Model Based Vision 1:
From Active Contours (ACMs) to Statistical Shapes (SSMs)

Spring 2021

Slides by: Dr Carole Twining

Presented by: Terence Morley

Handouts & Lecture Notes

- **Report in Scientific American (June 2014):**

  "In each study, however, those who wrote out their notes by hand had a stronger conceptual understanding and were more successful in applying and integrating the material than those who used [sic] took notes with their laptops."

  The Pen Is Mightier Than the Keyboard

  P. A. Mueller, D. M. Oppenheimer, *Psychological Science,* Vol 25,Issue 6, pp. 1159 – 1168, April-23-2014.

- **Handouts are to aid note taking, not a total replacement for note taking**
- **Podcasts, slides, pdfs etc on BlackBoard**

Model-Based Vision 1 & 2: Summary

- **Motivation for model-based approach**

  Simplest Freeform Models:

  Active Contour Model (ACM), Snakes

  Parametric, Learnt Models:

- **Representing Shape mathematically:**

  Statistical shape model (SSM), medial, distance maps

- **Modelling Shape Variability (SSM & PCA)**
- **Finding a shape in an image:**

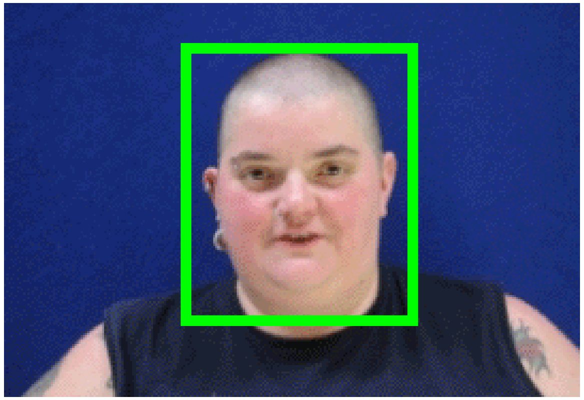
  Profile modelling & Active Shape Model (ASM)

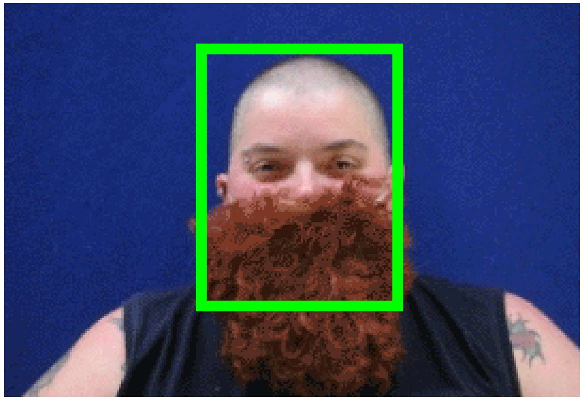
- **Modelling whole regions/whole images:**

  Active Appearance Model (AAM)

