Research Article

Model-based Segmentation of CT Images

O-C Marte P Marais

Collaborative Visual Computing Laboratory,
Department of Computer Science,
University of Cape Town, Rondebosch 7701, South Africa.
omarte@cs.uct.ac.za, patrick@cs.uct.ac.za

Abstract

This paper presents preliminary work on the segmentation of Computed Tomography data using a model-based approach. Conventional image processing of CT data is aimed at the production of simple iso-surfaces for surgical planning or diagnosis — such methods are not suitable for the automated detection of fractures, which is the ultimate application of our work. To address these deficiencies a surface-based technique with appropriate constraints is introduced. The output of the segmentation phase is a triangulated surface representing the bone or bones of interest. We illustrate the method applied to low resolution CT test data and discuss its robustness and performance.

Keywords: *image processing, computer vision, registration, segmentation, medical imaging, radiology, CT*
Computing Review Categories: *I.4.6, I.3.5, J.3*

1 Introduction

The field of medical imaging provides an excellent test-bed for new algorithms in computer vision and image processing. Medical data is information rich (many bits per pixel) and often degraded due to patient movement or inherent limitations imposed by the imaging technology. Furthermore, clinicians place very high requirements on the accuracy of any algorithm since lives may be at risk.

This paper presents preliminary work on a system to detect a subset of common fractures of the human skeletal system. Here we concentrate on the development of robust, real-time algorithms which can yield a speedy segmentation of a volumetric CT (Computed Tomography) data set. In this context, segmentation requires that we produce an anatomically accurate model of the bone (or bones) of interest which can then be passed onto a subsequent analysis stages. This analysis subsystem will be the subject of a later paper. The data for which these algorithms are being developed will be generated by a new CT reconstruction technique which utilises severely undersampled x-ray data to minimise patient x-ray exposure. Preliminary indications suggest that the resulting image data may have a number of unique artifacts and that naive or simplistic implementations of "standard" methods are likely to fail. We have thus focused our efforts on ensuring that our algorithms can cope with a significant degree of image degradation.

The paper is presented as follows. Section 2 presents a brief survey of related work. Section 3 provides a description of the methodology we have employed, while Section 4 provides some preliminary results and visualisations. Future plans are discussed in Section 5 and the paper concludes in Section 6.

2 Background and Related Work

The segmentation of volumetric data has been tackled in many different ways. In general, methods may be voxel-based or surface-based. The former usually apply simple criteria to classify each voxel in the scene as being part of a specific component. The latter attempt to recover the boundary of each component within the volume. Voxel-based methods are often adequate if a simple visualisation of the data is all that is required. However, even in these cases, without an underlying intensity model, naive voxel classification [10] can yield a misleading representation. Surface-based techniques may be constrained in a number of ways, ranging from simple continuity constraints to complex deformable templates[8, 7, 5, 2, 6]. Many of these methods require a significant amount of run-time computation which is counter to our requirement for a real-time solution. For others, a collection of training models is assumed to be available to generate constraints for the underlying deformable model. In practise, obtaining a representative collection of anatomical models is difficult or impractical.

Most surface-based techniques for volume segmentation are variations on the concept of a "snake" [4]: a surface or curve that gravitates towards features of interest (usually points on the component boundary). The surface is initialised in the volume and allowed to move towards the feature set. Once the process has converged the final surface is taken to represent the object/component of interest. A voxel-based representation has to be processed further to yield a visually useful result — either by means of volume rendering or iso-surface extraction. The former is particularly time consuming. Furthermore, a voxel data structure does not lend itself to the analysis of component

geometry. Additional problems occur when the features of interest, for example, the surface of the brain in an MRI volume, are themselves barely discernible or occluded. In this cases, additional models must be applied to exploit knowledge of local object geometry [6]. Fortunately the bone interface is usually fairly well defined in a CT image.

