Model Based Vision

Dr A Galata

Many slides from Dr Carole Twining,
University of Manchester

Model-Based Vision: Summary

- **Motivation for model-based approach**

  Focus on: Parametric, Learnt Models

- **Representing Shape mathematically:**

  Statistical shape model (SSM)

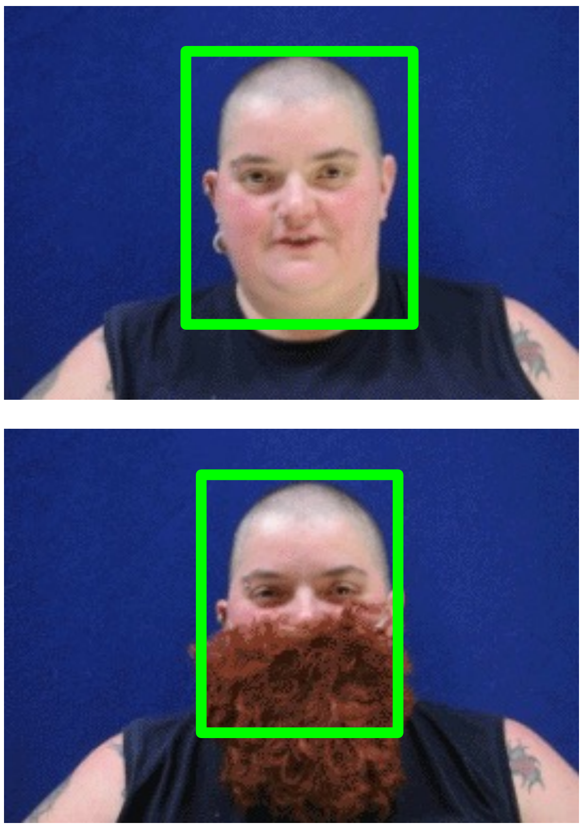
- **Modelling Shape Variability (SSM & PCA)**

- **Finding a shape in an image:**

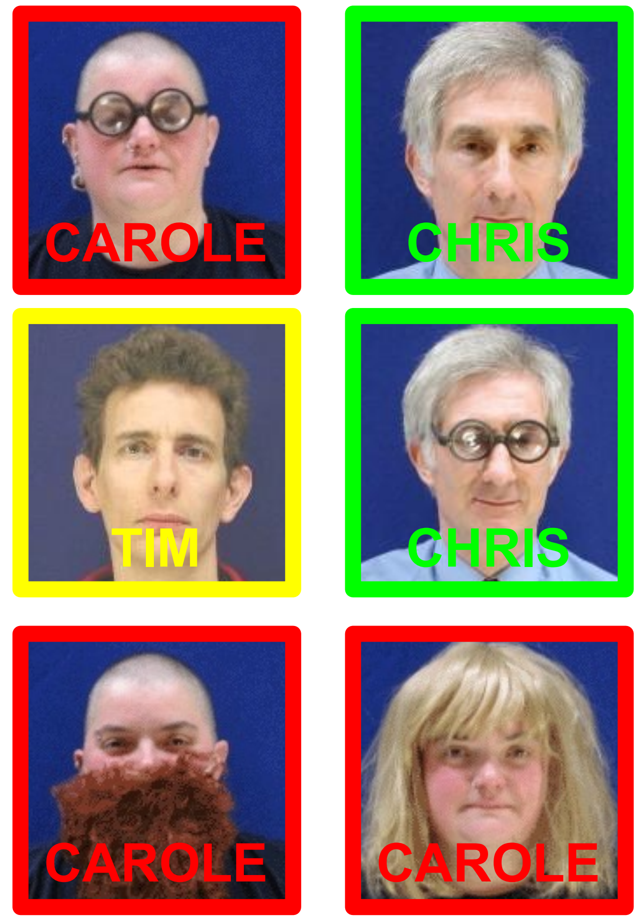
  Profile modelling & Active Shape Model (ASM)

Faces: Detection vs Recognition

- Viola & Jones: Face Detection (local feature detection)



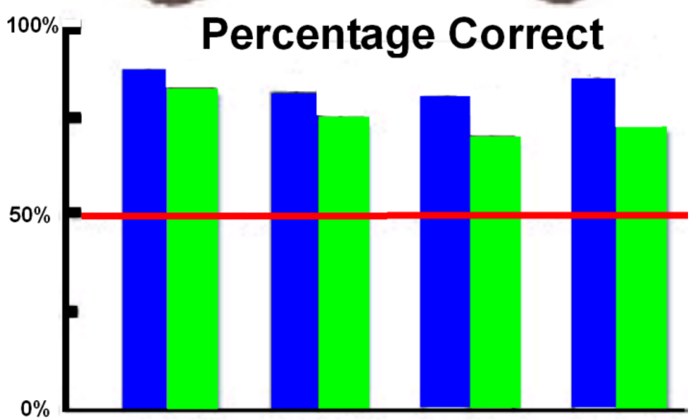
- Facial Recognition



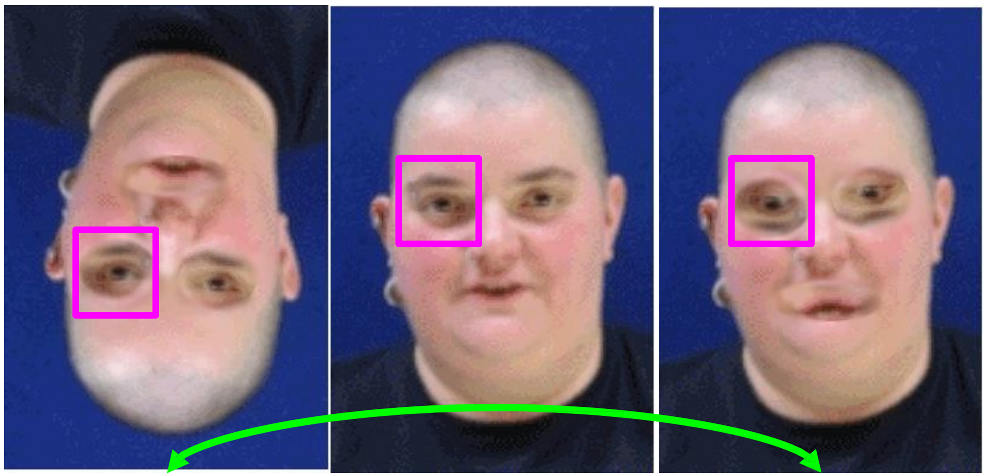
Facial Recognition: Sheep vs Humans

Kendrick et al., *Sheep don't forget a face,* Nature 414, 165-166 (2001)

