

Model Generation from Multiple Volumes using Constrained Elastic SurfaceNets

Michael E. Leventon and Sarah F. F. Gibson

Abstract

Three dimensional models of anatomical structures are currently used to aid in medical diagnosis, treatment, surgical guidance, and surgical simulation. Limitations on the resolution of medical scans can cause artifacts to appear in the models that do not exist in the patient's anatomy. The most severe artifacts occur due to the low sampling rate between image slices of a scan. This paper describes a method of combining two orthogonal scans to generate a model with higher resolution than models created from either of the scans alone. The two scans are first registered to each other and then a net of linked surface nodes is initialized for each of the scans. The nodes from the two nets are then merged and relaxed, subject to constraints set by the resolution of each scan. This generates a smooth surface representation which stays faithful to the original binary data.

1 Introduction

The generation of three-dimensional models of anatomical structures from medical imagery is important for applications such as surgical simulation, planning, and image-guided surgery. An internal scan typically consists of high-resolution data in the imaging plane and significantly lower resolution between imaging slices. The lack of high-resolution information

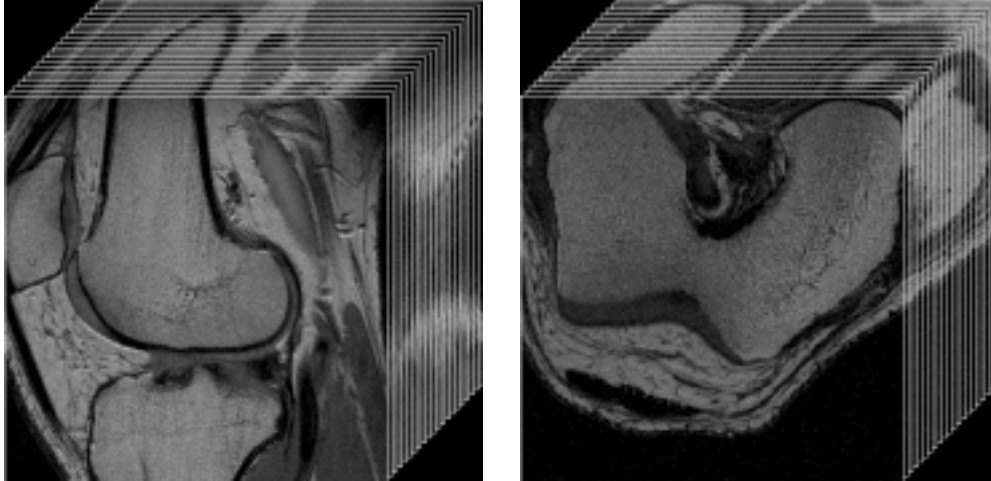


Figure 1: Two MR scans of a person’s knee. The scan on the left was acquired sagittally while the scan on the right was acquired axially. Both images have high resolution in-plane, but have about one quarter the resolution between planes.

along the scanning direction causes aliasing or terracing artifacts in anatomical surface models, which can be distracting or misleading to surgeons. For surgical simulation, the terraces subtract from the realism of the visualization and create very noticeable ridges when using haptics to feel the object’s surface.

These terracing artifacts can be reduced by increasing the resolution of the scan. However, for CT scans, higher resolution between imaging planes subjects patients to a higher dose of radiation. For MR scans, longer scan times are necessary to achieve higher resolution, which is more costly and is more difficult for the patient, who must remain absolutely still during image acquisition.

For clinical practice, scans are usually acquired in more than one orthogonal direction. For example, instead of acquiring a single very high resolution sagittal MR scan, lower resolution sagittal and axial scans may be acquired (see figure 1). Surgeons and radiologists use information from both acquisitions for diagnosis, surgical guidance, and treatment. Similarly,

we are interested in combining the information from two scans to produce three dimensional models of internal structures that have higher resolution than models created from either of the scans alone. The method proposed here is an extension of the Constrained Elastic SurfaceNet described in [10], which generates models from a single scan.

2 Previous Work

Two basic methods are commonly used to fit surfaces to binary data. In the first, the binary data is low-pass filtered, and an algorithm such as Marching Cubes is applied, where the surface is built through each surface cube at an iso-surface of the grey-scale data [18]. In order to remove terracing artifacts and reduce the number of triangles in the triangulated surface, surface smoothing and decimation algorithms can be applied. However, because these procedures are applied to the surface without reference to the original segmentation, they can result in loss of fine detail.

In the second general method for fitting a surface to binary data, the binary object is enclosed by a parametric or spline surface. Control points on the surface are moved towards the binary data in order to minimize an energy function based on surface curvature and distance between the binary surface and the parametric surface. McInerney and Terzopoulos used such a technique to detect and track the surface of the left ventricle in sequences of MRI data [20] and Takanahi *et al.* used a similar technique to generate a surface model of muscle from segmented data [22]. This approach has two main drawbacks for general applications. First, it is difficult to determine how many control points will be needed to ensure sufficient detail in the final model. Second, this method does not handle complex topologies easily.

Recently, Gibson [10] introduced Constrained Elastic SurfaceNets which fit an elastic net of nodes over the surface of a binary segmented dataset and move the node positions to reduce the surface curvature while constraining the net to remain within one voxel of the binary surface. This approach produces smooth surface models from binary segmented data that are faithful to the original segmentation.

