(+7.7)	71.9	(+2.9)
DBASM-KDE (our method)	**64.6**	(+10.0)	**54.5**	(+18.8)	**76.5**	(+7.5)
DBASM-KDE-H (our method)	**64.6**	(+10.0)	53.5	(+17.8)	**76.5**	(+7.5)

Tracking Performance - FGNET Talking Face Sequence





INSTITUTE OF SYSTEMS AND ROBOTICS
UNIVERSITY OF COIMBRA
FCTUC FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

This work was supported by the Portuguese Science Foundation (Fundação para a Ciência e Tecnologia - FCT), under the research project "Dinâmica Facial 4D para o Reconhecimento de Identidade"(grant PTDC/EIA-CCO/108791/2008). Pedro Martins, Rui Caseiro, and João F. Henriques also acknowledge the FCT through the PhD grants SFRH/BD/45178/2008, SFRH/ BD74152/2010 and SFRH/BD/75459/2010, respectively.