MESH MODELS FOR THREE-DIMENSIONAL ULTRASOUND IMAGING

M.H-M. Syn and R.W. Prager

CUED/F-INFENG/TR 210

December 1994

Cambridge University Engineering Department
Trumpington Street
Cambridge CB2 1PZ
England

E-mail: `mhs@eng.cam.ac.uk`, `rwp@eng.cam.ac.uk`

Abstract

An iteratively refined shape recovery procedure for use in 3D freehand ultrasound imaging is presented. This procedure utilises models of both 3D shape and shape variation, by using vibration modes from finite element analysis in the initial shape model, and updating the model using principal components analysis (PCA). The difficulty in applying common PCA schemes, which require prior segmentation of training examples, is thus avoided. This is very important, since manual segmentation in 3D ultrasound is impractical for a reasonably sized training set, particular it has to be performed for each new class of shapes. A scale and rotation invariant shape registration procedure is also presented. Such a procedure allows the quantitative comparison of segmented shapes from ultrasound with other modalities for the purposes of validation. It also allows us to address the issue of automated specification of homologous landmarks, by registering a densely specified example which has just been segmented with the current model. The issues of model construction and volume measurement are also discussed.

This report is a full copy of our 1st Dec 1994 submission to IPMI95 which appears in the proceedings as a 2 page poster.

Keywords: shape models, vibration modes, principal components, automated registration, model construction, volume measurement

Chapter 1

Introduction

Three-dimensional ultrasound imaging systems using‘ a freehand probe or mechanical rig have now had most engineering issues resolved. Much of the work done so far has been in calibration of position sensors, fast acquisition and fast rendering for real-time clinical use. Ultrasound images are particularly noisy, and volume rendered images of many internal organs are mostly unusable in their raw form due to noise artifacts obscuring details of interest. Although foetal images have high contrast due to amniotic fluid, foetal motion can make volume acquisition very difficult. In comparison, motion artifacts in other animate structures usually have a significant cardiac component, and gating image acquisition with the cardiac cycle can be effective.

In general however, the nature of noise processes in this modality make automated segmentation procedures, such as those commonly used with MR images, inadequate. Drop-out can occur towards the bottom of the scan, and the signal-to-noise ratio remains fairly constant throughout the image [24]. In the system we currently use[1] [11], which spools images to video tape and logs position data on the audio track, an unknown amount of image preprocessing and interpolation is performed by the video and scanner hardware. A prototype real-time system spools frame-grabber and proprioception data directly to workstation memory, but ideally image data should be captured as early as possible along the scanner output pipeline.

These problems suggest that, in order to move on to examining useful clinical applications for 3D ultrasound, the examination of detailed image processing techniques for recovering structural detail may be justifiably put aside for the moment. It is suggested that a more practical approach would recognise the fact that the ultrasonographer must have already visualised the structure of interest when performing a 3D scan, and that a synergistic combination of the ability of a trained clinician to segment and visualise the 3D structure, with a computer's ability to quantify structure can lead to more immediate dividends in clinical application. Initial feedback has been encouraging, regarding the usability of a system which requires the ultrasonographer to spend approximately 10 minutes[2] specifying correspondences between the volume image and a 3D shape model of the structure of interest.

Section 2 introduces a reduced-basis deformable model using vibration modes from a finite element approximation of the model. Section 3 develops a method of stable shape recovery using iteratively refined principal variation modes as degrees of freedom in least squares fitting from shape model to volume image. Section 4 presents a rotation and scale invariant method of shape registration, which allows mesh landmarks to be resampled and differences between shapes of arbitrary mesh density to be quantified. Section 5 presents an approximate method for mesh

[1] EPSRC project GR/H74032

[2] A straightforward 3D scan usually takes up to 10 minutes.

node (surface and volume) specification, and also develops a general procedure for establishing 3D connectivity between nodes. Methods of volume computation of mesh models are discussed briefly. Finally Section 6 discusses the advantages of a model-based approach to 3D ultrasound imaging, and the work still needed to develop a clinically usable system.

Chapter 2

Model variation

Deformable shape models need to be constrained in the ways they are allowed to deform, the standard approach being to minimise bending energy [15] so as to regularise local varations in the deforming contour. Template approaches which are essentially optic flow computations [2], also need to use some form of regularisation to propagate displacement information away from feature points.

Alternatively, deformation vectors used as degrees of freedom (dof) in shape variation can be derived analytically using Fourier bases [22, 6], wavelet bases [25], qualitatively defined polynomial modes [16], or vibration modes [17]. Prior knowledge about the way shapes vary (linearly) can be derived using principal components analysis (PCA)[1] of a training set of shape vectors [7].

2.1 Vibration modes

A linear elastic structure can be discretised into surface or volume elements, after which the Finite Element Method (FEM) can be used in modelling and solving for its dynamic behaviour

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f}(t) \tag{2.1}$$

where **M**,**C** and **K** are the structural mass, damping and stiffness matrices, **u** is the nodal *displacement* vector and $\mathbf{f}(t)$ represents (time-varying) applied nodal forces.

A *modal* or *spectral* decomposition [3] provides an alternative parameteristation of the dof [2] of the structure, into its intrinsic *vibration modes* and corresponding *natural frequencies* (see Appendix A). The low frequency modes of vibration form an excellent set of suitably "common" shape deformation vectors which can be used to fit a prior shape model to deformed shape examples.

Low frequency mode shapes tend to be whole-body, with localised variations becoming more evident with higher natural frequencies. Figure 2.1 shows some low-frequency modes of an ellipsoidal surface model.

Rigid-body vibration modes (translation and rotation) and non-rigid vibration modes can be used as **degrees of freedom** in constraining least squares fitting to new objects. Figure 2.2 shows the fit to a box, obtained using the modes from Figure 2.1.

Note that these are linear modes that have been derived under the assumption of small peturbations about a mean model. Thus only small modal amplitudes should be allowed when performing the linear least squares fit. Pentland [17] remarks that low frequency modes tend to be similar

[1] Alternatively the Karhunen-Loeve transform or expansion (of the covariance matrix or function).

[2] Only nodal translations in x, y and z are allowed.

for objects which are roughly "blob"-like , so that the moments of inertia about each axis are of the same order of magnitude (see Figures 4.1–4.4). This then allows us to justify the use of low frequency mode vectors for larger deformations, since a shape deformed using a mode vector can be deformed further using the recomputed mode vector which is very similar.

2.1.1 FEM modelling

A common criticism of many shape modelling schemes is that shapes are discretised at fairly *ad hoc* landmark positions. Bookstein [5] suggests that landmarks should have homologous ("biologically corresponding") correspondences in every sample in the class. It is probably unreasonable in the context of 3D ultrasound modelling, however, to expect that landmarks are the most statistically "efficient" in representational power, or are restricted to biologically meaningful features. A computational cost is therefore incurred by sampling more densely than needed, compared to sampling only meaningful features, although our computational bottleneck will still be in real-time volume rendering.