Thompson, P. (1980). Margaret Thatcher: A new illusion. Perception



- After Training & one year later



Motivation for Modelling

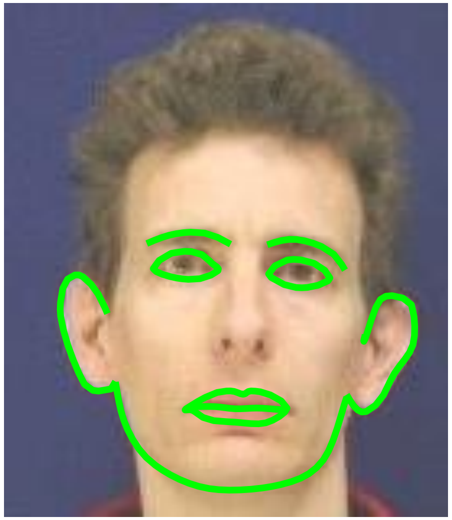
- To ‘understand’ images:
  - Not just image edges, but a face
  - Not just a face, but whose face
  - Not just me, but me with a beard
- Complex, multi-part structures, scope for confusion
- Noisy/missing data (e.g., glasses, facial hair)
  - Can’t interpret using image alone
- Model organises image evidence into a coherent whole

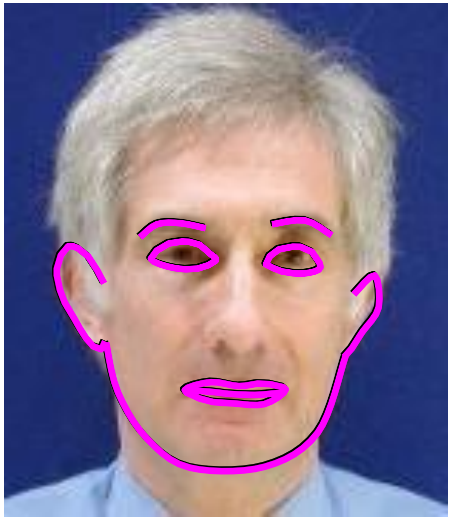
*Nose not included because I couldn't draw it!

Simplest Case: Image Contours

- Identifying basic facial features*
- Shape of individual features
- Relationships between features
- Encodes identity & expression
- First Basic Task:

  extracting suitable contours from images





Parametric Deformable Models

Incorporating Prior Information

- Training set of annotated images:

  Nature of the shape and expected shape variation

  Gray-scale patterns that define the shape location

- Build a model which can synthesize such shapes
- Use this generative model to search unseen images

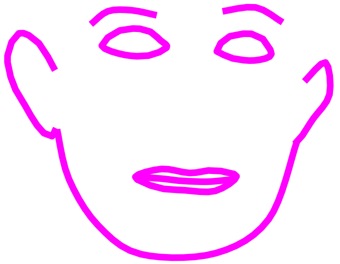
    Locate structures

    Transfer labels

    Compare new shape to those already seen

    Interpretation by synthesis

Modelling Issues



- Represent complicated multi-part shapes
- Represent variation of shapes across a population
- **Generalisation**:

  Need to represent all possible examples
- **Specificity**:

  Need to represent only ‘legal’ examples
- **Compactness**:

  A model with as few parameters as possible

Parametric Deformable Modelling Approach

Statistical Shape Models (SSMs):

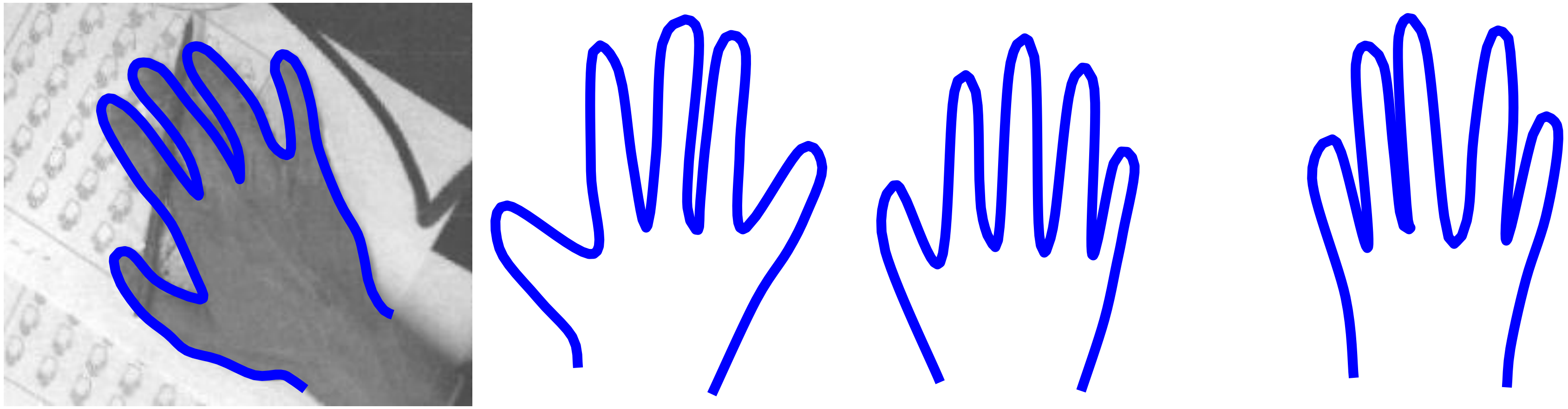
- **Representing complicated, multi-part shapes**
- **Modelling shape across a population**

Active Shape Models (ASMs):

- **Modelling expected image features near shape**
- **Search algorithm on a new image**