The primary objective of our approach is to develop a real-time method for segmentation of bone structures within CT, which can then serve as input to an expert system aimed at analyzing bone fractures. Clearly issues of robustness and accuracy are of prime concern: we need to ensure our segmentation extracts a good (physiologically accurate) model of anatomy — thus the requirement for a more sophisticated model-based approach. To satisfy clinical practise the results of the segmentation/analysis must be available when the scan terminates or shortly thereafter, ruling out computationally intensive segmentation schemes.

3 Methodology

In this section we outline the most important algorithms developed and used in our work.

3.1 The Bone Model

A model for the bone we wish to extract can be acquired from a high resolution CT or MRI scan of "normal" (non-pathological) anatomy[1]. The bone can be extracted by manual or partially automated segmentation tools. For high quality image data there is usually a clear delineation between different tissue types within the data set and segmentation becomes significantly easier. Human intervention (via editing tools) is required to correct segmentation errors and to ensure that the model is anatomically correct. An iso-surface extractor can be used to generate a triangulated surface which can be re-meshed to any desired vertex density by simple triangle sub-division. The model can then be used to extract similar anatomy from poor resolution data or data with many artifacts.

3.2 Segmentation

In order to evaluate the integrity of the bone we are interested in, we need to extract a surface model of the bone from the CT scan. This is accomplished by first isolating those parts of the data that are bone through heuristics and standard image processing techniques. Following this, a coarse to fine alignment of the template surface bone and CT scan voxel data is achieved through coarse alignment, rigid alignment, and finally local deformation.

[1] Several models may be required to accommodate sex and age differences. This is simply an additional *a priori* parameter. The algorithms remain unchanged.

3.2.1 Image Processing

Instead of dealing with the CT scan as a volumetric data set, we first consider the 3D volumetric data set as a stack of 2D components, and operate independently on each of these image slices.

Such a simplification is justified for a number of reasons:

1. The spatial resolution between pixels in the image slices will in the vast majority of cases be greater than the voxel resolution between consecutive slices in the volumetric data. It therefore make sense to take as much advantage of this superior resolution as we can before moving to the coarser volumetric domain.

2. The output from a CT scan is a stack of image slices. Without any data transformation we can make use of numerous well developed image processing techniques.

3. Working with volumetric data sets has high memory requirements. The simplification of the volumetric data set, after image processing on each image slice has been completed, will result in greater memory efficiency.

The aim of our image processing is to identify those pixels in each image slice that form part of the bone we are interested in. This can be achieved in two logical steps. First, we need to classify those pixels that we consider to be bone. Intensity thresholding based on values obtained from histogram analysis results in a classification of the bone segments contained in the image. Then, the pixel grouping that represents the target bone needs to be isolated from the other bone segments that have been found in the image slice. Simple heuristics based on knowledge of the anatomical area under investigation, combined with region growing and largest connected components can be used to isolate the target bone [9].

3.2.2 Registration of Target Voxel Model and Reference Mesh Model

Once we have isolated the target bone from muscle, fat, other bones and noise within each CT image slice, this voxel bone model must be matched against a reference bone model, represented as a triangulated surface mesh.

In order to use the *a priori* high level knowledge that is of benefit in a model-based segmentation approach, we need to ensure that the geometric features of the voxel bone model and the reference surface model are in close proximity. This requires registration of the voxel and surface models, so that the geometric information inherent in the reference model can be used to guide the segmentation process.

The alignment of the voxel model to the surface model is achieved through a coarse to fine approach. Firstly, the voxel and surface models are mapped into a common coordinate system and an initial crude orientation is estimated

for each. This is followed by global registration under a similarity transform (scaling, rotation and translation). Finally, a local deformation of the surface model ensures that a smooth constrained approximation of the voxelised bone surface is obtained. The accuracy of this fit is limited by the resolution of the voxel data, and to a lesser extent, the vertex density of the mesh model. Using prior information, the vertex density can be set to ensure sufficient flexibility in the surface model.

3.2.3 Coarse Registration

The voxel data is transformed into the reference mesh coordinate system. In this common coordinate system, the voxel bone model and the reference mesh model are both oriented arbitrarily. One of the primary benefits of model-based segmentation is to allow the reference model to guide the segmentation of the voxel data. In this way, *a priori* knowledge concerning the anatomical structure of the bone (which is encoded in the reference model) can be used to eliminate uncertainty in the segmentation process.