3 Dual SurfaceNets

Dual SurfaceNets extend the original SurfaceNet approach by combining information from two orthogonal volume image scans. The use of Dual SurfaceNets requires a number of preprocessing steps. First, the object of interest is segmented or extracted from each of the scans. The scans are then registered into a common coordinate frame, and a SurfaceNet is initialized for each of the segmentations. The dual relaxation process then relaxes the shapes of the SurfaceNets subject to constraints based on the resolution of the two scans.

3.1 Segmentation

The anatomical structures that appear in an internal scan such as MR and CT must be explicitly extracted or segmented from the scan before they can be directly used for surface model generation. By segmentation, we refer to the process of labeling individual voxels in the volumetric scan by tissue type, based on properties of the observed intensities as well as known anatomical information about normal subjects.

The segmentation is performed using automated techniques and semi-automated techniques. With CT data, segmentation can be performed relatively automatically using inten-

sity thresholding or other low-level image processing. However, with MRI, image segmentation is challenging and generally requires more sophisticated algorithms and significant human input. An example of a semi-automatic segmentation method developed by the Surgical Planning Lab at Brigham & Women's Hospital, uses real-time ray tracing running on a Thinking Machines Inc. CM-2 to render segmented tissues as parameter values are dynamically varied. The parameters of interest are the intensity thresholds used to define intensity-based tissue classifiers. These thresholds are interactively modified and both 2D slices and 3D renderings of the resultant classifiers are viewed in real time. If necessary, further structural correction of the segmentation is performed by interactively editing the 2D slices. Such editing consists of highlighting regions of interest with the mouse and specifying the desired label for the contained pixels.

3.2 Registration

To combine the information from two scans of the patient into one three-dimensional model, both datasets must be in the same coordinate system. Registering rigid objects requires solving the six degree of freedom pose of the object in one scan with respect to the object in the other scan. While sophisticated methods exist to register the scans directly from the image data [30, 31], a straightforward registration algorithm is used here because the binary segmentation is available. The best alignment of the models is found by performing stochastic gradient descent on the sum of squared differences of the smoothed segmented images. In practice, this straightforward registration method performs well on well-segmented binary image data.

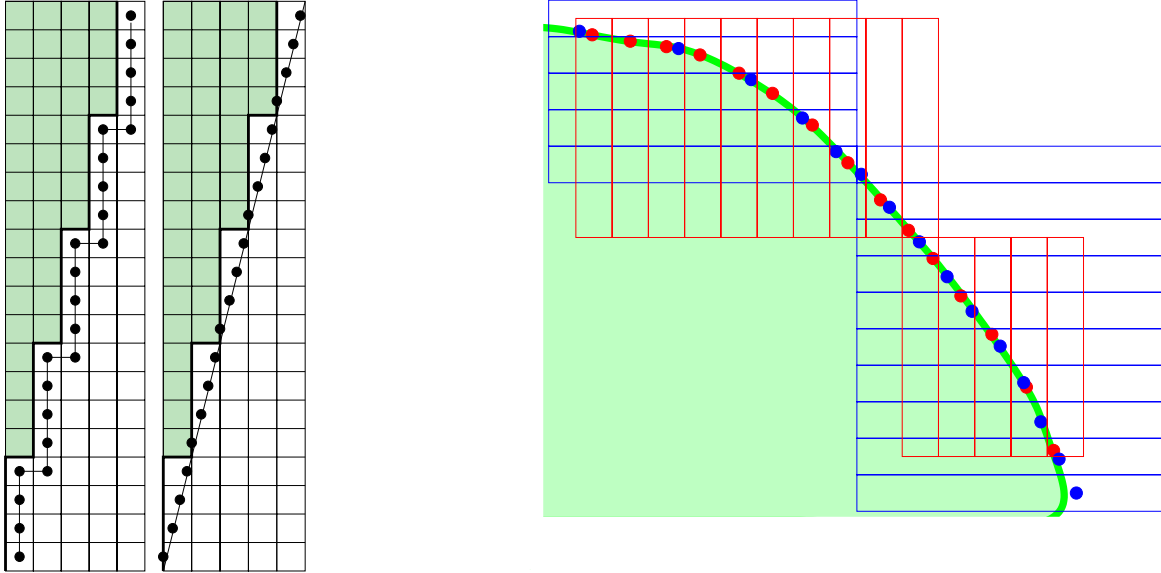


Figure 2: LEFT: A single SurfaceNet is built around an object and then relaxed, producing a smooth surface free of terracing. RIGHT: Two nets are built for the two orthogonal image scans. In this example, the red coded image and its node points have high horizontal resolution, but low vertical resolution, while the blue coded net has opposite properties. Nodes of both nets are initialized at the center of each cell and then relaxed subject to constraints to generate the smooth object surface shown in green.

3.3 Generating the SurfaceNet

Given a segmented image of an object, an initial SurfaceNet is created as described in [10].

The first step in generating a SurfaceNet is to locate cells that contain surface nodes. A cell is defined by 8 neighboring voxels in the binary segmented data, 4 voxels each from 2 adjacent planes. If all 8 voxels have the same binary value, then the cell is either entirely inside or entirely outside of the object. If at least one of the voxels has a binary value that is different from its neighbors, then the cell is a surface cell. The net is initialized by placing a node at the center of each surface cell and linking nodes that lie in adjacent surface cells. Each node can have up to 6 links, one each to its right, left, top, bottom, front, and back neighbors. Figure 2a illustrates the creation of a net from a binary image.