A finite difference solution to Equation 2.1 would only be defined at nodal points of the mesh, whereas nodal parameters in the FEM define piecewise interpolated functions defined *throughout* the structure. Thus a *continuous* elasto-mechanical formulation can be applied to models of any dimension, topology, shape and uniformity of material properties. This is certainly an advantage in modelling for 3D ultrasound imaging, compared to using thin-plate splines [5], superquadrics [25, 17] or simple geometric primitives such as tori and tubes [22]. The use of an intrinsically dynamic model in Equation 2.1 allows the extension into 4D modelling of moving structures [18].

Points which should be addressed in FEM model construction include the order of interpolation (or *shape*) functions to be used within each element, the shape of element, size of element, and whether surface or volume elements are used. For an ellipsoid example, there is little difference in the lowest 20% of derived vibration mode vectors using 3-node triangular elements, 6-node triangular elements, 9-node quadrilateral elements, or 8-node wedge (volumetric) elements.

The 3-node triangular and 4-node tetrahedral elements are therefore used for surface and volume meshing. Both have the computational advantage of a constant Jacobian matrix for transforming between element and world coordinates, are conforming because of linear interpolation functions, and can be tessellated directly using the Delaunay triangulation with compatibility checks in Section 5.2

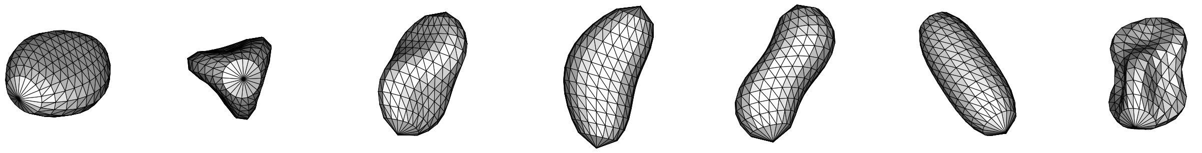
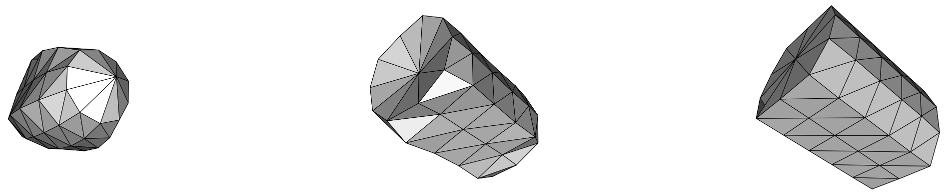


Figure 2.1: Some low frequency vibration modes of an ellipsoid surface model



(a) (b) (c)

Figure 2.2: Least squares fit to a box using (a) 21 modes; (b) 60 modes; (c) 180 modes

Chapter 3

PCA by iterative estimation

PCA of a training set of shape vectors operates by linearly projecting successive principal modes to maximise the projected variance, resulting in an ordered[1] set of linear shape deformation vectors. These vectors are optimal in the sense that they account for the maximum class variance for a given number of components or modes[2].

The application of PCA to 3D ultrasound imaging is particularly awkward since the aim is to use a set of principal modes for shape recovery, and yet we need to have recovered a set of shapes in order to derive these principal modes. Cootes and Taylor [8] attempt to merge estimates of the principal modes and intrinsic vibration modes in 2D, because of the same problem: that it is unreasonable to expect a large training set of segmented examples for each new shape model.

This is done for a small training set by generating random hybrids of vibration modes of training shapes and principal component modes of the training set, and then resampling. This involves an *ad hoc* rule to limit the amplitudes of the generated hybrids so that a “sensible” population is presented to a further PCA training step. Another weakness of their PCA scheme is that training examples are “normalised” by translation, rotation and scaling to minimise a least squares metric. The scaling normalisation step removes useful class information about “size” variation. A better approach would be to allow this scale information to be appropriately projected amongst the principal modes, so that any discarding can be justified in terms of optimality. Figure 3.1 shows that “scale variation” and “shape variation” are not generally separable

Figure 3.1: (a) original shapes; (b) scale normalised to minimise a least squares fit metric. These show that scale normalisation corrupts useful class information about shape variations between the three shapes.

A more elegant approach to deriving the principal variation modes (PVMs) of a constantly growing set of segmented examples, utilises the current PVM estimate to recover a new shape example, which is then fed back to update the current PVM estimate. Figure 4 describes an iterative scheme which bootstraps the initial estimate of PVMs using the intrinsic vibration modes of the initial model. Alternative bases which can be used instead of relatively computationally expensive

[1] Principal modes are ordered by variance.

[2] Biologically plausible (self-organising) neural networks using simple Hebbian learning rules have been shown to perform PCA [19] in much the same way described here.

vibration modes were introduced in Section 2. It is felt however that the estimates provided by the vibration modes are sufficiently good that for the models currently used, computational cost is not an issue.

1. FIRST SAMPLE $\mathbf{x}_0$

$$\begin{aligned}(\mathbf{\Phi}_0, \mathbf{\Lambda}_0) &= modes(\mathbf{x}_0)\\ \overline{\mathbf{x}}_0 &= \mathbf{x}_0\\ \mathbf{S}_0 &= \mathbf{\Phi}_0^T \mathbf{\Lambda}_0 \mathbf{\Phi}_0\end{aligned}$$

GOTO 2.

2. UPDATING WITH $\mathbf{x}_{n+1}$

$$\begin{aligned}\overline{\mathbf{x}}_{n+1} &= update(\overline{\mathbf{x}}_n, \mathbf{x}_{n+1})\\ \mathbf{S}_{n+1} &= update(\mathbf{S}_n, \mathbf{x}_{n+1})\\ (\mathbf{\Phi}_{n+1}, \mathbf{\Lambda}_{n+1}) &= eig(\mathbf{S}_{n+1})\end{aligned}$$

$$n = n + 1$$

GOTO 3.

3. LINEAR LEAST SQUARES FIT

$$\mathbf{x}_{n+1} = min(\|(\mathbf{x}_n + \alpha\mathbf{\Phi}_n) - \mathbf{x}_{n-1}\|)$$

GOTO 4.

4. FREE FORM LOCAL FITTING WITHIN EACH SCAN PLANE TO GIVE DECOUPLED $\mathbf{x}_{n+1}$

GOTO 2.

Figure 3.2: Schematic flowchart for iterative estimation of PVMs. $(\mathbf{\Phi}, \mathbf{\Lambda})$ are defined in Appendix A, $(\mathbf{x}, \mathbf{S})$ are defined in Appendix B and C, and α is defined in Equation 6.

This initial bias of the PVMs is weighted as being worth (say) 5 new sample vectors. The PVMs are updated as new samples are recovered, and these converge towards the principal components of the set of all segmented shapes in the current database. Another advantage of this formulation[3] is that our model of *surface shape* and *shape variation* is being constantly updated, compared with the static model of Cootes *et al.*

Each new shape is recovered using a linear least squares fit of correspondences defined by the clinician between volume image and shape model. This provides a very good shape estimate and model registration in the noisy volume image, and *further refinement can be performed in each scan plane* by a free-form deformation of the intersected contour to fit image evidence more closely. See Figure 3.3.

3.0.2 2D segmentation

This allows us to draw from the extensive 2D segmentation literature, is far less computationally expensive than a full 3D segmentation, and is the appropriate way use images which are acquired

[3] Appendix B and C describe the update rules for the mean vector and covariance matrix.

