Linear models of shape and appearance

Jan Kybic

Goals

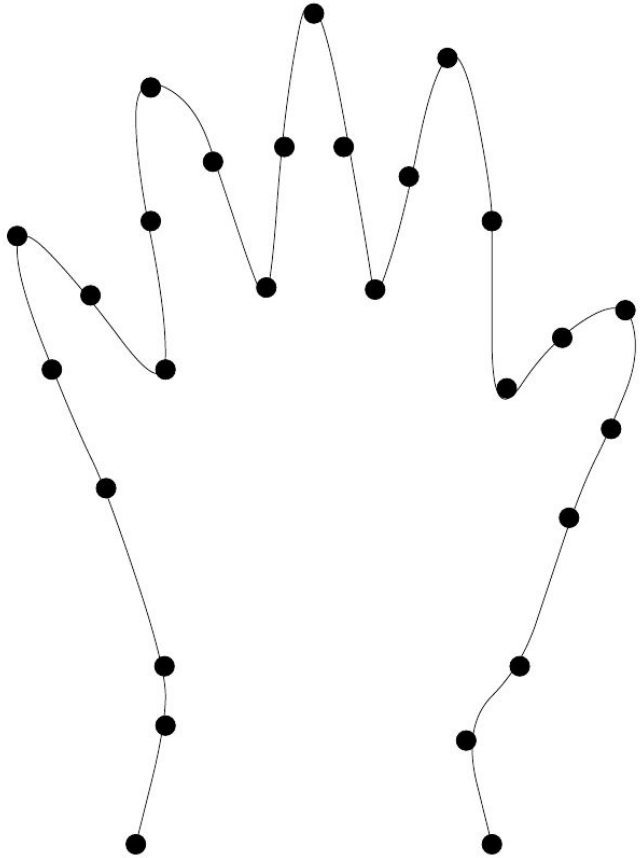
- Describe likely shapes and appearances
- Low-dimensional description
- A priori model for segmentation
- Improve segmentation results
- Simplify optimization

Models and techniques

- Point distribution models (PDMs)
- Active shape models - active contours + PDMs
- Active appearance models

- Sonka, Hlavac, Boyle book, Chapter 10
- Svoboda, Kybic, Hlavac companion book, Chapter 10
- Cootes et al: Active Appearance Models
- Cootes et al: Active Shape Models

Point model example



Training set

M examples, N landmarks

$$\mathbf{x}^1 = (x_1^1, y_1^1, x_2^1, y_2^1, \ldots, x_N^1, y_N^1)^T$$

Alignment

Suppose we have a moving shape and a reference shape described by landmark coordinates (x_i, y_i) and (x'_i, y'_i), respectively. We need to find a transformation consisting of rotation, translation, and scaling that transforms the moving shape onto the reference shape in the 'best' way [Section :10.3], defined as minimizing a sum of squared distances

$$E = \sum_{i=1}^{M} w_i \left\| s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} - \begin{bmatrix} x'_i \\ y'_i \end{bmatrix} \right\|^2 \tag{10.2}$$

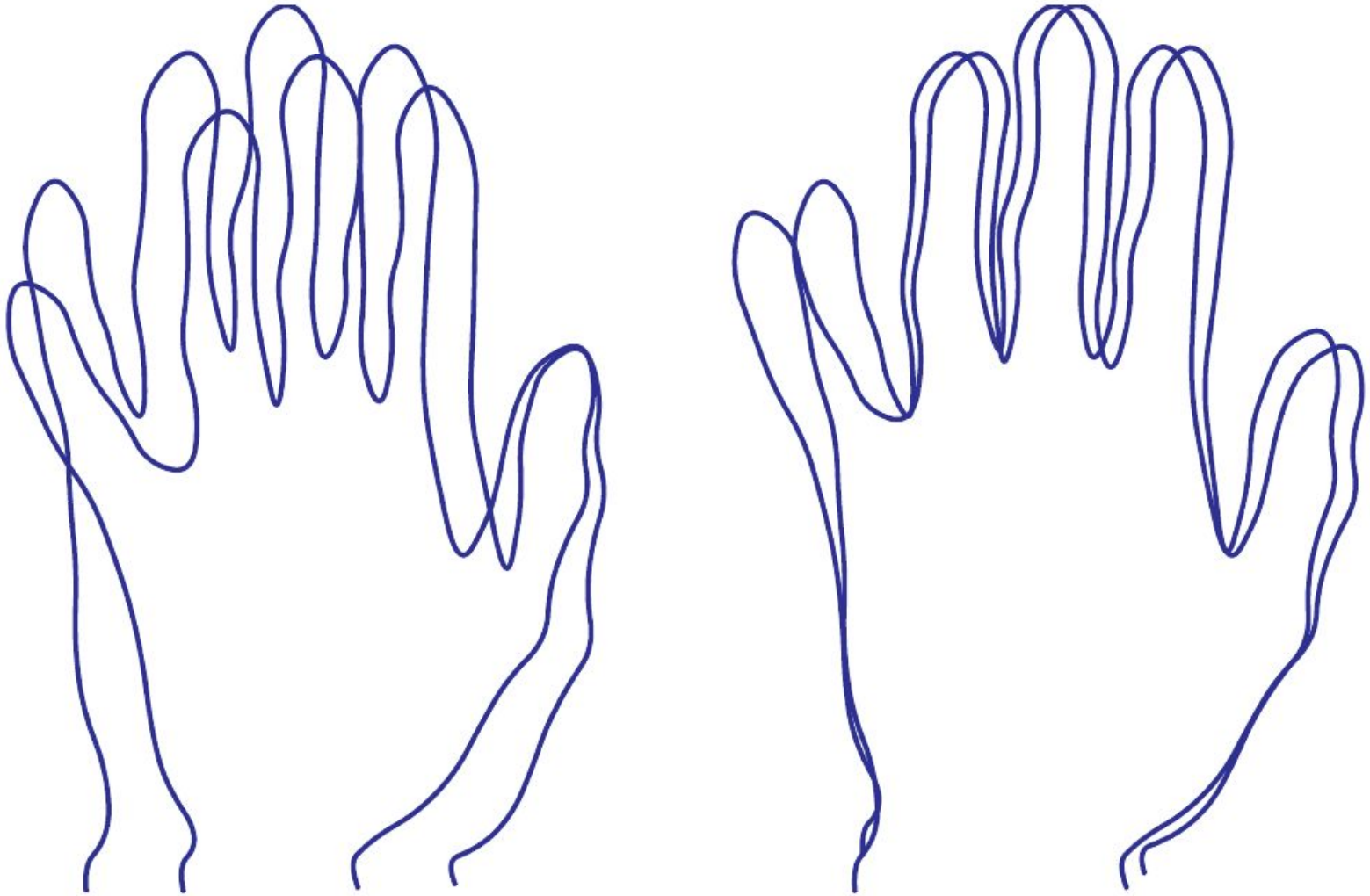
We decompose the minimization of $E(\theta, s, t_x, t_y)$ to an outer minimization with respect to θ and inner minimization with respect to s, t_x, t_y. Minimization with respect to s, t_x, t_y is performed by setting the corresponding partial derivatives to zero

$$\frac{\partial E}{\partial t_x} = 0\,, \qquad \frac{\partial E}{\partial t_y} = 0\,, \qquad \frac{\partial E}{\partial s} = 0\,,$$

which leads to the following system of linear equations

$$\begin{aligned}
s \sum_{i=1}^{M} w_i\, q(\;\; y_i, -x_i, \theta) - N\, t_x &= -\sum_{i=1}^{M} w_i\, x'_i \\
s \sum_{i=1}^{M} w_i\, q(-x_i, -y_i, \theta) - N\, t_y &= -\sum_{i=1}^{M} w_i\, y'_i \\
s \sum_{i=1}^{M} w_i^2 \Big(q^2(y_i, -x_i, \theta) + q^2(x_i, y_i, \theta) \Big) - t_x \sum_{i=1}^{M} w_i\, q(y_i, -x_i, \theta) - t_y \sum_{i=1}^{M} w_i\, q(-x_i, -y_i, \theta) &= \\
- \sum_{i=1}^{M} w_i\, x'_i\, q(y_i, -x_i, \theta) + \sum_{i=1}^{M} w_i\, y'_i\, q(x_i, -y_i, \theta)\,, &
\end{aligned} \tag{10.3}$$

where $q(a, b, \theta) = a \sin\theta + b \cos\theta$. The dependency $E(\theta) = \min_{(s,t_x,t_y)} E(\theta, s, t_x, t_y)$ is non-linear, so the outer minimization with respect to θ is performed numerically. This normally only needs a few iterations, as the function is smooth and one dimensional.



before and after alignment

Algorithm 10.5: Approximate alignment of similar training shapes

1. In a pairwise fashion, rotate, scale, and align each $\mathbf{x}^i$ with $\mathbf{x}^1$, for $i = 2, 3, \ldots, M$ to give the set $\{\mathbf{x}^1, \hat{\mathbf{x}}^2, \hat{\mathbf{x}}^3, \ldots, \hat{\mathbf{x}}^M\}$.
2. Calculate the mean of the transformed shapes (the details of this procedure are outlined in Section 10.3).
3. Rotate, scale, and align the mean shape to align to $\mathbf{x}^1$.
4. Rotate, scale, and align $\hat{\mathbf{x}}^2, \hat{\mathbf{x}}^3, \ldots, \hat{\mathbf{x}}^M$ to match to the adjusted mean.
5. If the mean has not converged, go to step 2.

mean shape is given by

$$\bar{\mathbf{x}} = \left[\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2, \ldots, \bar{x}_N, \bar{y}_N\right]$$

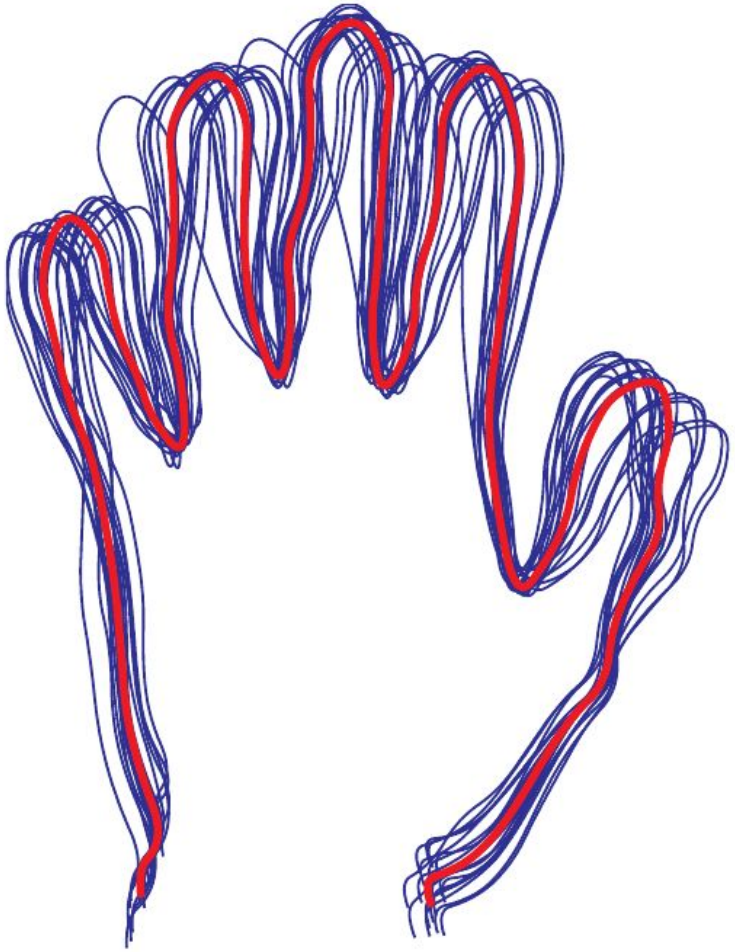
$$\bar{x}_j = \frac{1}{M}\sum_{i=1}^{M} \hat{x}_j^i \qquad \text{and} \qquad \bar{y}_j = \frac{1}{M}\sum_{i=1}^{M} \hat{y}_j^i \,.$$

