Today

Image registration

Announcement

Quiz #7 is available in Blackboard.

Due date: 11:59pm EST, Tuesday, April 8th

Open book and open notes

Recall: Image Registration

Problem: given two or multiple images (or 3D volume), compute the transformation between images (or 3D volume)

$$\mathbf{I_1} = \mathbf{H}\mathbf{I_2}$$

Major steps:

- Find interest points
- Compute features such as SIFT features
- Match the interest points and establish correspondences
- Compute transformations

Recall: Image Registration

$$\mathbf{I_1 = HI_2}$$

Different forms of transformation

- Rigid object (building, car, etc.)
  - Homography for planar object
    - Rigid transformation (e.g., aerial images)
      - Translation, rotation, and scaling
    - Affine homography

- Non-rigid object (people, medical image, animals, etc.)
  - Free-form deformation
  - Model-based (e.g., Active Appearance Model)

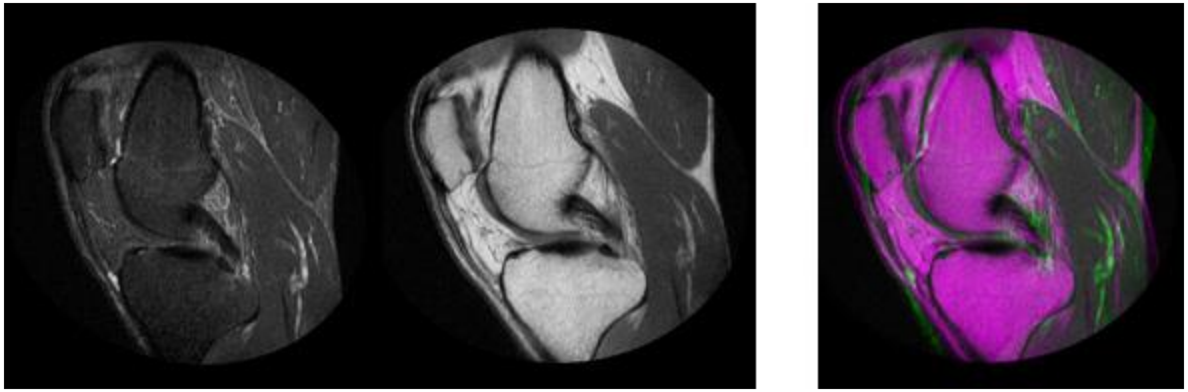
Examples of Image Registration

Moving image   Reference image   Registered image

Rigid object

Moving image   Reference image   Registered image

Non-Rigid object



http://www.mathworks.com

Homography for Planar Object

Homography defines the relationship between

- a 3D planar object and its corresponding image or
- two images capturing the same planar object from different views

Homography for Planar Object

Recall: perspective projection

Full perspective projection model

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \mathbf{r}_3^t & f_u t_X + u_0 t_z \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_z \\ \mathbf{r}_3^t & t_z \end{bmatrix}$$

Homography for Planar Object

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u \mathbf{r}_1^t + u_0 \mathbf{r}_3^t & f_u t_X + u_0 t_z \\ f_v \mathbf{r}_2^t + v_0 \mathbf{r}_3^t & f_v t_Y + v_0 t_z \\ \mathbf{r}_3^t & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \\ 1 \end{bmatrix}$$

Planar object

$$= \begin{bmatrix} f_u r_{11} + u_0 r_{31} & f_u r_{12} + u_0 r_{32} & f_u t_X + u_0 t_z \\ f_v r_{21} + v_0 r_{31} & f_v r_{22} + v_0 r_{32} & f_v t_Y + v_0 t_z \\ r_{31} & r_{32} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Homography for Planar Object

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Where **H** is a 3x3 matrix

$$\mathbf{H} = \begin{bmatrix} \boldsymbol{h_{11}} & \boldsymbol{h_{12}} & \boldsymbol{h_{13}} \\ \boldsymbol{h_{21}} & \boldsymbol{h_{22}} & \boldsymbol{h_{23}} \\ \boldsymbol{h_{31}} & \boldsymbol{h_{32}} & \boldsymbol{h_{33}} \end{bmatrix}$$

