ABSTRACT

Sona B. Kalra

Evaluation of a Medial Model for Segmentation of Kidneys in CT Images
(Under the direction of Edward L. Chaney, PhD)

The purpose of this study was to evaluate the use of a statistically trainable, medially based deformable model for automatic segmentation of kidneys in CT images. Manual segmentation is plagued by inter-operator and intra-operator variability, as well as time inefficiency, which would decrease significantly with automatic segmentation. A set of 44 CT images was selected for training. The images were manually segmented using a slice-by-slice contouring tool. Mean 3D medial models were calculated along with their principal geodesic modes of variation from a population of models fit to this data. Using an intrinsic coordinate system, the models were referenced to obtain templates based on the intensity profiles at different points across the boundary. The resulting data was used to automatically segment target CT images and a comparison was made between the manual segmentations made by radiation oncologists and the automatic segmentations obtained based on the geometric and intensity statistics provided. The final models showed marginal segmentation improvement and better internal correspondence.
Acknowledgments

I would like to acknowledge Dr. Ed Chaney for his valuable guidance and encouragement, without whose support this work would not have been possible; Dr. Sarang Joshi and Dr. Julian Rosenman, members of my committee, for their support; Dr. Steve Pizer, Joshua Stough, Gregg Tracton, and Manjari Rao for their willingness to help and patience in answering questions, and all the researchers comprising the Medical Image Display and Analysis Group.
Contents

List of Tables ................................................................. vi
List of Figures ................................................................. vii
List of Abbreviations ......................................................... ix

Chapter 1- Introduction ....................................................... 1
  1.1 Radiation Therapy and Treatment Planning ....................... 1
  1.2 Treatment Planning for External Beam Radiotherapy ........... 2
  1.3 Segmentation in Treatment Planning .............................. 4

Chapter 2- Background ....................................................... 7
  2.1 Shape Representation via M-reps ................................ 7
  2.2 Segmentation using M-reps ...................................... 11
  2.3 Objective Function ................................................. 14
  2.4 Kidney segmentation .............................................. 18

Chapter 3- Materials and Methods ....................................... 20
  3.1 Training and Target Images ...................................... 20
  3.2 Generation of Geometric Statistics ............................. 23
  3.3 Mean model, PGA, Intensity Template Formation. ............. 29
  3.4 Intensity Statistics ................................................ 34
3.4 Evaluation ...................................................... 37

Chapter 4- Results ............................................ 40
  4.1 Comparison of Segmentation Results .................. 45

Chapter 5- Conclusions and Discussion .................... 52
  5.1 Analysis of the Results .................................. 52
  5.2 Future Directions ........................................ 56

References ..................................................... 58
LIST OF TABLES

4.1 Results for Segmentations Based on Current Statistics ................. 42
4.2 Rao's Results ......................................................... 43
4.3 Difference in Volume Overlap Values ............................... 44
# LIST OF FIGURES

1.1 A Virtual Simulation Beams-Eye-View of a Lung Tumor ................. 4

2.1 2D figure represented by Medial Axis and Branch in the Medial Axis ... 8

2.2 Medial Atom and Boundary Implied by a Medial Atom ..................... 9

2.3 M-rep Coordinate System Representation ................................. 10

2.4 Medial Mesh of a Kidney, 3D Surface Rendering of Boundary ............ 10

2.5 Geometric Transformations Available at Model Deformation Stage ...... 13

2.6 Mean Kidney Model and Principal Modes of Variation ................... 14

2.7 Image-Model Error Depiction ............................................. 16

2.8 Regularity Error Depiction ................................................. 17

2.9 Reference Model Error Depiction ......................................... 18

3.1 CT Image with Contrast .................................................... 22

3.2 CT Image of Kidneys with Strong Motion Artifact ......................... 23

3.3 Anastruct Editor Used for Manual Segmentation .......................... 24

3.4 Binary Image of Kidney .................................................... 25

3.5 M-rep Showing Corresponding Points to Landmarks on Image ............ 27

3.6 Landmark Error Depiction ................................................ 28

3.7 Graphical View of Data with Mean Centering and Principal Vectors ... 30

3.8 Left Mean Kidney Model .................................................. 31

3.9 Principal Geodesic Modes for Left Kidney ................................ 31

3.10 Right Mean Kidney Model ............................................... 32

3.11 Principal Geodesic Modes for Right Kidney .............................. 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.12</td>
<td>Goodness of Fit of Principal Modes</td>
<td>33</td>
</tr>
<tr>
<td>3.13</td>
<td>Eleven Points on the Normal to Each Point on the Boundary</td>
<td>34</td>
</tr>
<tr>
<td>3.14</td>
<td>Image Regions for the Different Types of Filters</td>
<td>35</td>
</tr>
<tr>
<td>3.15</td>
<td>Intensity Profile for One Point on Surface of M-rep</td>
<td>35</td>
</tr>
<tr>
<td>3.16</td>
<td>Intensity Profiles for Left and Right Kidneys</td>
<td>36</td>
</tr>
<tr>
<td>3.17</td>
<td>Visualization of Profiles on the Mean Kidney</td>
<td>37</td>
</tr>
<tr>
<td>3.18</td>
<td>Hausdorff Distance</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Case 642R – Current vs. Previous Segmentation</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>Case 634R – Current vs. Previous Segmentation</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>Case 639R – Current vs. Previous Segmentation</td>
<td>48</td>
</tr>
<tr>
<td>4.4</td>
<td>Case 639R – Current vs. Previous M-reps</td>
<td>49</td>
</tr>
<tr>
<td>4.5</td>
<td>Case 642R – Hand Initialization Study</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>Case 633L – Current vs. Previous Mesh</td>
<td>53</td>
</tr>
<tr>
<td>5.2</td>
<td>Case 639R – 3D Representations of Segmentation by Human Rater A</td>
<td>54</td>
</tr>
<tr>
<td>5.3</td>
<td>Case 639R – 3D Solid Representation of Current vs. Previous Models</td>
<td>55</td>
</tr>
</tbody>
</table>
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of Interest</td>
</tr>
<tr>
<td>M-rep</td>
<td>Medial Representation</td>
</tr>
<tr>
<td>3D/2D</td>
<td>Three Dimensional/Two Dimensional</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PGA</td>
<td>Principal Geodesic Analysis</td>
</tr>
</tbody>
</table>
CHAPTER 1 - Introduction

The purpose of this thesis is to evaluate the use of a statistically trainable kidney model for the purpose of automatic segmentation in CT images for radiation treatment planning. This chapter provides an introduction to radiation therapy, treatment planning, and segmentation of anatomical objects.