3.4 Relaxation of a Single SurfaceNet

Once defined, the SurfaceNet can be relaxed to reduce terracing artifacts while remaining faithful to the input segmentation [10]. To relax the net, each node is repositioned to reduce an energy measure in the links. In the examples presented here, SurfaceNets were relaxed iteratively by considering each node, $p[i]$, in sequence and moving that node towards a position equi-distant between its linked neighbors.

$$\hat{p}[i] = \frac{1}{\#\{\mathcal{N}(i)\}} \sum_{j \in \mathcal{N}(i)} p[j] \quad (1)$$

where $\mathcal{N}(i)$ is the set of linked neighbors of point i . The energy is computed as the sum of the squared lengths of all of the links in the SurfaceNet¹. Defining the energy and relaxation in this manner without constraints will cause the SurfaceNet to shrink to a single point. Hence, to remain faithful to the original segmentation, a constraint is applied that keeps each node inside its original surface cell. This constraint favors the original segmentation over smoothness and forces the surface to retain thin structures and cracks.

3.5 Dual Relaxation of SurfaceNets

Relaxing a single SurfaceNet of an object significantly reduces the artifacts contained in the model. However, if the resolution in the scan is low in one direction, there may not be enough information in one scan to fully constrain the model and remove the terraces. We

¹Alternative energy measures and relaxation schemes are also feasible. For example, a system that adjusts node positions to reduce local curvature would produce smoother surfaces and with less sharp corners than the method used here.

therefore consider using two scans, where one has higher resolution along the direction where the other has lower resolution, as illustrated in figure 1.

To relax two models of an object together, the individual SurfaceNets are built as described above. The two SurfaceNets, once aligned in the same coordinate frame, are iteratively relaxed towards one another with the constraint that each node must lie within its surface cell. In one relaxation step, each point $p[i]$ in the first net is updated by taking an average (weighted by distance) of the points $q[j]$ in the other net.

$$\hat{p}[i] = \frac{\sum_j w(p[i], q[j])q[j]}{\sum_j w(p[i], q[j])}, \quad (2)$$

where

$$w(u, v) = e^{\frac{-1}{2\sigma^2}\|u-v\|^2} \quad (3)$$

The point $\hat{p}[i]$ could violate its constraint by lying outside its cell, $c[i]$. The new position of the point, $p'[i]$ is $\hat{p}[i]$ if it lies inside the cell and the closest point on the cell boundary if $\hat{p}[i]$ lies outside the cell. In the next iteration, the second net is relaxed towards the first. After each full dual relaxation step, the nets are each relaxed individually for one iteration. The individual relaxation keeps each net smooth as they merge. The iteration progresses until the positions of some user-defined fraction of the nodes have converged, at which time one of the two nets is chosen to generate the final triangle model.

If the segmentation and registration were ideal, then the true surface would always lie in the intersection of the surface cells of the two images. In this case, the two nets would converge on the identical surface with all surface cell constraints satisfied. Figure 2b shows

a 2D example of a surface passing through the surface cells of two nets.

In general, the surface cells of the two scans do not overlap perfectly due to imaging, segmentation, and registration errors. We therefore provide a means of relaxing the constraints to allow the nets to merge more closely. After a few iterations, any point that is pulled outside its constraining cell cannot meet a corresponding point in the other net. This signifies discrepancies between the two models. In these instances, the constraining cell of every such point is dilated (preserving aspect ratio) by a small amount at the end of the iteration, allowing those points to move closer to the other net in the next iteration. Although the resultant net can move more than one voxel from the segmentations, the final model is guaranteed to be between the two initial models.

4 Results

Results of the dual relaxation are shown in figures 6 through 9. The first test dataset is of a femur. One scan was acquired axially at a resolution of $0.27\text{mm} \times 0.27\text{mm} \times 1.00\text{mm}$. The other scan (of the same person) was acquired sagittally (one year later) at a resolution of $0.25\text{mm} \times 0.25\text{mm} \times 1.40\text{mm}$. The femur was segmented manually from both images². Figure 6 shows the results of running Marching Cubes [18], individual SurfaceNets [10], and Dual SurfaceNets on the images. No decimation was performed on any of the models. Notice the terracing artifacts in the models generated with Marching Cubes and individual SurfaceNets along the direction that the scans were acquired. The model generated using Dual SurfaceNets on both scans represents the fine details in the original scans well but does

²These datasets were provided by the Surgical Planning Lab of Brigham and Women’s Hospital.

not contain the terraces.

In the second example, we consider building a model from extremely low resolution scans. Figure 7 shows results of model generation from subsampled versions of the original segmentations. The axial and sagittal scans were subsampled by a factor of 4 to resolutions of $1.09\text{mm} \times 1.09\text{mm} \times 4.00\text{mm}$ and $1.00\text{mm} \times 1.00\text{mm} \times 5.60\text{mm}$ respectively. The model generated using Dual SurfaceNets at the low resolution contains slightly less detail than the high resolution version, but it is remarkably smooth and free of terracing artifacts, while remaining faithful to the original segmentations.

The surface models can be visually verified by superimposing the relaxed net on the image data. Figure 8 shows the input segmentations to the Dual SurfaceNet algorithm and the final result of the net. The top row consists of the full resolution image, and the bottom row shows the subsampled images. Despite the blockiness evident in all the input segmentations (especially in the subsampled images), the final models are very smooth and capture the details of the femur.

The Dual SurfaceNet algorithm was also tested on a subsampled MR skin model. One image was acquired sagittally and one axially. Both were subsampled to $2.0\text{mm} \times 2.0\text{mm} \times 6.0\text{mm}$. The results of running the various model generation methods are shown in figure 9. Notice that the axial scan is cropped below the nose and on the top of the head. The model generated using both SurfaceNets is smooth above the nose, but contains the same terracing artifacts as the sagittal SurfaceNet below the nose (where there was no information present from the axial scan).