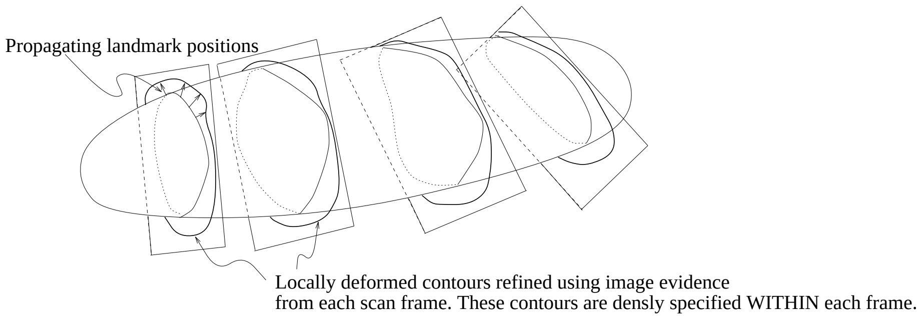


Figure 3.3: Shape recovery procedure provides intersected contours which act as prior templates in 2D segmentation

along a set of arbitrarily oriented scan planes and do not constitute a true 3D volume image. This final point means that voxel sampling concerns [4] are irrelevant in utilising image information, since the volume rendered image is then used only for visualisation purposes.

A 2D segmentation scheme is described in [23] which deforms an initial contour using simulated annealing to optimise over intensity edge evidence. But any 2D segmentation scheme which utilises the very good prior provided by the intersected contour can be used, and the weak membrane model [9] is currently being investigated for this purpose, together with deterministic solution methods[5]. The appropriate fusion of a *structural* prior with an *intensity* field is an important concern. Another important issue in allowing a free form refinement of the model, is that this decouples the final recovered shape vector from the PVMs with which it was initially recovered. Thus innovation is fed back into the mean and covariance estimates, and PVMs can be updated appropriately.

3.1 User interface design

Graphical user interface (GUI) design for clinical use will be depend very much on usability feedback from clinicians. Our proposed GUI has one window with the rendered volume image just acquired during the scan. A second window shows a scan frame, the outline of which is shown superposed on the volume rendered image. With appropriate controls to scroll through image sequences (the usual method of 3D visualisation in 2D imaging), the clinician should be able to position landmarks by mouse in the volume image with some accuracy. The corresponding point on the shape model can then be picked out on the surface of a shape model in another window, using an existing GUI.

Stereo viewers are not useful for ultrasound volume images, because of noise and structural artifacts obscuring internal detail[6]. The suggested solution combines the use of a volume image incorporating the scan frame outline (allowing visualisation of 3D position) and scan frame images (allowing internal visualisation).

[4] When converting pixels from arbitrary scan planes into a voxel volume.

[5] Blake [4] questions the flexibility of Markov Random Field formulations in incorporating prior models, especially when coupled with stochastic solution methods.

[6] Translucency rendering does not help either.

Chapter 4

Shape registration

When comparing two models, the aim is to identify unique properties or invariants of the models' shapes so that they can be matched and registered; the ability to perform such a comparison allows for quantification in clinical applications. The free vibration modes of a shape model represent one such ordered set of unique (eigen!) orthogonal description vectors.

The use of these modes for object recognition has been presented by Sclaroff and Pentland [20], based on work by Shapiro and Brady [21]. Sclaroff and Pentland use feature points which have not had connectivity established, applying Gaussian interpolation functions to approximate the shape model's silhouette and low-frequency vibration modes, after which *modal matching* is performed.

The method is not rotation invariant however, since the (x, y) modal components at each node vary with structural orientation. They suggest the use of polar coordinates instead, but utilising the same affinity metric in (r, θ) as in (x, y) leads to better matching performance using either only r or θ. They also suggest that using only θ affords scale invariance in comparing structures, but the implicit assumption is that the structures are similar enough and "blob"-like so that the centroid of each structure is in the same place relative to the feature points. These problems would be compounded for a (r, θ_1, θ_2) parametrisation in 3D.

4.1 Modal matching

Consider two surface (or volume) meshes, each with an *arbitrary* number of nodes. The problem is then to find corresponding nodes in both meshes, given that they are discretised differently. The use of a modal description, which is an alternative and fully equivalent parametrisation of the mesh, is therefore attractive. Regardless of surface sampling, the overall *shape* of each mode is the same if the two structures are similar in shape, regardless of scale, translation or rotation. A shape comparison problem in Cartesian space is therefore exchanged for a series of shape comparison problems in modal space.

4.2 Establishing correspondences

The rigid-body modes ($\phi_1 \to \phi_6$), and most of the high frequency modes are discarded, leaving 20–40[1] of the low frequency modes ($\phi_7 \to \phi_{modes}$) for use in modal matching.

Each node has associated with it a triplet which defines its deformation in 3D Cartesian space for each mode ϕ_i. Table 4.1 above shows the modes for both structures Q and R, one with n nodes,

[1] See Appendix A for computational aspects of mode derivation.

	ϕ_7	ϕ_8	ϕ_9	$\ldots$
node 1				
node 2				
node 3				
$\vdots$				
node n				

	ϕ_7	ϕ_8	ϕ_9	$\ldots$
node 1				
node 2				
node 3				
$\vdots$				
node m				

Table 4.1: Mode vectors for (a) structure Q; (b) structure R.

and one with m nodes. To test each pair of nodes for correspondence, one each on Q and R, the strength of similarity is assigned on the basis of a distance metric, $d()$, between the deformation vectors at both nodes, summed over mode vectors $\phi_7 \rightarrow \phi_{modes}$.

$$d() = \sum_{i=7}^{modes} \|\phi_Q - \phi_R\| \tag{4.1}$$

A similarity metric is required which gives higher values with better correspondences, for use in weighting the linear least squares registration procedure in Section 4.7. For computational purposes, a threshold similarity value, d_{thresh}, can be defined to filter out the weakest correspondences. Given d_{min}, the value of $d()$ for the best matching node pair, the similarity metric can then be defined as

$$sim() = \frac{d_{thresh} - d()}{d_{thresh} - d_{min}} \tag{4.2}$$

which varies linearly with the distance metric $d()$.

4.3 Mode ordering

Figures 4.1–4.4 shows the 7th to 16th modes of ellipsoids and ellipsoidal variants. Original shapes are on the extreme left. Note that orientation is not consistent between views in each figure.

The mode shapes are similar between the variants although the deformation field is sampled only at mesh nodes. It can therefore be seen that the modal decomposition gives mode vectors ranked in increasing order of eigenvalue, which can be directly compared between shapes. A small subset of low-frequency modes is retained, but the remainder can be discarded regardless of the number of nodes (and therefore number of modes) each variant has.

The structural matrices generated by 3D mesh models are relatively ill-conditioned compared to 2D, and the power method (see Appendix A) fails to derive eigenvalues with sufficient accuracy, so that mode ordering is incorrect for Figures 4.1–4.4 (see Table 4.2), as well as being computationally expensive (726×726 structural matrices for Figure 4.1).