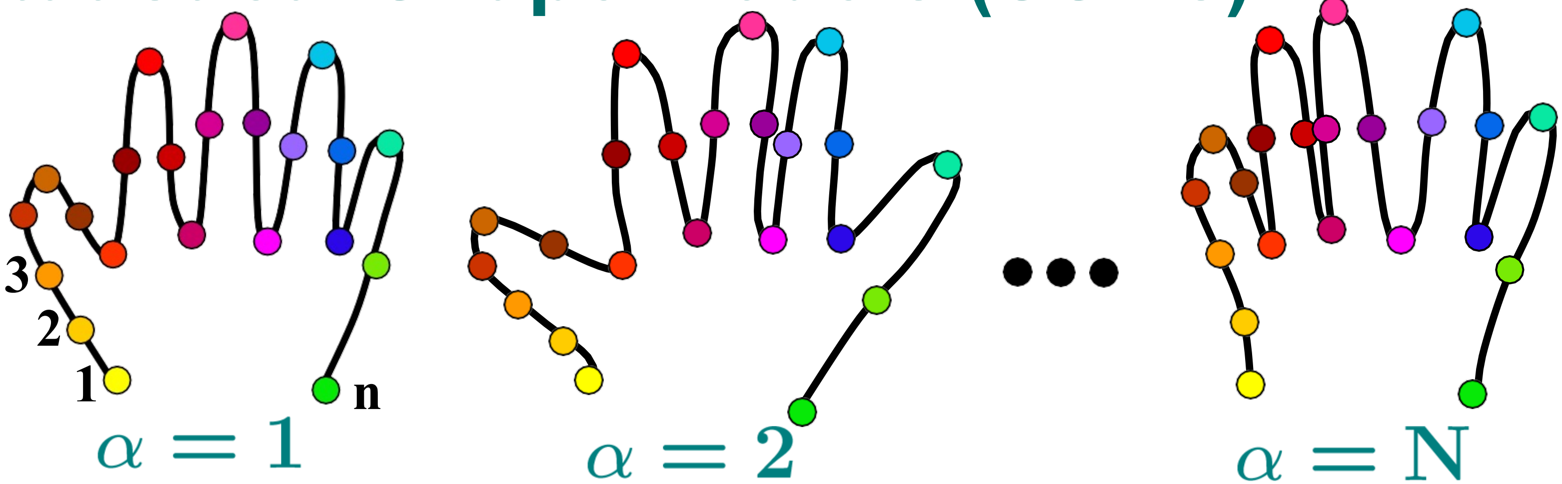
Representing a Population of Shapes

Training Data



- Set of images, containing object of interest
- Shape annotation on each image
- Training set of shapes, including required variation

Representing Shape: Statistical Shape Models (SSMs)



- Single shape, set of n points (& spline to join them)
- Shape vector: $\underline{x}^{\alpha} = \underline{x}^{1} = (\mathrm{x}_1, \mathrm{y}_1, \mathrm{x}_2, \mathrm{y}_2, \ldots, \mathrm{x}_\mathrm{n}, \mathrm{y}_\mathrm{n})$
- Corresponding points on all shapes
- Entire training set, set of shape vectors:

$$\{\underline{x}^{\alpha} : \alpha = 1, 2, \ldots, \mathrm{N}\}$$

Shape Alignment

- What do we mean by shape?
- Shape: what is unchanged by similarity transformation:

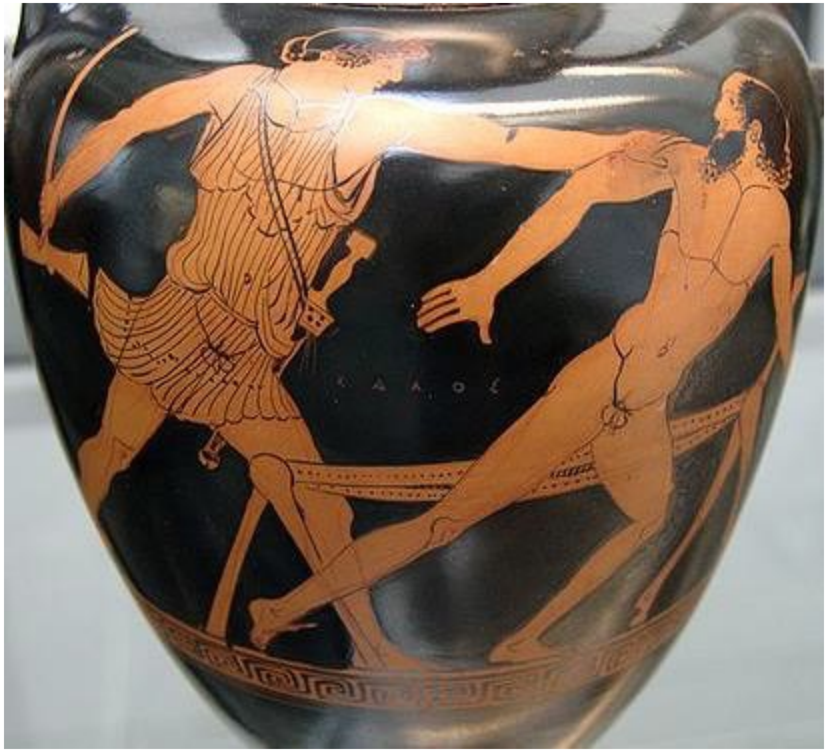
  Scaling

  Translation

  Rotation

- Align set of shapes, uniform position, scale and orientation (Generalized Procrustes analysis)

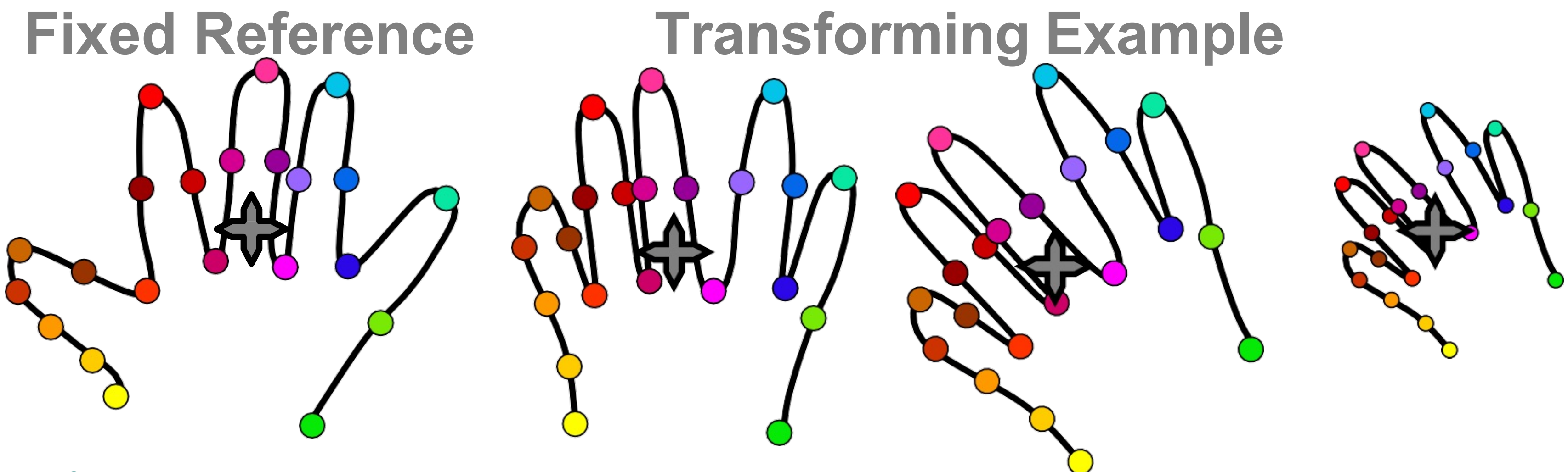
Shape Alignment: Why “Procrustes” Analysis?