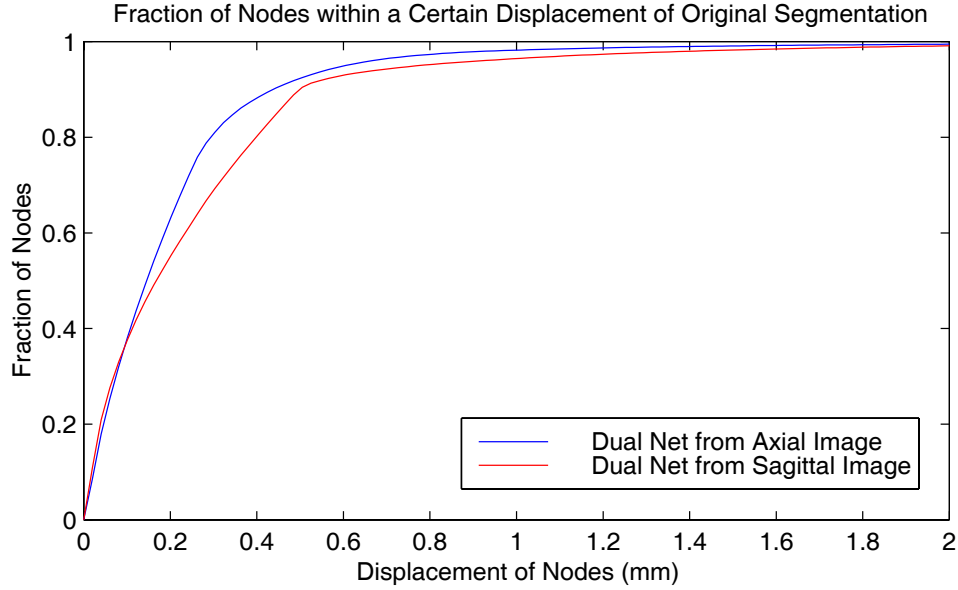


Figure 3: A plot of the amount that the nodes of each net are displaced from their initial positions (at the centers of their cells). Over 97% of the nodes lie within one voxel of the starting point.

5 Validation

Three dimensional models of anatomical structures are now routinely used by surgeons when treating patients [13]. Thus, there is a clear need to validate the process by which such models are generated. One method of validating the result of relaxing Dual SurfaceNets is by visual inspection. The 3D model can be superimposed onto the original grayscale image, as shown in figure 8. The borders of the model can be confirmed by examining each slice of the image.

In SurfaceNets, each node of the model is guaranteed to lie within one voxel of the original binary segmentation. Dual SurfaceNets can uphold the same constraints, but in practice these constraints need to be relaxed slightly to effectively combine the information in both nets. The distance that a node strays from its initialization point (the center of its cell) can be constrained during the relaxation. Furthermore, upon convergence, the distribution of displacements can be viewed to determine the goodness of the fit. Figure 3 graphs the

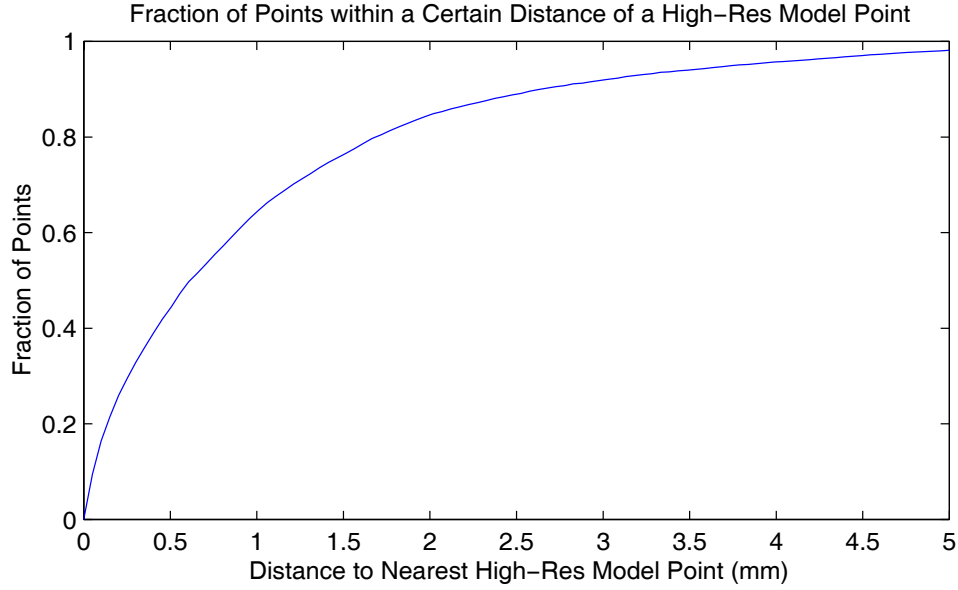


Figure 4: A plot of the distance of each node from the low resolution nets to the nearest node in the high resolution (ground-truth) segmentation.

fraction of nodes that lie within a certain distance of their original positions. In these scans, the distances from one corner to the other in a voxel are 1.07 mm and 1.44 mm for the axial and sagittal scans respectively. In both nets, over 97% of the points lie within one voxel of their starting position.