4.4 Scale invariance

In order to describe the modes in an entirely object-centred fashion, each mode shape is considered relative to the *fundamental mode*. For two objects of the same shape, made of the same material, but of arbitrary relative scale, the ratio of each harmonic frequency to the fundamental frequency is the same. Figure 4.5 illustrates this.

		Modes													
Figures 4.1–4.4	(subspace iteration)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Figure 4.1	(power method)	7	8	10	11	12	13	17	18	9	15	16	14		
Figure 4.2	(power method)	7	8	10	11	13	14	19	20	9	15	16		12	
Figure 4.3	(power method)	7	8	11	12	13	17	18	19	9	15	16	14		
Figure 4.4	(power method)	12	13	15	16	10	11	20	21	14	18	19	17		

Table 4.2: Mode ordering generated using power method

For example, an arpeggio formed by a set of harmonics on a piece of string, will be transposed in key if played on a piece of string of different length but with the same physical properties. The relative amplitudes of each mode vector, $\|\phi_i\|$, to the fundamental mode vector, $\|\phi_7\|$ will also be the same. This is because the eigenvectors derived in Equation A.1 are $\mathbf{M}$-orthonormalised.

Thus by normalising the amplitude of each harmonic mode to the fundamental mode, mode vectors (both shape and amplitude) can be compared from each structure directly. Table 4.3 and Figures 4.6 and 4.7 show this to be the case for the structures in Figures 4.1–4.4.

Modes	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Figure 4.1	1.00	1.00	2.42	2.42	2.82	2.82	3.22	3.22	3.99	3.99	4.00	4.43	4.81	4.81
Figure 4.2	1.00	1.00	2.62	2.62	2.66	2.66	2.83	2.83	3.00	3.27	3.27	3.69	5.09	5.09
Figure 4.3	1.00	1.00	2.63	2.71	2.77	2.77	2.84	2.84	3.05	3.48	3.50	3.85	5.25	5.25
Figure 4.4	1.00	1.00	2.87	2.87	2.97	2.97	3.46	4.04	4.04	4.07	4.08	4.08	6.77	6.78

Table 4.3: Mode eigenvalues normalised by fundamental (subspace iteration)

4.5 Relative weighting of modes

The lowest frequency mode vectors change very little in shape and amplitude, for structures which are similar in shape. This sensitivity increases with frequency, providing us with an ordered scale of mode vector stability. This is reinforced by the fact that low frequency eigenpairs are derived with greater accuracy (see Appendix A) in eigenproblem solution methods.

Higher frequency mode shapes exhibit more rapid[2] and localised surface variation (see Figure 2.1), providing prominent features which can be used more effectively in determining correspondences. On the other hand, the more rapid rate of change in shape over the mesh results in loss of accuracy in the ability of the mesh discretisation to describe these changes. For higher frequency mode shapes used in modal matching, there are more lower frequency modes onto which it can be projected, thus increasing the confidence with which nodes can be compared using the similarity metric in Equation 4.2.

The overall effect, confirmed by experiments with different weighting schemes, is that the similarity matrices generated by each mode are generally of the same importance.

[2] In the sense of change per unit surface area.

4.6 Rotation invariance

Ideally a coordinate system defined relative to the body of the structure would be used, but this is not generally possible since it first requires knowledge of 3D shape and orientation[3]. Instead, in view of their stability and relative accuracy, the lower-frequency mode shape vectors are used as a set of reference vectors relative to which higher frequency mode vectors can be described. In other words, the translational and rotational degrees of freedom in rigid-body registration are subtracted *by using the lower frequency mode vectors as a body-centred coordinate system.*

Each mode vector ϕ_i is projected on to the lower-frequency reference vectors $(\phi_7, ..., \phi_{i-1})$, giving a new set of $(i-7)$ coordinate axes at each node. These axes vary with node location, but remain orientated in a consistent way relative to the body of the structure. The coordinate values are given by the projected amplitudes relative to the reference mode amplitudes. Note again the points made in Section 4.4. These coordinate values, ϕ', can now be directly compared between structures because they are already scale invariant. The revised distance metric is then

$$\begin{aligned} d() &= \sum_{i=10}^{modes} (\phi'_Q - \phi'_R) \\ \phi'_i &= \sum_{j=7}^{i-1} \phi_i.\phi_j \end{aligned} \tag{4.3}$$

where the summation starts at the 10th mode, because at least three modes, the 7th (fundamental), 8th and 9th, are required as scale-compensated and body-centred reference vectors in 3D.

The smaller the distance between modal displacement vectors of each harmonic frequency, the stronger the correspondence between two nodes, giving a new similarity metric $sim()$.

4.7 Weighted least squares fit

An overconstrained least squares fit procedure can then be performed. A fit metric $fit(Q, R)$ is minimised, which comprises the sum of squared distances between every corresponding node pair, weighted by the strength of correspondence, $sim()$.

$$fit(Q, R) = \sum_{corr(Q,R)} sim().\|(\mathbf{x}_Q - \mathbf{x}_R)\| \tag{4.4}$$

where $corr(Q, R)$ denotes the set of corresponded node pairs between Q and R which have not been filtered by d_{thresh}. The degrees of freedom allowed in minimising $fit(Q, R)$ comprise the PVMs of the model if they are available, or the vibration modes of either Q or R, as well as rigid body translation and rotation modes for pose recovery. The problem can now be reformulated so that it fits a standard least squares framework. Consider all nodes n_1 on Q which have at least one correspondence $\{n_2\}$ established to R, in $corr(Q, R)$. This is effectively one overall correspondence to a point in space near R, whose coordinates are the weighted mean of points in $\{n_2\}$. They are weighted by the similarity metric $sim()$. $\{n_2\}$ is replaced with the resultant correspondence $n_{2,res}$, allowing use of the least squares formulation

$$\mathbf{\Delta} - \alpha.\mathbf{\Phi} \tag{4.5}$$

[3] It is possible to use estimated axes of inertia for ellipsoidal shapes [17]

where $\mathbf{\Delta}$ contains the difference vectors, $||n_{2,res} - n_1||$, which are to be minimised in the least squares sense, α contains the modal amplitudes to be determined, and $\mathbf{\Phi}$ comprises the modal degrees of freedom. Equation 4.5 can be minimised using the Householder QR procedure [10]. The recovered modal amplitudes α_i in α then describe the modal amplitudes needed to deform structure Q to fit structure R.

Translation modes are linear, and can be solved for in one step. Similarly the other non-rigid body modes which were derived under a linear peturbation assumption. If the amplitude of rotation modes exceeds the bounds of the linearisation approximation $\sin(x) = x$, of about 10 degrees, then the least squares recovery is limited to this linearisation bound, and registration is recovered in two or more steps consecutively.

The assumption implicit in this approach is that the structures being registered are asymmetric enough so that the net correspondence is determined mainly by the correct correspondence. This would not be the case for the artifically generated ellipsoidal meshes in Figures 4.1 and 4.2, for instance.

An alternative least-squares approach is to use "gravitational forces" defined by the inverse square distance and weighted by each correspondence in $corr(Q, R)$. The inclusion of the actual nodal separations at each time step in square distances allows the registration to converge on the optimum fit local to the starting configuration. For the example cited above, this method converges to only one registration solution, at the expense of having to use more than one time steps to avoid overshoot. As before, the dof in registration are limited to the PVMs of the model.