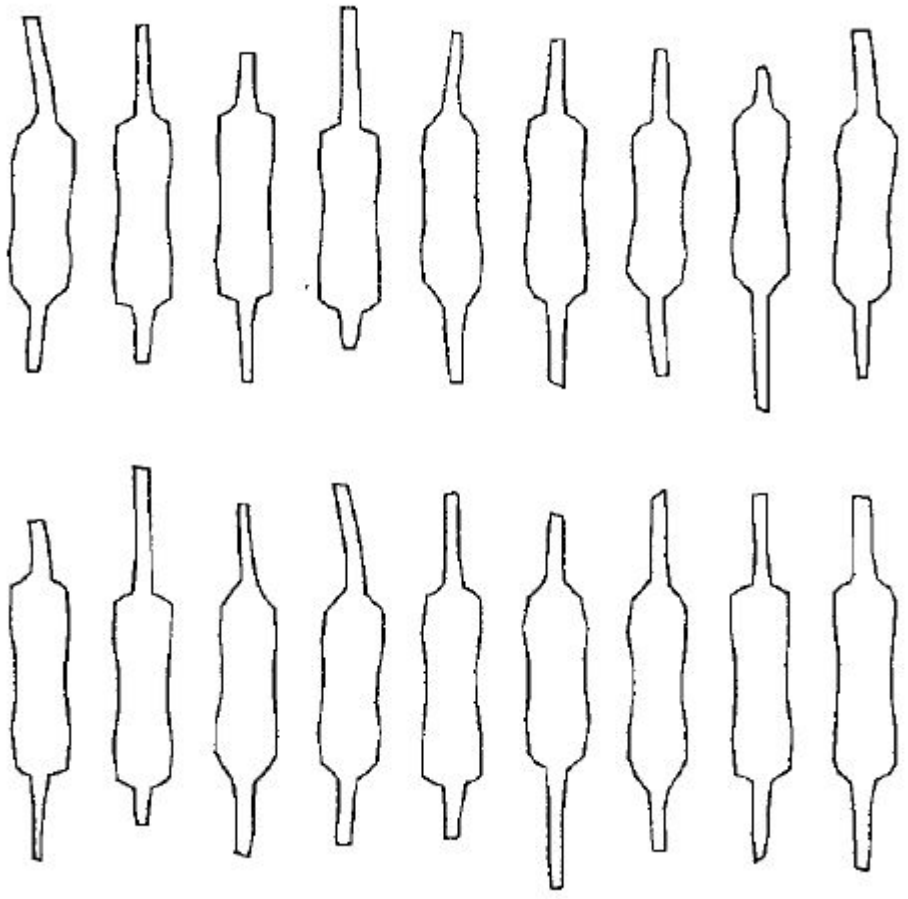
Statistical model

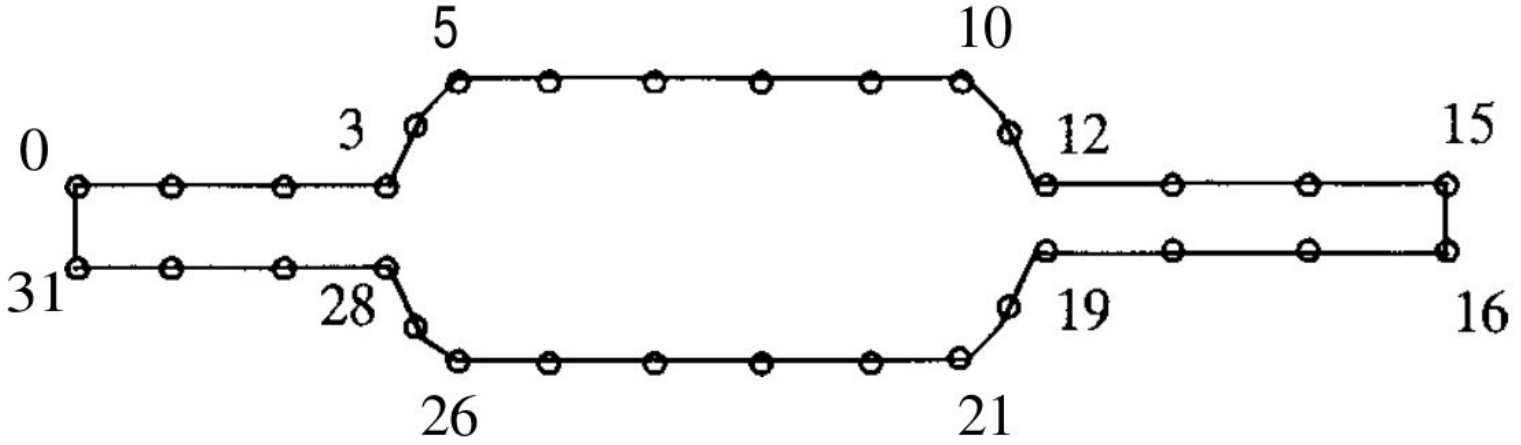
$$\delta \mathbf{x}^i = \hat{\mathbf{x}}^i - \bar{\mathbf{x}} \, .$$

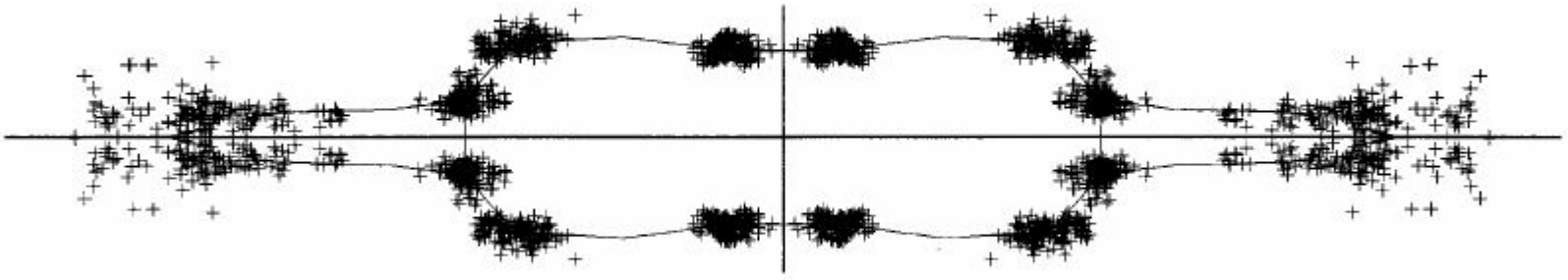
is a zero-mean multivariate Gaussian random vector of dimension *Md.*
Covariance matrix:

$$S = \frac{1}{M} \sum_{i=1}^{M} \delta \mathbf{x}^i \, (\delta \mathbf{x}^i)^T$$






5
10
0
3
12
15
31
28
19
16
26
21



Dimensionality reduction

Find a low dimensional model for
x

$$\mathbf{x} \approx \bar{\mathbf{x}} + P_t\, \mathbf{b}_t$$

$\delta\mathbf{x}$

$$P_t = \left[\mathbf{p}^1 \mathbf{p}^2 \mathbf{p}^3 \ldots \mathbf{p}^t\right]$$

$$\mathbf{b}_t = \left[b_1, b_2, \ldots, b_t\right]^T$$

Best projection

$$\min \mathrm{E}\left[\|\delta\mathbf{x} - \mathbf{p}b\|^2\right]$$

$$\text{assume } \|\mathbf{p}\| = 1 \quad \text{then} \quad b = \langle \mathbf{p}, \delta\mathbf{x} \rangle$$

$$\max \mathrm{E}\left[\langle \mathbf{p}, \delta\mathbf{x} \rangle^2\right] = \mathbf{p}^T \mathrm{E}\left[\ \delta\mathbf{x}\, \delta\mathbf{x}^T\ \right] \mathbf{p} = \mathbf{p}^T S \mathbf{p}$$

$$L = \mathbf{p}^T S\mathbf{p} - \quad \lambda(\|\mathbf{p}\|^2 - 1)$$

$$\nabla_{\mathbf{p}} L = 2S\mathbf{p} - 2\lambda\mathbf{p}$$

$$S\mathbf{p} = \lambda\mathbf{p}$$

$$\mathbf{p}^T S\mathbf{p} = \lambda$$

The basis **p** is the eigenvector of *S* corresponding to the largest eigenvalue.

PCA

$$S\mathbf{p}_i = \lambda_i \mathbf{p}_i \ .$$

Use *K* largest eigenvalues.

$$\mathbf{x} \approx \bar{\mathbf{x}} + P_t \, \mathbf{b}_t$$

Further, it can be shown that the variance of b_i over the training set will be the associated eigenvalue λ_i; accordingly, for 'well-behaved' shapes we might expect

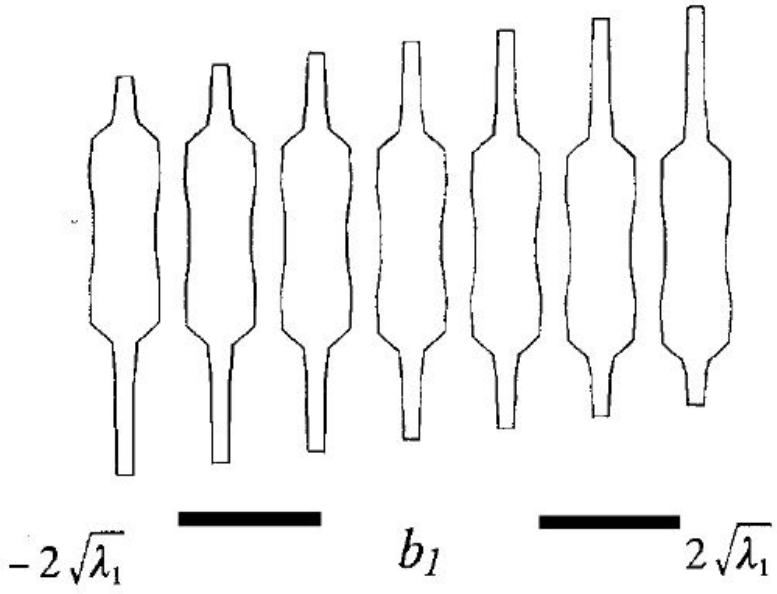
$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i}$$

—that is, most of the population is within 3σ of the mean. This allows us to generate, from knowledge of P and λ_i, plausible shapes that are not part of the training set.