The validation process is often hindered by the difficulty in obtaining ground truth. While we do not have explicit ground truth, we generated the low resolution femur model and then compared the result with the high resolution femur segmentation. Ideally, each point of the low resolution model should fall near the high resolution surface. Figure 4 plots the fraction of nodes that are within a certain distance of a point on the high resolution model. The voxel extents of the axial and sagittal scans used in generation of the nets are 4.28mm and 5.78mm respectively. The majority of model points fall within a millimeter of the high resolution model, and 98% lie within one sub-sampled voxel of the original data.



Figure 5: An illustration of the difficulty in merging features that are not registered well. If (a) and (b) are the two input models of the same object, the bump is most likely the same feature, in which case the ideal merge would produce one bump in the middle (d), versus (c) where two bumps result.

The final model is not only very smooth, but also faithful to the input segmentation.

6 Future Work

The promising initial results described herein provide various directions of future work. The Dual SurfaceNets currently rely heavily on the segmentation and registration of the anatomical structures being accurate. Nodes are currently attracted to nodes of the other model based solely on distance (see definition of w , equation 3). Slight registration or segmentation errors could cause the closest point on the other net to be different from the corresponding point. Thus sub-structures in the objects do not overlap, resulting in the merging of different features. Figure 5 illustrates this problem. These assumptions can be relaxed by adding a richer feature correspondence measure. We are exploring methods of including normal information and/or curvature information in the weighting function. Consider a new weighting function,

$$w(u, \hat{n}_u, \kappa_u, v, \hat{n}_v, \kappa_v) = f_{dist}(u, v) + \lambda_1 f_{norm}(\hat{n}_u, \hat{n}_v) + \lambda_2 f_{curv}(\kappa_u, \kappa_v) \quad (4)$$

where f_{dist} , f_{norm} , and f_{curv} are strictly positive and decrease monotonically as the arguments diverge. For a node to attract another, it must be close and have similar normal and

curvature. One major difficulty of this approach is determining how to weight these various factors. The addition of richer features will help in merging non-rigid structures such as cortex, cartilage, or meniscus. In effect, the local, non-rigid registration can be computed during relaxation of the nets.

Another interesting extension involves accounting for shape change. A SurfaceNet does not change topology during relaxation. However, in some cases, the two nets may have different topologies, preventing them from merging completely. After relaxing both nets, it is not clear which net should be used as the result. We plan to explore methods of adapting the topology of the net during the relaxation process to account for the differences in the models.

References

- [1] R. Avila and L. Sobierajski. A haptic interaction method for volume visualization. In *Proc. Visualization'96*, pages 197–204. IEEE, 1996.
- [2] M. Bro-Nielsen and S. Cotin. Real-time volumetric deformable models for surgery simulation using finite elements and condensation. In *Proc. Eurographics*, volume 15, pages 57–66, 1996.
- [3] D. Chen and D. Zeltzer. Pump it up: computer animation of a biomechanically based model of muscle using the finite element method. In *Proc. SIGGRAPH 92*, pages 89–98., 1992.
- [4] M. Desbrun and M-P Gascuel. Animating soft substances with implicit surfaces. In *Proc. SIGGRAPH 95*, pages 287–290, 1995.
- [5] J. Foley, A. vanDam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley, 1992.
- [6] S. Gibson. 3D chainmail: a fast algorithm for deforming volumetric objects. In *Proc. Symposium on Interactive 3D Graphics*, pages 149–154. ACM SIGGRAPH, 1997.