- The 3D - 2D mapping is reduced to 2D - 2D mapping
- **H** matrix is invertible ⇨ We can recover 3D from 2D if **H** is known

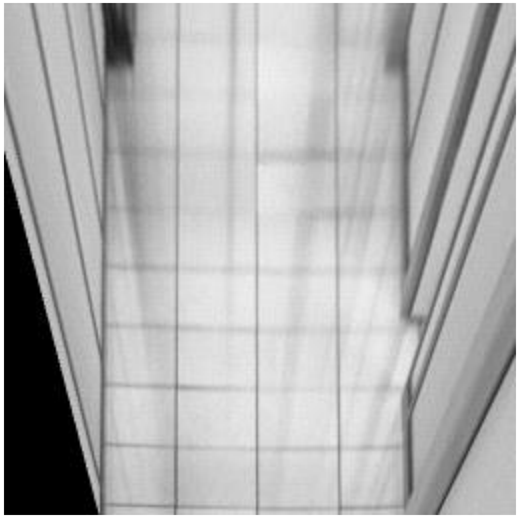
$$\frac{1}{\lambda} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{H}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Examples of 3D Reconstruction for Planar Objects – Removing Perspective Distortion

4 pair of point correspondences are sufficient to remove the perspective distortion

Courtesy of Hartley & Zisserman 1999

Examples of 3D Reconstruction for Planar Objects – Removing Perspective Distortion



Courtesy of Hartley & Zisserman 1999

Estimate the Homography Matrix

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad u = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}$$

$$v = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$

How many degrees of freedom for **H**?

8 DoFs. Why?

Estimate the Homography Matrix

H has 8 DoFs

$$u = \frac{\frac{h_{11}}{h_{33}}x + \frac{h_{12}}{h_{33}}y + \frac{h_{13}}{h_{33}}}{\frac{h_{31}}{h_{33}}x + \frac{h_{32}}{h_{33}}y + 1} \qquad v = \frac{\frac{h_{21}}{h_{33}}x + \frac{h_{22}}{h_{33}}y + \frac{h_{23}}{h_{33}}}{\frac{h_{31}}{h_{33}}x + \frac{h_{32}}{h_{33}}y + 1}$$

➡ **H can be solved up to a scaling factor**

➡

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11}' & h_{12}' & h_{13}' \\ h_{21}' & h_{22}' & h_{23}' \\ h_{31}' & h_{32}' & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Estimate the Homography Matrix

Two solutions:

- **Enforce** $h_{33} = 1$
- **Enforce** $\|\mathbf{h}\| = 1$

First Solution

By enforcing $h_{33} = 1$

$$u = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + 1} \qquad v = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + 1}$$

$$\Rightarrow \quad \begin{aligned} h_{11}x + h_{12}y + h_{13} - h_{31}\,xu - h_{32}yu = u \\ h_{21}x + h_{22}y + h_{23} - h_{31}\,xv - h_{32}yv = v \end{aligned}$$

The 8 unknowns can be solved given N (>=4) non-collinear points

$$\mathbf{Ah} = \mathbf{b}$$

2Nx8 8x1 2Nx1

$$\Rightarrow \quad \mathbf{h} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b}$$

Second Solution

By enforcing $\|\mathbf{h}\| = 1$

$$u = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}} \qquad v = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$

$$\Rightarrow \quad h_{11}x + h_{12}y + h_{13} - h_{31}\,xu - h_{32}yu - h_{33}u = 0$$
$$h_{21}x + h_{22}y + h_{23} - h_{31}\,xv - h_{32}yv - h_{33}v = 0$$