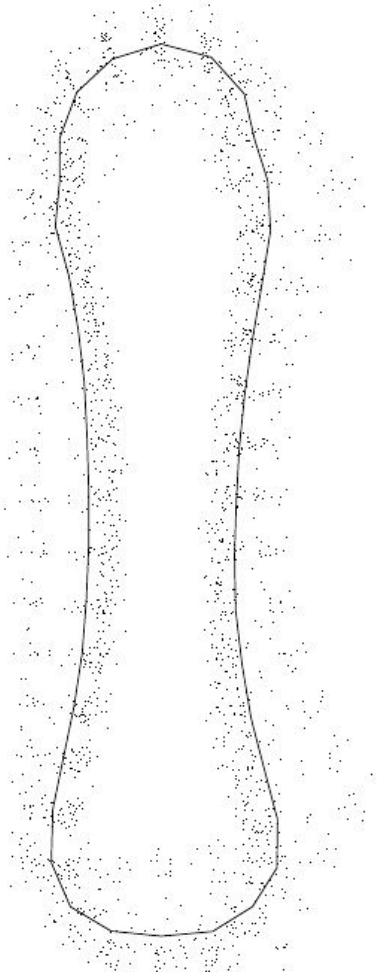
TABLE 1

Eigenvalues of the Covariance Matrix Derived from a Set of Resistor Shapes

Eigenvalue	$\frac{\lambda_i}{\lambda_T} \times 100\%$
λ_1	66%
λ_2	8%
λ_3	5%
λ_4	4%
λ_5	3%
λ_6	3%



Metacarpal bone



Index i	$\lambda_i/\lambda_{\text{total}}$ [%]	Cumulative total
1	63.3	63.3
2	10.8	74.1
3	9.5	83.6
4	3.4	87.1
5	2.9	90.0
6	2.5	92.5
7	1.7	94.2
8	1.2	95.4
9	0.7	96.1
10	0.6	96.7
11	0.5	97.2
12	0.4	97.6
13	0.3	97.9
14	0.3	98.2
15	0.3	98.5
16	0.2	98.7

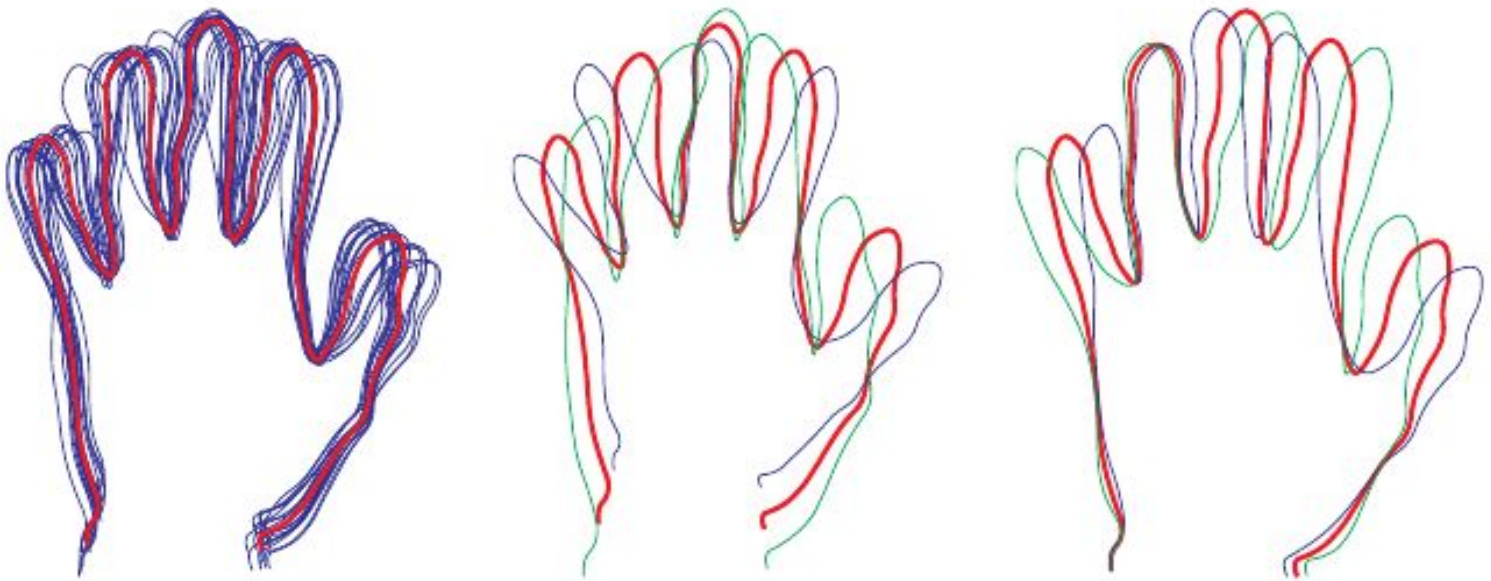
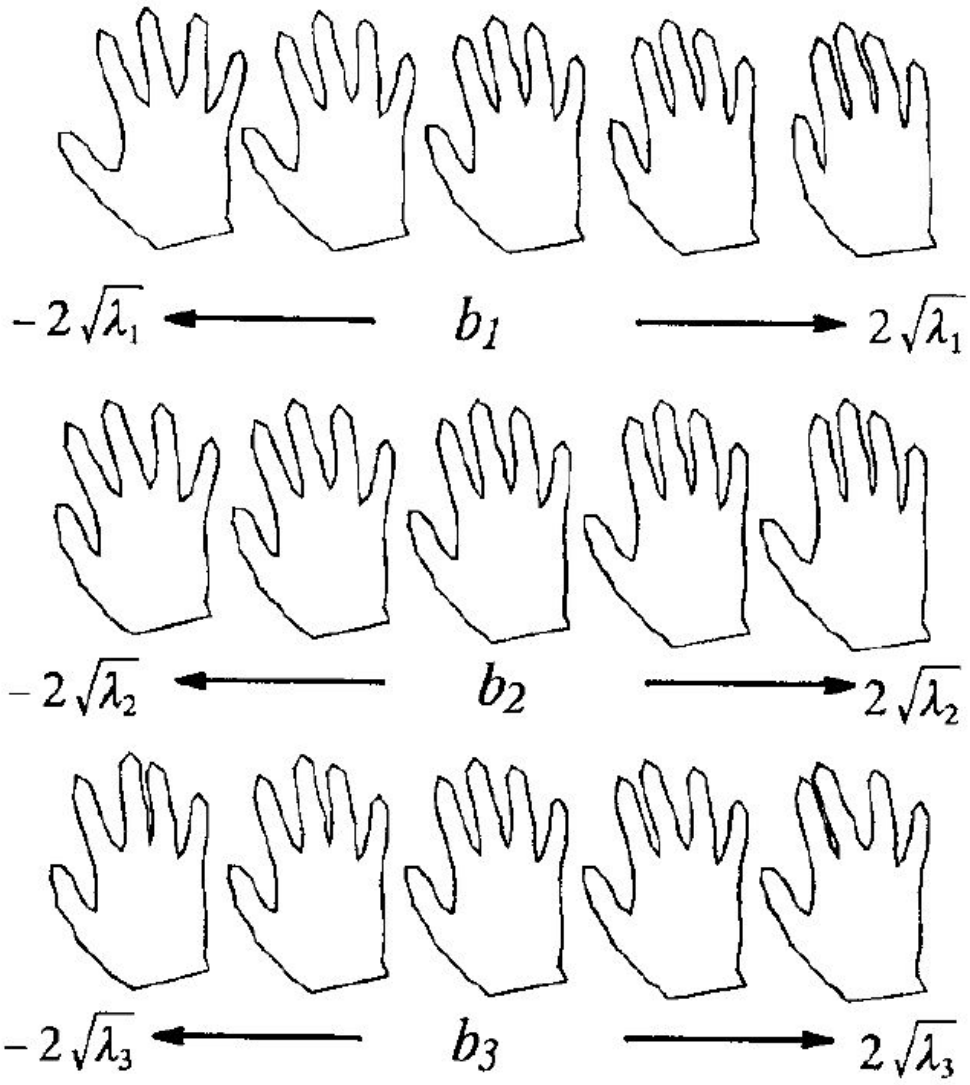


Figure 10.5: (a) The mean shape (in red) superimposed over all shapes after alignment. Changes corresponding to the (b) first and (c) second mode. The mean shape is in red, the shape corresponding to $-3\sqrt{\lambda}$ in blue and the shape corresponding to $+3\sqrt{\lambda}$ in green.


$-2\sqrt{\lambda_1}$
b_1
$2\sqrt{\lambda_1}$
$-2\sqrt{\lambda_2}$
b_2
$2\sqrt{\lambda_2}$
$-2\sqrt{\lambda_3}$
b_3
$2\sqrt{\lambda_3}$

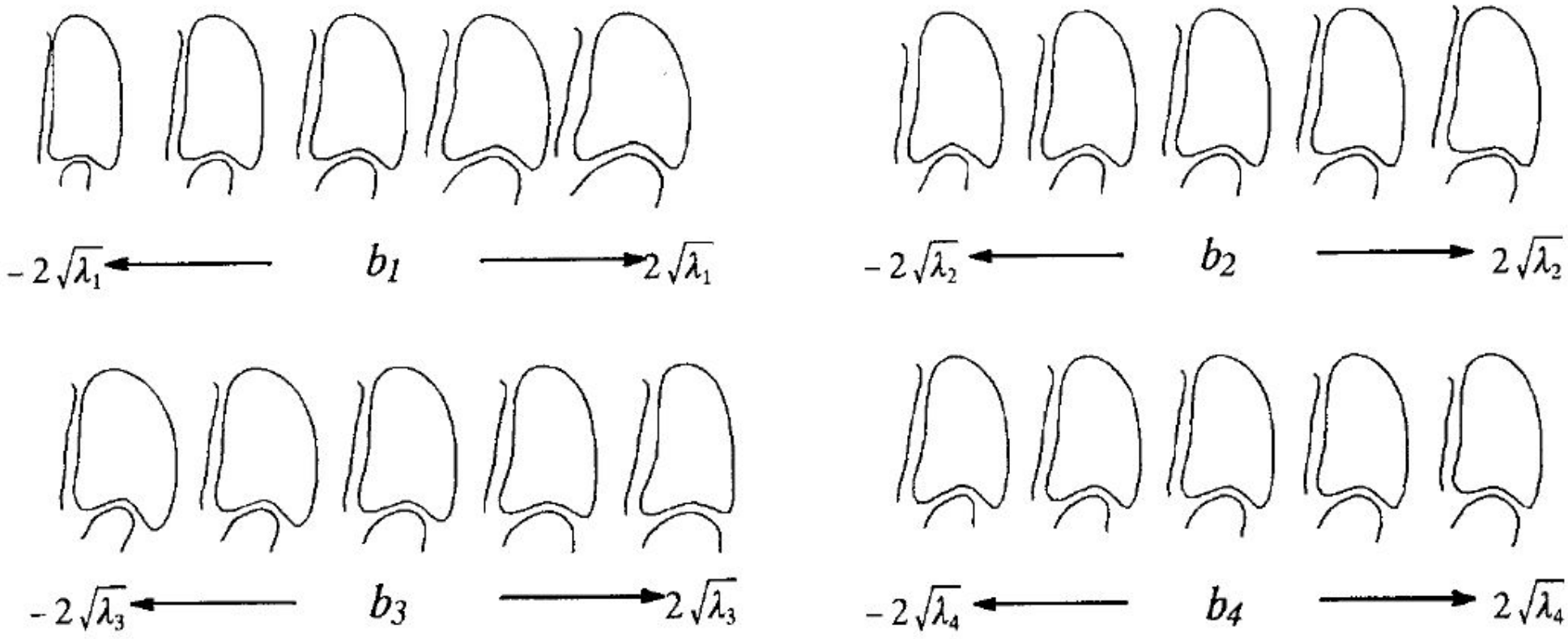
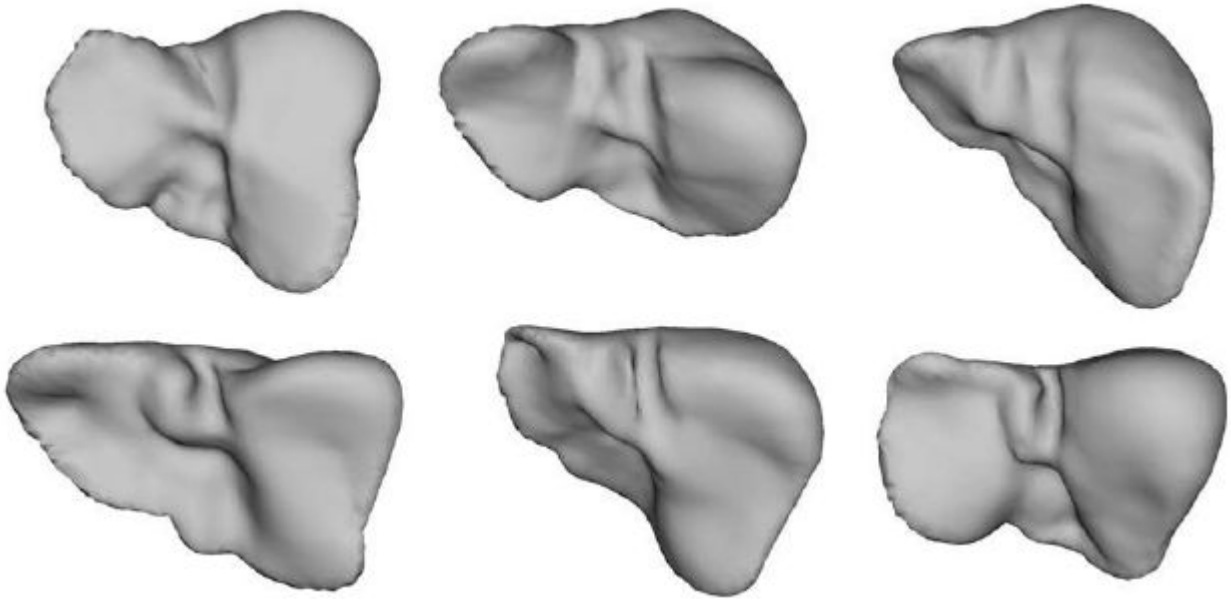