1.1 Radiation Therapy and Treatment Planning

Tumors are rapidly dividing cells that can invade their surroundings and spread to other organs. During radiation therapy, precisely controlled doses of radiation are delivered to a well-defined volume of tumor inside the cancer patient. One method of radiation therapy involves the insertion of radioactive implants into the tumor, which then delivers a dose for a specific amount of time. Another method of radiation therapy treatment involves the use of an external beam of ionizing radiation that penetrates the patient’s skin and targets the tumor region. This ionizing radiation destroys the ability of these cancerous cells to replicate, destroying their ability to grow.

The main goal of radiation therapy is to uniformly deliver radiation doses throughout the volume of the tumor while minimizing the irradiation of normal tissue. Accomplishing this goal is dependant upon the location of the tumor as well as the
characteristics of the radiation treatment plan. Treatment of these tumors may be complicated by the presence of radiosensitive organs, which include most organs in the human body.

To enable the precise irradiation of the tumor volume while minimizing the harmful doses to surrounding normal tissue, several external ionizing beams are directed towards the patient at different angles. The intersection of these beams is roughly the centroid of the tumor region. The ability to accurately localize this tumor from the surrounding regions is an important element in the success of radiation therapy. Therefore, to maximize the success of radiation therapy, radiation treatment planning procedures are performed prior to delivering treatment.

Radiation treatment planning defines the tumor and normal tissue regions through the use of clinical examinations, operative findings, and imaging techniques such as computerized tomography (CT), magnetic resonance imaging (MRI), and radiography that may or may not contain contrast agents. A good radiation treatment plan takes into account patient positioning, beam arrangement and dose requirements, shielding, and localization of tumor and vital anatomical objects.

1.2 Treatment planning for External Beam Radiotherapy

During the first part of conventional treatment planning for external beam radiotherapy, the patient is placed on an x-ray machine or a CT scanner and scans
are taken of the body. The radiation oncologist views these images and identifies the tumor volume and the areas that require treatment. Along with the radiation oncologist, the dosimetrist determines the optimal angles and number of ionizing radiation beams necessary to maximize tumor destruction and minimize the effects on surrounding tissue. From this calculation, marks are placed on the patient’s body, which are used to guide daily treatment. After this simulation, the second step of treatment planning begins. Special computers are used to determine the amount of energy that will be directed towards the patient’s body, and the treatment time for each beam.

The modern technique of virtual simulation combines the initial simulation and the treatment planning into a single computer-based process. 3D scans of the patient are taken and viewed graphically via a specialized computer program. This allows the radiation oncologist to view the patient and the tumor from many different angles. Virtual simulation enhances the ability of the radiation oncologist to design the optimal shape and angle of entry for the radiation beams, as well as calculate dose distribution without the physical presence of the patient. Convenience and the ability to view the patient graphically in 3 dimensions have made virtual simulation the method of choice in external beam treatment planning. For further discussion of virtual simulation, please see [Sherouse and Chaney 1991].
1.3 Segmentation in Treatment Planning

The use of computer models of the patient in radiation therapy has allowed radiation oncologists to deliver treatment to patients more accurately. One of the critical steps in this computerized treatment planning process is the segmentation of tumors and anatomical structures on each CT/MRI slice.

The goal of segmentation is to identify crucial structures in an image. By extracting this information from the image, radiation physicists can more clearly define the tumor area and identify radiosensitive organs which may be harmed by the application of radiation.
Typical planning systems depict tumors and normal tissue in 3D space for each patient. Segmentation of this image data is a difficult task in the field of image processing and human ability is currently superior in this area compared to the ability of the computer. However, *manual segmentation*, the segmentation of anatomical structures by a human, has its own limitations. Although it is the primary method used in treatment planning, the accuracy of segmentation depends on the operator’s experience and working conditions during the segmentation procedure. Segmentation is also prone to inter-operator and intra-operator variability. In addition, the identification of three-dimensional objects on two-dimensional CT slices is difficult and prone to error. Finally, anatomical organs continuously change shape due to breathing as well as bladder and bowel changes. For further discussion on the problems of manual segmentation, please refer to [Chaney et al 2005].

Applying automatic segmentation techniques to the process of radiation treatment planning could save doctors several hours for each plan and significantly reduce the inter-operator and intra-operator variability. In addition, because of the significant time savings it could achieve, automatic segmentation could enhance the ability of radiation oncologists to adjust their treatment on a more regular basis by adapting treatment planning before each dose is delivered. Because organs and tumors are not static objects, the ability to alter the treatment plans in real time is an exciting prospect that could significantly enhance the success of radiation therapy. By attempting to automatically segment kidneys in CT images to closely resemble manual segmentations, this study is a step in the direction towards full body automatic segmentation.
CHAPTER 2 – Background

This study uses a particular class of medial models, called m-reps, that are effective in describing solid objects and thus can be used in segmentation of anatomical structures. The specifics of how m-reps work and how they can be used in kidney segmentation are described in this chapter.

2.1 Shape representation via M-reps

This study is based largely around the use of “M-reps” [Pizer 1999 and Joshi 2001] to portray shape, specifically anatomical structures, in three-dimensions. M-reps are medially based solid models that are a class of medial models [Blum 1962, 1967, 1973] which effectively capture geometric information in 3-dimensional space.