Faces: Detection vs Recognition

- Viola & Jones: Face Detection

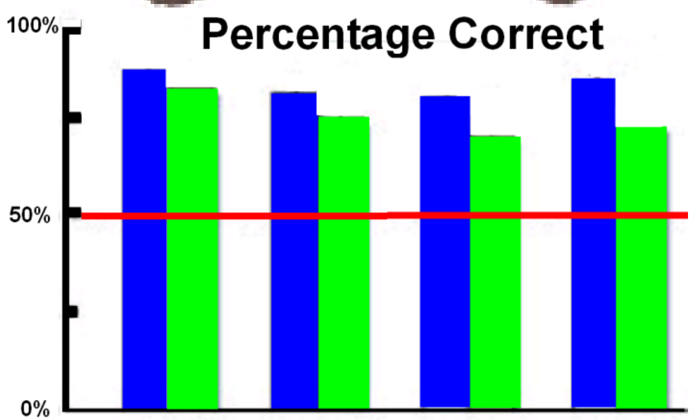




- Facial Recognition

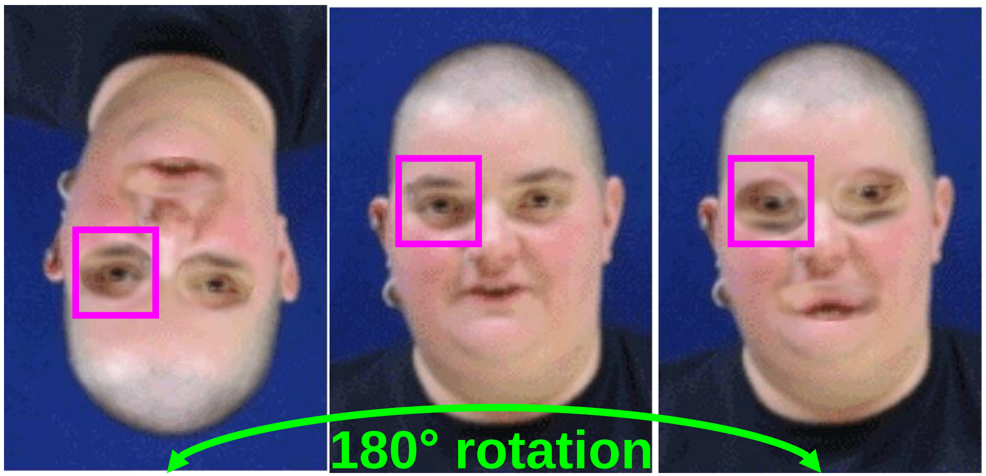
Facial Recognition: Sheep vs Humans

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



- **After Training & one year later**

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



- **Same reaction in other primates**

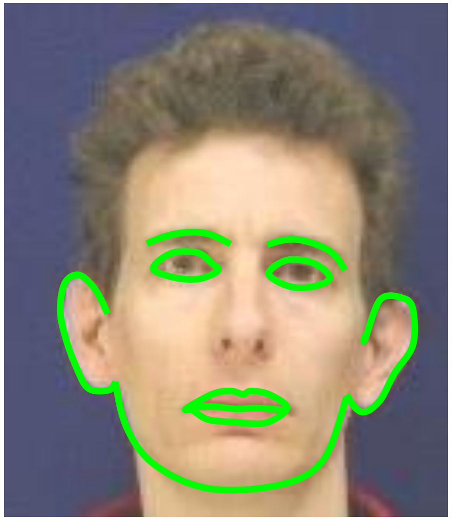
Motivation for Modelling

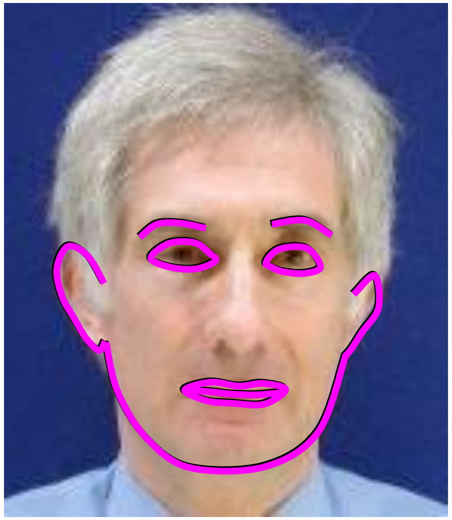
- To ‘understand’ images:
  - Not just image edges, but a face
  - Not just a face, but whose face
  - Not just me, but me with a beard
- Complex, multi-part structures, scope for confusion
- Noisy/missing data (e.g., glasses, facial hair)
  - Can’t interpret using image alone
- Model organises image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

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



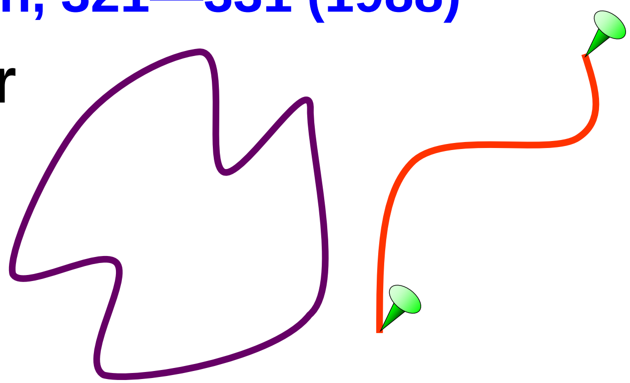


Freeform Deformable Models

Active Contour Models (ACMs)

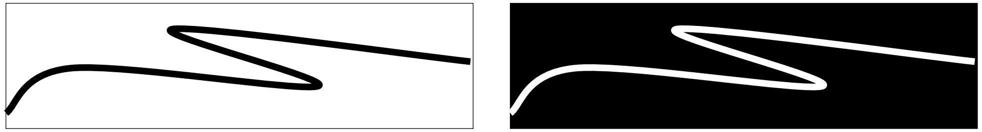
Kass, Witkin & Terzopoulos, *Snakes: Active Contour Models*, International Journal of Computer Vision, 321—331 (1988)

- **Simple closed or open contour**
- **User-interaction:**
  - **Initialising closed contour**
  - **Fixing ends of open contours**

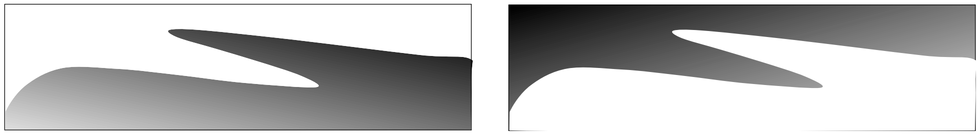


- **Contour attracted towards relevant image feature:**

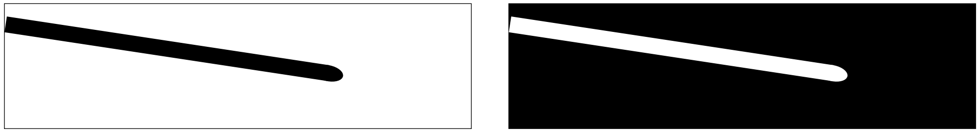
Lines



Edges



Terminations

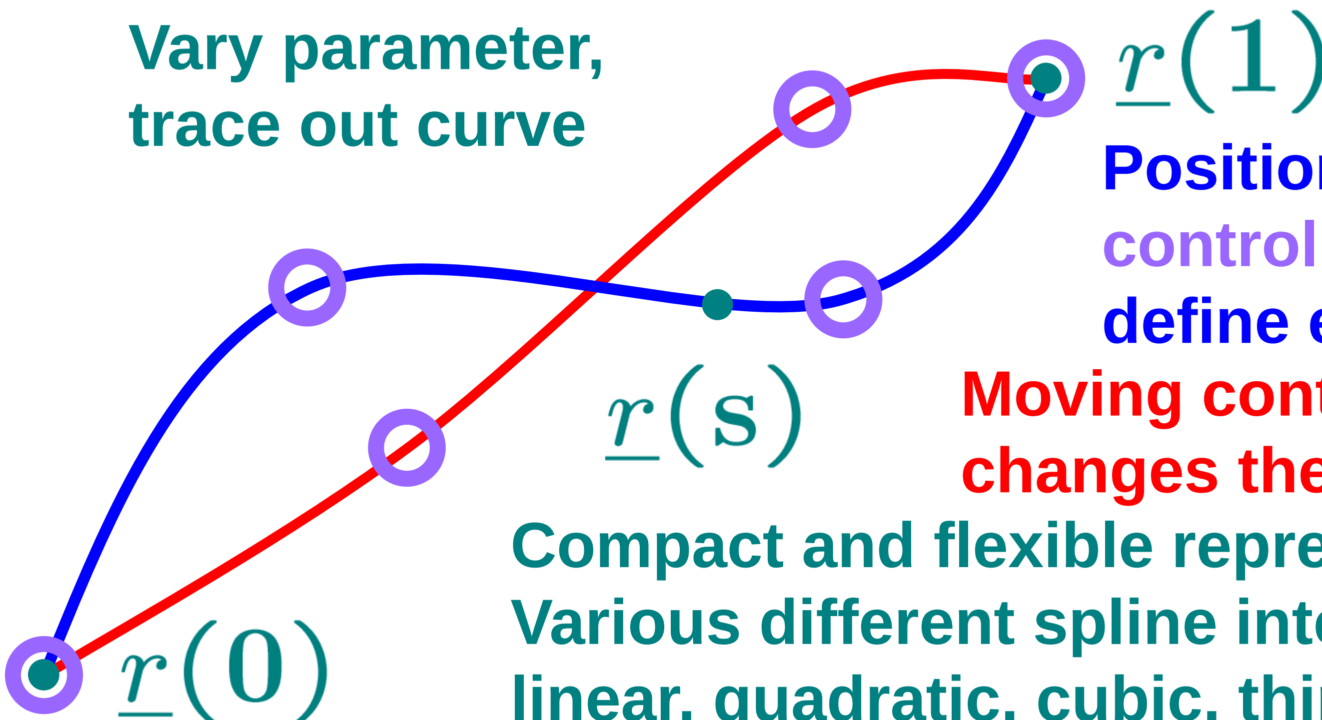


Representing a Contour: Splines

- **Contour: continuous, parametric curve**

$$\underline{r}(\mathrm{s}) = (\mathrm{x}(\mathrm{s}), \mathrm{y}(\mathrm{s})),\ s = 0 \text{ to } s = 1$$