4.8 Results

Simple experiments with the synthetic ellipsoids in Figures 4.2–4.4 were performed. Similarity matrices were constructed for the 10th to 20th modes, incorporating $sim()$ for each pair of nodes in Q and R. The net similarity matrix was formed by summing each modal similarity matrix (see Section 4.5), and incorporated the similarity metric and distance metric from Equation 4.3. The node ordering and discretisation of these ellipsoids are the same, so the net similarity matrix was predominantly diagonal. Registration was correctly recovered using the gravitational fit procedure.

The stability of the procedure is currently being investigated for more noisily specified meshes (see Section 5) which result in noisy eigenpairs. Other metrics of similarity and overall shape fit are also being examined.

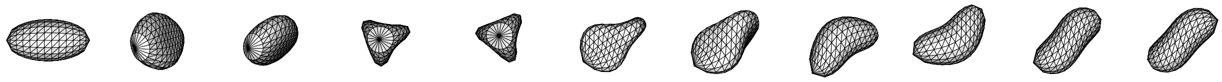


Figure 4.1: Ellipsoid surface with 242 nodes and 480 elements

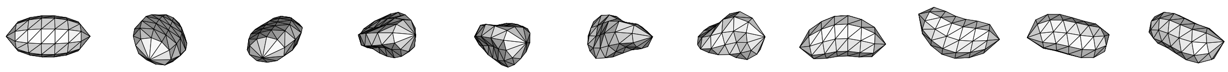


Figure 4.2: Ellipsoid surface with 62 nodes and 120 elements

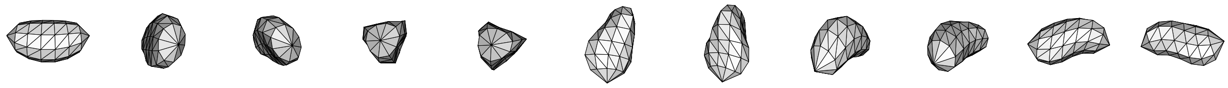


Figure 4.3: Ellipsoid surface (bent) with 62 nodes and 120 elements

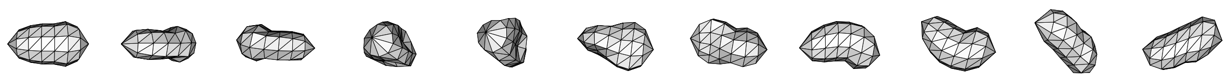


Figure 4.4: Ellipsoid surface (stubbed) with 62 nodes and 120 elements

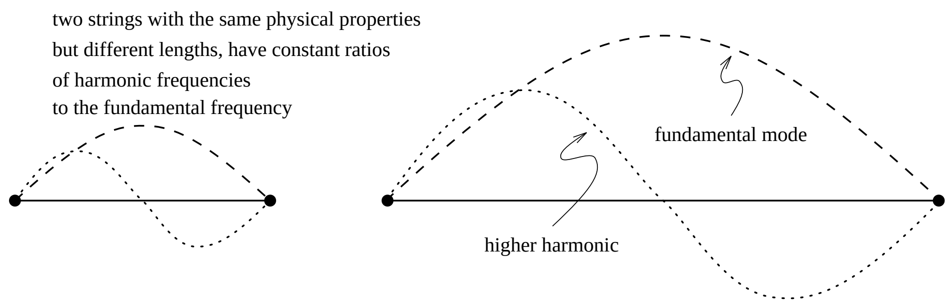


Figure 4.5: Fundamental and harmonic modes of two strings

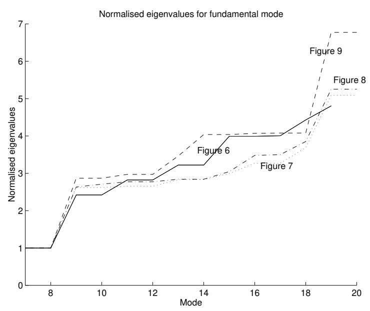


Figure 4.6: Mode eigenvalues normalised by fundamental (subspace iteration)

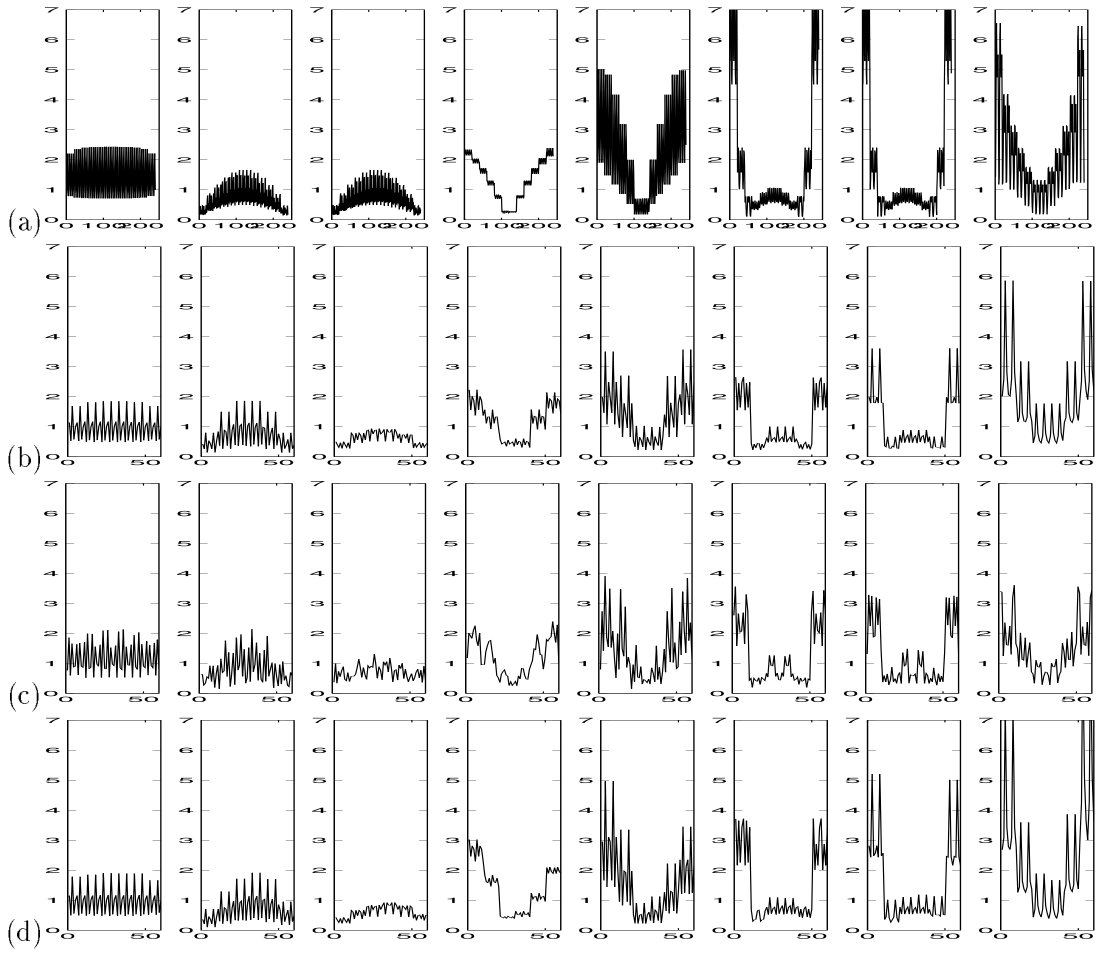


Figure 4.7: Nodal amplitudes of 8th to 15th mode vectors, normalised by fundamental (subspace iteration). (a) Figure 6; (b) Figure 7; (c) Figure 8; (d) Figure 9.