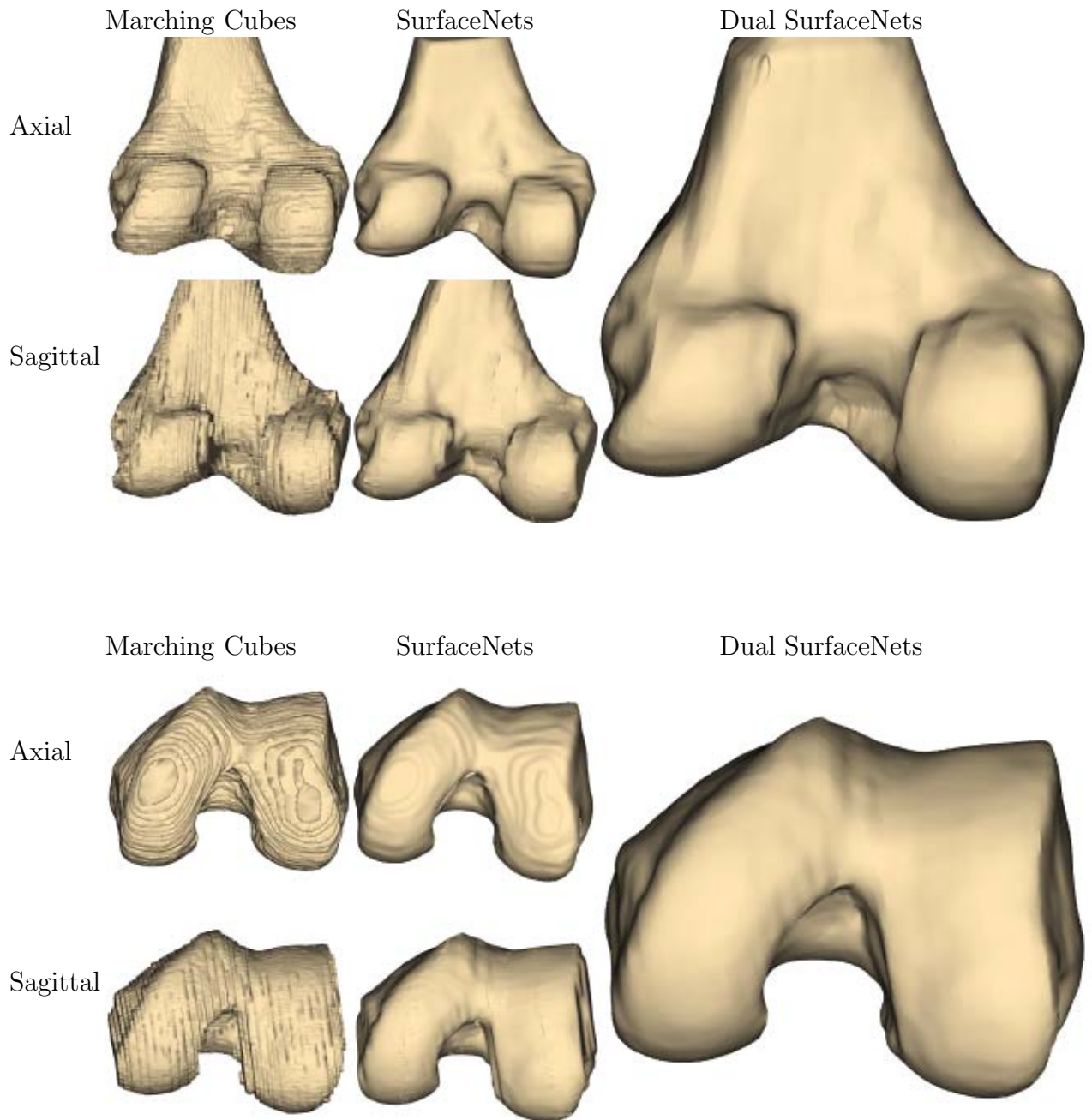


Figure 6: Surface models of the high resolution scan of the femur. The top and bottom sets of images show two views of the same models. For each scan, the model has been generated using Marching Cubes and SurfaceNets. The model generated using Dual SurfaceNets combined the information from both scans.

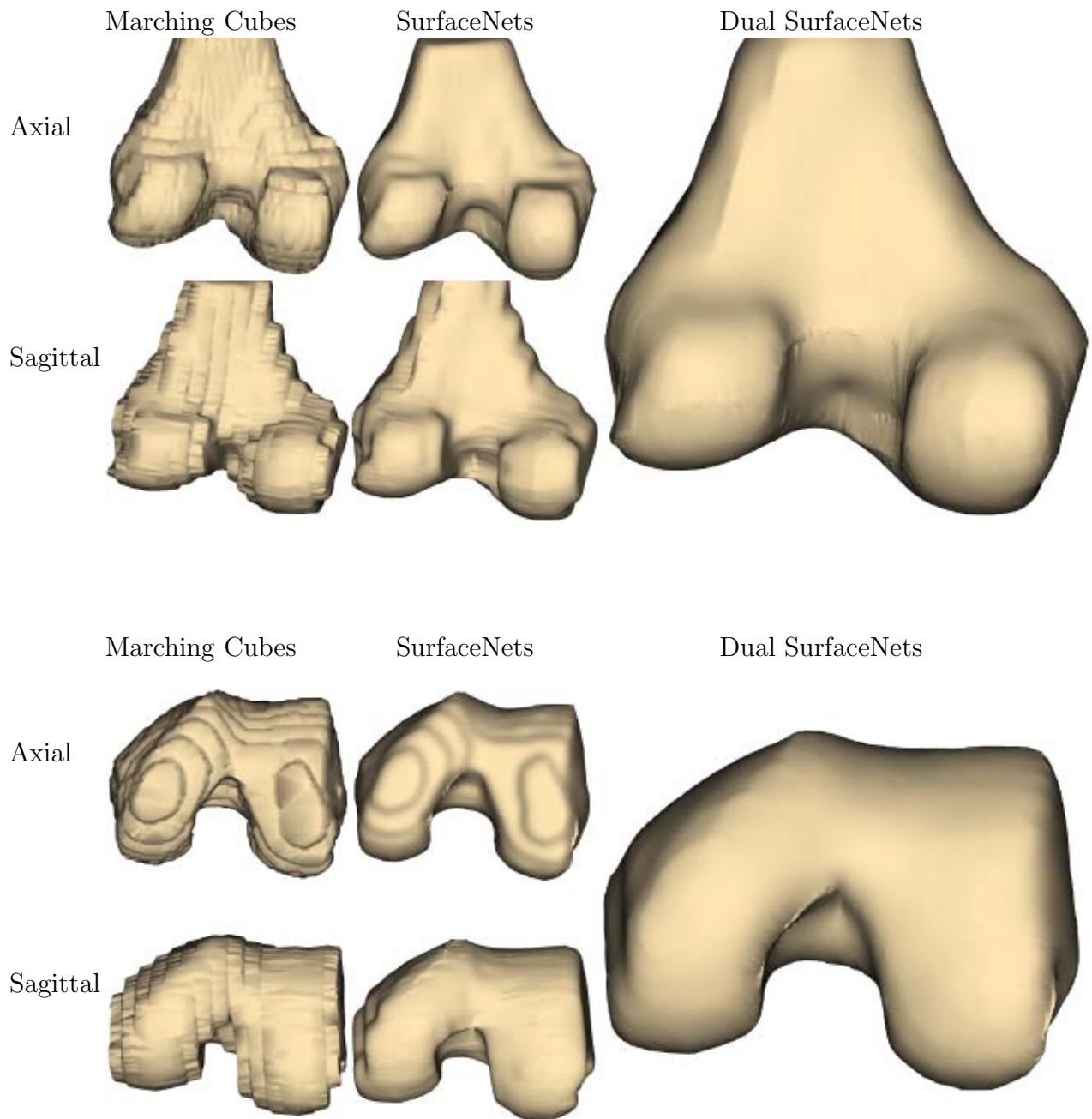


Figure 7: Surface models of a low resolution (subsamped) scan of the femur. The top and bottom sets of images show two views of the same models. For each scan, the model has been generated using Marching Cubes and SurfaceNets. The model generated using Dual SurfaceNets combined the information from both scans.