Vary parameter, trace out curve



Positions of spline control points define entire curve

Moving control points changes the curve

Compact and flexible representation

Various different spline interpolants: linear, quadratic, cubic, thin-plate etc

Snakes: Energies and Forces

- **Energy Minimization:**

integrated along curve — internal energy of curve — curve/image interaction energy — user-defined constraints

$$\mathcal{E} = \int_0^1 \mathrm{E}(\underline{r}(\mathrm{s}))\mathrm{ds} = \int_0^1 \left[\boxed{\mathrm{E}_{\mathrm{int}}(\underline{r}(\mathrm{s}))} + \boxed{\mathrm{E}_{\mathrm{image}}(\underline{r}(\mathrm{s});\mathcal{I})} + \boxed{\mathrm{E}_{\mathrm{con}}(\underline{r}(\mathrm{s}))}\right] \mathrm{ds}$$

$$\mathrm{E}_{\mathrm{int}}(\underline{r}(\mathrm{s})) = \alpha(\mathrm{s})|\underline{r}_{\mathrm{s}}(\mathrm{s})|^2 + \beta(\mathrm{s})|\underline{r}_{\mathrm{ss}}(\mathrm{s})|^2$$

elasticity term: resists stretching

thin-plate term: resists bending

$$\underline{r}_{\mathrm{s}}(\mathrm{s}) = \frac{\mathrm{d}\underline{r}(\mathrm{s})}{\mathrm{ds}}$$

$$\underline{r}_{\mathrm{ss}}(\mathrm{s}) = \frac{\mathrm{d}^2\underline{r}(\mathrm{s})}{\mathrm{ds}^2}$$

$$\mathrm{E}_{\mathrm{image}}(\underline{r}(\mathrm{s});\mathcal{I}) = \pm\mathcal{I}(\underline{r}(\mathrm{s}))$$

contour favours dark/light regions of the image

$$\mathrm{E}_{\mathrm{image}}(\underline{r}(\mathrm{s});\mathcal{I}) = -|\vec{\nabla}\mathcal{I}(\underline{r}(\mathrm{s}))|$$

contour favours regions with high image gradient

- **Forces: Gradient of energy**
- **Solution -- Active contour:**

Minimize total energy OR find where forces cancel

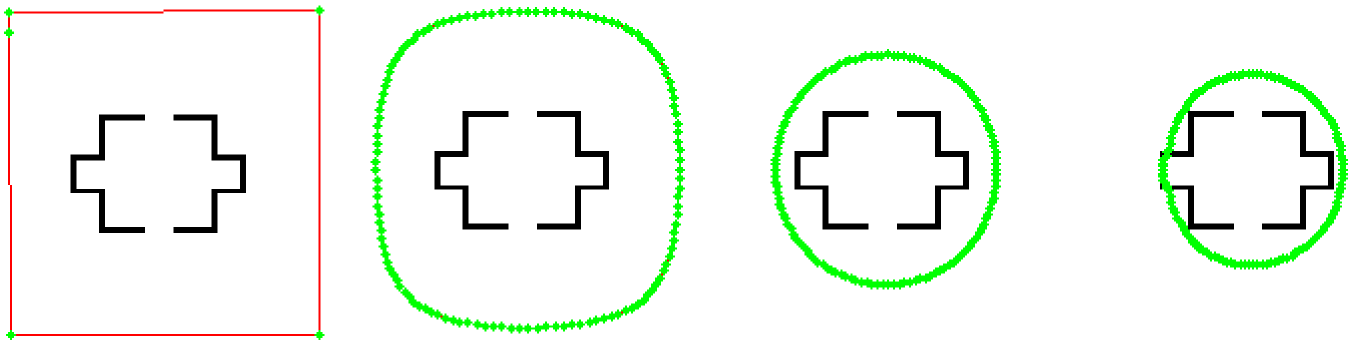
MATLAB Software:

http://www.iacl.ece.jhu.edu/static/gvf/snakedemo/snake_demo.zip

Example 1: Snake finding lines

$$E_{\mathrm{image}}(\underline{r}(s); \mathcal{I}) = +\mathcal{I}(\underline{r}(s))$$