For this to be possible, we require that the reference mesh model and voxel model are closely aligned in space as well as in orientation/pose. In other words, we require that the discriminating features of the particular bone under consideration in the voxel and mesh model are in close proximity.

The initial rough registration can be approached in a number of different ways. We have explored two methods:

1. **Bounding Box Registration:** This method involves registering the bounding box enclosing the target voxel model and the bounding box enclosing the mesh reference model.

2. **Longest Line Segment Registration:** This method involves registration of the line segment representing the longest Euclidean distance through the target model with the line segment representing the longest Euclidean distance through the reference model. This technique works particularly well for elongated bones.

It is of course possible that the rough registration will result in an inconsistent orientation of target and reference models. The most common problem that can occur is an inversion of the orientation of the two models. In other words, the front-end of the target model is aligned with the back-end of the reference model, and vice versa. Fortunately, this ambiguity can be detected and corrected in our case by referring to the orientation of surrounding bones.

3.2.4 Rigid Registration

The target and reference models now share a common orientation and are spatially adjacent. A global alignment of the two models is now required. We used Iterated Closest Points (ICP) [1, 11] to achieve a rigid alignment between voxel and surface models.

ICP iteratively finds the rigid transformation (translation and rotation) that will map one set of points (reference points) onto another set of points (target points). For each iteration of the algorithm, the target and reference point sets are brought closer to one another through successive transforms. After a number of iterations convergence is achieved when successive iterations are no longer able to bring the two point sets any closer. At this point, the ICP iterations terminate.

In its basic form, the ICP algorithm requires knowledge of the one-to-one mapping between the two point sets (source points with corresponding points). We do not have such correspondence between voxel and surface models. Section 3.2.5 explains how the matching between these two point sets is achieved.

3.2.5 Correspondences

Finding the matching points between the voxel and surface models plays an important role in the segmentation process. Inaccurate correspondences will result in a poor global registration of the two models. Feature points in the two model will not match and the surface model used to guide the segmentation will be incorrectly applied and will thus yield a warped estimate of the underlying bone surface.

We assume that the rough registration phase brings the surface and voxels models into close proximity. The search for a corresponding points can thus be limited to a small subset of candidate points surrounding the reference surface model point.

A spatial subdivision of the voxel model is achieved through the use of a kd-tree [3]. Using this data structure, the above mentioned candidate points can be efficiently extracted from the voxel model with a rectangular range query. Kd-trees have the properties that for n points:

- they can be constructed in $O(n \log n)$ time,
- require $O(n)$ storage, and
- a rectangular range query can be completed in $O(\sqrt{n}+k)$ time for k points found in the range.

We conduct a search along the surface normal of the reference model point and find the candidate point that is closest in Euclidean distance terms to this ray. This is the point that we select as our correspondence. This process is repeated for each point of the reference model. The result is that each point on the reference model has a matching point in the target model.

3.2.6 Local Deformation

After global alignment we need to account for the variation that occurs between bones of the same kind. To address this deviation requires a local alignment of target and reference models.

We define a *local patch*, the area of effect of the local registration, as a predefined proportion of the total surface area about a surface point. For example, each local patch

on the surface model might constitute 1% of the total surface area of the mesh.

This definition of a local mesh patch has the advantage of accommodating both mesh connectivity and Euclidean distance while preserving locality (a local patch will not include disjoint regions). It also provides an obvious way to change the scale at which local alignments occur. A "curvature hierarchy" for each reference bone model is constructed as a pre-process for use in subsequent bone analysis. So, in effect we gain this local patch definition for free.

The magnitude of the local alignment is determined by a weighted average distance between every corresponding point in the local patch on the target and reference model.

The direction of deformation is calculated as a weighted average of the unit surface normals of the local patch on the reference model. What this signifies is that the local patch is *pulled* towards its target based on the direction in which the surface faces.