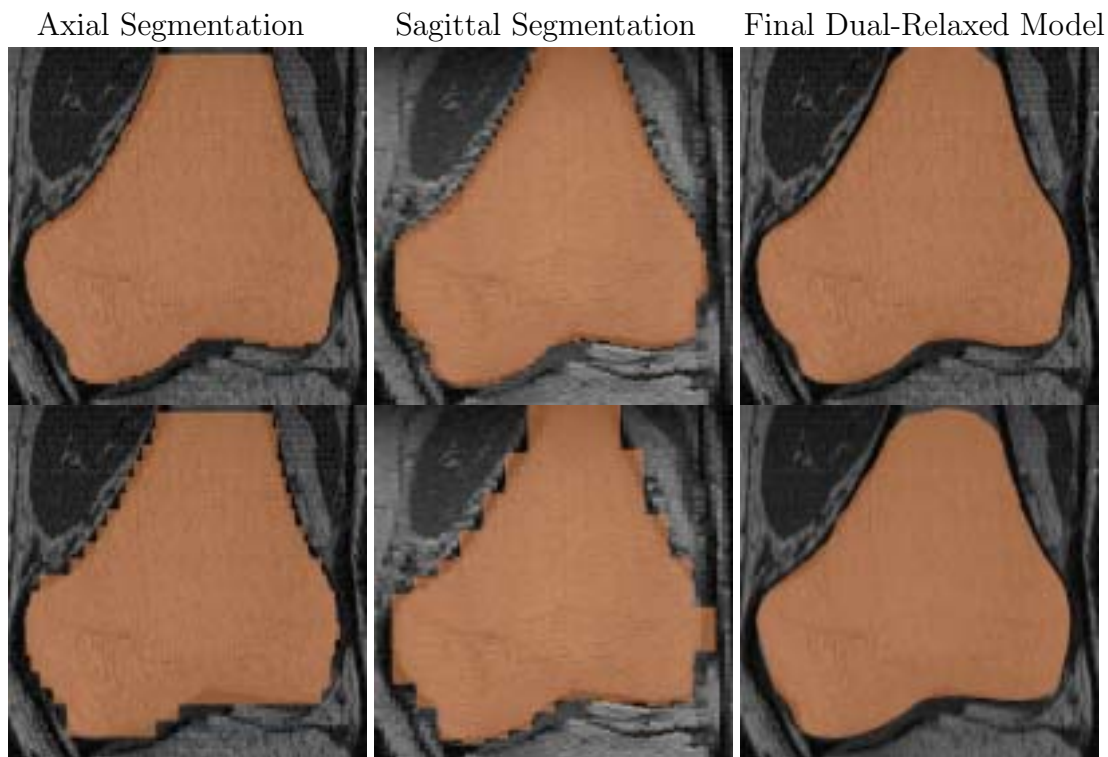


Figure 8: Models of the femur overlaid on the original grayscale data. The top row of images are the original high resolution scans. The bottom row of images are the subsampled scans. The first two images in each row are the segmentation input to the Dual SurfaceNets. The result of relaxing the nets is shown in the third image.

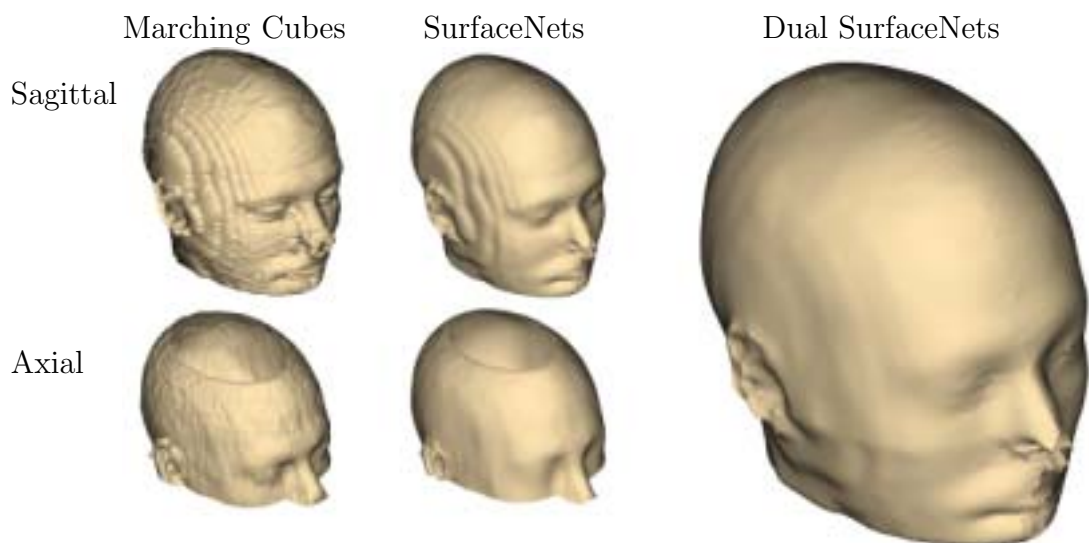


Figure 9: Surface models of a low resolution (subsampled) scan of a head. The model has been generated using Marching Cubes, SurfaceNets, and Dual SurfaceNets. Note the artifacts in the Dual SurfaceNet model beyond the extent of the axial scan.