FIG. 12. Effects of varying each of the first four parameters of the heart ventricle model individually.

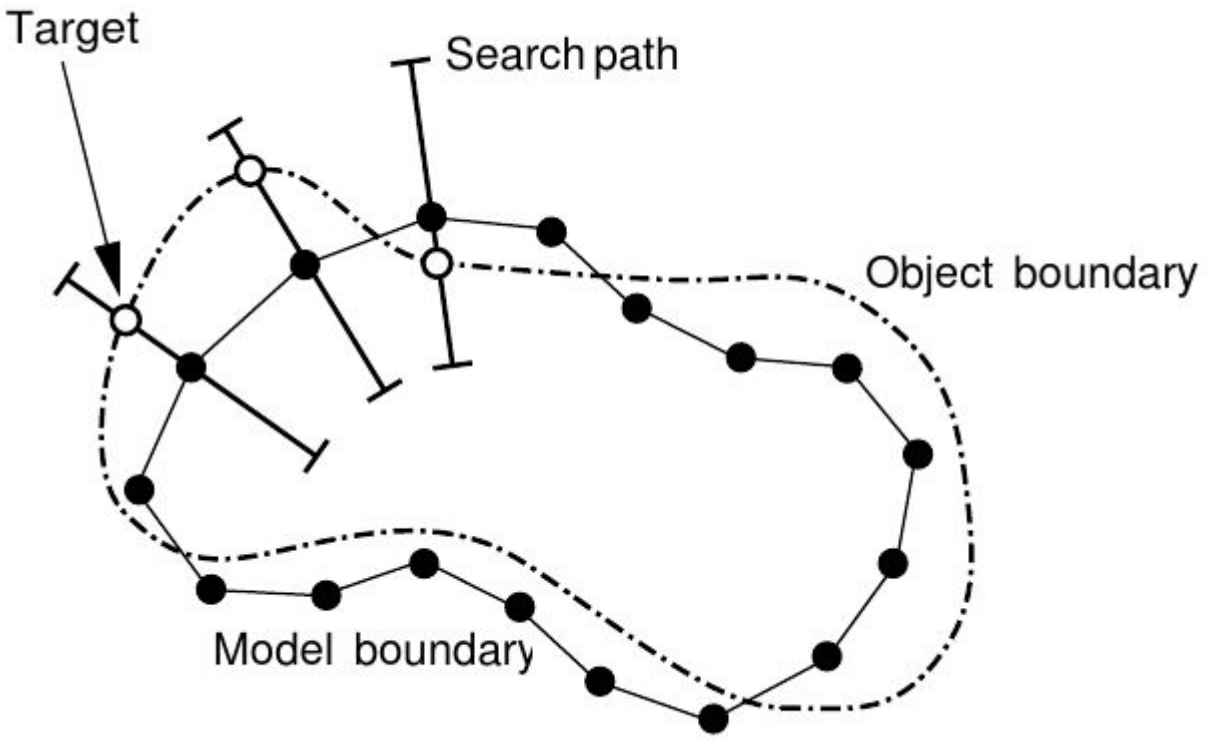


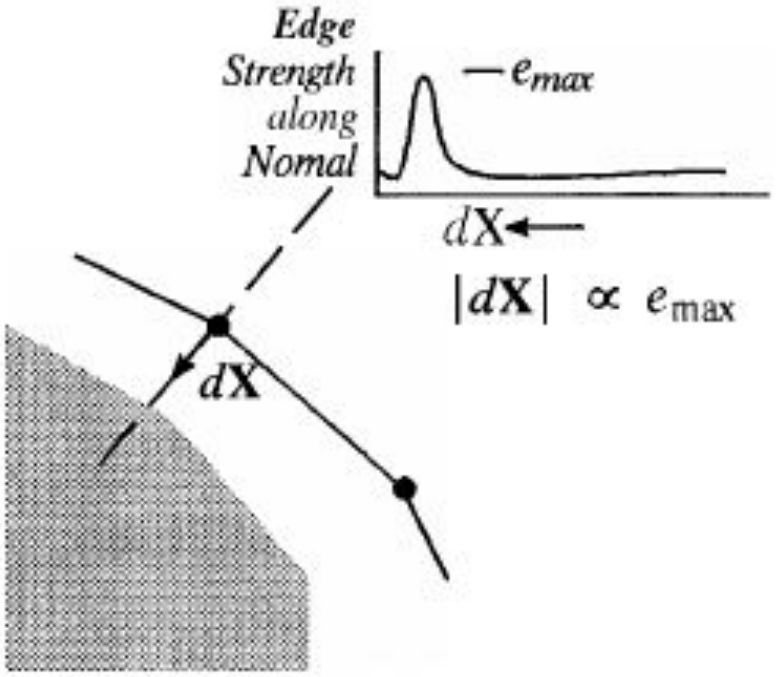
Liver shape variations. Columns = eigenmodes.

Fiting PDM to data
Active Shape models

Given an image and an initial position find

- Pose (translation, rotation, scale)
- Shape parameter vector **b**


Target
Search path
Object boundary
Model boundary


Edge
Strength
along
Nomal
e_{max}
dX
$|d\mathbf{X}| \propto e_{max}$
$d\mathbf{X}$

Fitting an ASM

- Find target positions for each landmark
- Adjust pose
- Adjust model parameters **b**
- Iterate until convergence

5. Determine the model adjustment $\delta \mathbf{b}_t$ that best approximates $\delta \widetilde{\mathbf{x}}$. From equation (10.5) we have

$$\widetilde{\mathbf{x}} \approx \bar{\mathbf{x}} + P_t \, \mathbf{b}_t$$

and we seek $\mathbf{d}b_t$ such that

$$\widetilde{\mathbf{x}} + \delta \widetilde{\mathbf{x}} = \bar{\mathbf{x}} + P_t(\mathbf{b}_t + \delta \mathbf{b}_t) \, .$$

Hence

$$\delta \widetilde{\mathbf{x}} \approx P_t \, \delta \mathbf{b}_t \, .$$

With the properties of eigen-matrices, we can deduce

$$\delta \mathbf{b}_t = P_t^T \, \delta \widetilde{\mathbf{x}}$$

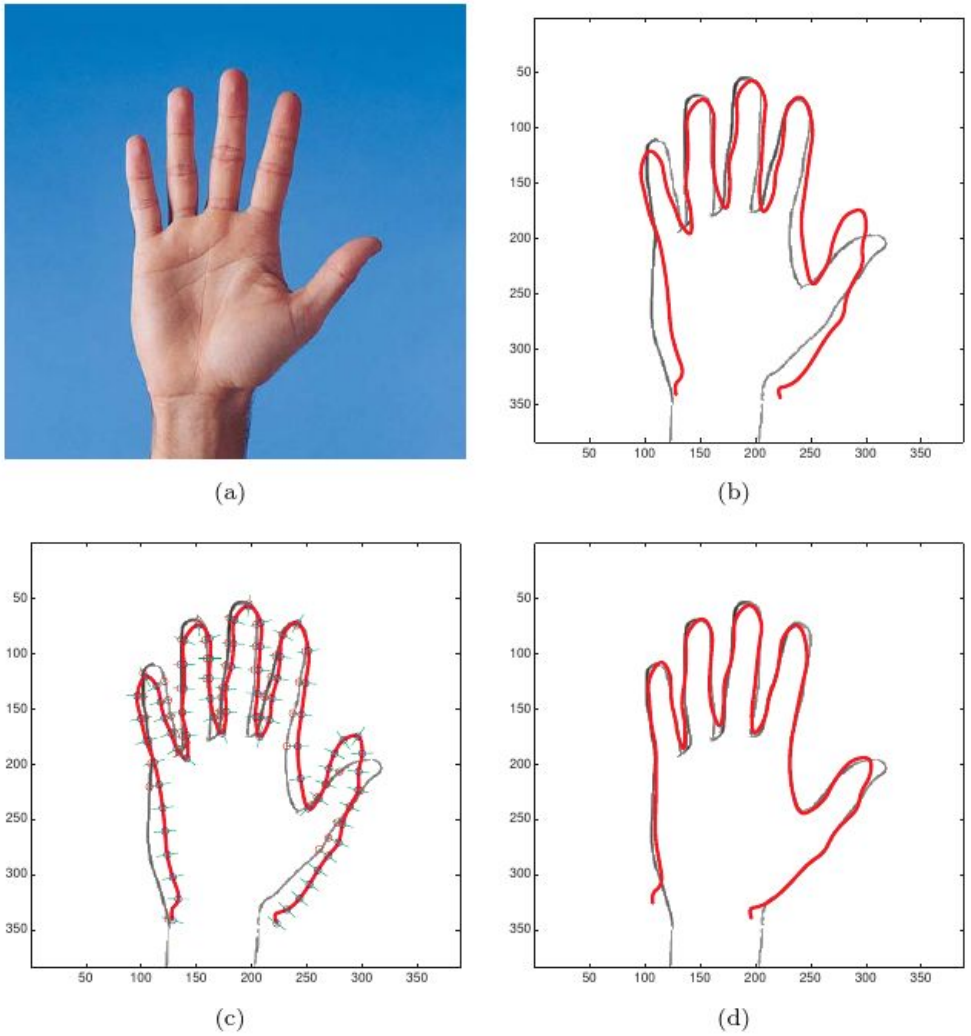


Figure 10.6: (a) Hand image. (b) The corresponding edge map with superimposed initial shape. (c) First iteration of the fitting process with the shape contour in red, current landmarks positions as blue circles, search lines in green, and new landmark positions as red circles. (d) Final fit.