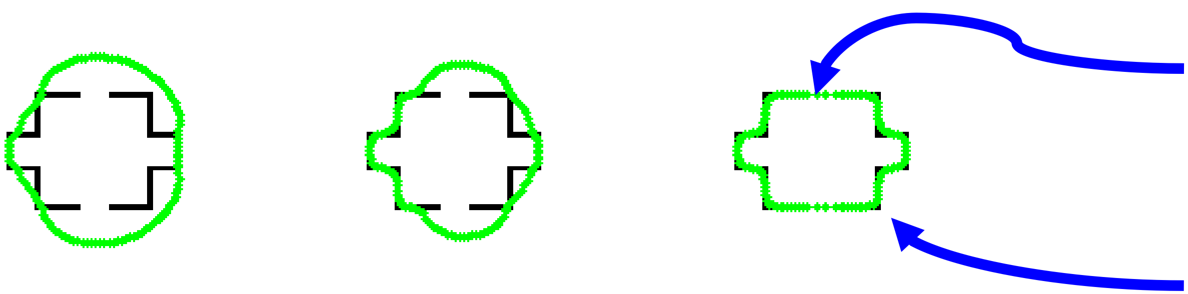
contour favours dark lines



initial contour

no image forces, elasticity only, pulling contour into contracting circle

contour contacts line and is held by image force



elasticity: pulls contour into straight line across gap

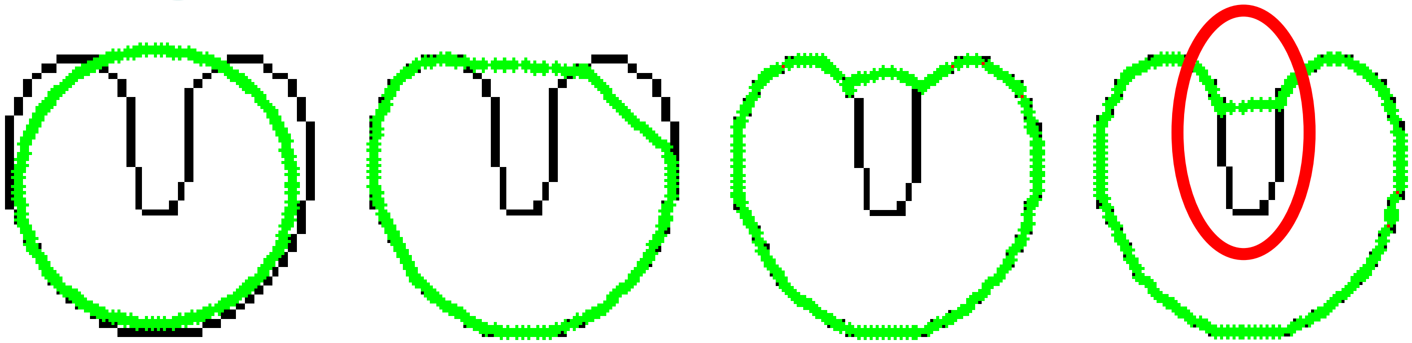
elasticity: contour gets pulled out of corners a little

contour gradually pulled in until covers whole line

Example 2: Snake finding lines

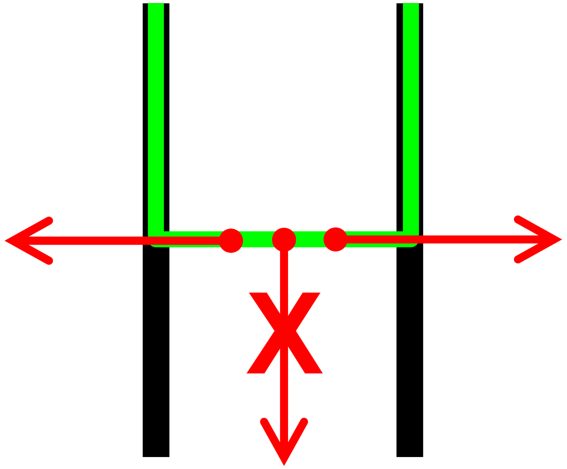
$$E_{image}(\underline{r}(s);\mathcal{I}) = +\mathcal{I}(\underline{r}(s))$$

contour favours dark lines



- Snakes have a problem with boundary concavities & projections
- Can add external pressure forces or internal 'balloon' forces
- Additional forces can cause additional problems, push contour off actual lines

Image Forces

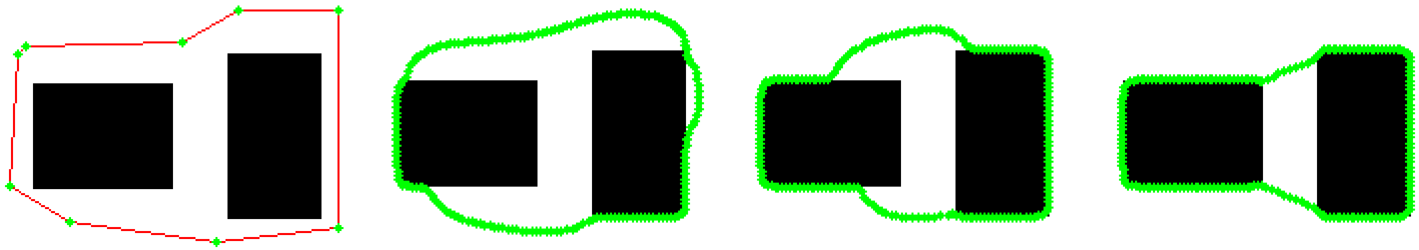


parallel lines

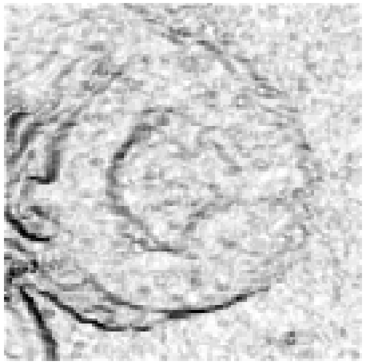
Example 3: Snake finding edges

$$\mathrm{E_{image}}(\underline{r}(\mathrm{s});\mathcal{I}) = -|\vec{\nabla}\mathcal{I}(\underline{r}(\mathrm{s}))|$$

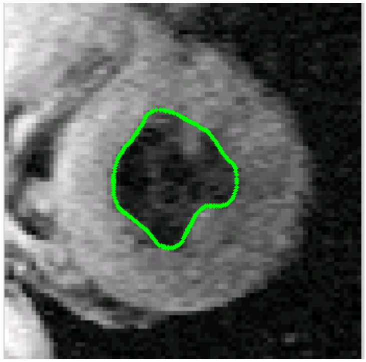
contour favours edges



- Edges hold contour but don't attract it
- No mechanism for splitting the contour
- Results depend on initialisation
- What about real images?