M-reps are based around the premise of describing shape as a combination of several spheres that are tangent to the boundary of the object at exactly two distinct points. Each of these spheres can then in turn be described by the coordinates of its center point and its corresponding radius. To simplify the representation of the object, we can think of it as having a ‘medial axis,’ which is the connection of the centers of all the bitangent spheres. A key element in shape representation using
m-reps is the ability to recreate the object knowing only the medial axis and the corresponding radius for each sphere on that axis.

![Diagram](image_url)

**Fig. 2.1.** A 2D figure represented by (a) boundary (b) bitangent circles in the figure (c) medial axis (d) protrusion represented by a branch in the medial axis.

The medial axis in two-dimensional space is a curve, while in three-dimensional space it is a sheet. Because this study was based in three-dimensional space, the medial axis of all the objects were represented by sheets. The sheet is represented by an m-rep as a grid of atoms with two associated spokes. Each atom is the center point of a bitangent sphere while the spokes are the radii to the surface of the object.

Each medial atom of an m-rep contains information regarding the coordinates of the medial sheet at a sample point $x$, the distance from $x$ to the object boundary or
radius $r$, the two vectors of length $r$, representing the spokes which point to the points of tangency of the inscribed sphere, the angle $\Theta$ formed between each of the spokes and the angular bisector vector $\mathbf{b}$, and a frame $F = (\mathbf{n}, \mathbf{b}, \mathbf{b}^\perp)$ that defines the tangent plane of the medial sheet at $\mathbf{x}$. Another parameter $\eta$, elongation, is defined by atoms which lie on the outer edges of a mesh and specifies the curvature of the object about that point (Fig.2.2) [Rao 2003].

M-rep models are represented by an object based coordinate system defined by $(u,v,t,\tau)$ (Fig.2.3), where $u$ and $v$ represent the row and column corresponding to the position of the medial atom, $t \in [-1,1]$ indicates which side of the locus the point lies on, i.e., $t = -1$ or $+1$ for internal medial points and runs continuously around the crest region from $-1$ through $0$ at the boundary through $+1$, and $\tau$ measures the distance along the spokes from the boundary with $\tau > 0$ outside the boundary, $\tau < 0$ inside the boundary and $\tau = -1$ at the medial locus. This object-intrinsic coordinate system
provides spatial and orientational correspondence between an m-rep in two different states of deformation [Rao 2003].

Fig. 2.3. M-rep coordinate system represented by \((u,v,t,\tau)\). The figures illustrate the variation of these coordinates along the medial axis.

Through m-reps the surface of the object can be approximated by sampling the medial axis and determining the 'implied boundary,' which is defined as the surface through the tips of all the spokes on the model.

Fig. 2.4. (a) Example of a mesh of medial atoms for a kidney (b) Wire frame surface rendering of the boundary implied by the medial mesh (c) 3D surface rendering of the boundary implied by the medial mesh.
2.2 Segmentation using M-reps

One of the main advantages of m-reps is that they allow classification of object deformations into along-object deviations, namely elongations and bendings, and across-object deviations, namely bulgings and attachment of protrusions or indentations (multifigure m-reps). An additional advantage is that distances can be expressed as a fraction of medial width. These properties allow positions and orientations to be followed through deformations of elongation, widening, or bending [Pizer et al, IJCV 2001].

The process of m-rep segmentation operates from large to small-scale levels, at each level *deforming* the m-rep model by optimizing in a Bayesian framework, an objective function over a set of geometric transformations available at that scale level [Rao 2003]. The objective function is the sum of four terms:

Objective Function = aI + bL + gR + dM

I = Image Match
L = Landmark Match
R = Regularity Match
M = Model Match

Where a, b, g, d are scaling values that can be altered to give different weights to each term. The landmark match term is only used during the training process for initialization and to gather statistical information, and will be discussed in the next chapter.
Initialization of the kidney model is performed manually by positioning the model near the target object in the image data. From there, automatic segmentation commences in a two step process, using the objective function as the deformation heuristic.

1. Deformation of the entire kidney model via a similarity transform and principal modes of variation. The similarity transform involves translation, rotation, and scaling of the entire model, while PGA produces shape changes based on the principal modes determined by the statistical analysis of training shapes.

2. Deformation of each individual atom recursively until the objective function is optimized or until each atom has been deformed a specific number of times. This process involves translation, rotation, and scaling of each individual atom and its associated spokes.

Principal geodesic analysis is a way of analyzing shape variability that is similar to principal components analysis (PCA) in Euclidean space [Styner 2001]. PCA however, is only applicable when the parameters of the model are in Euclidean space, which is not the case with medial models. Medial model parameters have been shown to be part of a symmetric space [Fletcher et al 2003], however, and so shape variation can be described as principal components of a high dimensional geodesic space. These principal geodesic modes are found by using a training set to define the dominant shape variations along orthogonal axes in a set of kidney models. The initial steps in m-rep segmentation use principal modes from PGA.
along with the similarity transform to deform the entire model in a way that optimizes the objective function.

Fig.2.5. Set of geometric transformations available at the first stage of model deformation include translation, rotation, and scaling.
Fig. 2.6. (a) The mean kidney model (b) The first three principal modes of deformation of the kidney m-rep. Each row displays the models corresponding to 
\[
\{-3\sqrt{\lambda_i}, -1.5\sqrt{\lambda_i}, 0, +1.5\sqrt{\lambda_i}, +3\sqrt{\lambda_i}\},
\]
along the \(i^{th}\) principal component.

The next step in m-rep segmentation is the deformation of each individual atom. This deformation can involve translation, rotation, and scaling of each individual atom and elongation, angle extension/reduction, and rotation of its corresponding spokes to optimize the objective function. Each atom is optimized until convergence or until a specific number of optimizations has been performed.