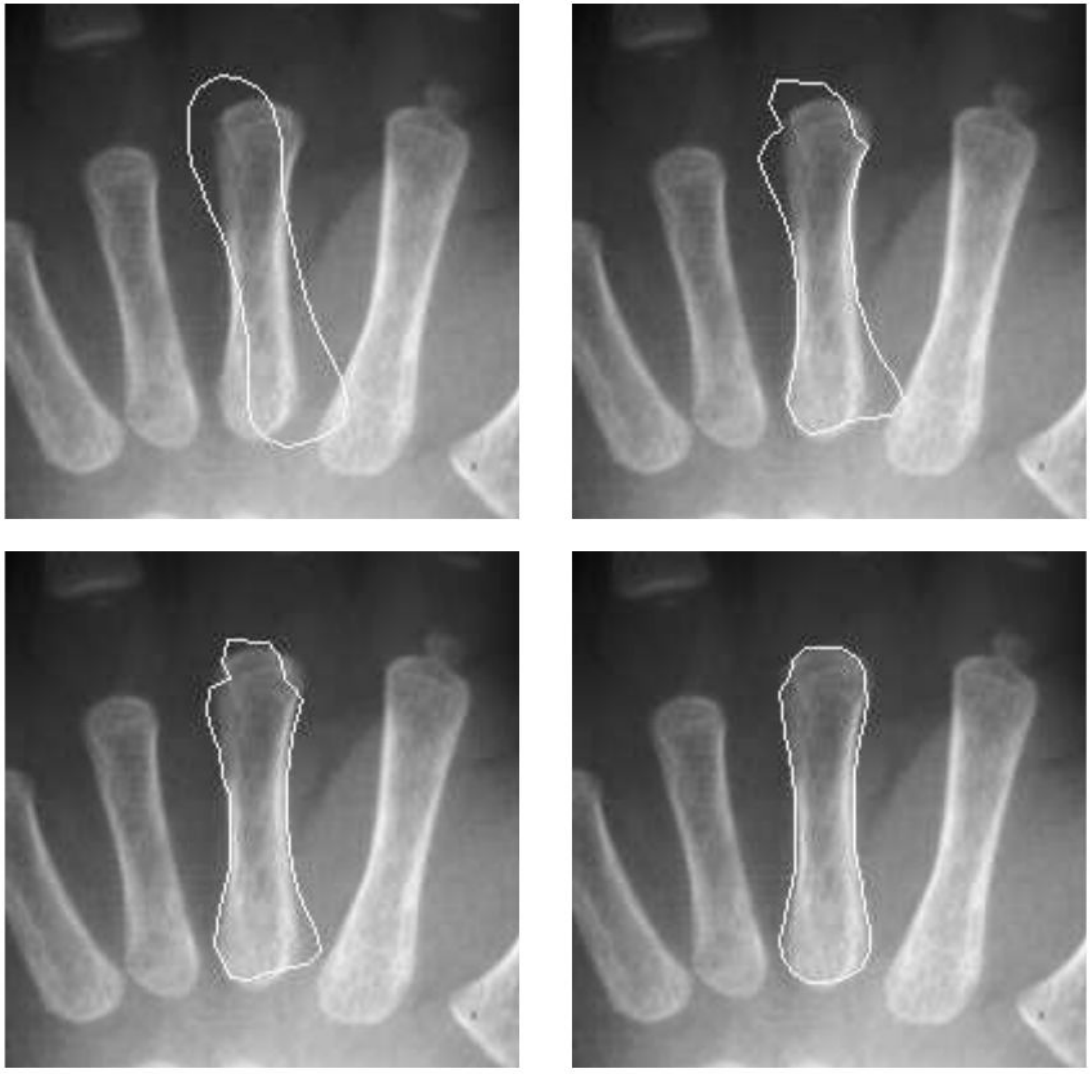
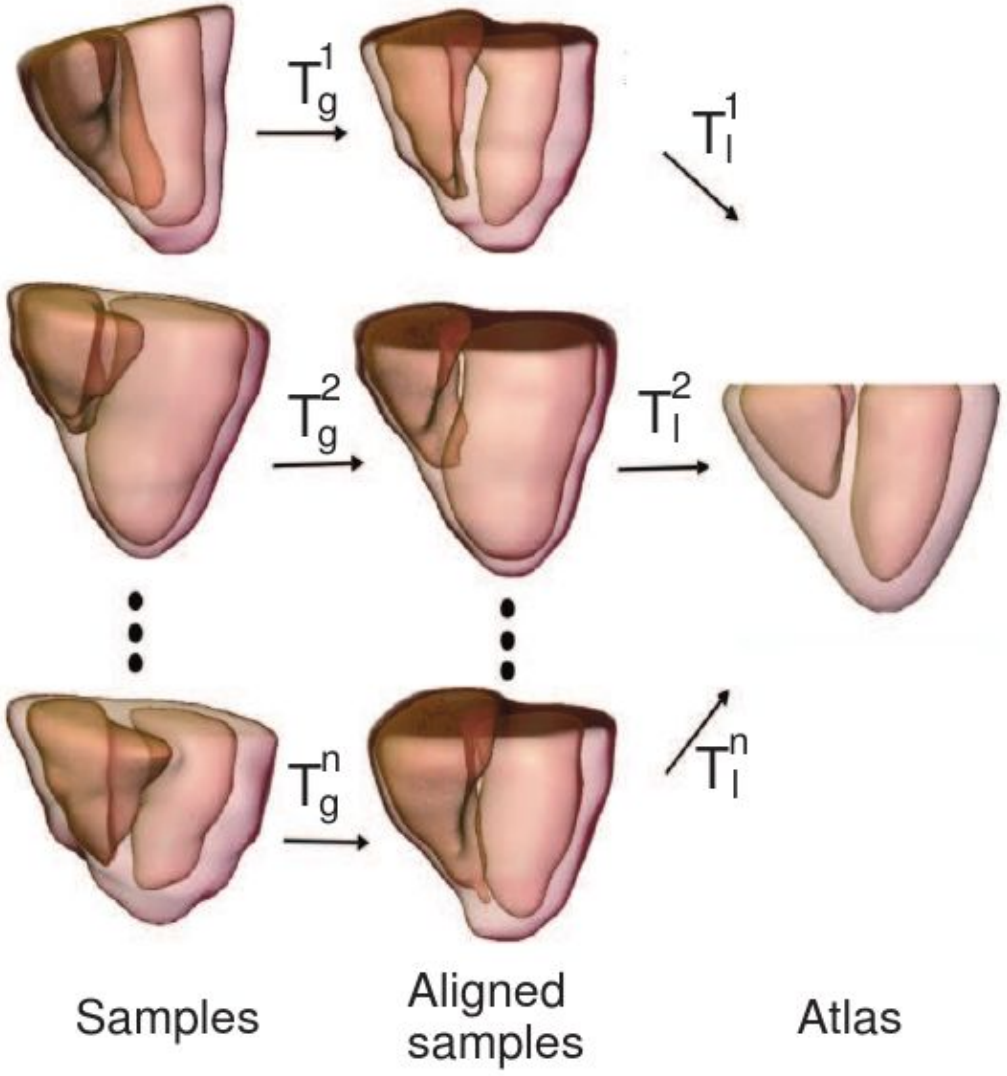
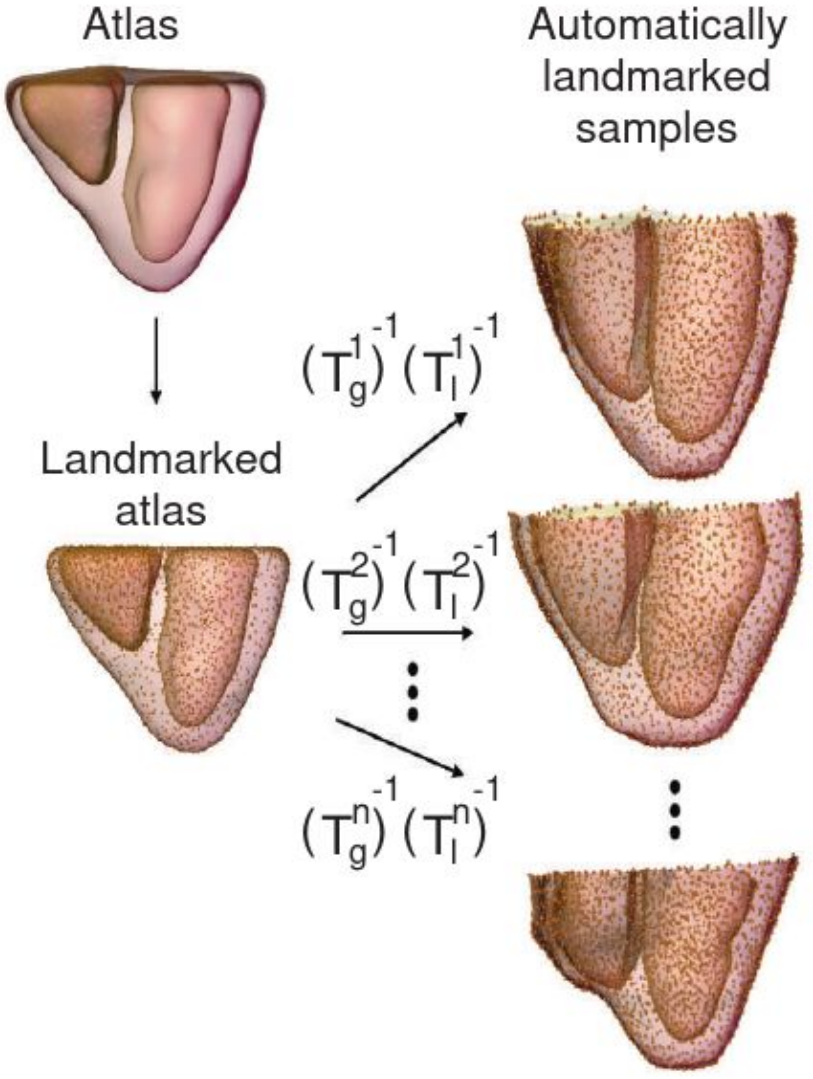


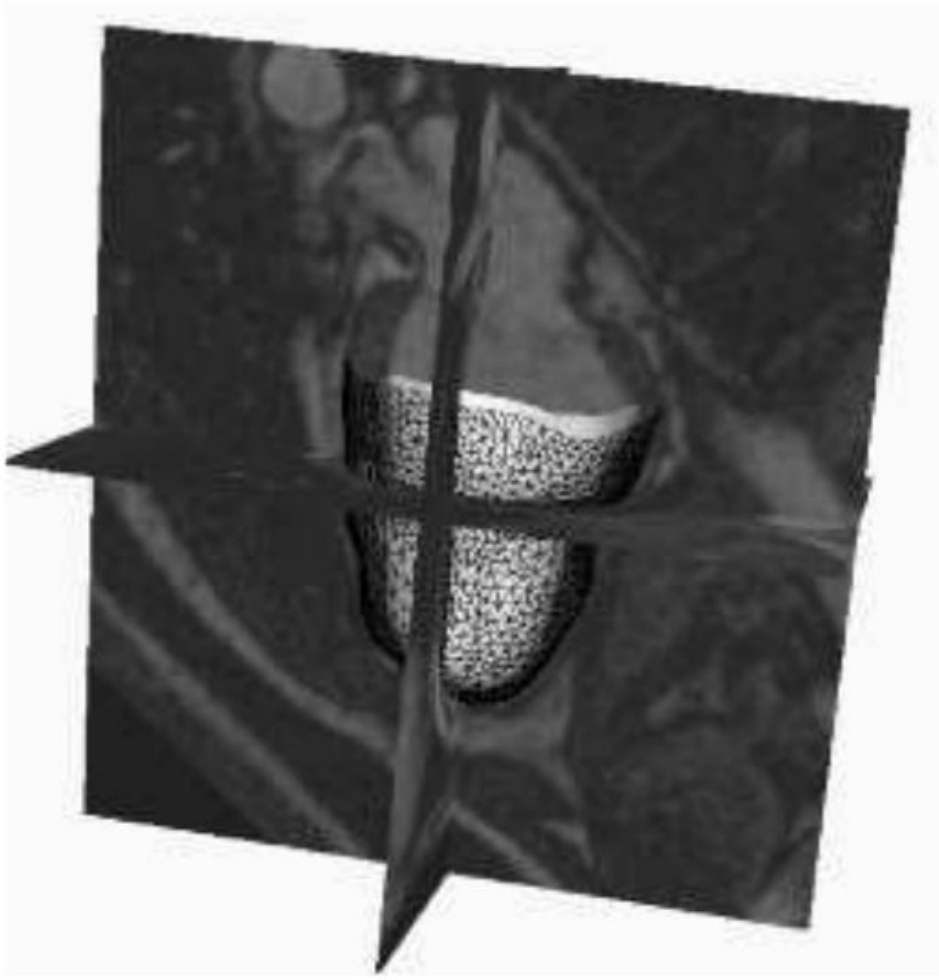
Figure 10.12: Fitting an ASM to a metacarpal; various stages of convergence—initialization, 3, 6, and 10 iterations. *Courtesy of N. D. Efford, School of Computer Studies, University of Leeds.*

