IEEE International Conference on Image Processing 2014

UNIVERSIDADE DE COIMBRA
UNIVERSITY OF COIMBRA

Likelihood-Enhanced Bayesian Constrained Local Models

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

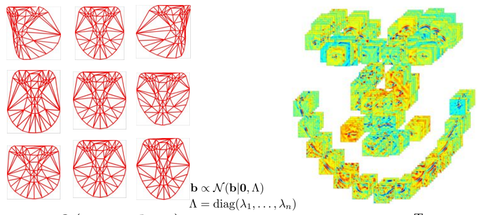
Institute of Systems and Robotics - University of Coimbra - Portugal

Overview:

- **Goal:** Non-rigid face alignment in unseen images.
- The Constrained Local Models (CLM): combine a set of local detectors and a global optimization strategy that constrains the feature points to be in the subspace spanned by a linear shape model (Point Distribution Model - PDM).
- CLM two step fitting approach:
  - (1) Local search using the detectors (likelihood map for each landmark).
  - (2) Global optimization strategy that estimates the shape parameters that jointly maximize all the detections.
- New constrained clustering stage designed to learn multiple sets of local detectors per landmark.

CLM: Shape (PDM) and Local Detectors

$\mathbf{s} = (x_1, y_1, \dots, x_v, y_v)^T$ $\quad \{\mathbf{h}_1^{(m)}, \dots, \mathbf{h}_v^{(m)}\}_{m=1,\dots,M}$

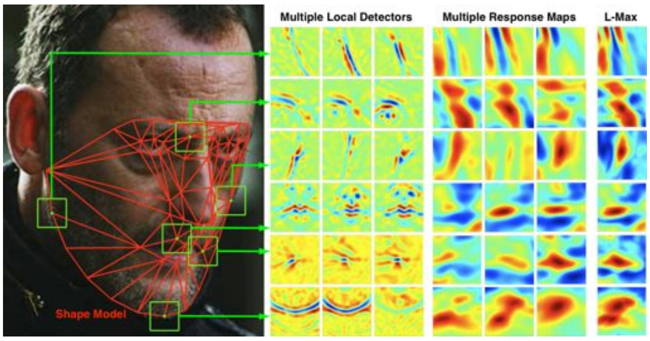


$\mathbf{b} \propto \mathcal{N}(\mathbf{b}|\mathbf{0}, \Lambda)$, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$

$$\mathbf{s} = \mathcal{S}(\mathbf{s}_0 + \Phi\mathbf{b}, \mathbf{q}) \qquad \mathcal{D}_i(\mathbf{I}(\mathbf{x}_i)) = \mathbf{h}_i^T \mathbf{I}(\mathbf{x}_i)$$

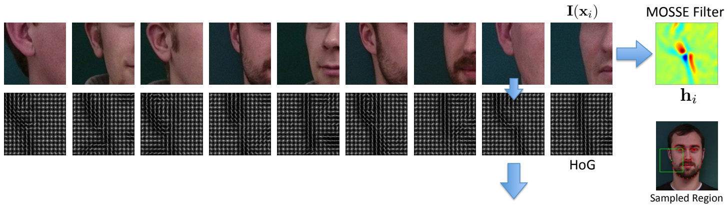
- (**b**) shape and (**q**) pose parameters

Multiple Local Detectors



$\mathbf{h}_i^{(1)} \ \dots \ \mathbf{h}_i^{(M)} \quad \mathbf{p}_i(\mathbf{z}_i)^{(1)} \ \dots \ \mathbf{p}_i(\mathbf{z}_i)^{(M)} \ \mathbf{p}_i(\mathbf{z}_i)_\infty$

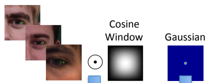
Learning Multiple Local Detectors



Local Detector (MOSSE)

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

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

Multiple Response Maps

$$p_i(\mathbf{z}_i)^{(m)} = \frac{1}{1 + e^{-a_i \beta_1 \mathcal{D}_i(\mathbf{I}(\mathbf{z}_i)) + \beta_0}}$$