Noisy edge-strength image



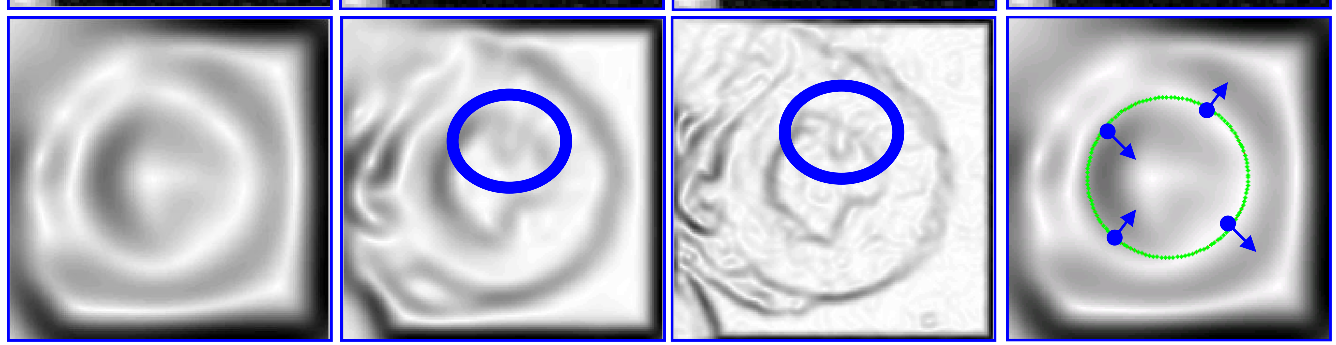
Scale Space:

As in Edge Detection (Edge Based Vision): $\mathrm{G}_\sigma * \mathcal{I}$

- **Convolve/Blur image with gaussian:** or $\vec{\nabla}(\mathrm{G}_\sigma * \mathcal{I})$
- **Gives larger basin of attraction for** $\mathrm{E}_{\mathrm{image}}(\underline{r}(\mathrm{s}); \mathcal{I})$
- **Coarse to fine search to find significant edges**

$\sigma = 10$ $\sigma = 5$ $\sigma = 2$

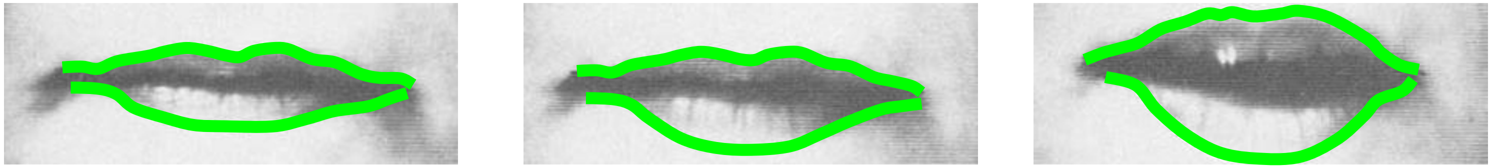
Edge Strength Images



Traditional Snakes: Summary

Advantages:

- Can work well on certain images
- Relatively simple, few parameters (elasticity etc)
- General: can fit to a wide range of shapes
- Useful in interactive applications such as image mark-up
- In tracking, where little motion between frames



Traditional Snakes: Summary

Problems:

- **Doesn't deal well with concavities and projections**
  - **Gradient Vector Flow snakes (GVF), balloon forces**
- **Poor capture range. Blurring helps, but can obscure the very detail you want to find**
  - **GVF etc**
- **Because general, multiplicity of minima on even relatively simple images**
- **Can't deal with change of topology/splitting**
  - **Level set implementation of snakes allows this**

Solution:

- **Incorporate detailed prior knowledge of shape**

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**

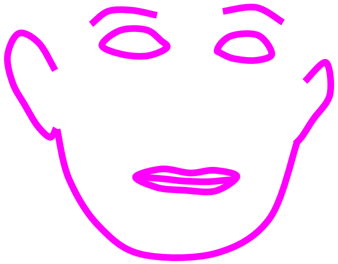
  Locate structures

  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
- Contour models: Failures

  Only simple (closed/open) contours, (too) general, concavities -- not general enough, not specific

Parametric Deformable Modelling Approach

Statistical Shape Models (SSMs):

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

Active Shape Models (ASMs):

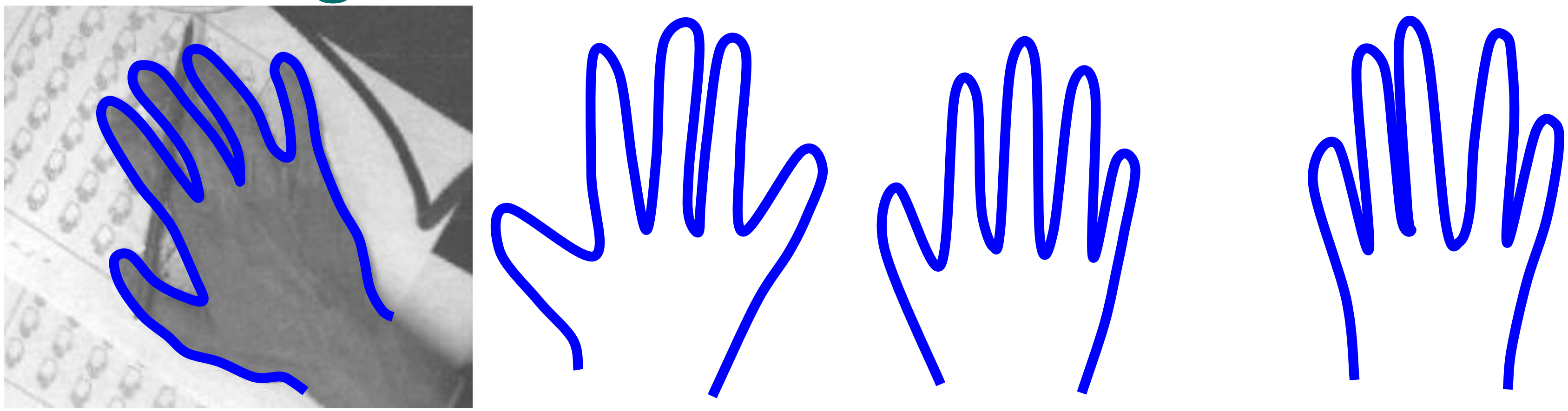
- **Modelling expected image features near shape**
- **Search algorithm on a new image**

Active Appearance Models (AAMs):