Chapter 5

Building shape models

5.1 Surface and volume discretisation

An ideal mesh surface and volume sampler produces the most accurate shape sampling (depending on fit metric) for a given number of sample points. Since FEM is being applied to the meshed structure, the ideal tessellation results in elements which have the same accuracy in modelling material properties within each elemental domain. For instance, surface patches of high curvature will produce higher stress concentrations, and hence require denser sampling.

An approximate but efficient method of sampling the isosurface of a binary voxel volume uses the *octree decomposition* of the volume to describe the surface, which can then be seeded with surface sampling points. Mesh simplification algorithms exist which adapt to surface curvature e.g. the Wrapper [12]. Mesh spacing is usually refined using iterative Laplacian smoothing.

5.2 Mesh generation and compatibility

The Delaunay triangulation is commonly used to establish connectivity between mesh nodes because it satisfies the empty circumsphere criterion and thus expresses adjacency relationships between mesh nodes. It is generally unique, or can be made so, and the triangular elements generated maximise the minimum value of the internal angles [1, 26].

Meshing on the basis of spatial adjacency without reference to the model being meshed can lead to compatibility problems between triangulated mesh and model. An algorithm is therefore required which adapts the mesh so that it is compatible with the model, and also to label mesh components in the process. Mesh elements are defined to be nodes, edges, faces and regions[1], and they can either be `IN`, `ON` or `OUT` of the model. Figure 5.1 shows an example of mesh incompatibility, and also of possible labelling problems in narrow regions.

5.3 Resolving compatibility and labelling problems

1. Label mesh nodes.
    - Surface nodes are classified as `ON`.
    - Internal nodes are classified as `ON`.

[1] In this case faces are triangles and regions are tetrahedra

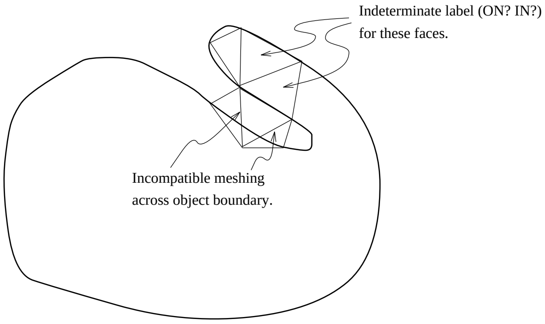


Figure 5.1: Mesh incompatibility; labelling problem.

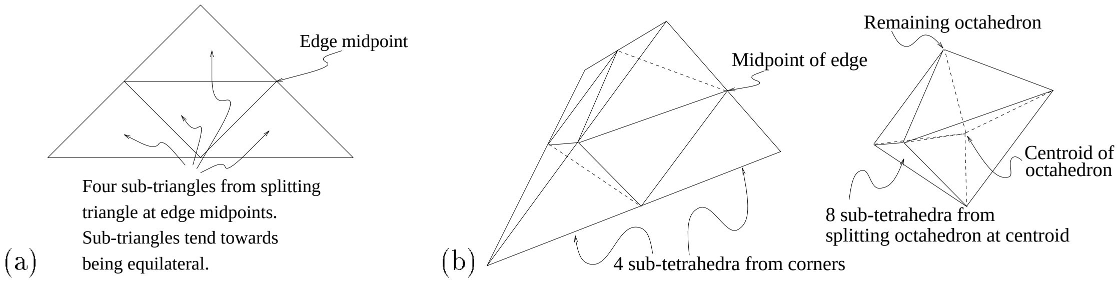


Figure 5.2: Subdivision of (a) face; (b) region.

2. If any edge intersects the structure's boundaries, subdivide the edge by generating a new `ON` node.

3. Label edges, faces and regions.
   - Label as `IN` if at least one node is `IN`.
   - Label as `OUT` if at least one node is `OUT`.
   - Otherwise label `ON`

4. Delete `OUT` regions, and update mesh lists.
   - Derive new face list from remaining regions.
   - Derive new edge list from remaining faces.
   - Derive new node list from remaining edges.

5. Subdivide `ON` regions (see Figure 5.2), classify new nodes created and go to 2;
   Otherwise exit.

5.4 Model volume

Gauss' theorem states that

$$\int\int\int_V \nabla \cdot \Phi . dV = \int\int_A \Phi \cdot \mathbf{n} . dA \tag{5.1}$$

(a) (b)

Figure 5.3: (a) Octree sampled sphere; (b) manually sampled thyroid from 3D ultrasound frames. Both meshed using Delaunay triangulation with compatibility checks.

where ∇ is the gradient operator $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$, and $\mathbf{n}$ is the unit outward normal. By setting $\nabla \cdot \Phi = 1$ (using $\Phi = [x/3, y/3, z/3]^T$), the volume enclosed in V is computed by evaluating the surface integral on the right hand side

$$\int\int\int_V dV = \int\int_A f(\mathbf{x}).dA \tag{5.2}$$

The model is composed of m surface elements, with each surface patch defined parametrically within the element using interpolation functions. The surface integral can then be rewritten as a sum of integrals over all surface elements

$$\int\int\int_V dV = \sum_{j=1}^{m} \int\int_{A_j} f(\mathbf{x}).dA_j = \sum_{j=1}^{m} \int\int_{A_j} f(r,s)|\mathbf{J}|.dr.ds \tag{5.3}$$

The integrals can be evaluated using Gaussian quadrature rules when using Bezier triangles, giving an exact polynomial integration. For a piecewise polyhedral model however, the same answer can be arrived at by summing elemental volumes using the scalar triple product.

Chapter 6

Conclusion

We have presented a snapshot of work in progress which aims towards a procedure for extracting shape information from a sequence of 3D ultrasound scans. The heart of the procedure is the iterative shape recovery scheme and a 2D segmentation scheme. The procedure can be further improved with the use of a shape registration algorithm.

An elasto-mechanical interpretation of model shape and shape variation has been used throughout, although a more generalised Bayesian formulation is possible[1]. The particular advantage of using a linear elastic model is the linearly approximate set of "growth" vectors which are used in stable shape recovery; the aim is not to build an accurate growth model (e.g. reaction-diffusion processes) which properly describes the differences between biological structures. There is in fact no profound reliance on mechanical models in any of our procedures, except in the weak-membrane approach to 2D segmentation, and that can be replaced by something more sophisticated once better growth models are in common currency in the medical imaging field.

Least squares metrics have been used in our shape registration procedure simply because $\|Ax - b\|_2^2$ is a differentiable function of x and the minimisers are easily found. The 2-norm is also preserved under orthogonal transformation.