### 2.3 Objective Function

The heuristic used to optimize the model deformation is the objective function (from p.12):

Objective Function = \(aI + bL + gR + dM\)

\(I = \) Image-Model Match
L = Landmark Match
R = Regularity Match
M = Model Match

Except for the landmark match term, which is only used during training, each of these will be described below.

Image-Model Match

During the training process, a Danielsson Distance Map is calculated for each image. Six connected neighbors are used to find the object’s boundary voxels. This allows quick determination of how far the model’s implied boundary is from the image data’s boundary. The distance between a point on the model’s boundary to the closest point on the image boundary is squared and the sums of these measurements are calculated. During the segmentation process, image match is measured by normalized correlation of the target image intensities with template intensities. An intensity offset also is applied for each intensity profile [Stough 2005]. This gives a value for the image-model match.
Fig. 2.7. A close up view of the error region between the image boundary and the implied model boundary.

**Regularity Match**

The regularity term penalizes deformation that causes the atoms to become irregularly spaced in the medial grid. The term computes the predicted placement of the atom as the average of its nearest neighbors and calculates the difference between this predicted point and the actual position of the atom in question. A correction is applied that relocates atoms closer to their predicted position.

Each atom on the grid has four neighbors directly connected to it, with the exception of edge atoms which have three and corner atoms which have two. For the purposes of this study, the regularity penalty does not take into account corner atoms. For the other atoms, the nearest neighbors average their positions in symmetric space to determine the optimal position for that atom on a perfectly
aligned grid. The differences between this position and the actual position are summed together for each atom and the result is the regularity penalty.

![Diagram of aligned grid with theoretical location and actual location]

Fig.2.8. The distance between the actual position of an atom and its theoretical location as defined by its nearest neighbors is the regularity error.

**Model Match**

The model match term defines how well a deformed model matches a reference model. This is useful when close correspondence with a reference model is a key point, as for segmenting the same structure in sequential images of the same patient. The model match measures how much work was done during the atom stage by summing the atom for atom symmetric space distance across the two models.
2.4 Kidney Segmentation

Automatic kidney segmentation has many challenges including the need to obtain CT images of good quality with little motion artifact. Another challenge is that the crowded tissue environment around the kidney often has similar intensities, making it difficult to distinguish between the kidney and other organs. This is particularly true of the adrenal gland, which is often in close contact with the kidney.

This particular study focuses on the kidney for several reasons. First, the kidneys have a high risk of radiation exposure during treatment for patients with pelvic and abdominal cancer. For this reason, kidney segmentation is common for patients suffering from these types of tumors. Secondly, the kidney has a well-defined shape that can be easily modeled with m-reps. The kidney shape is simple and can
effectively be represented using a single figure m-rep model. Thirdly, the kidney is located in a consistent, easily identifiable region. Locating the kidney is not difficult due to the unique intensity representations of surrounding organs in CT scans such as the liver and spinal cord.

The aim of the present study is to extend upon Rao's work on intensity templates that showed the effectiveness of having a locally varying template as opposed to using a Gaussian derivate based template [Rao 2003]. This study uses many of the same techniques applied by Rao to develop a mean kidney model with principal modes and intensity templates to aid in automatic segmentation of the kidney. The goal, therefore, is to evaluate the effectiveness of geometric and intensity statistics that were developed through the use of statistical training from a large population of known kidneys.
CHAPTER 3 – Materials and Methods

This chapter describes the methods that were used to develop the mean model and its corresponding principal geodesic modes and intensity template. The first part of the chapter describes the CT images that were used in training the model. The next section goes into the details of how the mean model and principal geodesic modes were determined, and the final part of the chapter describes the testing of the model in specific target images.

3.1 Training and Target Images

This study required a set of images for training to provide statistical information of the shape of the kidney as well as its intensity statistics, and a set of testing images to provide an evaluation of how well kidney segmentation was performed using the geometric and intensity statistics as the kidney model. All the images were obtained from the Department of Radiation Oncology, from the University of North Carolina at Chapel Hill. Most of the images had a raster resolution of 512x512 pixels per slice, while the pixel size ranged from 0.098mm X 0.098mm to 0.156 X 0.156 for each image.

Careful selection of the training images had to be applied, as these images would provide the statistical basis for the shape and intensity information of the kidney. A
database of over 700 images was examined, with CT scans of various sections of the body. Each of these images had to pass a selection criterion to be an eligible part of this study. The images chosen for training purposes in this study are the same images used in Rao’s work [Rao 2003]. Each training image had these characteristics:

1. The presence of a full kidney
2. The patient was lying on their back (supine position) with the head towards the gantry
3. No contrast agent was used
4. No tumor, disease, or kidney stones existed in any image
5. Little or no motion artifact
6. At least a small margin after the edges of the kidney
7. Slice thickness was less than 5mm per slice

Because the information in these images would be used to compute geometric and intensity statistics, the importance of consistency between the images was stressed in determining the characteristics of images to be used in the study. For instance, if an image did not have an entire kidney in its CT scan, it would be difficult to understand the shape of the whole kidney in that particular patient. For this reason, each image in this study had a full kidney present.
Fig. 3.1 shows a CT image with contrast. The contrast artificially enhances the intensity of the kidney compared to its surrounding areas. In understanding the intensity changes along the boundary of kidney, it is important that the intensity of all the organs are neither enhanced nor diminished in any way. For this reason, CT images that used contrast were not used.

The right and left kidneys were treated as completely separate anatomical objects and geometric statistics were developed for each of them. During training, 43 left kidneys and 44 right kidneys were used. Some of the training images contained a left kidney but not a right kidney and vice versa. The result of using CT scans with only one kidney is that unequal numbers were used during training for the left and right kidney.
Figure 3.2 shows a CT image with high motion artifact. Motion artifact makes it difficult to ascertain the boundaries of the anatomical structures in the image. For this reason, images with high motion artifact can not be used for statistically training of the kidney shape.