The weighting is determined in such a way that the centre of the local patch has the largest influence and there is a gradual fall-off in weighting values as we move towards the edges of the patch. The central points of a patch therefore dominate the local deformation while peripheral points have less influence.

The local alignment between target and reference model proceeds iteratively until the registration cannot be significantly improved. Iteration terminates when the Mean Squared Error between the two models has fallen below some model specific threshold or a preset maximum number of iterations has been reached. The choice of Mean Square Error and maximum number of iterations affect the time taken to achieve local alignment as well as the quality of local alignment. If enough iterations are executed, a point is reached in the alignment process where an additional iteration will not improve the alignment of target and reference models. This means that a Mean Square Error and optimal number of iterations can be chosen for the model under consideration.

It should be noted that the local deformation is constrained. There is a limit on the magnitude of the deformation that can occur at any iteration of the algorithm. This is to ensure that a local patch does not get wrenched away from the model. If the magnitude of a local deformation is large, it is preferable that this deformation is completed incrementally, in successive iterations, as this helps preserve the smoothness of the surface curvature.

4 Preliminary Results

Results are presented for the segmentation of the *ulna* and *radius* (two bones within the forearm) of an adult female subject who underwent an axial CT scan. The axial spatial resolution of 10mm (see Table 1) meant that the final 3D CT scan was very coarse and thus provided a good test for the robustness of the segmentation algorithms.

4.1 Speed

The speed with which segmentation is achieved (see Table 3) reveals that the radiologist can analyse the results directly after the CT scan without much delay. The correspondences (see Section 3.2.5) account for the majority of processing time. Analysis reveals that significant speed improvements can be achieved by using an optimised version of the kd-tree.

Data	**Measurements**
image dimensions	512x512
pixel spatial resolution	0.46875mm
slice spatial resolution	10mm
number of slices	21

Table 1: **CT Acquisition Information:** The spatial data for the CT scan acquired from the DICOM header file.

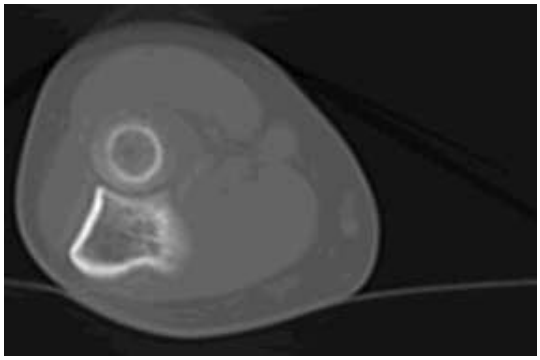
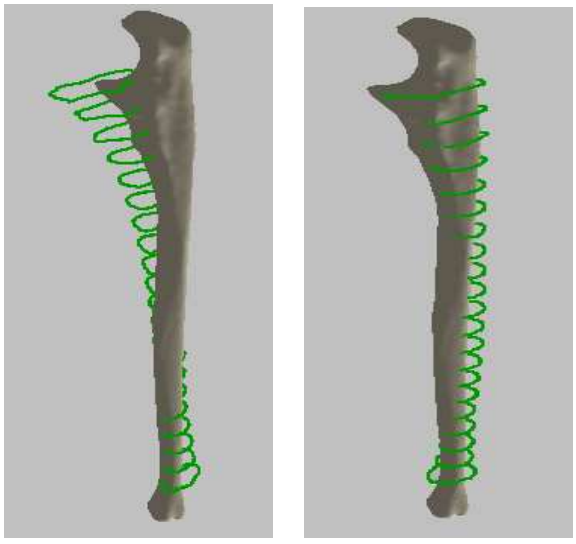


Figure 1: **CT Data:** An axial slice from the CT scan.



(a) (b)

Figure 3: **Alignment results:** The dark rings represent the voxel model while the surface model represents the reference model. (a) After coarse alignment, (b) rigid alignment of voxel and reference models.