- [7] S. Gibson. Linked volumetric objects for physics-based modeling. *submitted to IEEE Trans. on Visualization and Computer Graphics*, 1998.
- [8] S. Gibson. Using distance maps for accurate surface representation in sampled volumes. *submitted to SIGGRAPH'98*, 1998.
- [9] S. Gibson, C. Fyock, E. Grimson, T. Kanade, R. Kikinis, H. Lauer, N. McKenzie, A. Mor, S. Nakajima, H. Ohkami, R. Osborne, J. Samosky, and A. Sawada. Volumetric object modeling for surgical simulation. *Medical Image Analysis*, 1998. to appear in spring, 1998.
- [10] S. Gibson. Constrained Elastic SurfaceNets: generating smooth surfaces from binary segmented data. In *MICCAI*, 1998.
- [11] S. Gibson, J. Samosky, A. Mor, C. Fyock, E. Grimson, T. Kanade, R. Kikinis, H. Lauer, N. McKenzie, S. Nakajima, H. Ohkami, R. Osborne, and A. Sawada. Simulating arthoscopic knee surgery using volumetric object representations, real-time volume rendering and haptic feedback. In *Proc. CVRMed-MRCAS'97*, pages 368–378. Springer, 1997.
- [12] J.P. Gourret, N. Magnenat-Thalmann, and D. Thalmann. Simulation of object and human skin deformations in a grasping task. In *Proc. SIGGRAPH 89*, pages 21–30, 1989.
- [13] E. Grimson, M. Leventon, G. Ettinger, A. Chabrierie, F. Ozlen, S. Nakajima, H. Atsumi, R. Kikinis, P. Black. Clinical Experience with a High Precision Image-guided Neurosurgery System In *MICCAI 98*, 1998.
- [14] K. Hohne, M. Bomans, A. Pommert, M. Riemer, C. Schiers, U. Tiede, and G. Wiebecke. 3D visualization of tomographic volume data using the generalized voxel model. *The Visual Computer*, 6(1):28–36, February 1990.
- [15] A. Kaufman. *Volume Visualization*. IEEE Computer Society Press, Los Alamitos, CA, 1991.
- [16] R. Koch, M. Gross, F. Carls, D. von Buren, G. Fankhauser, and Y. Parish. Simulating facial surgery using finite element models. In *Proc. SIGGRAPH 96*, pages 421–428, 1996.
- [17] Y. Lee, D. Terzopoulos, and K. Waters. Realistic modeling for facial animation. In *Proc. SIGGRAPH 95*, pages 55–62., 1995.
- [18] W. Lorensen and H. Cline. Marching cubes: a high resolution 3D surface construction algorithm. In *Proc. SIGGRAPH 87*, pages 163–169, 1989.

- [19] S. Lu, D. Cui, R. Yagel, R. Miller, and G. Kinzel. A 3Dcontextual shading method for visualization of diecasting defects. In *Proc. Visualization'96*, pages 405–407. IEEE, 1996.
- [20] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.
- [21] A. Mor, S. Gibson, and J. Samosky. Interacting with 3-dimensional medical data: Haptic feedback for surgical simulation. In *Proc. Phantom User Group Workshop'96*, 1996.
- [22] I. Takanahi, S. Muraki, A. Doi, and A. Kaufman. 3D active net for volume extraction. In *Proc. SPIE Electronic Imaging'98*, pages 184–193, 1998.
- [23] R. Osborne, H. Pfister, H. Lauer, N. McKenzie, S. Gibson, W. Hiatt, and T. Ohkami. EM-Cube: An architecture for low-cost real-time volume rendering. In *Proc. SIG-GRAPH/Eurographics Workshop on Graphics and Hardware*, pages 131–138, 1997.
- [24] J. Russ. *The Image Processing Handbook*. CRC Press, Boca Raton, FL, 1992.
- [25] W. Schroeder, W. Lorensen, and S. Linthicum. Implicit modeling of swept surfaces and volumes. In *Proc. Visualization'94*, pages 40–45. IEEE, 1994.
- [26] D. Terzopoulos and K. Waters. Physically-based facial modeling, analysis, and animation. *Journal of Visualization and Computer Animation*, 1:73–80, 1990.
- [27] G. Thurmer and C. Wurthrich. Normal computation for discrete surfaces in 3D space. In *Proc. Eurographics'97*, pages C15–C26, 1997.
- [28] S. Wang and A. Kaufman. Volume sampled voxelization of geometric primitives. In *Proc. Visualization'93*, pages 78–84. IEEE, 1993.
- [29] S. Wang and A. Kaufman. Volume-sampled 3D modeling. *IEEE Computer Graphics and Applications*, 14:26–32, 1994.
- [30] W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, R. Kikinis. “Multi-Modal Volume Registration by Maximization of Mutual Information”. *Medical Image Analysis*, 1(1):35–51, 1996.
- [31] J. West, J. Fitzpatrick, *et al.* “Comparison and evaluation of retrospective intermodality image registration techniques.” In *Medical Imaging: Image Processing*, volume 2710 of *Proc. SPIE*, Newport Beach, California, February 1996.
- [32] R. Yagel, D. Cohen, and A. Kaufman. Discrete ray tracing. *IEEE Computer Graphics and Applications*, 12:19–28, 1992.