Procrustes (Προκρούστης) or:

"the stretcher [who hammers out the metal]"

Procrustes Alignment for SSMs:

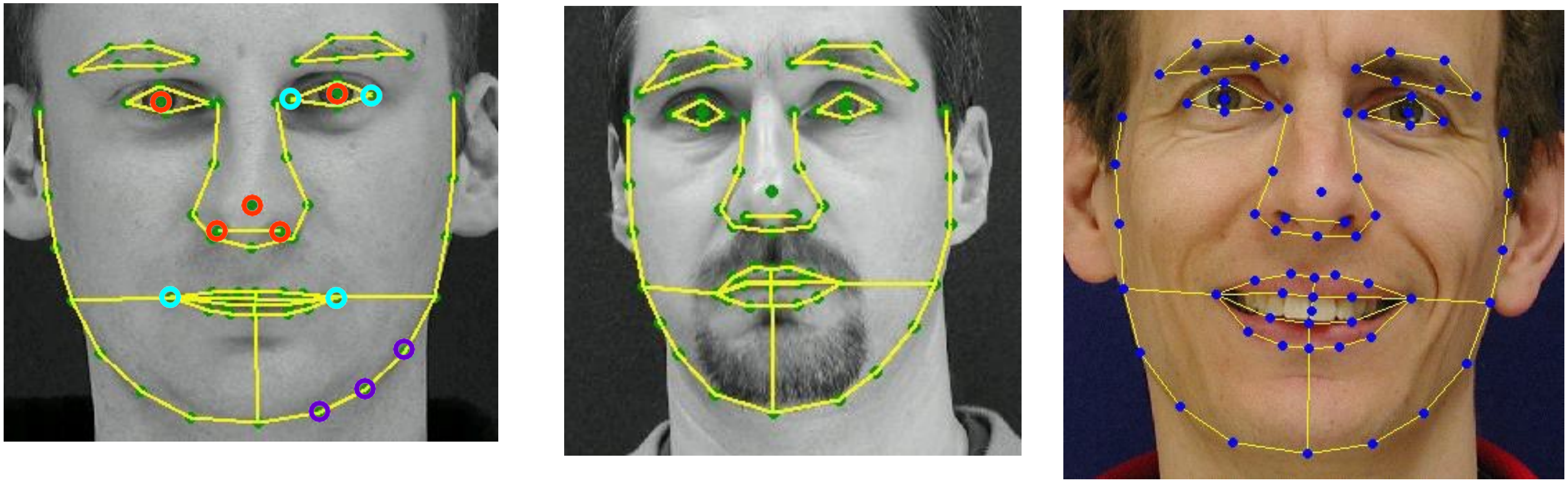


- Misalignment: distances between corresponding points

$$\sqrt{\mathrm{SSD}} = \sqrt{(x_1 - x_1')^2 + (y_1 - y_1')^2 \ldots + (y_n - y_n')^2}$$

- Match centre of mass (solves for translation)
- Match scale
- Solve for rotation
- Repeat for all shapes in training set
- Variants on the algorithm (e.g., iterative alignment to evolving mean)

SSM Training Examples:



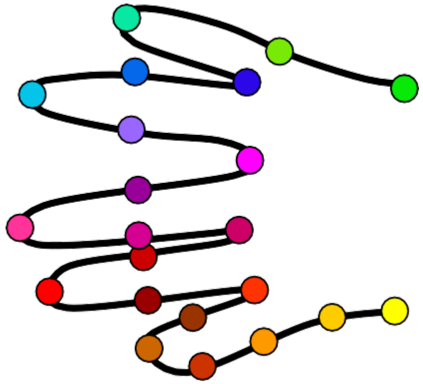
- **Need good identifiable landmarks:**

  Points (nostrils, tip of nose, pupils), Corners (eyes, mouth), Junctions

- **Other points can be equally-spaced along boundary**

  Use as many points as you need to define the shape

Statistical Shape Models:



$\alpha = N$

Advantages:

- Simple, intuitive shape representation
- Add as many points as required
- Corresponding points on different shapes:

  Points move in correlated fashion as parts move or shape changes

Disadvantages:

- Mark-up time-consuming, error-prone
- Correspondence hard to define on some objects
- Surfaces: hard to do & equal-spacing doesn't work!