3.2 Generation of Geometric Statistics

The process computing geometric statistics with its associated involves the following steps:

1. Manual segmentation of kidneys from the training images
2. Conversion of manual segmentations to binary images
3. Deformation of m-rep model into each binary image
4. Production of mean model with its associated principal geodesic modes
The first step in the process of geometric statistical training is the hand segmentation of images. As was discussed in Chapter 1, prior to treatment planning, CT scans are taken of the patient, and these images are segmented manually by the radiation oncologist. This procedure is performed on software that utilizes a point and click operation to outline the boundary of the anatomical object in question for each cross sectional slice.

Fig.3.3. (Software: Anastruct_editor - part of the PLanUNC suite of RadioTherapy Tools, developed at UNC Radiation Oncology (1986 by Sherouse, Chaney, Cullip, GTracton, Julian Rosenman, et al)

Fig 3.3 shows the manual segmentation of the right and left kidneys in one cross sectional slice of a CT scan. To segment the entire kidney, each slice must be viewed and the boundary outlined in similar fashion. As described in Chapter 1, hand segmentation is a tedious process, subject to inter-operator and intra-operator
variability. In spite of this, the manual segmentations are the best estimate as to the actual location and shape of the kidney, and therefore used for statistical geometric training.

The output of the hand segmentations is a set of two-dimensional contours which represent the kidney. This *contour stack* becomes the basis for the binary image, which is an image reduced to two intensity values: 1 (representing a voxel inside the contour) and 0 (representing a voxel outside the contour). The binary image, therefore, is a computer representation of the hand segmented contours.

Based on the generally known shape of the kidney, a single figure medial mesh of 15 atoms (3 columns and 5 rows) was fit into each of the training binary images via the automatic segmentation method described in Chapter 2, to generate an m-rep segmentation for each kidney in the training data set [Dam 2003]. For training, a
Danielsson Distance Map is used for measuring image match instead of normalized correlation of intensities. The starting model used here was a mean model produced in a similar fashion to what is being discussed in this section. A model was fit to each image and geometric statistics developed from these deformed m-reps. This new model then became the starting model for the next round of optimizations in the binary images. The process was iterated twice until the mean model converged in such a way that it was not being significantly altered from the previous mean model.

Fitting the initial m-rep into the binary images is done by automatic segmentation, with two distinct differences from what was previously discussed in Chapter 2. The first difference is the initial model fitting stage. Instead of manually placing the model into the general vicinity of the kidney, the computer calculates the centroid and volume of the starting model and image and aligns them. This allows for pure automatic segmentation of binary images without any manual intervention. The second difference is the use of landmarks in the objective function. Landmarks are placed on training images to create correspondence with anatomically important reference points.
Fig. 3.5. An m-rep showing the points on the model that correspond to the landmarks on the image.

Landmarks were placed on three points on each training image. They are:

1. Top of the Kidney
2. Bottom of the Kidney
3. Pelvis of the Kidney

Each landmark has a central point and a tolerance, defined by its radius. The error used for the landmark penalty is proportional to the square of the distance between the central point of the sphere to the corresponding point on the model.
Fig.3.6. Close up view of the distance between the landmark centroid and the corresponding model point. Because the corresponding model point is inside the sphere of tolerance, the landmark error for this case was less than 1.

Once the initial m-rep was fit to each binary image, the next step was to develop a mean model with its corresponding principal modes of variation.
3.3 Mean model, PGA, Intensity Template formation

In order to understand the shape variation of the kidneys, an extension to the traditional means of principal components analysis (PCA) must be applied. PCA is a useful method when describing variability in Euclidean space, but results in unacceptable shapes when applied to curved symmetric spaces as for m-rep models. The same approach, however, can be used through principal geodesic analysis (PGA) which can be applied to curved spaces. A short description follows.

When analyzing information using PCA, data first needs to be mean centralized. Once this has been completed, the largest variations from the mean need to be calculated. In Fig. 3.7, the axis for the most significant variation from the mean is noted by the ‘1.’ This indicates that this direction describes the first principal component in variability. The axis for the second principal mode of variation is noted by the ‘2.’ Notice how more of the data is described by the first mode than the second. Also notice that the two axes of variation are orthogonal to each other.
The variation in the data can now be represented using a mean value and its principal eigenvectors (represented by the arrows in the figure above). This method of describing data is effective in linear space. When describing data in curved symmetric space, PGA needs to be used. This allows a non-linear, multi-dimensional space to be described by a mean and \( n \) number of orthogonal eigenvectors, where \( n \) equals the number of dimensions minus one. For further discussion on Principal Geodesic Analysis please refer to [Fletcher et all 2003].

After fitting the initial model to all the binary images, the procedure of principal geodesic analysis was used to extract a mean and its principal modes for a manifold consisting of a 3x5 matrix of atoms. The results of this shape analysis are illustrated in Fig. 3.8-3.11.
Fig. 3.8. Left Mean Kidney Model in different views. (a) atoms & spokes (b) wire frame, and (c) solid.

Fig. 3.9. Changes in shape of the first and second principal eigenvectors for the left kidney. -3, 0, and 3 standard deviations pictured here. ROI is indicated.
Fig. 3.10. Right Mean Kidney Model in different views. (a) atoms & spokes (b) wire frame, and (c) solid

Fig. 3.11. Changes in shape of the first and second principal eigenvectors for the right kidney. -3, 0, and 3 standard deviations pictured here. ROI is indicated.
In our data set, it took approximately 40 principal modes to describe almost all of the shape variability in the training kidneys. To simplify the process while still maintaining significant description of data variation, this study implemented only the first five eigenvectors, which described 65-75% of the shape variability at the global stage.

How well does it describe data?

First mode -37%
First 5 modes -75%

First mode -30%
First 5 modes -65%

Fig.3.12. Graph depicting the amount of data described by each eigenmode.
Intensity Statistics

Statistically characterizing the intensity characteristics at each point along the implied boundary enhances the ability to find the edge of the kidney in CT images. Rao showed in her work that a locally varying intensity template was more effective than using a Gaussian derivate based template for automatic segmentation [Rao 2003]. The next step in the process was to develop an intensity template for the left and right kidney models discussed above.