The 8 unknowns can be solved given N (>=4) non-collinear points

$$\mathbf{A}\mathbf{h} = \mathbf{0}$$

2Nx9   9x1   2Nx1

$\Rightarrow$ **h** can be solved by performing SVD on **A**

Homography between Two Images

If two images ($\mathbf{I}_1$ and $\mathbf{I}_2$) capturing the same planar object from different views, then a point $\mathbf{x}_1 = [u_1 \quad v_1 \quad 1]^T$ in $\mathbf{I}_1$ is corresponding to a point $\mathbf{x}_2 = [u_2 \quad v_2 \quad 1]^T$ in $\mathbf{I}_2$ such that

$$\mathbf{x_1} = \mathbf{H_1 x} \qquad \mathbf{x}_2 = \mathbf{H_2 x}$$

The same 3D point

$$\mathbf{x_1} = \mathbf{H x_2} \qquad \mathbf{H} = \mathbf{H_1 H_2^{-1}}$$

$\mathbf{H}$ can be estimated using at least 4 pairs of point correspondence

Special Case of Homography

Pure rotation

$$\mathbf{H} = \begin{bmatrix} cos\theta & -sin\theta & 0 \\ sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Pure scaling

$$\mathbf{H} = \begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Pure translation

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Pure shearing

$$\mathbf{H} = \begin{bmatrix} 1 & s_x & 0 \\ s_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Affine transformation

$$\mathbf{H} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

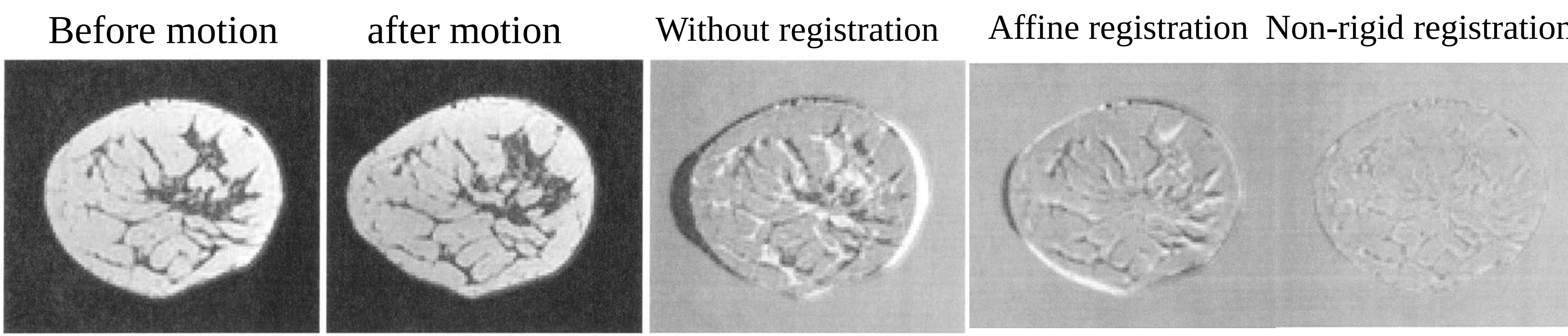
Image Registration

Problem: given two or multiple images (or 3D volume), compute the transformation between images (or 3D volume)

$$\mathbf{I_1} = \mathbf{H}\mathbf{I_2}$$

Different strategies for registration

- Model-free registration
  - Homography for planar object
  - Free-form deformation for nonrigid object

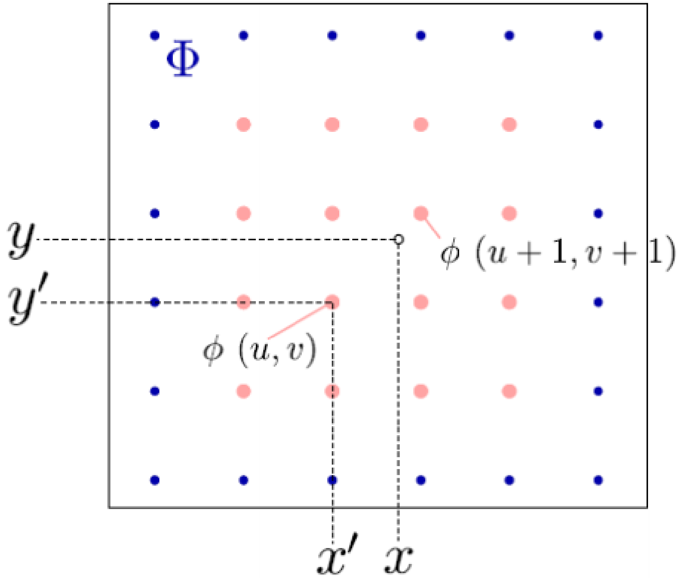


Rueckert et al, "Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images", IEEE Trans. on Medical Imaging, 1999

