**Image Segmentation using Statistical Shape Models**

Prof Tim Cootes
Centre for Imaging Sciences
The University of Manchester

---

**Aim of Research**

Extract useful information from images
- often using statistical models
- ‘explain’ the image

---

**Why is it difficult?**

- Wide variation in shape and appearance

---

**Model Building**

Example Images

Statistical Analysis

Shape Models

Appearance Models

---

**Building Shape Models**

- Require labelled training images
  – landmarks represent correspondences

\[ \mathbf{x} = (x_1, \ldots, x_n, y_1, \ldots, y_n)^T \]

---

**Aligning Shapes**

- Generalised Procrustes Analysis
  – Align each shape to the mean

**Sample Images**

- Unaligned
- Align with translation
- Align with similarity
Building Shape Models
- Given aligned shapes, \( \{x_i\} \)
- Apply PCA
  - Compute mean and eigenvectors of covar.
  \[
x = \bar{x} + b_1p_1 + b_2p_2 + b_3p_3 + \cdots
\]
- \( P \) – First \( t \) eigenvectors of covar. matrix
- \( b \) – Shape model parameters

Spine Shape Model
\[x = \bar{x} + b_1p_1 + b_2p_2 + b_3p_3 + \cdots\]
\( \leftarrow b_1 \rightarrow \leftarrow b_2 \rightarrow \leftarrow b_3 \rightarrow \)

Face Model Modes
- 100 people, 4 images each
\[
x = \bar{x} + b_1p_1 + b_2p_2 + b_3p_3 + \cdots
\]

Spine Shape Model

Knee shape model

Shape Model of the Pelvis
\[x = \bar{x} + b_1p_1 + b_2p_2 + b_3p_3 + \cdots\]
\( \leftarrow b_1 \rightarrow \leftarrow b_2 \rightarrow \leftarrow b_3 \rightarrow \)

3D Models
Generated by Imorphics Ltd
### Appearance Models

- Model both shape and texture
- Use all image information

\[ \text{Appearance Model of Pelvis} \]

### Global Face Model

- 400 images from 100 different people

Note: Mixes ID, expression, head pose etc

### Automatic Annotation

- Accurate correspondences required to build shape/appearance models
  - Slow/hard to manually annotate data
Goal: Fully Automatic System

- Build models from unlabelled images

**Approach**

Find the correspondences across the set of images

Construct model in an ‘average’ reference frame

**Representation of Deformation**

- Use piece-wise linear warp field
  - Triangular/Tetrahedral mesh
  - Simple, compact, efficient
  - Trivially invertible

**Groupwise Algorithm**

- Initialisation: Affine registration
- Generate control points
- Repeat
  - Build model from current data
  - For each image
    - Optimise positions of control points
- Until happy/dead/conference deadline etc

**Registering Faces**

- 293 individuals from XM2VTS
- Evolution of robust estimate of mean:

**Example Modes**

Varying appearance model parameters
(Note shadowing caused by glasses)

Modes of model built from images without glasses:
3D Registration

- Registering 270 3D MR Brain Images
- Evolution of the model mean:

Initialisation

- Affine registration not always sufficient
- Problem:
  - Many similar structures, so many local minima

Improved Initialisation

- Parts+Geometry models can efficiently consider many candidates
- Automatically construct parts+geometry model to initialise dense registration

Results of Parts+Geometry Matching

- Model
- Matches on three images

Automatic Model building

- Challenge is to select good set of parts
- Use Genetic Algorithm style optimisation:
  - Create set of different models
    - Each a random selection of parts
  - Repeat
    - Evaluate each model
    - Discard worst 50%
    - Combine good examples to form new set
  - Until happy

Best 3 part model

- Model
- Initial Mean
- Mean after dense registration
**Shape Models**

- Model from automatic registration

**Image Interpretation**

- Match model to new image
- Allows analysis of shape of object

**Matching Algorithms**

Typical approach:
- Global search to find approx. position
- Local search:
  - Optimise pose/parameters of model
  - Maximise fit of model to image data

**Model Matching using Regression**

- Estimate parameter update from current image sample

\[
\delta p
\]

\[
\hat{\delta p}
\]
Estimating updates

• Assume functional relationship
  \[ \delta \mathbf{p} = F(s) \]
  \( s \) = normalised intensities, or residuals