Combining Multiple Detections

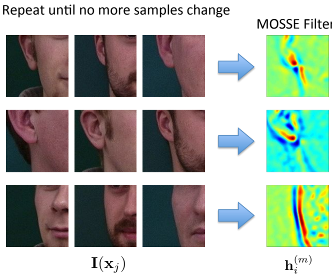
$$p_i(\mathbf{z}_i)_\infty = \max_{\mathbf{z}_i}\{p_i(\mathbf{z}_i)^{(1)}, \dots, p_i(\mathbf{z}_i)^{(M)}\}$$

Constrained Clustering

$$\arg\max_{\mathbf{h}_i^{(m)}} \sum_{j=1}^{N} \sum_{m=1}^{M} \mathbf{I}(\mathbf{x}_j) * \mathbf{h}_i^{(m)}$$

M - clusters
N - examples

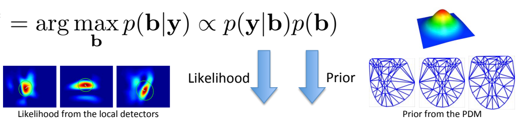
Solve (for each landmark i) using a two step approach:
- Initial clustering by k-means
① Build basic detectors using the current clustering estimate
② Move samples to the cluster with highest correlation
- Repeat until no more samples change



The Alignment Goal

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

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



$$p(\mathbf{y}|\mathbf{b}) \propto \mathcal{N}(\Delta\mathbf{y}|\Phi\mathbf{b}, \Sigma_{\mathbf{y}}) \qquad p(\mathbf{b}_k|\mathbf{b}_{k-1}) \propto \mathcal{N}(\mathbf{b}_k|\mathbf{0}, \Lambda)$$

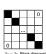
$$p(\mathbf{y}|\mathbf{b}) \propto \exp\left(-\frac{1}{2}(\underbrace{\mathbf{y} - (\mathbf{s}_0}_{\Delta\mathbf{y}} + \Phi\mathbf{b}))^T \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - (\mathbf{s}_0 + \Phi\mathbf{b}))\right)$$

Observed shape (**y**)

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

Uncertainty Covariance

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



Posterior Distribution

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

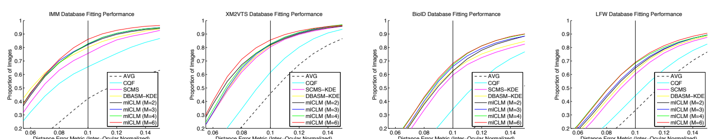
Inference by a Linear Dynamic System (LDS)

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

$$\mathbf{K} = \mathbf{P}_{k-1}\Phi^T(\Phi\mathbf{P}_{k-1}\Phi^T + \Sigma_{\mathbf{y}})^{-1}$$
$$\boldsymbol{\mu}_k^{\mathbf{F}} = \mathbf{A}\boldsymbol{\mu}_{k-1}^{\mathbf{F}} + \mathbf{K}(\mathbf{y} - \Phi\mathbf{A}\boldsymbol{\mu}_{k-1}^{\mathbf{F}})$$
$$\mathbf{\Sigma}_k^{\mathbf{F}} = (\mathbf{I}_n - \mathbf{K}\Phi)\mathbf{P}_{k-1}$$

Qualitative Results - Labeled Faces in the Wild (LFW)

Evaluation - Fitting Performance Curves



INSTITUTE OF SYSTEMS AND ROBOTICS
UNIVERSITY OF COIMBRA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Work supported by the Portuguese Science Foundation (Fundação para a Ciência e Tecnologia - FCT), under the research project "DG2CVPR - Differential Geometry for Computer Vision and Pattern Recognition" (grant PTDC/EIA-CRO/122812/2010). Pedro Martins, Rui Caseiro, and João F. Henriques also acknowledge the FCT through the grants SFRH/BDP/90200/2012, SFRH/ BD74152/2010 and SFRH/BD/75459/2010.