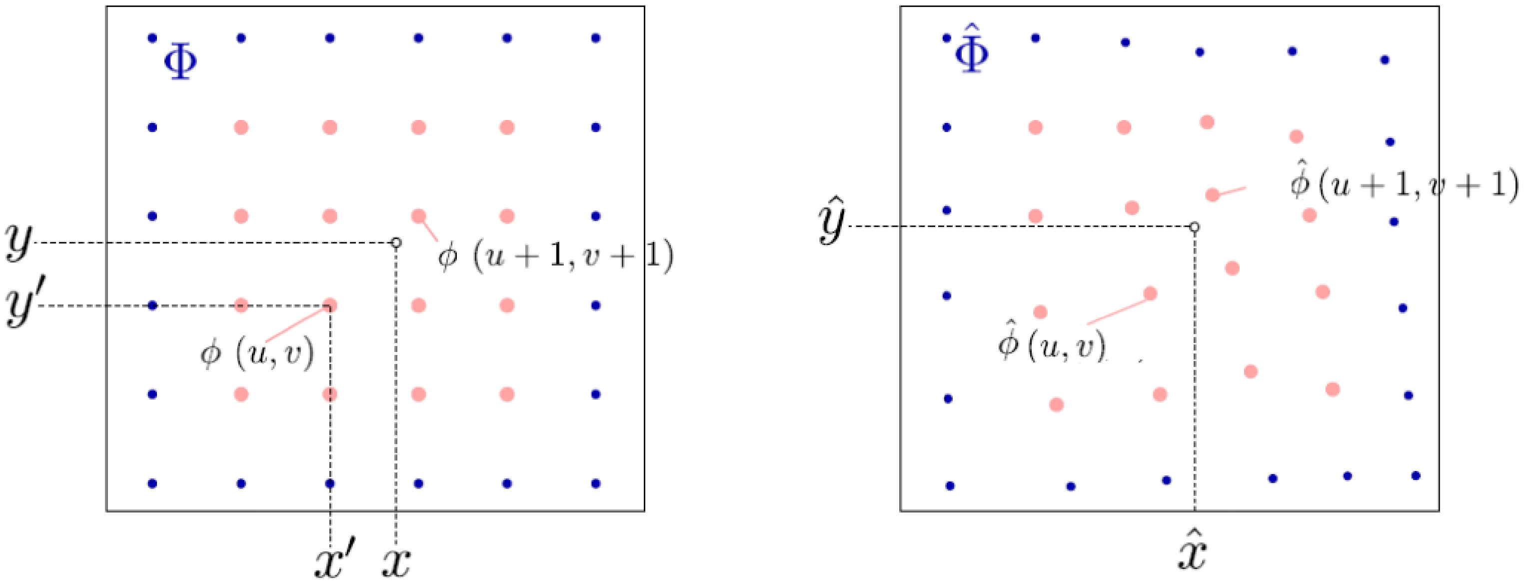
Free-form Nonrigid Transformation

The correspondence between each pair of points is not individually estimated.