(a) (b) (c)

Figure 2: **Image Processing:** In (a) the bone is isolated from the rest of the forearm. Then, in (b) the Ulna is separated from the Radius, and thinned (c) to reduce computational overhead.

(a) (b)

Figure 4: **Segmentation results:** The final surface model (a) Radius (b) Ulna produced by the segmentation process.

Data	**Ulna**	**Radius**
Vertices	4320	5675
Faces	8636	11346
Voxels	1915	1936

Table 2: **3D Information:** Breakdown of mesh and voxel attributes for the target and reference models.

4.2 Visualisation

Figures 1, 2, 3, 4 show the various stages of the segmentation process, from the input of CT image data to the output of a segmented surface model. Observe that the final surface model is rough in certain places. This is due to the coarse nature of the CT scan and can be eliminated by subsequent geometry-preserving smoothing. In practice, a large bone would not be scanned in its entirety, only the local area under investigation. This allows for a higher resolution scan to be performed.

Task	**Ulna**	**Radius**
Image Processing	0.23929s	0.23869s
Coarse Alignment	0.00035s	0.00040s
Rigid Alignment	0.01050s	0.03165s
Local Deformation	0.00611s	0.00771s
Correspondences	8.72194s	10.59366s
Total Time	8.97820s	10.87212s

Table 3: **Timing measurements:** Total time taken for the segmentation of the Ulna and Radius as well as for each individual task in the segmentation process. The timings were conducted on a AMD Athlon 700Mhz with 768MB of RAM.

5 Future Work

We intend using the results of the segmentation (the surface model) together with a *curvature hierarchy* to detect a subset of commonly occurring bone fractures. The geometry of the surface model will be analysed to determine the integrity of the bone under inspection. A full system will require the construction of a “fracture knowledge base” in order to classify the different types of bone fractures that occur.

6 Conclusion

We have introduced a model-based segmentation scheme for degraded volumetric CT data. This work constitutes the first phase of a system to detect fractures within the human skeletal system. The model is generated from real medical data and is represented as a mesh surface with associated curvature information. The model is mapped into the data volume using prior knowledge of patient alignment and iteratively aligned with the voxel data corresponding to the bone of interest using a coarse to fine approach. Preliminary results on low resolution CT data show that the algorithm is robust and recovers a good approximation of the underlying bone. Processing times indicate that even without any optimisation, the algorithm can be run in real-time and thus form part of a useful tool for clinicians.

7 Ackowledgements

We would like to thank African Medical Imaging (AMI) Inc. for their generous support, as well as the MRC/UCT Medical Imaging Research Unit under which this project falls. We would also like to thank our colleagues and co-workers for useful suggestions and comments.

References

[1] P.J Besl and N.D McKay. A method for registration of 3-d shapes. *IEEE. Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[2] T.F Cootes, A. Hill, C.J Taylor, and J. Haslam. The use of active shape models for locating structure in medical images. *Image and Vision Computing*, 12(6):276–285, July 1994.

[3] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer Verlag, 2nd edition, 2000.

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

[5] A. Kelemen, G. Szekely, and G. Gerig. Three-dimensional model-based segmentation. In *Proc. of International Conference on Computer Vision*, January 1998.

[6] P.C Marais. *The Segmentation of Sparse MR Images*. PhD thesis, Oxford University, 1998.

[7] T. McInerney and D. Terzopolous. Deformable models in medical image analysis:a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[8] G. Szekely, A. Kelemen, A. Brechbuhler, and G. Gerig. Segmentation of 2-d and 3-d objects from mri volume data using constrained elastic deformations of flexible fourier surface models. *Medical Image Analysis*, 1(1):19–34, 1996.

[9] J. T. Tou and R. C. Gonzalez. *Pattern Recognition Principles*. Addison-Wesley, Reading, Massachusetts, U.S.A., 1974.

[10] K. Vincken, A. Koster, and M. Viergever. Probabilistic segmentation of partial volume voxels. *Pattern Recognition Letters*, 15:477–484, May 1994.

[11] Z. Zhang. Iterative point matching for registration of free-form curves. Technical Report 1658, INRIA, March 1992.