- **Model & synthesize image rather than just shape**

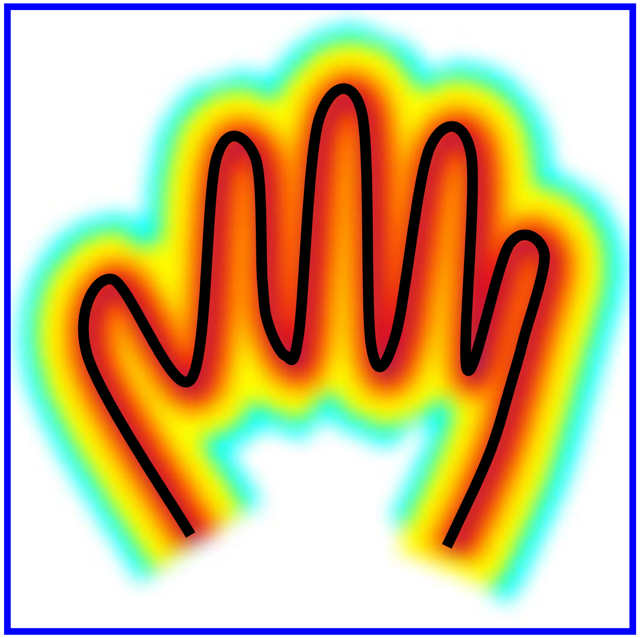
Representing a Population of Shapes

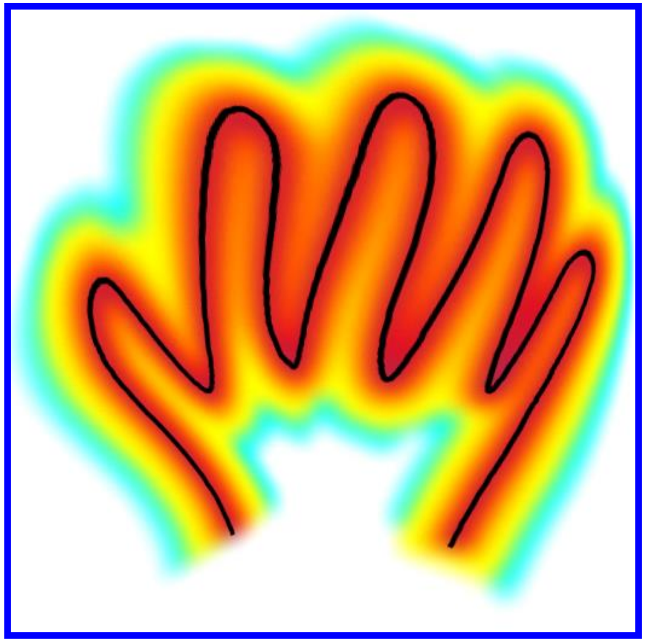
Training Data



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

Representing Shape: Distance Maps

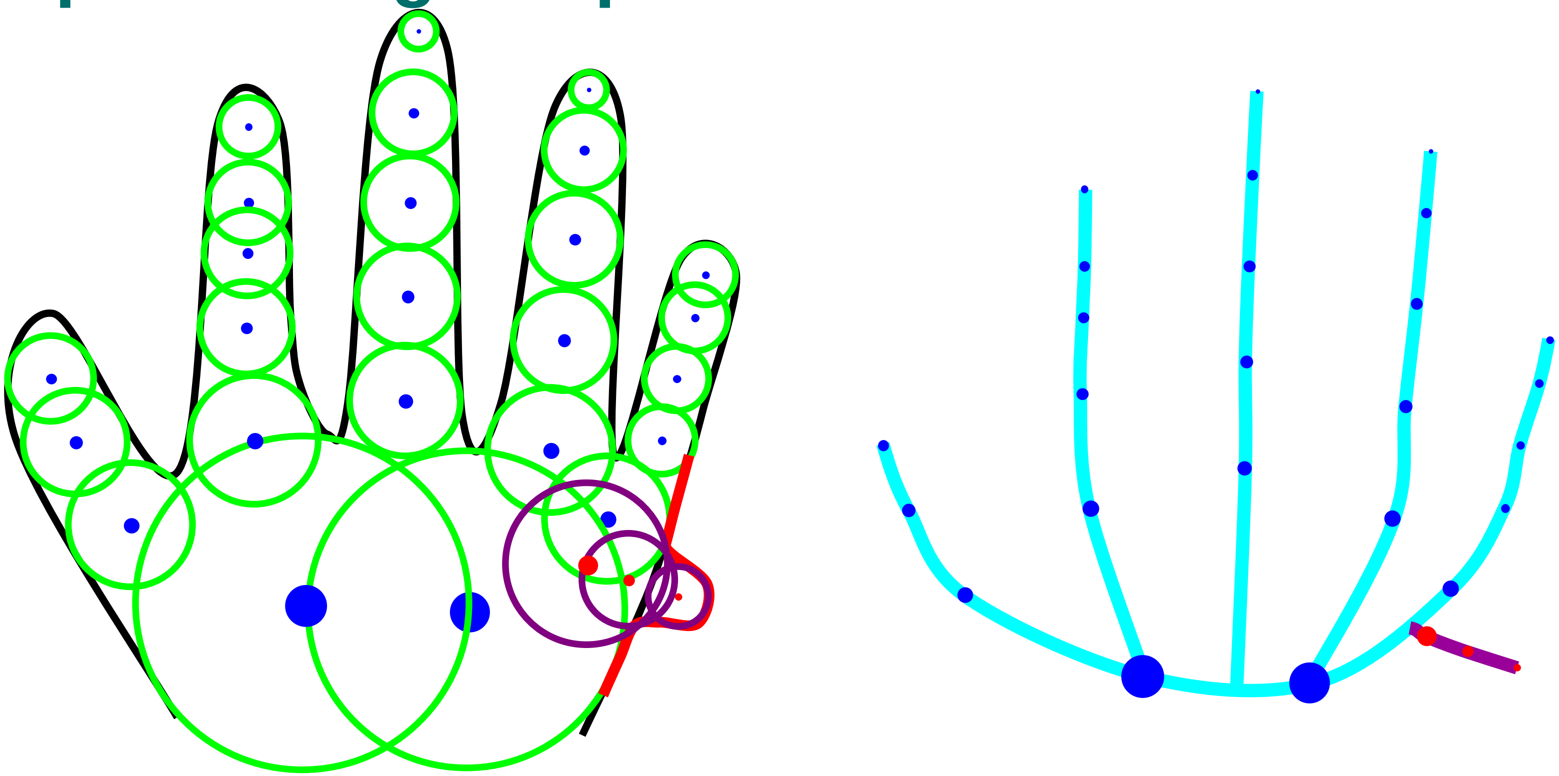




- Take shape, and compute distance from shape
- Set of images to model, rather than shapes