Heart segmentation (3D)

- Non-rigid registration for alignment
- No need to mark landmarks in all images
- Landmarks propagated from the atlas shape
- or use keypoint matching
- Segmentation of new images with ASM


T_g^1
T_l^1
T_g^2
T_l^2
T_g^n
T_l^n
Samples
Aligned samples
Atlas

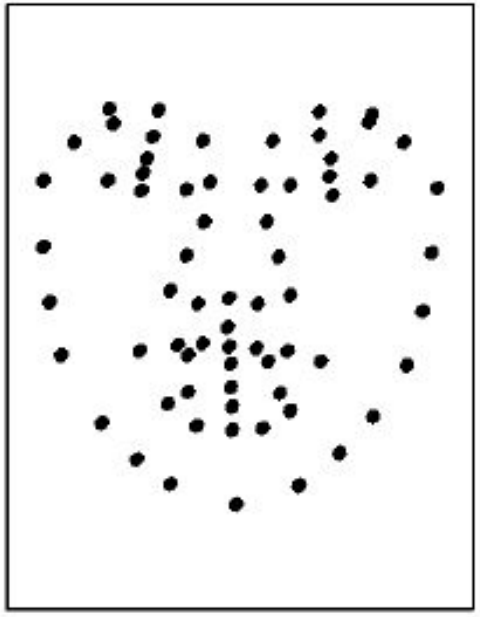

Atlas
Automatically landmarked samples
Landmarked atlas
$(T_g^1)^{-1}(T_l^1)^{-1}$
$(T_g^2)^{-1}(T_l^2)^{-1}$
$(T_g^n)^{-1}(T_l^n)^{-1}$



Active appearance models

- ASM + linear appearance model
- Low-dimensional model
- Goals:
  - Model appearance, generate synthetic shapes
  - Segmentation
  - Classification and diagnostics

Labelled image



Points

Shape-free patch

Simultaneous change of appearance and shape

c_1: 

c_2: 

c_3: 

c_4: 

Find landmarks for training images.

Algorithm 10.7: AAM construction

1. Compute an ASM and approximate each shape sample as a linear combination of eigenvectors, where $\mathbf{b}_s = P_s^T(\mathbf{x} - \bar{\mathbf{x}})$ represents the sample shape parameters (equation 10.3).
2. Warp each image to the mean shape using a linear or non-linear image interpolation.
3. Normalize each image to the average intensity and unit variance $\bar{\mathbf{g}}$.
4. Perform a PCA on the normalized intensity images.

5. Express each intensity sample as a linear combination of eigenvectors, where $\mathbf{b}_g = P_g^T(\mathbf{g} - \bar{\mathbf{g}})$ represents the sample gray-level parameters.
6. Concatenate the shape coefficient vectors $\mathbf{b}_s$ and gray-level intensity coefficient vectors $\mathbf{b}_g$ in the following manner

$$\mathbf{b} = \begin{bmatrix} W\,\mathbf{b}_s \\ \mathbf{b}_g \end{bmatrix} = \begin{bmatrix} W\,P_s^T(\mathbf{x} - \bar{\mathbf{x}}) \\ P_g^T(\mathbf{g} - \bar{\mathbf{g}}) \end{bmatrix}, \tag{10.6}$$

where W is a diagonal weighting matrix that relates the different units of shape and gray-level intensity coefficients.

7. Apply PCA to the sample set of all $\mathbf{b}$ vectors, yielding the model

$$\mathbf{b} = Q\mathbf{c}\,, \tag{10.7}$$

where Q is a matrix consisting of eigenvectors (from equation 10.6) and $\mathbf{c}$ are the resulting model coefficients characterizing how the model instance differs from the mean shape and mean appearance. In other words, with the zero vector $\mathbf{c} = 0$, the modeled instance corresponds to the mean shape and appearance.

AAM fitting

- Input: image
- Output:
  - geometrical transformation (pose)
  - shape parameters
  - appearance parameters

Residual minimization

model gray-level intensity vector $\mathbf{g}_m$ from appearance coefficients **c**

target image patch warped to the mean shape

$$\mathbf{g}_s = T\mathbf{u}^{-1}(g_{im})$$

global intensity transformation $\breve{\mathbf{u}}$

Minimize residual $\mathbf{r}(\mathbf{p}) = \mathbf{g}_s(\mathbf{p}) - \mathbf{g}_m(\mathbf{p})\,,$

Assume linearity by Taylor expansion

$$\mathbf{r}(\widetilde{\mathbf{p}} + \delta\mathbf{p}) \approx \mathbf{r}(\widetilde{\mathbf{p}}) + \frac{\partial \mathbf{r}(\widetilde{\mathbf{p}})}{\partial \mathbf{p}}\delta\mathbf{p}\,,$$

Jacobian matrix is

$$\frac{\partial \mathbf{r}(\widetilde{\mathbf{p}})}{\partial \mathbf{p}} = \frac{\partial \mathbf{r}}{\partial \mathbf{p}} = \begin{bmatrix} \frac{\partial \mathbf{r}_1}{\partial \mathbf{p}_1} & & \frac{\partial \mathbf{r}_1}{\partial \mathbf{p}_M} \\ & \ddots & \\ \frac{\partial \mathbf{r}_N}{\partial \mathbf{p}_1} & & \frac{\partial \mathbf{r}_N}{\partial \mathbf{p}_M} \end{bmatrix}$$

Optimal update

$$\delta \mathbf{p} = \underset{\delta \mathbf{p}}{\operatorname{argmin}} \left\| \mathbf{r}(\widetilde{\mathbf{p}} + \delta \mathbf{p}) \right\|^2$$

Least-squares solution using the Taylor linearization

$$\delta \mathbf{p} = -\left(\frac{\partial \mathbf{r}^T}{\partial \mathbf{p}} \frac{\partial \mathbf{r}}{\partial \mathbf{p}} \right)^{-1} \frac{\partial \mathbf{r}^T}{\partial \mathbf{p}} \mathbf{r}(\widetilde{\mathbf{p}}) = -\mathbf{R}\mathbf{r}(\widetilde{\mathbf{p}}) \, .$$

Jacobian is assumed constant, R is constant

$$\frac{\partial \mathbf{r}(\widetilde{\mathbf{p}})}{\partial \mathbf{p}} \approx \frac{\partial \mathbf{r}(\mathbf{p}^*)}{\partial \mathbf{p}} \, .$$

Jacobian is estimated numerically from random perturbations on the training set.

Image multiresolution for speed-up.

Algorithm 10.8: Active Appearance Model matching

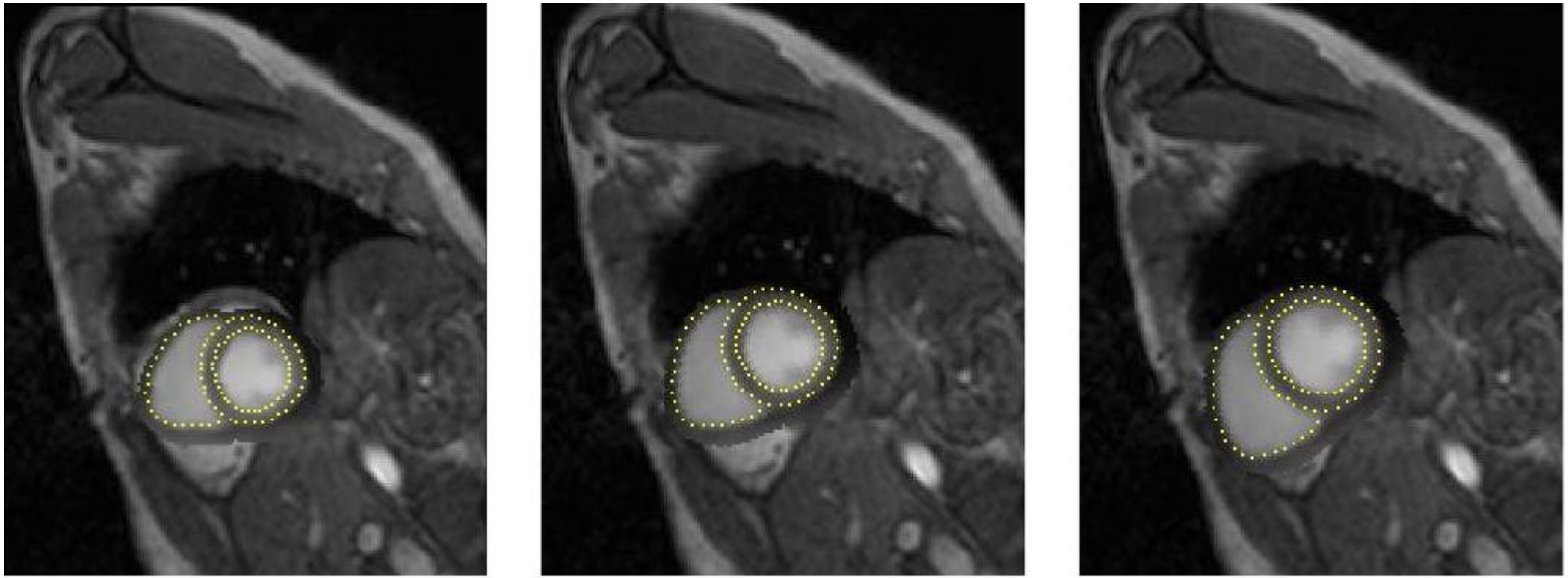
1. Place an appearance model roughly on the object of interest using the parameters $\mathbf{c}$, $\mathbf{t}$, and $\mathbf{u}$ and compute the difference image $\mathbf{g}_s - \mathbf{g}_m$.
2. Compute the RMS of the difference image, $E(\mathbf{r}) = \|\mathbf{r}\|^2$.
3. Compute the model corrections $\delta\mathbf{p}$ as derived above from the residual (equations 10.15).
4. Set $k = 1$.
5. Compute new model parameters as $\mathbf{c} := \mathbf{c} - k\delta\mathbf{c}$, $\mathbf{t} := \mathbf{t} - k\delta\mathbf{t}$, and $\mathbf{u} := \mathbf{u} - k\delta\mathbf{u}$.
6. Based on these new parameters, recompute $\mathbf{g}_s - \mathbf{g}_m$ and recalculate the RMS.
7. If the RMS is less than E, accept these parameters and go to step 3.
8. Else set k to 1.5, 0.5, 0.25, etc. and go to step 5. Repeat steps 5–8 until the error cannot be reduced any further.

