Bayesian Tangent Shape Model
For Face Alignment

Leon Gu
Carnegie Mellon University

Presented at CVPR 2003, Wisconsin
Face Alignment

Locate shape structure

Shape

Pose

Geometrical Transform

- Rotation
- Translation
- Scale
- ......
Observation and Regularization

Starting Approximation

Observation

Regularization
## Related Work

<table>
<thead>
<tr>
<th>Active Contour or Snakes</th>
<th>Active Shape Model</th>
</tr>
</thead>
</table>

- Spline curves
- Stretchness and smoothness
- General amorphous objects
- Landmark points
- Prior Density
- Object with specific structure
Our Method

• Prior knowledge: same as ASM

• Regularization rules

  • Shape: a weighted average of regularized shape and observed shape.

  • Shape parameters: continuously regularized by multiplying a shrinking factor.

  • Pose parameters: constrained by observation noise.

• Convergence guaranteed by EM
Results
Shape Spaces

\[ Y \rightarrow \text{Align} \rightarrow X \rightarrow \text{PCA-Proj} \rightarrow \Phi, b \]

- \( Y \): Shape Space
- \( X \): Tangent Space
- \( \Phi, b \): Principal Subspace
BTSM Formulation

- **b**: shape parameter
- **ε**: shape noise
- **X**: tangent shape
- **Y**: observed shape
- **θ**: pose parameter
- **η**: observation noise

**X**: Tangent Shape

**Y**: Observed Shape

**p(X)**: Prior Model

**p(Y|X)**: Observation Model
Prior Model $P(X)$

\[ x = \mu + \Phi_r b + \varepsilon \]

- **Isotropic shape noise**: 
  \[ \varepsilon \sim N(0, \sigma^2 I) \]

- **Variance**: 
  \[ \sigma^2 = \frac{1}{n} \sum_{i=r+1}^{n} \lambda_i \]

- **Geometrical view**:

Prior model:
- $b$: shape parameter
- $\varepsilon$: isotropic noise
- $X$: tangent shape

Diagram showing the relationship between $x$, $\mu$, $\Phi_r b$, and $\varepsilon$. The diagram illustrates the geometrical view of the prior model.
Why Model $\varepsilon$

Work with Principal Subspace Only?

- Distance outside Principal Subspace
- Compensate noise
Observation Model \( P(Y|X) \)

\[ y = T_\theta (x) + \eta \]

\[ \eta \sim N(0, \Sigma), \]
\[ \Sigma = diag(\rho_1^2, \ldots, \rho_N^2) \otimes I_2 \]
\[ \rho_i = D(T_{\theta^{old}}(x_{i^{old}}) , y_i) \]

**Observation Model**

- \( X \): tangent shape
- \( \theta \): pose parameter
- \( \eta \): observation noise
- \( Y \): observed shape
Posterior

$$p(b, c, s, \theta \mid y) \propto \exp\left\{-\frac{1}{2}[(\sigma^2 + s^{-2} \rho^2)^{-1}(\Phi_r T_{\theta}^{-1}(y) - b)^2 + (\Phi_r T_{\theta}^{-1})(y) \|^2) \right.$$ 
$$\left. + s^2 \rho^{-2} \| A^T T_{\theta}^{-1}(y) \|^2 + b^T \Lambda^{-1}b\} \cdot \frac{\text{const}}{(\sigma^2 + s^{-2} \rho^2)^{(N-2)} s^{-4} \rho^4}\right.$$ 

Q-Function

$$\log p(b, s, c, \theta \mid x, y) = \log p(b \mid x) + \log p(\gamma \mid x, y)$$
$$= -\frac{1}{2} \{b^T \Lambda^{-1}b + \sigma^{-2} \| x - \mu - \Phi_r b \|^2\} - \frac{1}{2} \rho^{-2} \| y - X \gamma \|^2 + \text{const}$$

where $\gamma = (c_1, c_2, s \times \cos \theta, s \times \sin \theta)^T$ and $X = (x, x^*, e, e^*)$.

Expectation-Maximization

E-Step: computing $\langle x \rangle, \langle \|x\|^2 \rangle$ 
M-Step: estimate $b$ and $\theta$
E-Step

Updating Tangent Shape

\[ X = (1 - p) \times (\Phi_r b) + p \times \Phi \Phi^T T_{\theta}^{-1} \]

Weight \( p \)

\[ p = \frac{E_{\text{shp noise}}}{E_{\text{shp noise}} + s^{-2} E_{\text{obs noise}}} \]

\[ \rho_i = D \left( T_{\theta}^{old} (x_i^{old}) - y_i \right) \]

\[ \sigma^2 = \frac{1}{n} \sum_{i=r+1}^{n} \lambda_i \quad s: \text{scaling factor} \]
M-Step

Update Shape Parameter

Shape para: \( b_i = \alpha_i \left( \Phi_r^T x \right)_i \)

Shrinking coeff: \( \alpha_i = \frac{\lambda_i}{\lambda_i + \sigma^2} \), \( \sigma^2 = \frac{1}{n} \sum_{i=r+1}^{n} \lambda_i \)

"continuous regularization by SNR"

Update Pose Parameter

Pose para: \( \theta = \arg \min \sum w_i \left( T_\theta(x_i) - y_i \right)^2 \)

"weighted procrustes analysis"

Weight: \( w_i = \frac{1}{\rho_i^2} \), \( \rho_i = D \left( T_{\theta^{old}}(x_i^{old}) - y_i \right) \)
More Results
Precision: BTSM vs ASM

- 870 manually labeled face (training: 599, testing: 271)
- X: the index of testing faces; Y: dist(BTSM) – dist(ASM)
Comparison

ASM

BTSM

ASM

BTSM
Numerical Stability

![Graph showing numerical stability comparison between BTSM and ASM](image)

- BTSM
- ASM

![Error vs. # of iter graph](image)

- b2 in ASM
- b2 in BTSM

Error

# of iter
Summary

• Two Simple Ideas
  • De-Noising by Shrinkage
    • Suppress noise
    • Preserve major shape deformations
  • Penalize outliers by iterative re-weighting

• Pro’s
  • Well generalized to novel, unseen faces
  • Robust to image noise
  • Fast (from 30ms to 170ms); fully automatic

• Con’s
  • Relies on face detector for good initialization
  • Limited to frontal faces