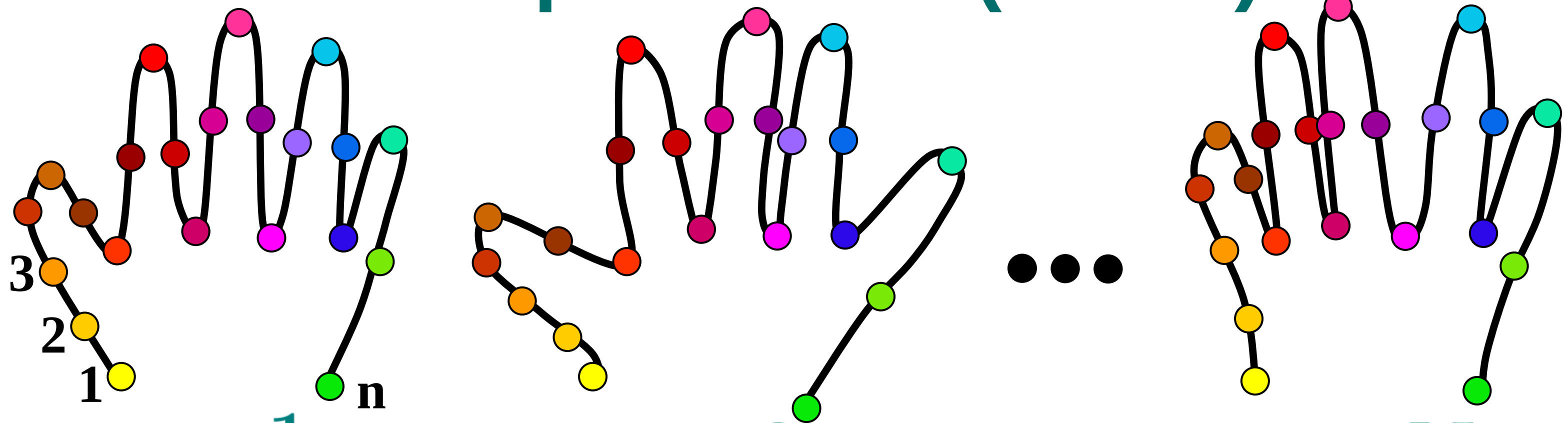
Made things more complicated!

Representing Shape: Medial Axis & M-REPS



- Shape as envelope of circles
- Centres lie on the medial axis/skeleton
- Flexing skeleton = bending fingers
- Altering radii = thicker or thinner fingers
- Sensitive to small changes in the shape

Representing Shape: Statistical Shape Models (SSMs)



$\alpha = 1$ $\alpha = 2$ $\alpha = N$

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

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

Shape Alignment

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

  Scaling

  Translation

  Rotation

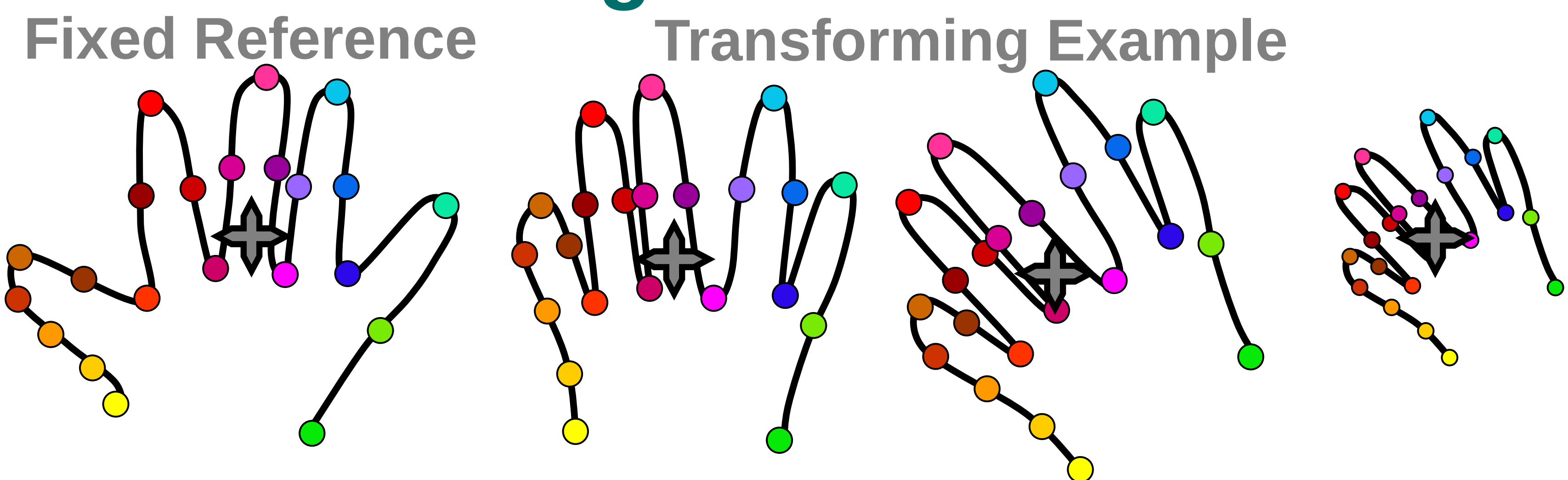
- Align set of shapes, uniform position, scale and orientation (Generalized Procrustes analysis)

Shape Alignment: Why “Procrustes” Analysis?