appearance

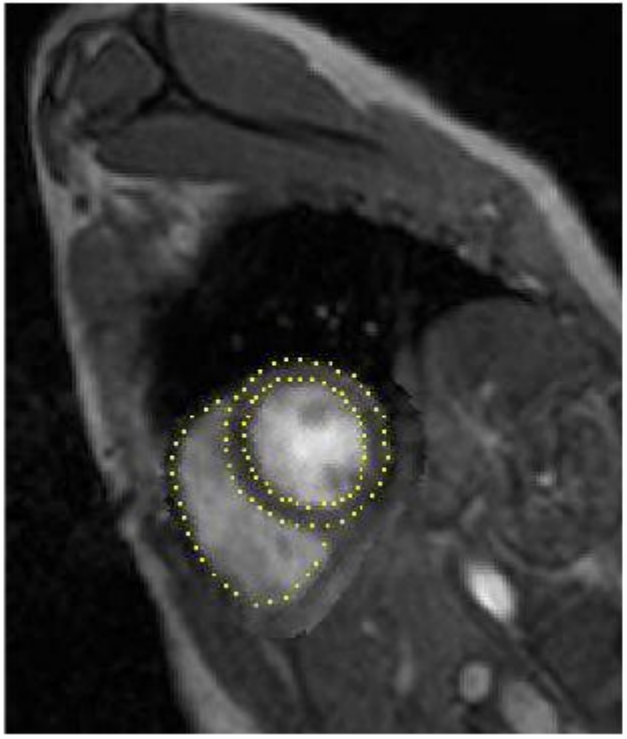
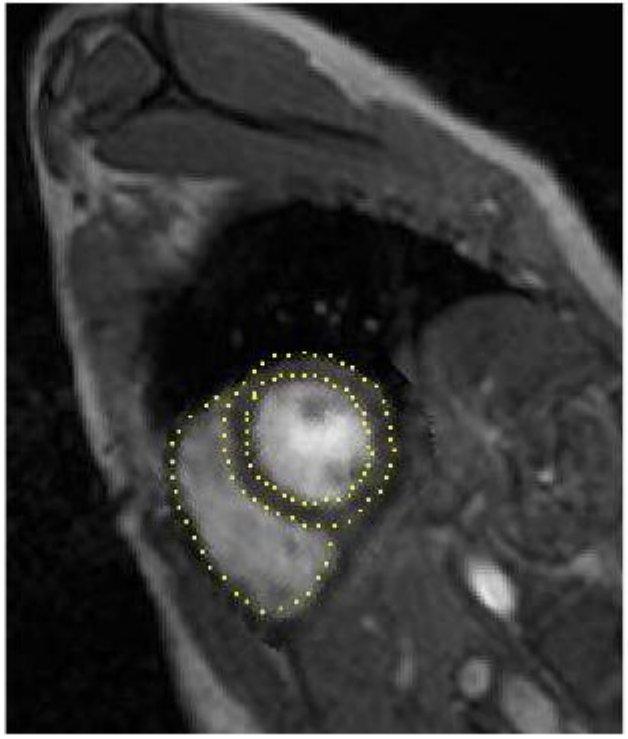
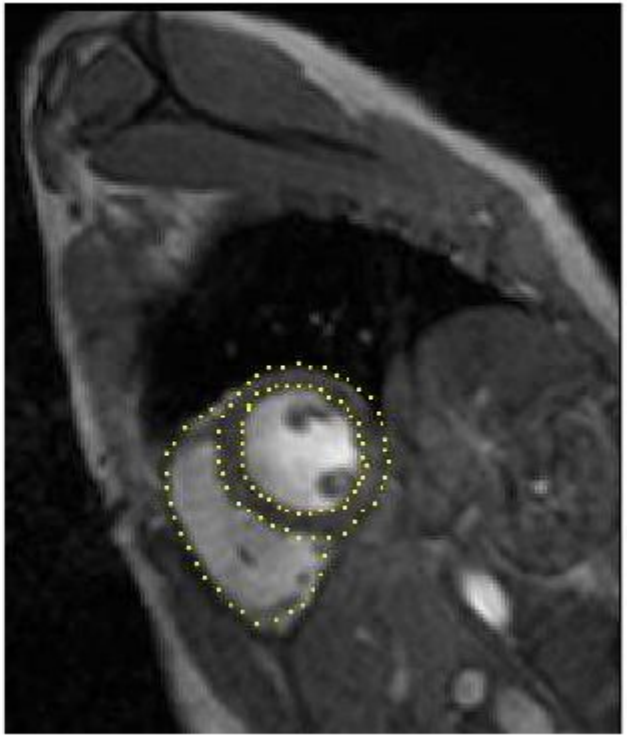
shape