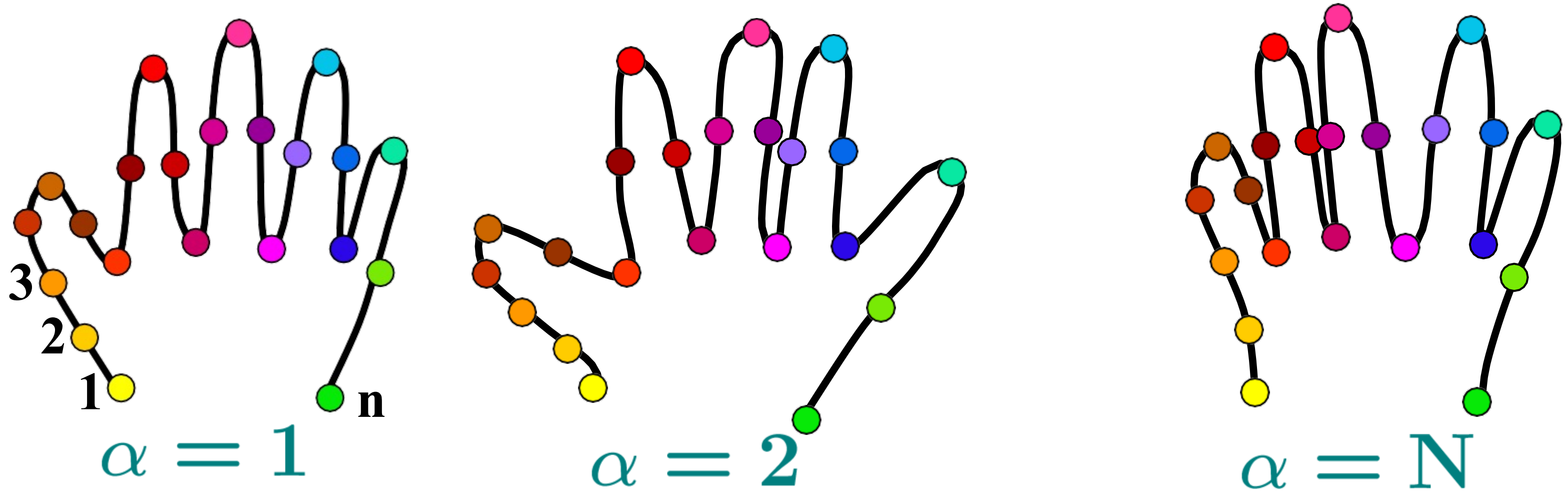
Progress so far:

- **Need to include prior knowledge**
- **Training data**
- **Representing sets of shapes**

Next :

- **Modelling distributions of shapes**
- **Modelling image appearance**
- **Search algorithm**

Statistical Shape Models (SSMs)

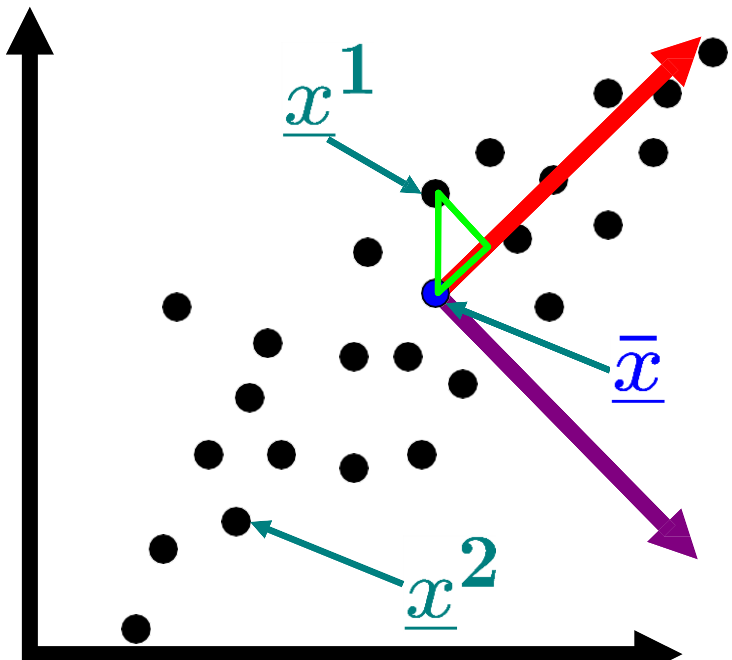


- Set of ***aligned*** shapes
- Shape vector: $\underline{x}^{\alpha} = (\mathrm{x}_1^{\alpha}, \mathrm{y}_1^{\alpha}, \mathrm{x}_2^{\alpha}, \mathrm{y}_2^{\alpha}, \ldots, \mathrm{x}_n^{\alpha}, \mathrm{y}_n^{\alpha})$
- Corresponding points on all shapes
- Entire **aligned** training set, set of shape vectors:

$$\{\underline{x}^{\alpha} : \alpha = 1, 2, \ldots, N\}$$

Principal Component Analysis (PCA)

$\{\underline{x}^{\alpha} : \alpha = 1, 2, \ldots, \mathrm{N}\}$



Shape Space:
axes = coordinates of every shape point

Mean shape:

$$\bar{\underline{x}} \doteq \frac{1}{\mathrm{N}} \sum_{\alpha=1}^{\mathrm{N}} \underline{x}^{\alpha}$$

$\hat{\underline{n}}$, unit axis vector

Maximize data projection:

$$\arg\max_{\hat{\underline{n}}} \sum_{\alpha=1}^{\mathrm{N}} \left(\hat{\underline{n}} \bullet (\underline{x}^{\alpha} - \bar{\underline{x}})\right)^2$$

Repeat:

$$\arg\max_{\hat{\underline{m}}} \sum_{\alpha=1}^{\mathrm{N}} \left(\hat{\underline{m}} \bullet (\underline{x}^{\alpha} - \bar{\underline{x}})\right)^2,$$

Constraint: $\hat{\underline{m}} \bullet \hat{\underline{n}} = 0$

PCA Solution: Covariance Matrix

Mean shape: $\underline{\bar{x}} \doteq \frac{1}{N}\sum_{\alpha=1}^{N}\underline{x}^{\alpha}$

n: number of points on each shape
2: number of spatial dimensions

Covariance matrix:

$\mathbf{C_{ij}} \doteq \frac{1}{N}\sum_{\alpha=1}^{N}(\underline{x}^{\alpha}-\underline{\bar{x}})_{\mathbf{i}}(\underline{x}^{\alpha}-\underline{\bar{x}})_{\mathbf{j}},\ \mathbf{i},\mathbf{j}=1,\ldots \mathbf{2n}$

Solve covariance matrix eigenproblem:

$$\mathbf{C}\underline{\hat{n}}^{\mu} = \lambda^{\mu}\underline{\hat{n}}^{\mu}$$

(matrix $\times$ vector $=$ number $\times$ vector)

Eigenvectors: $\{\underline{\hat{n}}^{\mu}\}$, directions of new axes

Ordered eigenvalues: $\{\lambda^{\mu} : \lambda^{1} \geq \lambda^{2} \ldots\}$,

how much variance in each direction

Generative Shape Models:

$$\underline{x} = \underline{\bar{x}} + \mathbf{P}\underline{b}$$

generated shape

mean shape

matrix, columns are eigenvectors

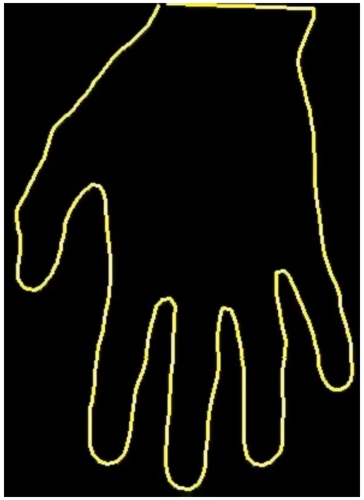
shape parameters, coords wrt PCA axes

- New shape = mean plus weighted sum of eigenvectors
- PCA automatically finds relevant modes of variation