After fitting the initial model to all the binary images, the fit models were then registered with the corresponding grayscale CT images. Along each of the 2562 points on the m-rep, the normal to the boundary was sampled at 11 points, five on the interior of the m-rep, five on the exterior. Intensity information was collected for each point along the normal, each of which were separated by approximately 11 pixels.

Fig.3.13. Normals drawn at points on the surface of the m-rep sampled at 11 points with the 6th sample being the point on the surface of the m-rep
The intensity profile along each normal was examined from point 1 to point 11 and placed in one of three different categories: light-to-dark, dark-to-light, and notch.

Fig. 3.14. Image region corresponding to light-to-dark filter, light-to-dark, and notch. ROI indicated.

Fig. 3.15. Intensity profile along the normal for one point on the surface of the m-rep across all training images. Although there is a discrepancy on some of the kidneys, the trend is towards a higher intensity on point 1 of the normal moving toward a lower intensity on point 11 on the normal. This
particular point on the m-rep surface then is likely most often indicated by a light-to-dark intensity change. Each of the 2562 point was examined the same way. For a more detailed explanation of intensity templates, please refer to [Rao 2003].

(a)       (b)

Fig.3.16. Intensity Profile for (a) left kidney and (b) right kidney

Notice in Fig 3.16 above that the most common profile seen on the kidney is the light-to-dark filter. This is not surprising by visually inspection of CT scans, as the majority of points on the boundary typically show a higher intensity for the kidney than the surrounding areas. The least common profile seen is the dark-to-light profile, which is completely absent in the left kidney. Again, this is not surprising, as the only time the dark-to-light pattern is expected is when the liver abuts the kidney.
Fig. 3.17. Images depicting the intensity profiles seen on the boundary of the (a) left kidney and (b) right kidney.

3.4 Evaluation

Each of the target images used in this study was segmented manually by two experts as well as by the computer. To evaluate the accuracy of the automatic segmentations, they were compared to the manual segmentations using a software package called Valmet [Gerig 2001, Stough 2002], which involved three separate comparison heuristics:

1. Volume overlap
2. Hausdorff distance
3. Mean surface distance
(1) Volume overlap

The volume overlap was calculated as being the intersection of the two volumes divided by the union of the two volumes. The scale for this comparison ranged from 0 to 1, with 1 indicating identical volumetric areas and 0 indicating that no intersection between the two regions existed. The surface of the manual segmentations was computed using the Marching Cubes [Lorensen and Cline 1987] method.

(2) Hausdorff distance

The Hausdorff distance defines the largest minimal distance between the two surfaces in question. In Fig. 3.18, two contours, Cuve1 and Curve 2 are given. The minimal distance from every point on Curve 1 to any point on Curve 2 is calculated. The maximum of all these minimal distances is known as the ‘worst case’ or Hausdorff distance. This distance is not necessarily symmetric since the minimal distance from the first object to the second object is not necessarily the minimal distance from the second object to the first. This is accounted for by finally calculating,

\[ \text{Hauss\_Dist} (A,B) = \max \{ \text{dist} (A,B), \text{dist}(B,A) \} \]
Fig. 3.18. The closest distance (a) from the first curve to the second is not the same as the closest distance (b) from the second curve to the first.

(3) Mean surface distance

The mean absolute surface distance describes how much on average the two surfaces differ. Point to point correspondence on the surfaces is required for this measure. The mean surface separation for a given kidney is defined in terms of closest points, i.e.,

\[
\frac{1}{2} \left[ \frac{1}{N_1} \left( \sum_{i=1}^{N_1} \min_{j=1, \ldots, N_2} |y_i^1 - y_j^2| \right) + \frac{1}{N_2} \left( \sum_{i=1}^{N_2} \min_{j=1, \ldots, N_1} |y_i^1 - y_j^2| \right) \right]
\]

where \( N_1 \) and \( N_2 \) are the respective numbers of boundary voxels in the two kidneys being compared, and \( y_i^1 \) and \( y_j^2 \) are the coordinates of the boundary voxel centers of the respective kidneys. This calculation is also not symmetric. A common average is therefore derived by combining the two averages [Rao 2003].
CHAPTER 4 – Results

This chapter summarizes the results of the automatic segmentations using the geometric and intensity statistics developed in this study. To evaluate the effectiveness of these models, 12 CT images were used as targets, each having properties similar to those described in Chapter 3 for the training images. These were first manually segmented by two human raters (A and B) and then compared to the automatic segmentation performed by the computer (C).

Manual segmentations by experts A and B were performed using a slice-by-slice contouring tool known as MASK [Tracton 1994]. Each of the target images contained both a left and right kidney, for a total of 24 segmentations performed by raters A, B, and C. The automatic segmentations (C) for each kidney were then compared to the two manual segmentations (A and B) using the evaluation metrics discussed in section 3.4. The tables below show the results of those comparisons.

The results in Table 4.1 were obtained from segmentations performed in this study using the mean kidney models as the initial models for automatic segmentations. The metrics described in Chapter 3 are shown in the table and used to evaluate the effectiveness of the computer segmentations versus the human raters A and B. Mean and standard deviations across all target kidneys are calculated for each metric.
The results in Table 4.2 were generated by using the models obtained during Rao's work in 2003. Since that time, the evaluation program has changed and the output is inconsistent with that which was generated previously. For this reason, Rao's numbers were re-calculated using the current evaluation program and the models used in her study. The output is tabulated in Table 4.2.