global appearance

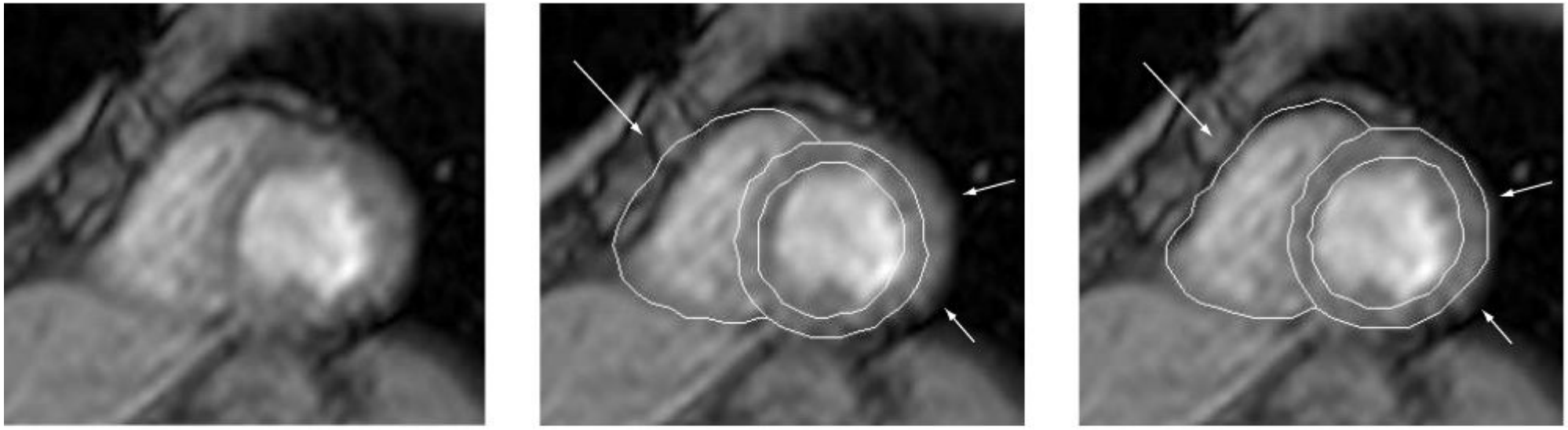

Initial
3 its
8 its
11 its
Converged
Original



Cardiac segmentation using AAM,

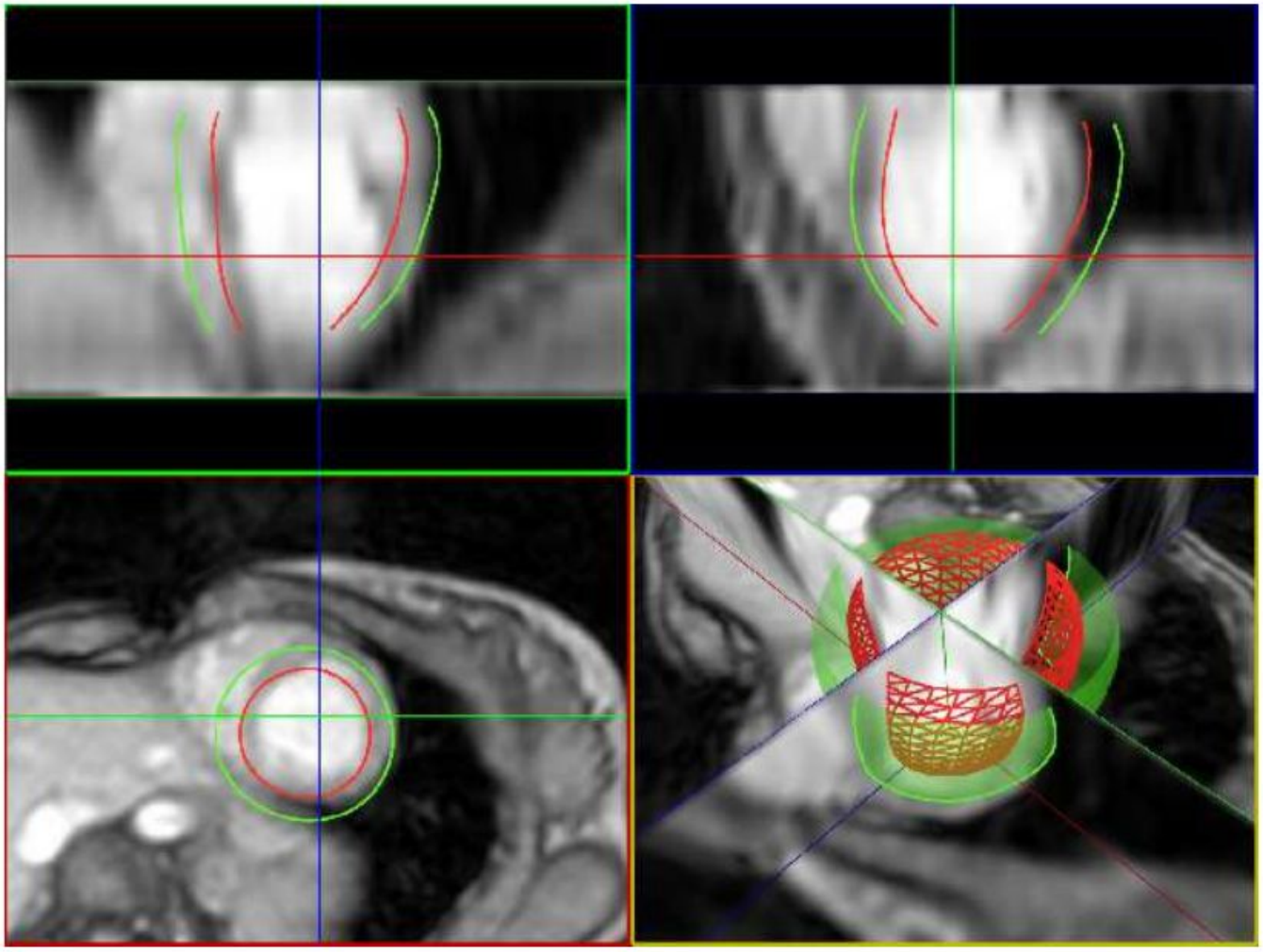
Final position



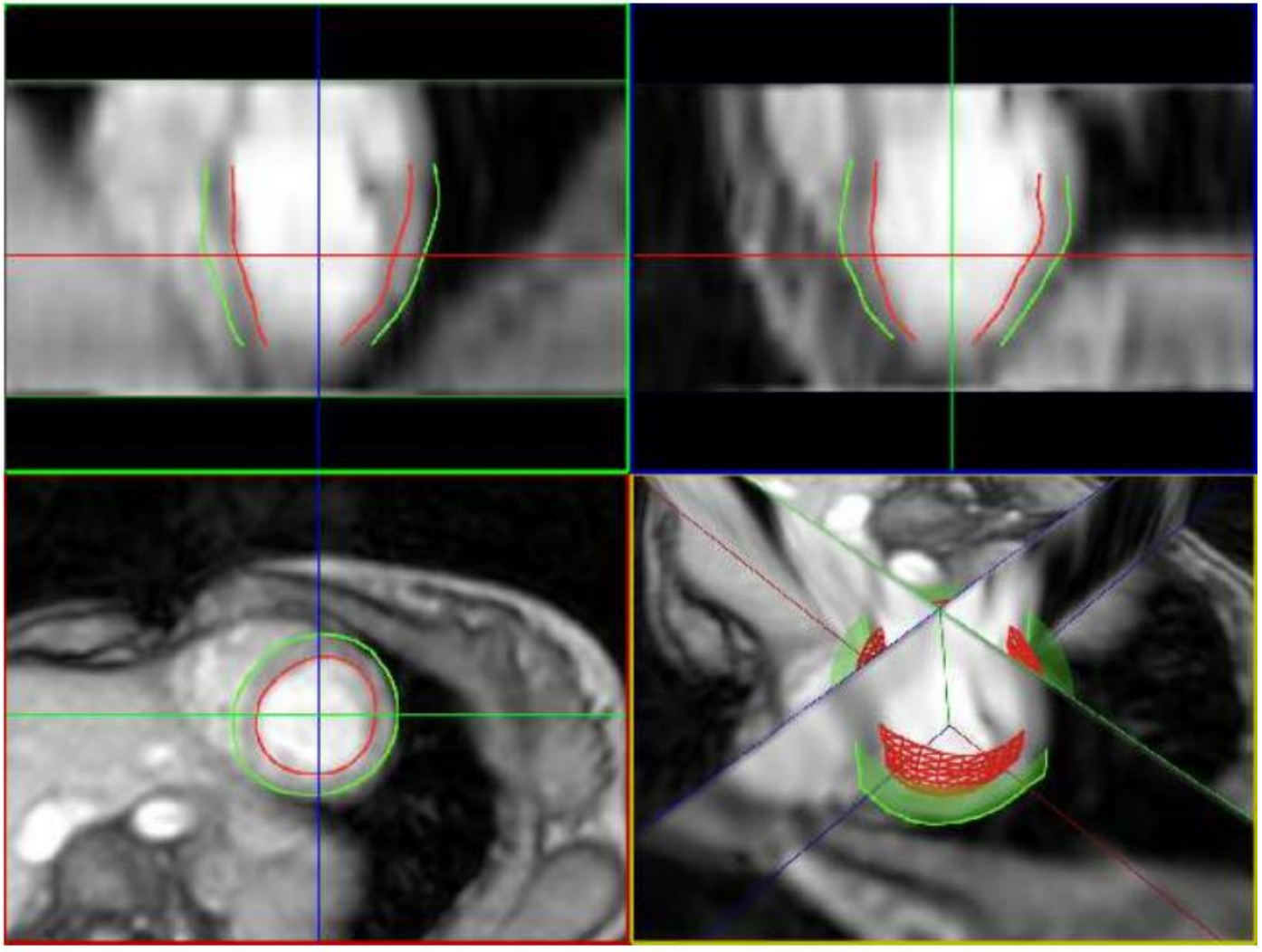
(a) (b) (c)

Figure 10.17: Comparison of 2D conventional AAM and multistage hybrid ASM/AAM segmentation of the left and right ventricles in a cardiac short-axis magnetic resonance image. (a) Original image. (b) Conventional AAM segmentation demonstrating a good gray level appearance fit but locally poor border positioning accuracy (arrows). (c) Hybrid ASM/AAM approach result shows substantial improvement in border detection positioning (arrows). *©2001 IEEE [Mitchell et al., 2001].*

hybrid = refine shape and appearance independently at the end



3D AAM initial position



final position

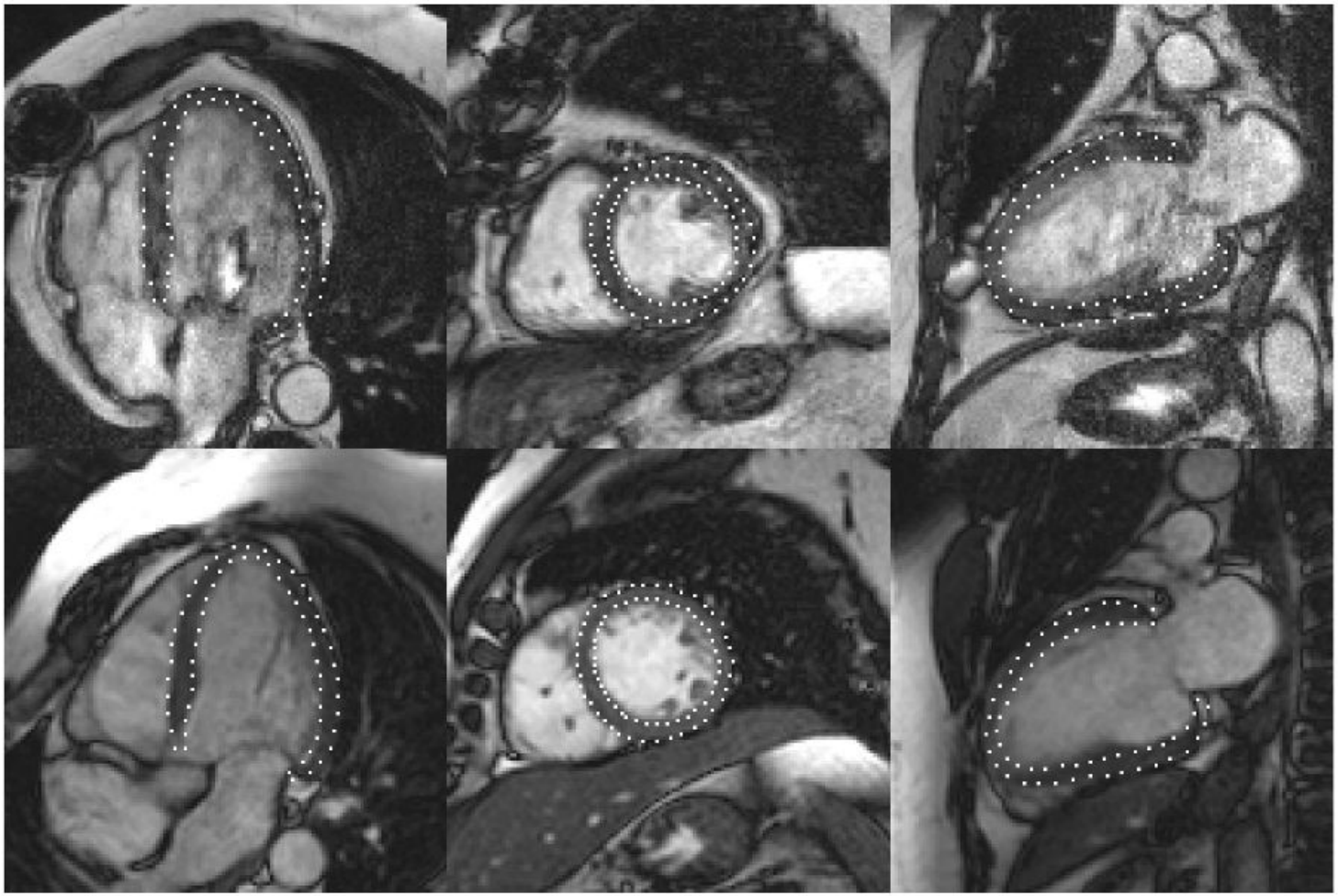
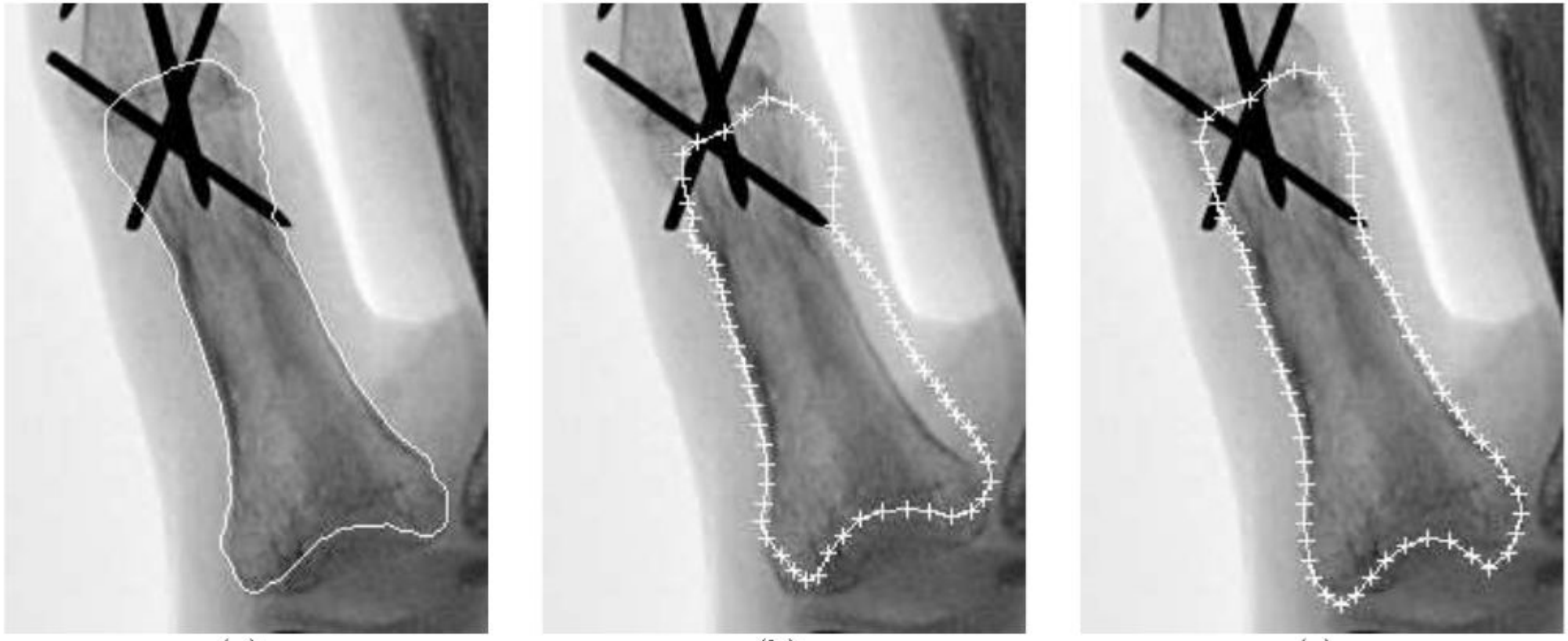


Figure 10.19: Multi-view AAM detected contours (white dotted lines) for two patients (top and bottom row) in a 4-chamber (left), short-axis (middle) and 2-chamber view (right). *©Springer Verlag [Lelieveldt et al., 2006], with kind permission of Springer Science and Business Media*



(a) (b) (c)

Figure 10.20: Robust AAM segmentation of proximal phalanx X-ray image with implants. (a) Manually determined bone contours. (b) Result of AAM segmentation—landmarks marked by '+'. (c) Robust AAM approach copes with the gray level disturbance caused by implants and provides acceptable segmentation. *©Springer Verlag [Beichel et al., 2005b], with kind permission of*

ASM, AAM - Conclusions

- + Powerful stochastical modeling of shape and appearance
- + Regularization for segmentation
- + Data driven descriptors
- + Fast fitting algorithms exist
- - Needs data
- - Needs annotations
- - Fitting may fail
- - Assumes linearity and normality (extensions exist)