The shape recovery algorithm uses iteratively refined estimates of principal variation modes to perform a stable, reduced-basis recovery from a noisy volume image, and bootstraps its estimates using the intrinsic vibration modes of the shape model.

The registration algorithm for surface or volume meshed models uses a body-centred modal decomposition of the model's structure to establish correspondence and hence registration. The modal transform allows us to use meshes with arbitrary numbers of nodes, by matching correspondences in modal space. There are no problem specific parameters to be established, the only free parameter being d_{thresh}, which is used to filter out weaker correspondences. The procedure can be made scale and rotation invariant using modal projection.

Our presentation addresses an important weakness in PCA training schemes. The placement of landmarks and the implicit establishment of correspondences between landmarked examples, is subject to errors in manual placement[2]. This is also the case for the initial 3D shape recovered by the clinician, after which free form 2D segmentation improves the local fit in each scan frame. This results in a new and decoupled mesh shape which is densely specified within each scan frame (see Figure 3.3). The subsequent registration of the densely sampled mesh with the current shape model then recovers the correspondences automatically, and allows us to resample the mesh by retaining the strongest one-to-one correspondences.

[1] The most straightforward approach is to map strain energy to a Gibbs function, and hence to probability space.

[2] Hill and Taylor [13] present a two stage process in improving landmark placement.

Eigenvalue and eigenvector results presented in our registration experiments were derived using the subspace iteration implementation in a commercial package, which does not allow specification of accuracy. The registration algorithm can be therefore be made more robust and accurate if we can compute error bounds on the derived eigenpairs (See Appendix [**?**]).

Model-building and volume measurement have also been briefly discussed, allowing a sufficient number of tools for the design of both clinically useful applications in 3D ultrasound, as well as experiments to validate the derived results with MRI derived ground truth. Current work is concentrated on designing a user-interface for clinical manipulation of ultrasound sequences, developing a robust shape registration algorithm, improving the 2D segmentation algorithm, and developing an efficient generalised eigenproblem solver for low frequency eigenpairs using Lanczos methods.

Acknowledgements

Many thanks to T.J. Cham and Jon Lawn of the Vision Group. Thanks also to Gene Golub for help with Lanczos methods. Mike Syn is supported by an EPSRC Research Studentship. Both authors gratefully acknowledge the support of EPSRC project GR/H74032.

Appendix A

Modal computation

The $\mathbf{M}$-orthonormalised vibration mode vectors, $\mathbf{\Phi}$, and corresponding natural frequencies[1] in the diagonal matrix, $\mathbf{\Lambda}$, are computed by solving the symmetric generalised eigenproblem

$$\mathbf{K\Phi} = \mathbf{M\Lambda\Phi} \tag{A.1}$$

where $\mathbf{M} > 0$ and $\mathbf{K} \geq 0$ are as in Equation 2.1. The matrices are of order $3n$, for a structure with n nodes and 3 translational dof at each node.

For Figure 4.2, the *power method*[2] returns all eigenpairs in 62 CPU seconds on a Sparcstation 10/30, though with limited accuracy. Conditioning in this context refers to the ratio λ_1/λ_2 which dictates the rate of convergence [10]. Since these are always rigid body modes which should have zero eigenvalue (since there is no strain energy in these modes), the power method tends to be unreliable for use with 3D mesh models. It fails to converge for the thyroid mesh in Figure 5.3, and is inaccurate in deriving the eigenvalues for the simple 3D meshes in Figures 4.1–4.4)[3]. Storage requirements scale as $3 \times 9n^2$ ($\mathbf{M}$, $\mathbf{K}$ and workspace), although it is possible to use sparse-matrix indexing methods if an appropriate matrix-vector product routine exists, this being the only procedure required to implement the power method.

Subspace iteration [3] is an efficient method for deriving p of the lowest-order eigenpairs of a large eigenproblem. Using $q > p$ iteration vectors increases computational effort, but also leads to faster convergence since the (asymptotic) convergence rate for each eigenvector is λ_i/λ_{q+1}. The initial subspace can be generated using the Lanczos procedure, and the iteration vectors are usually reorthogonalised as the procedure tends to make them more and more parallel. Convergence requires that

$$\frac{|\lambda_i^{(k+1)} - \lambda_i^{(k)}|}{\lambda_i^{(k+1)}} \leq \text{tol}; \qquad i = 1, ..., p \tag{A.2}$$

where tol should be 10^{-2s} for eigenvalues to be accurate to about $2s$ digits, with smaller eigenvalues being evaluated more accurately. Eigenvectors are accurate to about s or more digits. The implementation in the Abaqus FEM package on an HP 735 for 30 eigenpairs takes 50 CPU seconds.

Lanczos methods [14] can also be used to extract low-frequency eigenpairs. They have attractive convergence properties (comparable to if not better than subspace iteration), and such a procedure

[1] $\lambda_i = \omega_i^2$ are the elements of $\mathbf{\Lambda}$, where ω_i is the natural frequency of mode i.

[2] Double-precision implementations in `eispack` and `lapack`

[3] Refer back to Section 4.3 and Table 4.2

is currently being implemented. Preliminary timings on a Sparcstation 10/30 for the ellipsoid in Figure 4.1 returns 20 eigenpairs in 57 CPU seconds.

Using either subspace iteration or Lanczos methods on matrix sizes of the order of 200–2000, it is reasonable to aim to extract 20–40 eigenpairs accurately and place error bounds such that mode ordering for shape registration can be performed with confidence.

A.1 Error measure for eigenpairs

A useful error measure [3] which is used to test for accuracy and convergence of an eigenpair $(\hat{\phi}, \hat{\lambda})$ is

$$\epsilon = \frac{||\mathbf{K}\hat{\phi} - \hat{\lambda}\mathbf{M}\hat{\phi}||}{||\mathbf{K}\hat{\phi}||} \tag{A.3}$$

$\mathbf{K}\hat{\phi}$ and $\hat{\lambda}\mathbf{M}\hat{\phi}$ represent the elastic and inertial nodal point forces respectively, when the structure is vibrating in mode $\hat{\phi}$. We are therefore evaluating the ratio of the out-of-balance nodal point forces, and the elastic nodal point forces. This quantity should be small if $(\hat{\phi}, \hat{\lambda})$ are an accurate solution. Note that it may be necessary to use the Sturm sequence property to identify which specific eigenvalue is being approximated.

A.2 Rigid body modes

Eigenvalue shifts are frequently used for calculating rigid-body eigenpairs. A shift, μ, on $\mathbf{K}$ gives

$$\hat{\mathbf{K}} = \mathbf{K} - \mu\mathbf{M} \tag{A.4}$$

and solving the eigenproblem

$$\hat{\mathbf{K}}\phi = \gamma\mathbf{M}\phi \tag{A.5}$$

gives the same eigenvectors as for Equation A.1, but with eigenvalues decreased by μ[4]. For our purposes it is more efficient to separately calculate rigid-body translations and rotations during the registration or fitting procedures.