Table 4.3 is a comparison between the volume overlaps obtained in this study for each case versus the volume overlap obtained during Rao's study. A negative number indicates that the volume overlap using Rao's model is higher than that using the model obtained in this study.
Table 4.1. Results for Segments Based on Current Geometric Statistics

<table>
<thead>
<tr>
<th>Kidney code</th>
<th>Average surface distance (A to C)</th>
<th>Average surface distance (B to C)</th>
<th>Hausdorff distance (A to C)</th>
<th>Hausdorff distance (B to C)</th>
<th>Volume Overlap (A to C)</th>
<th>Volume Overlap (B to C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>630L</td>
<td>0.254</td>
<td>0.253</td>
<td>1.363</td>
<td>1.329</td>
<td>0.689</td>
<td>0.694</td>
</tr>
<tr>
<td>630R</td>
<td>0.249</td>
<td>0.233</td>
<td>1.259</td>
<td>1.234</td>
<td>0.712</td>
<td>0.726</td>
</tr>
<tr>
<td>633L</td>
<td>0.252</td>
<td>0.271</td>
<td>1.106</td>
<td>1.196</td>
<td>0.753</td>
<td>0.729</td>
</tr>
<tr>
<td>633R</td>
<td>0.253</td>
<td>0.298</td>
<td>1.24</td>
<td>1.457</td>
<td>0.753</td>
<td>0.717</td>
</tr>
<tr>
<td>634L</td>
<td>0.289</td>
<td>0.296</td>
<td>0.985</td>
<td>1.088</td>
<td>0.719</td>
<td>0.711</td>
</tr>
<tr>
<td>634R</td>
<td>0.347</td>
<td>0.299</td>
<td>1.364</td>
<td>1.942</td>
<td>0.671</td>
<td>0.712</td>
</tr>
<tr>
<td>635L</td>
<td>0.226</td>
<td>0.259</td>
<td>1.014</td>
<td>0.911</td>
<td>0.744</td>
<td>0.724</td>
</tr>
<tr>
<td>635R</td>
<td>0.205</td>
<td>0.208</td>
<td>0.885</td>
<td>0.899</td>
<td>0.782</td>
<td>0.771</td>
</tr>
<tr>
<td>636L</td>
<td>0.198</td>
<td>0.202</td>
<td>1.729</td>
<td>1.1</td>
<td>0.808</td>
<td>0.816</td>
</tr>
<tr>
<td>636R</td>
<td>0.25</td>
<td>0.267</td>
<td>1.193</td>
<td>1.008</td>
<td>0.733</td>
<td>0.732</td>
</tr>
<tr>
<td>637L</td>
<td>0.251</td>
<td>0.221</td>
<td>1.312</td>
<td>0.687</td>
<td>0.698</td>
<td>0.766</td>
</tr>
<tr>
<td>637R</td>
<td>0.26</td>
<td>0.251</td>
<td>1.181</td>
<td>1.035</td>
<td>0.752</td>
<td>0.766</td>
</tr>
<tr>
<td>638L</td>
<td>0.218</td>
<td>0.261</td>
<td>1.266</td>
<td>1.145</td>
<td>0.768</td>
<td>0.745</td>
</tr>
<tr>
<td>638R</td>
<td>0.213</td>
<td>0.213</td>
<td>1.357</td>
<td>1.164</td>
<td>0.775</td>
<td>0.774</td>
</tr>
<tr>
<td>639L</td>
<td>0.246</td>
<td>0.271</td>
<td>1.894</td>
<td>1.254</td>
<td>0.756</td>
<td>0.769</td>
</tr>
<tr>
<td>639R</td>
<td>0.191</td>
<td>0.209</td>
<td>1.24</td>
<td>0.88</td>
<td>0.829</td>
<td>0.815</td>
</tr>
<tr>
<td>640L</td>
<td>0.236</td>
<td>0.278</td>
<td>1.546</td>
<td>0.899</td>
<td>0.772</td>
<td>0.747</td>
</tr>
<tr>
<td>640R</td>
<td>0.176</td>
<td>0.197</td>
<td>1.585</td>
<td>0.676</td>
<td>0.825</td>
<td>0.804</td>
</tr>
<tr>
<td>642L</td>
<td>0.231</td>
<td>0.305</td>
<td>1.655</td>
<td>0.921</td>
<td>0.738</td>
<td>0.699</td>
</tr>
<tr>
<td>642R</td>
<td>0.194</td>
<td>0.225</td>
<td>1.01</td>
<td>1.131</td>
<td>0.798</td>
<td>0.769</td>
</tr>
<tr>
<td>646L</td>
<td>0.191</td>
<td>0.254</td>
<td>1.221</td>
<td>1.206</td>
<td>0.784</td>
<td>0.718</td>
</tr>
<tr>
<td>646R</td>
<td>0.201</td>
<td>0.243</td>
<td>1.851</td>
<td>1.976</td>
<td>0.783</td>
<td>0.739</td>
</tr>
<tr>
<td>648L</td>
<td>0.261</td>
<td>0.301</td>
<td>1.778</td>
<td>1.36</td>
<td>0.732</td>
<td>0.715</td>
</tr>
<tr>
<td>648R</td>
<td>0.201</td>
<td>0.173</td>
<td>1.606</td>
<td>1.021</td>
<td>0.817</td>
<td>0.834</td>
</tr>
<tr>
<td>Mean</td>
<td>0.235</td>
<td>0.248</td>
<td>1.382</td>
<td>1.154</td>
<td>0.757</td>
<td>0.752</td>
</tr>
<tr>
<td>SD</td>
<td>0.038</td>
<td>0.039</td>
<td>0.286</td>
<td>0.327</td>
<td>0.043</td>
<td>0.044</td>
</tr>
</tbody>
</table>
Table 4.2. Re-computed Results for Rao’s Study [2003]