• Learn function from displaced training examples
  \( \{ \delta \mathbf{p}_i \} \)
  \( \{ s_i = s(p_{true} + \delta \mathbf{p}_i) \} \)

  \( \text{Random perturbations} \)
  \( \text{Image samples} \)

• Simplest approach: Linear Regression
  \[ \delta \mathbf{p} = \mathbf{R} \mathbf{s} \]

Active Appearance Model Algorithm

• Initial estimate \( \mathbf{I}_{init}(\mathbf{p}) \)
• Start at coarse resolution
• At each resolution
  – Measure residual error, \( r(\mathbf{p}) \)
  – Predict correction \( \delta \mathbf{p} = \mathbf{R} r \)
  – \( \mathbf{p} \Rightarrow \mathbf{p} - \delta \mathbf{p} \)
  – Repeat to convergence

AAM Tracking

• Sequence of AAMs with increasing resolution
• Generic – trained on range of people
• Search each frame in a few ms

Non-linear Regression

• Boosted Haar features:
  \[ \delta \mathbf{p}_j = \sum_{i=1}^{n} f_i(H(s)) \]
• Random Forests:
  \[ \delta \mathbf{p} = \frac{1}{n} \sum_{i=1}^{n} f_i(H(s)) \]
• Boosted Ferns
• ...

Boosted Sequences of Regressors

• Extension of multi-resolution approach
  – No iterations
• Supervised Descent Method [Xiong:CVPR2013]
  – Linear regression on non-linear features
• Explicit Shape Regression [Cao:CVPR12]
  – Boosted ferns/trees on pixel difference features

Local Model Matching

• AAM and variants: Predict whole shape
  – Or parts of shape
• Alternative: Predict position of each point
  – Lower dimensional output space
  – Easier to train
  – Potential to be more robust
Local Model Matching

\[ Q(b, t) = \sum_{i=1}^{n} R_i(T_i(x_i + P_i b)) \]

Find shape and pose parameters to optimise:

Classification vs Regression

Classification:
- Is this the point?

Regression:
- Where is the point?

Finding Points using Regression Voting

- Train “Random Forest” to predict offsets
- Scan regressor across ROI
- Each patch produces 1 vote per tree
- Accumulate votes in an array

Hand Radiographs

Claudia Lindner, Steve Adeshina

Varying \( b_1 \)

Varying \( b_2 \)
BoneFinder

Running fully automatic search...

Lindner et al., IEEE Transactions on Medical Imaging, 2013.

Fully automated femur detection

- Tested on 839 radiographs (ARCOGEN)

Error <0.9 mm for 99% of cases

Claudia Lindner

Knee Radiographs
Claudia Lindner, Xenios Milidonis

Fully automatic search results using AP knee radiographs.

Mean point-to-curve error:
<1 mm for 99% of 500 images

Median
95%ile
99%ile

Vertebral Fractures
Paul Bromiley, Judith Adams

T7: Grade 3
T8: Grade 3
T9: Grade 1
T10: Grade 2
T11: Grade 1
T12: Grade 1
L1: Grade 3
L2: Grade 1
L3: Grade 1
L4: Grade 1

Manual
Automatic

Performance vs. Classification

Black = normal
Red = deformed
Green = Grade 1
Blue = Grade 2
Cyan = Grade 3
Cyan w. points = AAM, Grade 3
Landmarks on Cephalograms

Claudia Lindner

Mode 1

Mode 2

Mode 3

ISBI Grand Challenge 2015

Facial Feature Tracking

Robust Facial Tracking

Conclusion

• Shape models powerful tools for image analysis
• Fast, robust, accurate matching algorithms now available

Tools available: www.bone-finder.com

Appearance modeling/AAM tools: (see my website)
P.S. It works on faces too

Active Shape Model

• Local optimisation
• Initialise near target
  – Search nearby for best match, $X'$
  – Update parameters to match to $X'$.

$(X', Y')$
Matching 3D ASMs

- Search for each point independently
  - Local model depends on current pose
- Estimate shape parameters + projection

Face Tracking with a 3D Model