Active Shape Model Unleashed with Multi-scale Local Appearance

Qiang Zhang\textsuperscript{1}, Abhir Bhalerao\textsuperscript{1}, Emma Helm\textsuperscript{2}, Charles Hutchinson\textsuperscript{2}

\textsuperscript{1} Computer Science, University of Warwick, UK
qiang@dcs.warwick.ac.uk

\textsuperscript{2} University Hospitals Coventry and Warwick Share, Coventry, UK
Three modifications to ASM

- **Shape** optimisation by energy minimisation
  
  Enhance the performance with very limited training samples

- **Adopt the subspace inverse gradient descent for local feature searching**
  
  Model and remove the variations of local appearance presented in the training data

- **Multi-scale local features for a seamless coarse-to-fine fitting**
  
  Integrate landmark observations from multi-scale features
1. Retain first $t$ significant shape variations

$$P_s = [e_1, e_2, ...]$$

2. Bound the shape variations $b_s$ within certain s.d.

$$s = \bar{s} + P_s b_s$$
Active Shape Model

Constraint

1. Retain first $t$ significant shape variations

2. Bound the shape variations $b_s$ within certain s.d. of mean shape

Drawback:

- Over constraint with high dimensional variations and limited training data
Shape optimisation

Active Shape Model

Constraint

- 1. Retain first $t$ significant shape variations
- 2. Bound the shape variations $b_s$ within certain s.d. of mean shape

Causes: (why)

1. Does not count the cost of cutting off minor components
   
   ASM assume the minor components are noises, but they could also be the shape details.

2. Shape variations outside the rigid boundary could still represent reasonable instances
   
   Shape prior can not cover all the variations without adequate training samples, which is always the truth in medical images.
Shape optimisation by energy minimisation

Solution:

\[ E(s) = E_{\text{shape}}^\text{DIFS}(s) \]

Distance \textbf{In} feature space [1]:
Distances from the statistical mean shape

\[ + E_{\text{shape}}^\text{DFFS}(s) \]

Distance \textbf{From} feature space [1]:
Energy of the minor variations been cut off

\[ + \beta E_{\text{image}}(s) \]

\[ E_{\text{image}}(s) = \sum_{j=1}^{2N} \frac{(x_j - \hat{x}_j)^2}{2\sigma_j^2}, \]

Shape optimisation by energy minimisation

What is $\beta$

$$E(s) = E_{shape}^{DIFS}(s) + E_{shape}^{DFFS}(s) + \beta E_{image}(s)$$
Lucas & Kanade algorithm for feature searching

Basic idea:

During training, the large appearance variations of local features can be modelled in an eigen-space.

During testing, to remove the large appearance variations, the local features of the testing image is projected onto the orthogonal space of the eigen-space, which result in a more robust fitting.
Multi-scale local features for a seamless fitting

What is multi-scale local features

A landmark can be described by local features at multiple scales, we combine them for a more comprehensive description.
Convergence range and precision at multiple scales

Why multi-scale
Combine multi-scale local features

- **Product of Gaussian**

likelihood: \( p(\Delta x) \sim \{\mathcal{N}(\Delta x_1, \sigma_1^2), ..., \mathcal{N}(\Delta x_L, \sigma_L^2)\} \)
Combine multi-scale local features

How to combine the multi-level observations

likelihood: $p(\Delta x) \sim \{\mathcal{N}(\Delta x_1, \sigma^2_1), \ldots, \mathcal{N}(\Delta x_L, \sigma^2_L)\}$

$$E_{image}(s) = \frac{1}{2} (s - \hat{s})^T \Sigma^{-1} (s - \hat{s})$$

$$E_{image}(s) = \frac{1}{2} \sum_{l=1}^{L} (s - \hat{s}_l)^T \Sigma_l^{-1} (s - \hat{s}_l),$$

An optimal shape with respect to multiple observations of landmarks at multi-scales
Shape optimisation vs. standard shape constraint

![Table](image)

Performance of three methods on segmentation of vertebrae and knees, trained with reducing $m$ samples


Experiments

Multi-scale vs. coarse-to-fine

- Convergence speed

<table>
<thead>
<tr>
<th>n per level</th>
<th>Failure rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coarse to fine:</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>12.5%</td>
</tr>
<tr>
<td>3</td>
<td>10.7%</td>
</tr>
<tr>
<td>5</td>
<td>5.3%</td>
</tr>
<tr>
<td>Our Method:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.6%</td>
</tr>
</tbody>
</table>

Convergence speed and failure rate of multi-scale strategy compared with coarse-to-fine methods
Multi-scale vs. coarse-to-fine

- Local feature variances and robustness

(a) Initial position  (b) Converged  (c) Partial occlusion

- Cyan ellipses show the variance of observations, red crosses show the optimal shape. Salient features have smaller variances therefore larger weights.
Conclusion

- **Higher precision in local feature searching.**
- **Better performance with limited training samples.**
  
  Shape is optimal in a statistical sense taking both the shape prior and feature confidence into consideration. The detail of salient features is better preserved.

- **Seamless and rapid convergence process.**
  
  Converge faster with lower failure rate. Larger scale features keep the shape from local minima, while the smaller scales take effect as soon as it gets into the convergence range.
Active Shape Model Unleashed with Multi-scale Local Appearance

Qiang Zhang
University of Warwick
qiang@dcs.warwick.ac.uk

Thank you