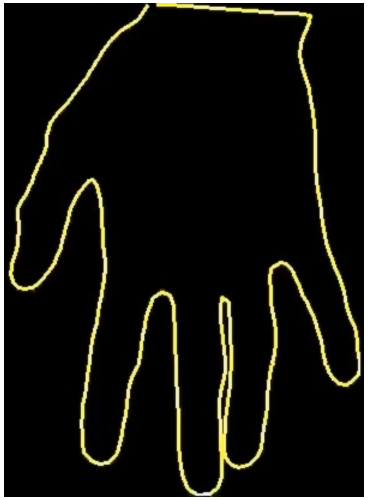
b可能是你关心多少特征值？可能几个特征值就可以确认是手了？只要符合那个线性关系？从而减少计算复杂度？

$b_1 \neq 0$ only

$b_2 \neq 0$ only



$b_3 \neq 0$ only

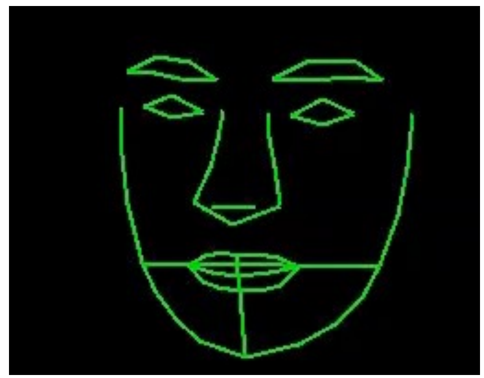


Generative Shape Models: Faces

First Mode

$b_1 \neq 0$ only

Second Mode

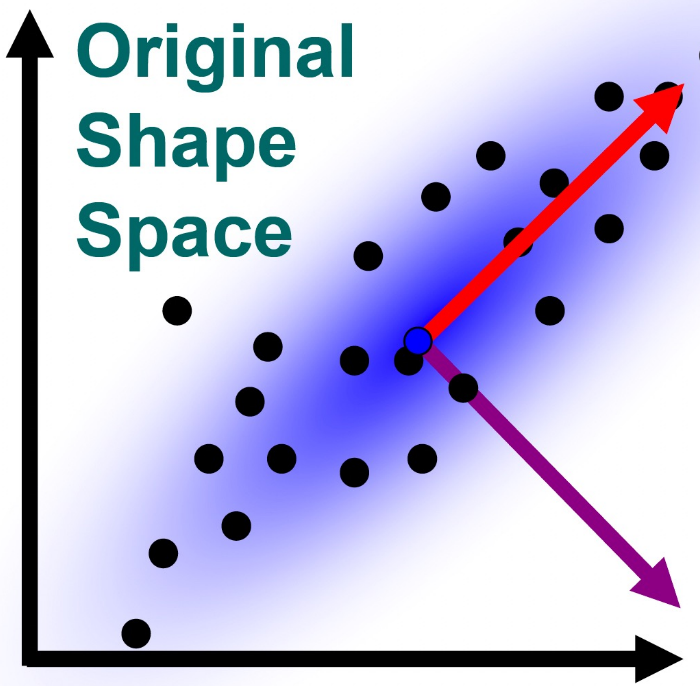


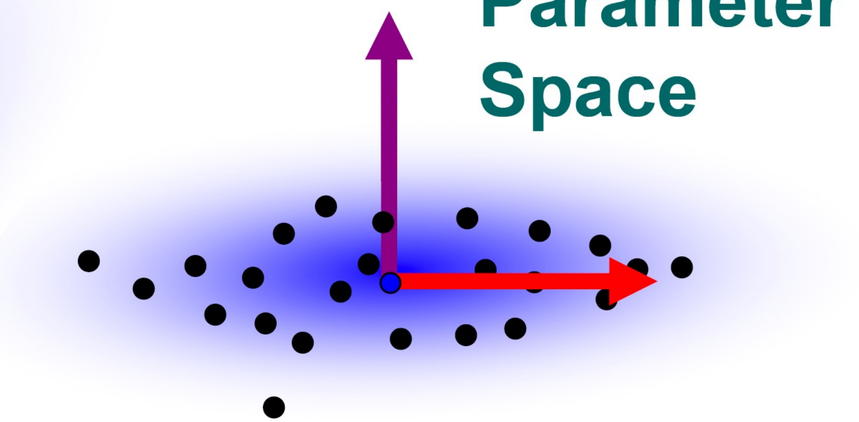
$b_2 \neq 0$ only

Third Mode

$b_3 \neq 0$ only

Statistical Shape Models:





no of shape modes retained

$$p(\underline{b}) \propto \prod_{\mu=1}^{\boxed{n_m}} \exp\left(-\frac{b_\mu^2}{2\lambda^\mu}\right)$$

- Multivariate Gaussian probability density function
- Aligned with PCA directions (eigenvectors)
- Matches variance seen in training set (eigenvalues)
- Product of Gaussians in parameter space

SSMs: Summary

Construction:

- Training set of shapes, corresponding landmarks
- Procrustes align shapes and compute mean shape
- Covariance matrix and solve PCA eigenproblem
- Shape parameters, modes of variation
- Construct gaussian probabilistic model

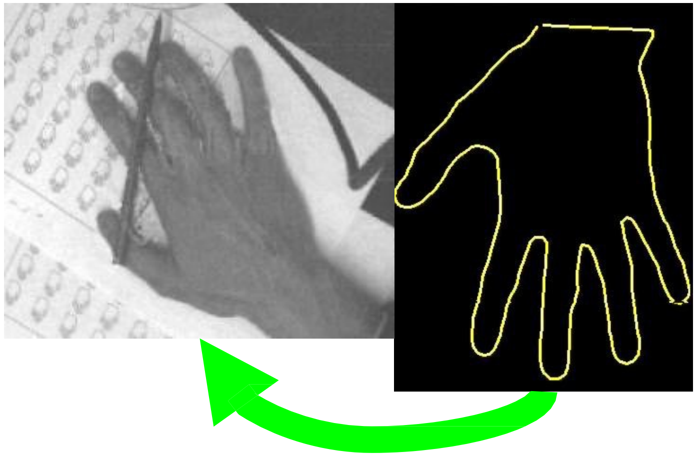
Results:

- General: modes of variation capture full variation
- Specific: modes capture only variation actually seen
- Assign probabilities to generated shapes

From SSMs to Active Shape Models

Task:

- Find shape in unseen image



Solution:

- Map from model frame to image frame
- Iterative localised search:

  Search in neighbourhood of current points for new points Fit

  model to suggested new shape

  Apply constraints to shape based on learnt variation

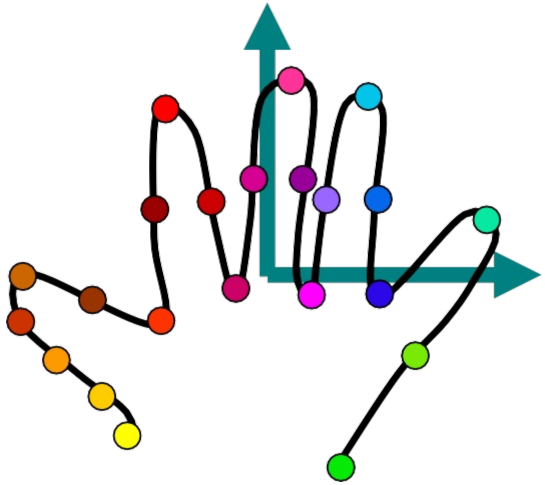
  Repeat until convergence
- This localised search

  - moves towards edges, lines etc

  - remains valid shape as it does so

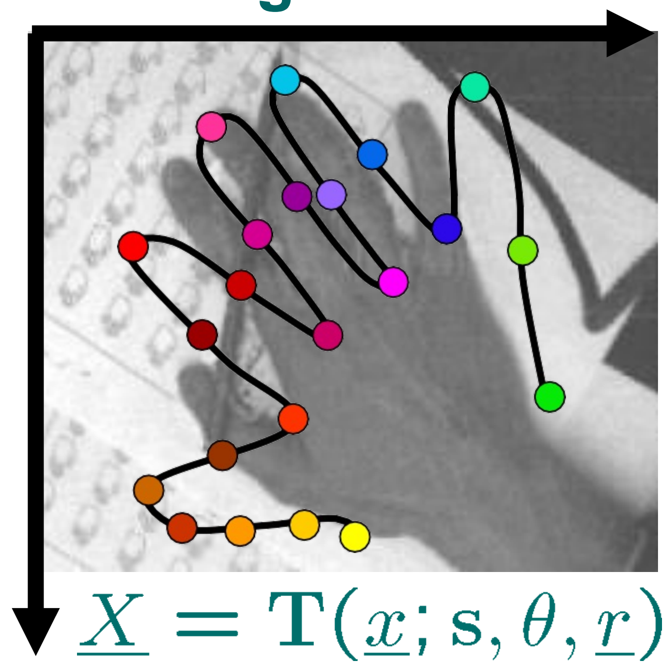
Placing the Model in an Image

Model Frame



$$\underline{x} = \underline{\bar{x}} + \mathrm{P}\underline{b}$$

Image Frame



$$\underline{X} = \mathrm{T}(\underline{x}; \mathrm{s}, \theta, \underline{r})$$

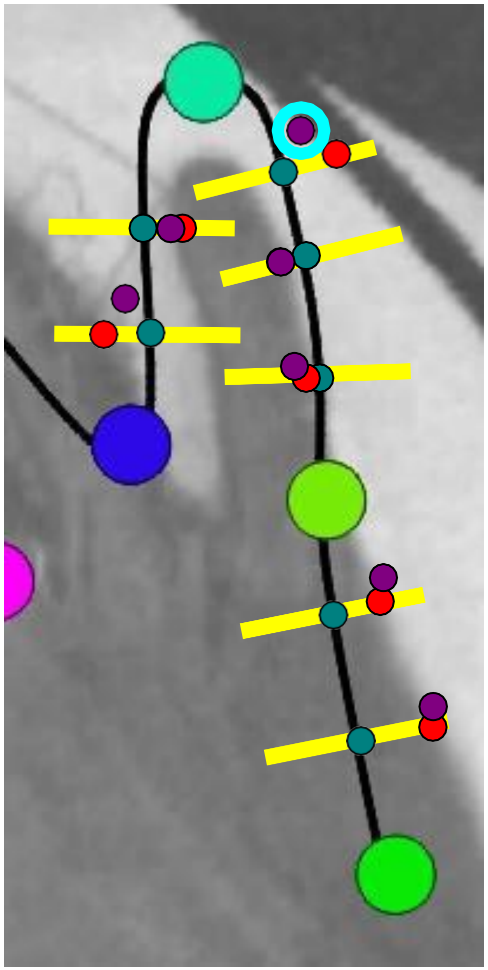
- **Scale, rotate, and translate to get to image frame:**

Pose transformation: $\mathrm{T}(\underline{x}; \mathrm{s}, \theta, \underline{r}) = \mathrm{sR}(\theta)\underline{x} + \underline{r}$

- **Total set of parameters to define shape in an image:**

Pose parameters: $\mathrm{s}, \theta, \underline{r}$, Shape parameters: $\underline{b}$

ASM Local Iterative Search



- Local search
- Initialise near target
- Search along normals
- Look for strongest edge
- Gives new set of suggested shape points

  $\underline{X}'$

- Best-Fit model to shape:

  $\underline{X}' \approx \mathrm{T}(\underline{x}(\underline{b}); \mathrm{s}, \theta, \underline{r})$

- Candidate shape:

  $\underline{X}'' = \mathrm{T}(\underline{x}(\underline{b}); \mathrm{s}, \theta, \underline{r})$

- Still leaves some errors

Problems:

Edge-Finding:

- Actual position not edge, or not strongest edge

Solution:

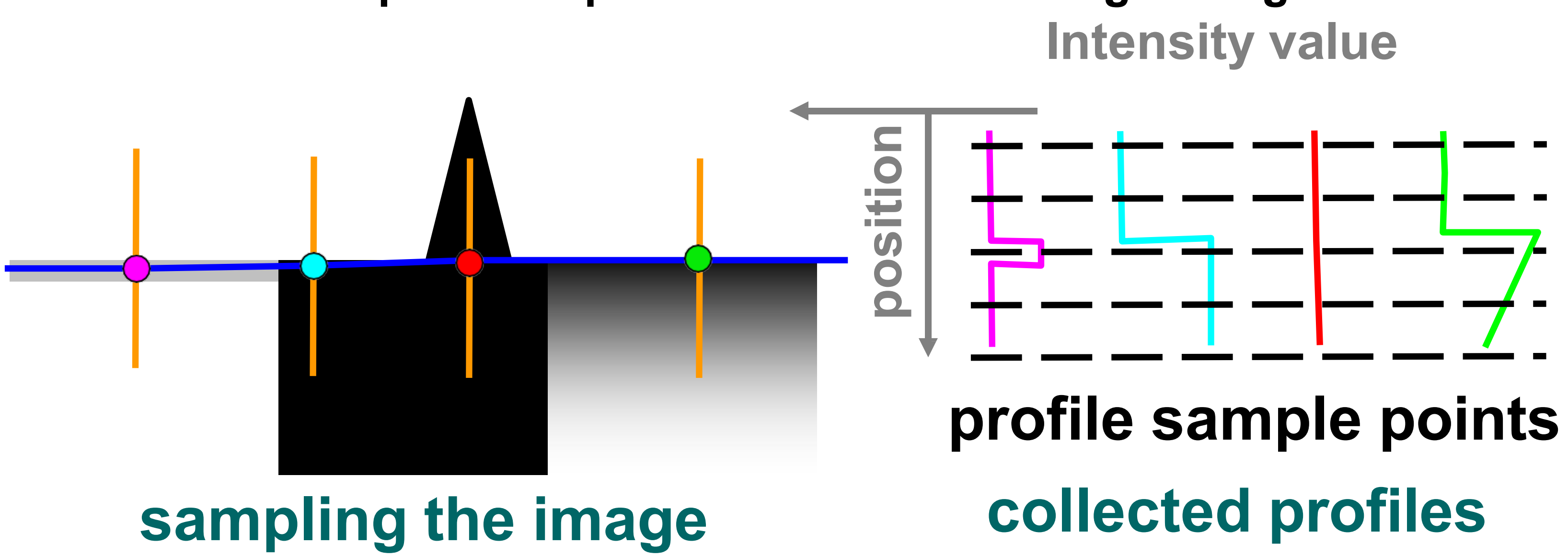
- Model profiles/appearance near points
- Apply shape and profile probability to control search

Profile Models

- For each shape point in each training image: Sample image values along normals to shape Normalise to eliminate illumination effects
- Build statistical model as for shape

  Profile vector like shape vector
- Model assigns probability to each possible profile

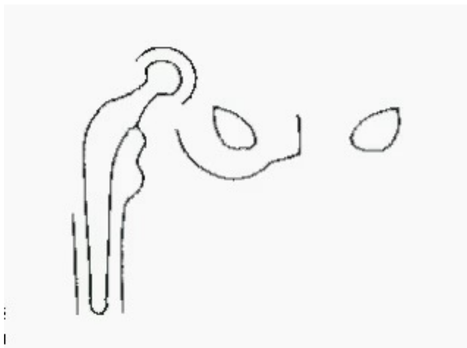
  Select most probable profile rather than strongest edge



ASM Multi-Resolution Search

- To increase basin of attraction, use multi-resolution
  - Gaussian pyramid of training images
  - Same shape points, but different profile models at each level
- Start search at coarse level, refine at finer level

ASM Search Example: Hip



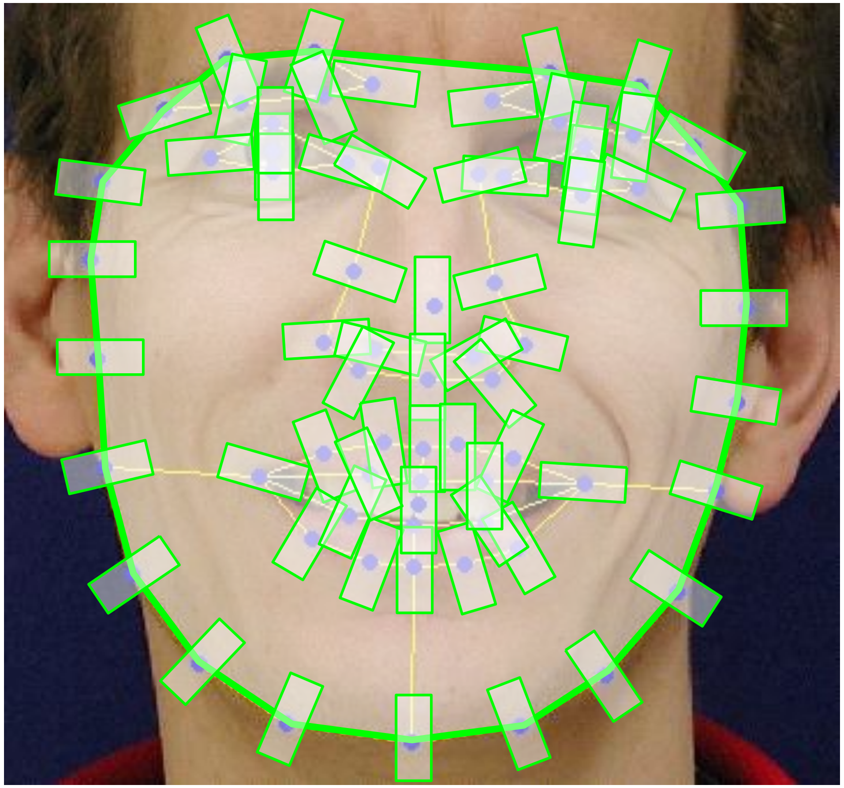
ASM: Summary

Advantages:

- Fast, simple, accurate
- Efficient extension to 3D

Disadvantages:

- Only sparse use of image information
- Treats local profiles as independent



Further Information:

- Mathematical details of PCA, correspondence problem

Electronic access via CAS/library to Springer ebooks:

Davies, Twining & Taylor, *Statistical Models of Shape*

- ASM, AAM etc
- Tim Cootes personal website
- Wikipedia: articles & links
- YouTube: many videos of ASM/AAM search & variants