A lattice with evenly spaced control points is used to compute the transformation



Free-form Nonrigid Transformation



$$\widehat{\Phi} = \Phi + \Delta\Phi$$

- $\widehat{\Phi}$ is estimated by minimizing SSD between two images
- Individual transformation between a pair of points is estimated from $\Delta\Phi$ by interpolation

Image Registration

Different strategies for registration

- Model-free registration
  - Good for
    - Registering two images viewing the same scene/object – small appearance difference
    - arbitrary scenes and objects
  - Suffer from
    - Large appearance/view changes
      - Identity, view point, illumination, background, etc.
    - Noise

Images from Helen Dataset http://www.ifp.illinois.edu/~vuongle2/helen/

- **Model-based registration (e.g., Active Appearance Model)**

Model-based Image Registration

Objective: register/fit an object model to a target object on an unseen image

Requirements:

- Object class is known
- A learned model from training data
  - Statistical models or statistical atlas

Resulting transformation:

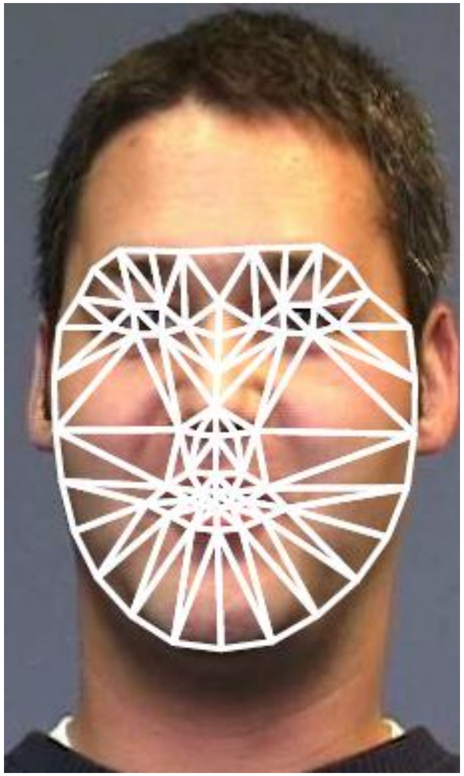
- Non-rigid
- Represented by a set of model parameters

The most popular model-based image registration is

Active Appearance Model and its variants

Active Appearance Model

- **First introduced by Edwards, Cootes, and Taylor for face analysis**
- **Revisited by Matthews and Baker with an efficient fitting algorithm**



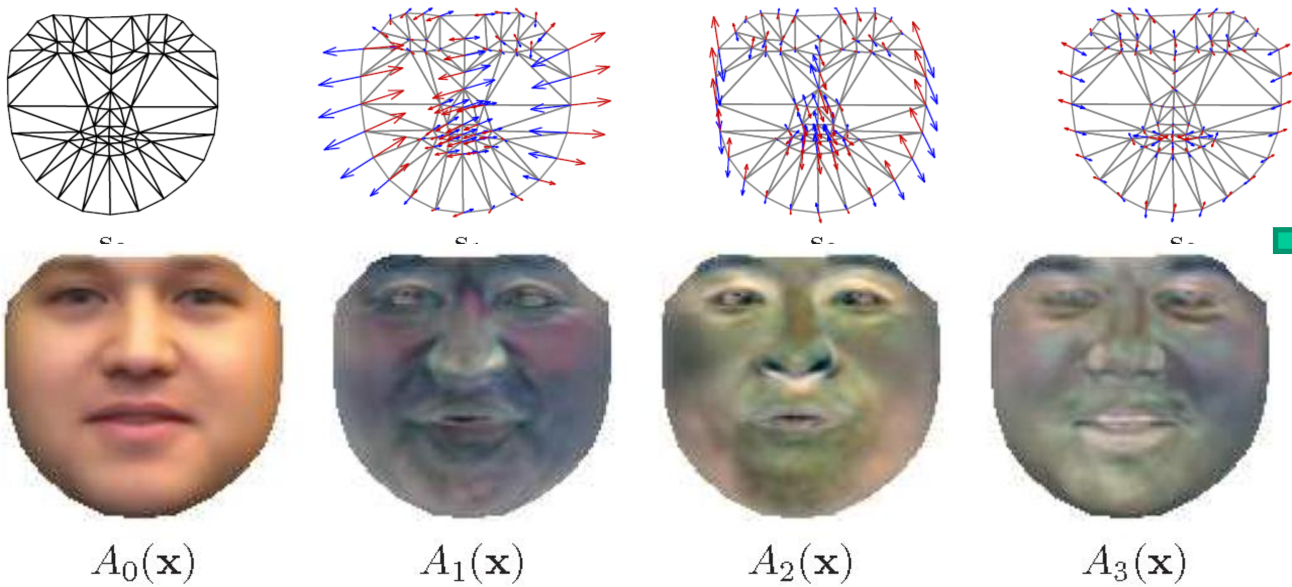
The target region/object is represented by a triangularized mesh