Procrustes (Προκρούστης) or:

"the stretcher [who hammers out the metal]"

Procrustes Alignment for SSMs:

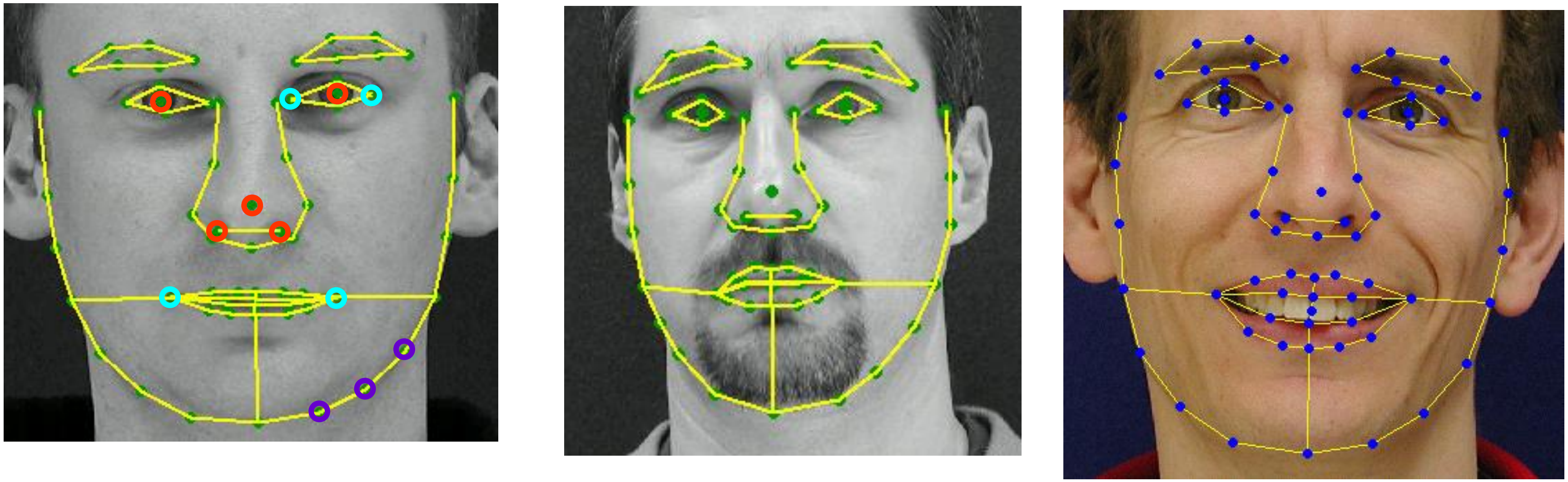


- **Misalignment: distances between corresponding points**

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

- **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:



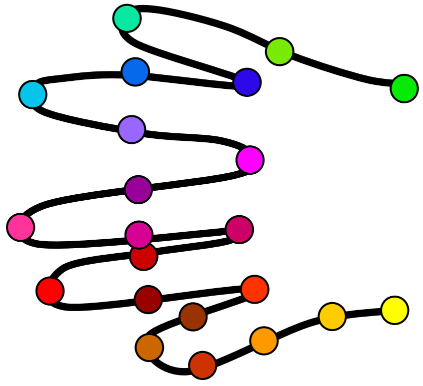
- **Need good identifiable landmarks:**

  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:



$\alpha = \mathrm{N}$

Advantages:

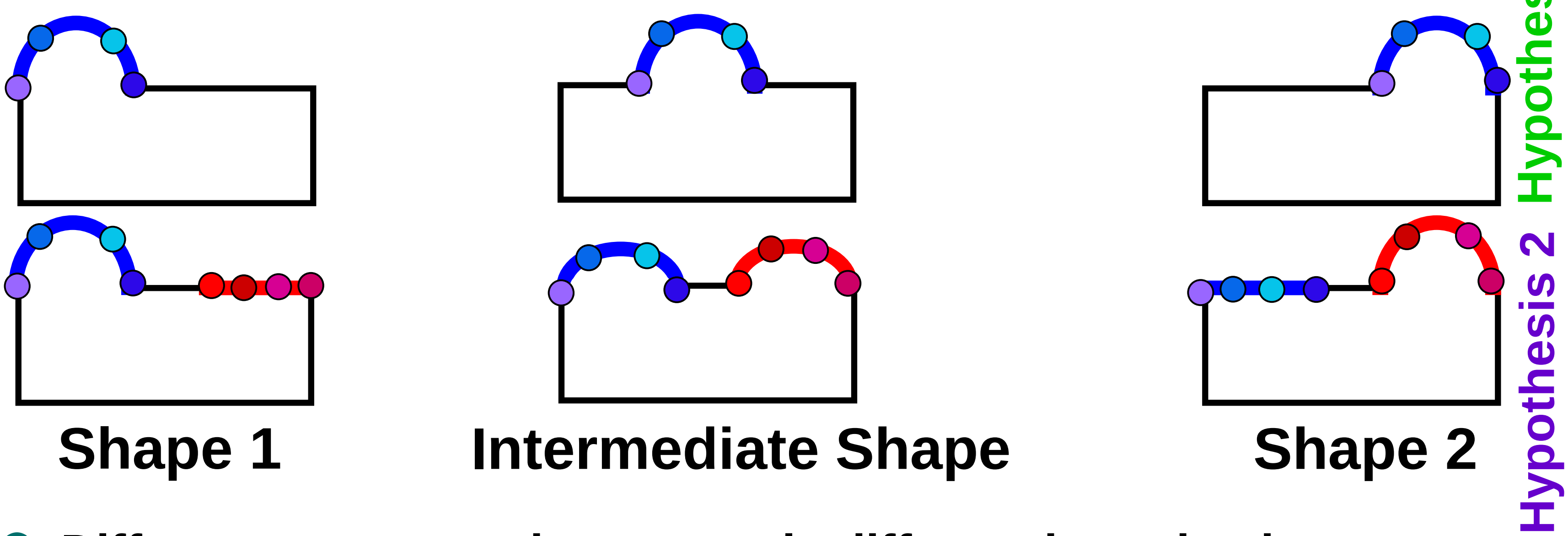
- Simple, intuitive shape representation
- Add as many points as required
- Corresponding points on different shapes:

  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!

Groupwise Correspondence:



- **Different correspondence equals different hypothesis as regards shape variation**
- **Need to use whole training set to decide on correct hypothesis**
- **Need flexible correspondence, & consider all hypotheses:**

  M-REPS: can manipulate, but need extra structure

  Distance maps: correspondence fixed

  SSMs: easy to manipulate correspondence

Summary:

Progress to date:

- Need to include prior knowledge
- Training data
- Representing sets of shapes
- Correspondence issues (MDL for groupwise case)

Next Lecture:

- Modelling distributions of shapes
- Modelling image appearance
- Search algorithm
- Modelling whole image – Active Appearance Models