[4] $\lambda = \gamma + \mu$

Appendix B

Updating $\overline{\mathbf{x}}_{n+1}$

Equation B.1 gives an iterative rule for updating the mean shape vector, $\overline{\mathbf{x}}_{n+1}$, which is used as the a-priori shape model.

$$
\begin{aligned}
\overline{\mathbf{x}}_n &= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \\
&= \frac{1}{n}\left(\sum_{i=1}^{n-1}\mathbf{x}_i + \mathbf{x}_n\right) \\
&= \frac{n-1}{n}\cdot\frac{1}{n-1}\cdot\sum_{i=1}^{n-1}\mathbf{x}_i + \frac{1}{n}\cdot\mathbf{x}_n \\
\overline{\mathbf{x}}_n &= \frac{1}{n}\cdot\left((n-1)\overline{\mathbf{x}}_{n-1} + \mathbf{x}_n\right) \\
\overline{\mathbf{x}}_{n+1} &= \frac{1}{n+1}\cdot\left(n\overline{\mathbf{x}}_n + \mathbf{x}_{n+1}\right) \qquad \text{(B.1)}
\end{aligned}
$$

Appendix C

Updating S_{n+1}

Equation C.1 gives the covariance matrix, $\mathbf{S}_n$,

$$
\begin{aligned}
\mathbf{S}_n &= \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \overline{\mathbf{x}}_n)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{x}_i^2 - n\overline{\mathbf{x}}_n^2\right) \\
\frac{n-1}{n}\cdot\mathbf{S}_n &= \frac{1}{n}\left(\sum_{i=1}^{n}\mathbf{x}_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}\mathbf{x}_i\right)^2\right) \qquad \text{(C.1)}
\end{aligned}
$$

and comparing to Equation C.2 which gives the covariance matrix, $\mathbf{S}_{n+1}$, for $(n+1)$ samples

$$
\begin{aligned}
\mathbf{S}_{n+1} &= \frac{1}{n}\left(\sum_{i=1}^{n+1}\mathbf{x}_i^2 - (n+1)\overline{\mathbf{x}}_{n+1}^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^{n+1}\mathbf{x}_i^2 - \frac{1}{n+1}\cdot\left(\sum_{i=1}^{n+1}\mathbf{x}_i\right)^2\right) \qquad \text{(C.2)}
\end{aligned}
$$

an iterative rule for updating the covariance matrix, $\mathbf{S}_{n+1}$ is obtained.

$$
\mathbf{S}_{n+1} - \frac{n-1}{n}\cdot\mathbf{S}_n = \frac{1}{n}\cdot\mathbf{x}_{n+1}^2 - \frac{n+1}{n}\cdot\overline{\mathbf{x}}_{n+1}^2 + \overline{\mathbf{x}}_n^2 \qquad \text{(C.3)}
$$

whence

$$
\mathbf{S}_{n+1} = \frac{n-1}{n}\cdot\mathbf{S}_n + \frac{1}{n}\cdot\mathbf{x}_{n+1}^2 - \frac{n+1}{n}\cdot\overline{\mathbf{x}}_{n+1}^2 + \overline{\mathbf{x}}_n^2 \qquad \text{(C.4)}
$$

Bibliography

[1] F. Aurenhammer. Voronoi diagrams - a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405, 1991.

[2] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Systems and experiment: Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[3] K.J. Bathe. *Finite element procedures in engineering analysis*. Prentice-Hall, 1982.

[4] A. Blake. Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstructino. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(1):2–12, January 1989.

[5] F. Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[6] Ch. Brechbuhler, G. Gerig, and O. Kubler. Surface parametrization and shape description. In R.A. Robb, editor, *Visualisation in Biomedical Computing 92*, volume 2359 of *Proc. SPIE*, Chapel Hill, North Carolina, October 1992.

[7] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. In H.H. Barrett and A.F. Gmitro, editors, *Proc. Information Processing in Medical Imaging*, volume 511 of *Lecture Notes in Computer Science*. Springer-Verlag, 1991.

[8] T.F. Cootes and C.J. Taylor. Combining point distribution models with shape models based on finite-element analysis. In *Proceedings 5th British Machine Vision Conference*, York, 1994. BMVA Press.

[9] D. Geiger and A. Yuille. A common framework for image segmentation. *International Journal of Computer Vision*, 6(3):227–243, 1991.

[10] G. Golub and C. van Loan. *Matrix computations*. The Johns Hopkins University Press.

[11] J.P.M. Gosling, R.W. Prager, and L.A. Berman. Proprioception accuracy in free-hand three-dimensional ultrasound imaging. Technical Report CUED/F-INFENG/TR169, Cambridge University Engineering Department, Trumpington Street, CB2 1PZ.

[12] A. Gueziec and D. Dean. The Wrapper: A surface optimization algorithm that preserves highly curved areas. In R.A. Robb, editor, *Visualisation in Biomedical Computing 92*, volume 2359 of *Proc. SPIE*, Chapel Hill, North Carolina, October 1992.

[13] A. Hill and C.J. Taylor. Automatic landmark generation for point distribution models. In *Proceedings 5th British Machine Vision Conference*, York, 1994. BMVA Press.

[14] M.T. Jones and M.L. Patrick. LANZ: Software for solving the large sparse symmetric generalized eigenproblem. Available on `netlib`.

[15] A. Witkin M. Kass and D. Terzopoulos. Snakes: Active contour models. In *Proc. First International Conference on Computer Vision*, pages 259–268, 1987.

[16] C. Nastar and N. Ayache. Non-rigid motion analysis in medical images: a physically based approach. In *IPMI93*, volume 687 of *Lecture Notes in Computer Science*, 1993.

[17] A. Pentland. Automatic extraction of deformable part models. *International Journal of Computer Vision*, 4:107–126, 1990.

[18] A. Pentland and B. Horowitz. Recovery of nonrigid motion and structure. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(7):730–742, 1991.

[19] T. D. Sanger. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks*, 2:459–473, 1989.

[20] S. Sclaroff and A. Pentland. Modal matching for correspondence and recognition. Technical Report 201, MIT Media Lab Perceptual Computing Section, May 1993.

[21] L. Shapiro and M. Brady. Feature-based correspondence: an eigenvector approach. *Image and Vision Computing*, 10(5):283–288, June 1992.

[22] L.H. Staib and J.S. Duncan. Deformable Fourier models for surface finding in 3D images. In R.A. Robb, editor, *Visualisation in Biomedical Computing 92*, volume 2359 of *Proc. SPIE*, Chapel Hill, North Carolina, October 1992.

[23] M. Syn, J.P.M. Gosling, R. Prager, L. Berman, and J. Crowley. Tracking the interframe deformation of structures in 3D ultrasound imaging. In R. Robb, editor, *Proc. Visualization in Biomedical Computing 94*, Rochester, Minnesota, October 1994. SPIE.

[24] T.A. Tuthill, R.H. Sperry, and K.J. Parker. Deviations from Rayleigh statistics in ultrasonic speckle. *Ultrasonic Imaging*, 10:81–89, 1988.

[25] B. Vemuri, A. Radisavljevic, and C.M. Leonard. Multi-resolution stochastic 3D shape models for image segmentation. In *Proceedings IPMI93*, volume 687 of *Lecture Notes in Computer Science*. Springer-Verlag, 1993.

[26] D.F. Watson. Natural neighbour sorting. *The Australian Computer Journal*, 17(4):189–193, 1985.