- Vertices of mesh represent a set of control points
- For each triangle, the location of any interior point can be defined by the 3 vertices

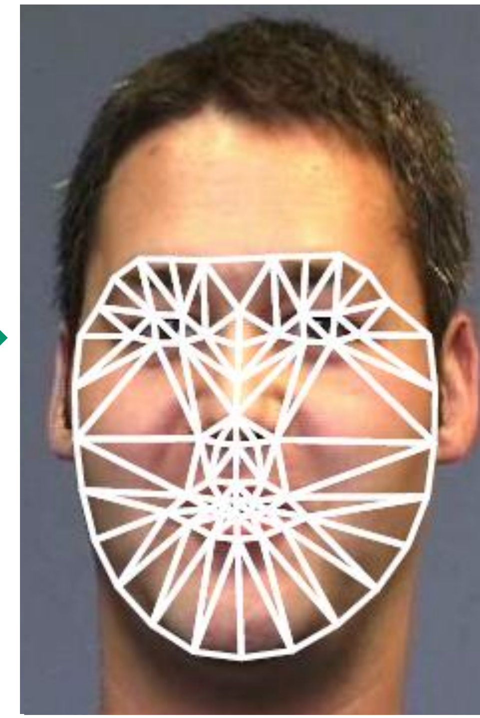
Adapted from Matthews and Baker, “Active Appearance Models Revisited”, IJCV 2004

Active Appearance Model

Model



Adapted from Matthews and Baker, “Active Appearance Models Revisited”, IJCV 2004

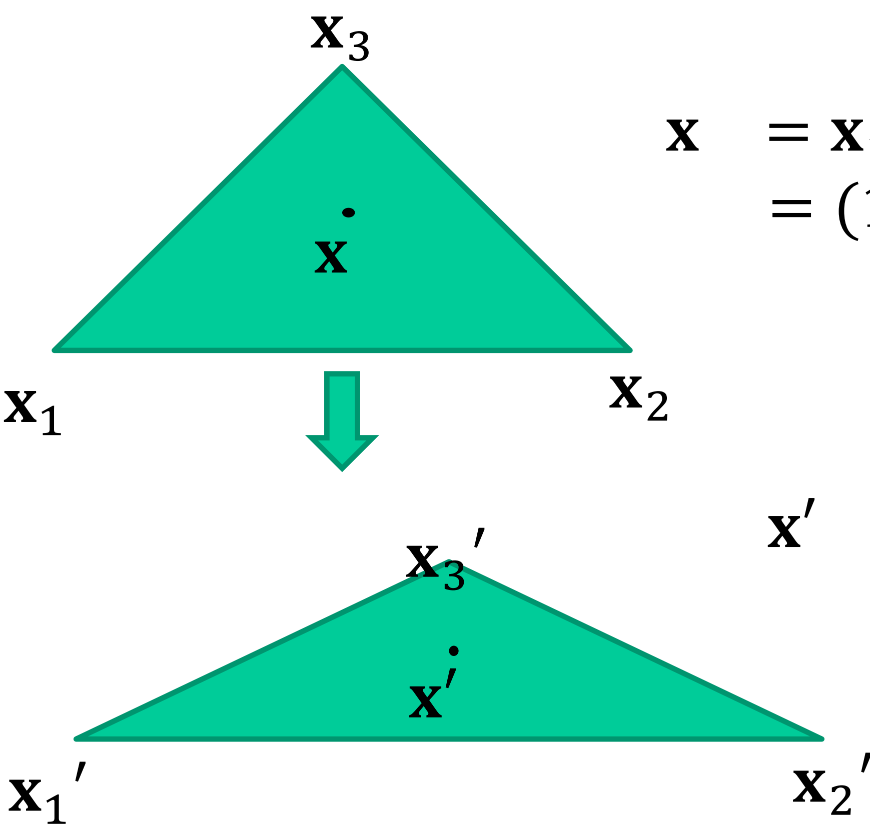


- For every triangle, an affine transformation is employed to established the correspondence between the model and the image
- A set of local affine transformation approximates nonrigid transformation → **Piecewise affine transformation**

Piecewise Affine Transformation

Assumption: the transformation in a local region (triangle) is linear

For any interior point in a triangle, it can be represented by



$$\begin{aligned}\mathbf{x} &= \mathbf{x_1} + \beta(\mathbf{x_2} - \mathbf{x_1}) + \gamma(\mathbf{x_3} - \mathbf{x_1}) \\ &= (1 - \beta - \gamma)\mathbf{x_1} + \beta\mathbf{x_2} + \gamma\mathbf{x_3}\end{aligned}$$

Transformation

$$\mathbf{x}' = (1 - \beta - \gamma)\mathbf{x_1}' + \beta\mathbf{x_2}' + \gamma\mathbf{x_3}'$$

Key problem is how to locate the vertices

Shape Model

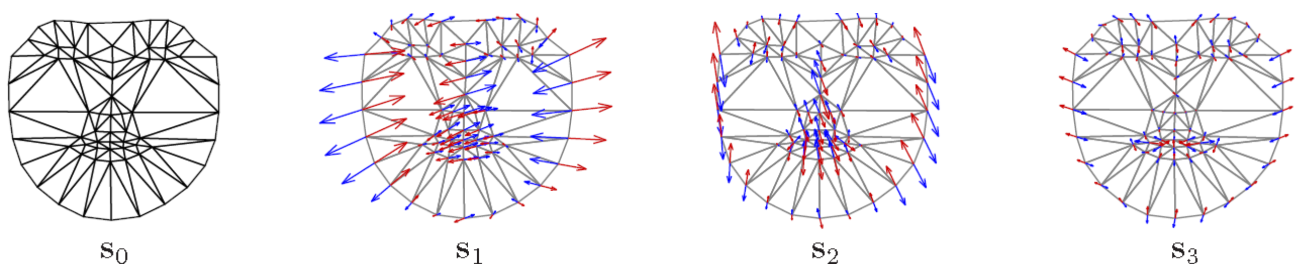
A shape of an AAM is defined as the coordinates of the N vertices forming the mesh

$$\mathbf{s} = [x_1, y_1, \cdots, x_N, y_N]^T$$

A shape can be represented as a linear combination of a mean shape and a set of shape variations

$$\mathbf{s} = \mathbf{s_0} + \sum_{i=1}^{M} p_i \, \mathbf{s_i}$$

- Major nonrigid variations found in a training set
- It is not allowed to deform shape arbitrarily

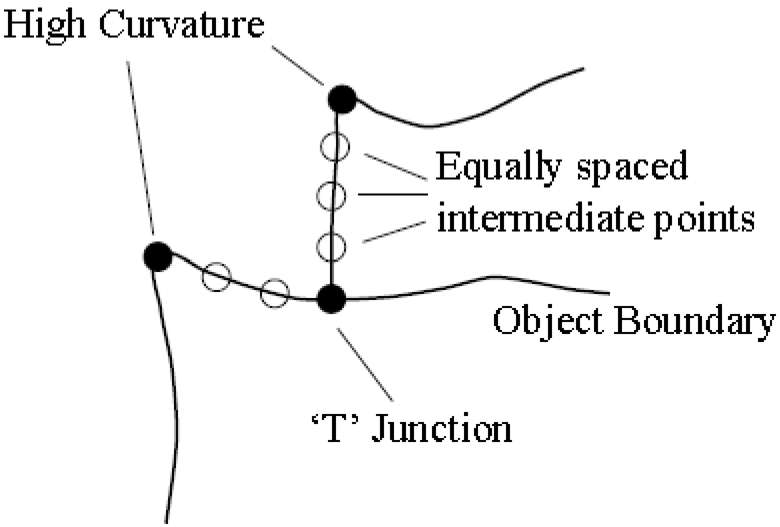


Adapted from Matthews and Baker, “Active Appearance Models Revisited”, IJCV 2004

Training a Shape Model

Major steps:

- **Label landmarks for each training image**

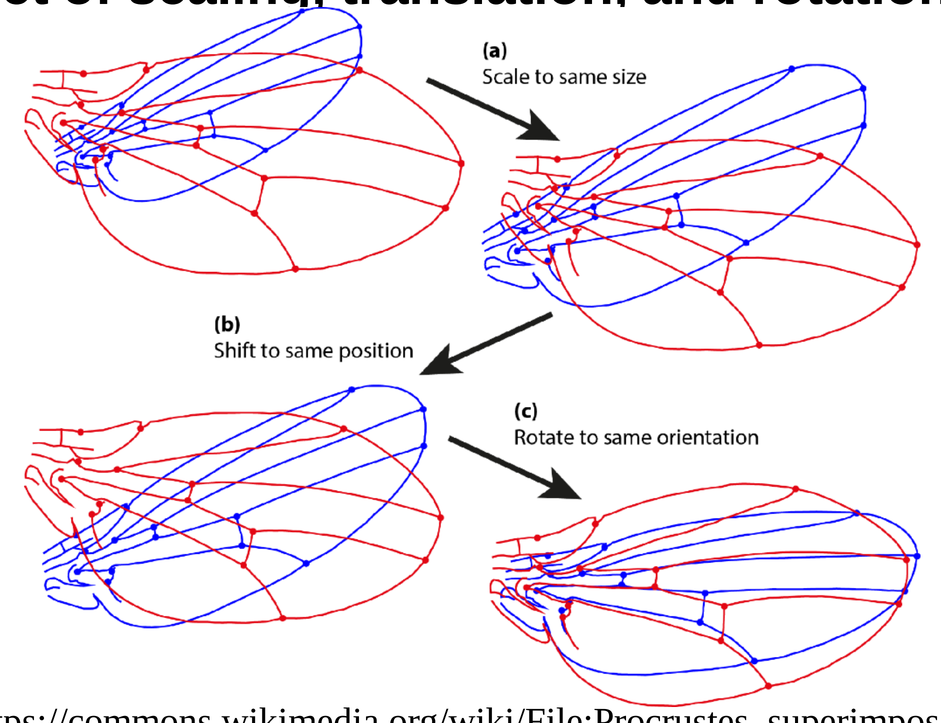


T. F. Cootes. Statistical models of appearance for computer vision

Training a Shape Model

Major steps:

- **Manually label landmarks for each training image**
- **Align training shapes by Procrustes Analysis to minimize the effect of scaling, translation, and rotation**



https://commons.wikimedia.org/wiki/File:Procrustes_superimposition.png

Training a Shape Model

Major steps:

- **Manually label landmarks for each training image**
- **Align training shapes by Procrustes Analysis to minimize the effect of scaling, translation, and rotation**

1. Translate each example so that its centre of gravity is at the origin.

2. Choose one example as an initial estimate of the mean shape and scale so that $|\bar{\mathbf{x}}| = 1$.

3. Record the first estimate as $\bar{\mathbf{x}}_0$ to define the default reference frame.

4. Align all the shapes with the current estimate of the mean shape.

5. Re-estimate mean from aligned shapes.

6. Apply constraints on the current estimate of the mean by aligning it with $\bar{\mathbf{x}}_0$ and scaling so that $|\bar{\mathbf{x}}| = 1$.

7. If not converged, return to **4**.

T. F. Cootes. Statistical models of appearance for computer vision

Training a Shape Model

Major steps:

- **Manually label landmarks for each training image**
- **Align training shapes by Procrustes Analysis to minimize the effect of scaling, translation, and rotation**
- **Perform PCA on the aligned shapes**
- **Retain the major variations**