<table>
<thead>
<tr>
<th>Kidney code</th>
<th>Average surface distance (A to C)</th>
<th>Average surface distance (B to C)</th>
<th>Hausdorff distance (A to C)</th>
<th>Hausdorff distance (B to C)</th>
<th>Volume Overlap (A to C)</th>
<th>Volume Overlap (B to C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>630L</td>
<td>0.29</td>
<td>0.307</td>
<td>1.431</td>
<td>1.488</td>
<td>0.639</td>
<td>0.642</td>
</tr>
<tr>
<td>630R</td>
<td>0.253</td>
<td>0.23</td>
<td>1.04</td>
<td>1.04</td>
<td>0.694</td>
<td>0.705</td>
</tr>
<tr>
<td>633L</td>
<td>0.229</td>
<td>0.263</td>
<td>1.167</td>
<td>1.658</td>
<td>0.748</td>
<td>0.713</td>
</tr>
<tr>
<td>633R</td>
<td>0.261</td>
<td>0.315</td>
<td>0.998</td>
<td>1.224</td>
<td>0.739</td>
<td>0.708</td>
</tr>
<tr>
<td>634L</td>
<td>0.259</td>
<td>0.299</td>
<td>1.286</td>
<td>1.369</td>
<td>0.696</td>
<td>0.673</td>
</tr>
<tr>
<td>634R</td>
<td>0.387</td>
<td>0.36</td>
<td>1.604</td>
<td>1.725</td>
<td>0.608</td>
<td>0.642</td>
</tr>
<tr>
<td>635L</td>
<td>0.201</td>
<td>0.236</td>
<td>0.749</td>
<td>0.911</td>
<td>0.780</td>
<td>0.754</td>
</tr>
<tr>
<td>635R</td>
<td>0.219</td>
<td>0.268</td>
<td>1.149</td>
<td>1.197</td>
<td>0.775</td>
<td>0.74</td>
</tr>
<tr>
<td>636L</td>
<td>0.182</td>
<td>0.161</td>
<td>1.356</td>
<td>0.786</td>
<td>0.794</td>
<td>0.825</td>
</tr>
<tr>
<td>636R</td>
<td>0.251</td>
<td>0.253</td>
<td>1.714</td>
<td>1.404</td>
<td>0.718</td>
<td>0.731</td>
</tr>
<tr>
<td>637L</td>
<td>0.186</td>
<td>0.147</td>
<td>1.42</td>
<td>0.81</td>
<td>0.771</td>
<td>0.839</td>
</tr>
<tr>
<td>637R</td>
<td>0.213</td>
<td>0.207</td>
<td>0.999</td>
<td>0.984</td>
<td>0.731</td>
<td>0.735</td>
</tr>
<tr>
<td>638L</td>
<td>0.193</td>
<td>0.23</td>
<td>1.297</td>
<td>1.134</td>
<td>0.776</td>
<td>0.747</td>
</tr>
<tr>
<td>638R</td>
<td>0.351</td>
<td>0.374</td>
<td>1.764</td>
<td>1.764</td>
<td>0.717</td>
<td>0.706</td>
</tr>
<tr>
<td>639L</td>
<td>0.205</td>
<td>0.203</td>
<td>2.046</td>
<td>1.489</td>
<td>0.819</td>
<td>0.829</td>
</tr>
<tr>
<td>639R</td>
<td>0.133</td>
<td>0.153</td>
<td>1.015</td>
<td>0.846</td>
<td>0.87</td>
<td>0.846</td>
</tr>
<tr>
<td>640L</td>
<td>0.179</td>
<td>0.211</td>
<td>1.893</td>
<td>2.018</td>
<td>0.767</td>
<td>0.745</td>
</tr>
<tr>
<td>640R</td>
<td>0.1</td>
<td>0.16</td>
<td>1.274</td>
<td>0.961</td>
<td>0.885</td>
<td>0.839</td>
</tr>
<tr>
<td>642L</td>
<td>0.227</td>
<td>0.396</td>
<td>1.521</td>
<td>1.588</td>
<td>0.718</td>
<td>0.617</td>
</tr>
<tr>
<td>642R</td>
<td>0.25</td>
<td>0.33</td>
<td>1.01</td>
<td>1.375</td>
<td>0.719</td>
<td>0.671</td>
</tr>
<tr>
<td>646L</td>
<td>0.188</td>
<td>0.219</td>
<td>1.465</td>
<td>1.459</td>
<td>0.783</td>
<td>0.741</td>
</tr>
<tr>
<td>646R</td>
<td>0.183</td>
<td>0.224</td>
<td>1.645</td>
<td>1.797</td>
<td>0.805</td>
<td>0.766</td>
</tr>
<tr>
<td>648L</td>
<td>0.206</td>
<td>0.291</td>
<td>1.413</td>
<td>1.602</td>
<td>0.782</td>
<td>0.727</td>
</tr>
<tr>
<td>648R</td>
<td>0.149</td>
<td>0.309</td>
<td>1.174</td>
<td>1.157</td>
<td>0.801</td>
<td>0.69</td>
</tr>
<tr>
<td>Mean</td>
<td>0.223</td>
<td>0.258</td>
<td>1.373</td>
<td>1.337</td>
<td>0.754</td>
<td>0.733</td>
</tr>
<tr>
<td>SD</td>
<td>0.066</td>
<td>0.073</td>
<td>0.305</td>
<td>0.456</td>
<td>0.065</td>
<td>0.067</td>
</tr>
</tbody>
</table>
### Table 4.3. Difference in Volume Overlap Values

<table>
<thead>
<tr>
<th>Kidney Code</th>
<th>Difference in Volume Overlap (A to C)</th>
<th>Difference in Volume Overlap (B to C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>630L</td>
<td>0.05</td>
<td>0.052</td>
</tr>
<tr>
<td>630R</td>
<td>0.018</td>
<td>0.021</td>
</tr>
<tr>
<td>633L</td>
<td>0.005</td>
<td>0.016</td>
</tr>
<tr>
<td>633R</td>
<td>0.014</td>
<td>0.009</td>
</tr>
<tr>
<td>634L</td>
<td>0.023</td>
<td>0.038</td>
</tr>
<tr>
<td>634R</td>
<td>0.063</td>
<td>0.07</td>
</tr>
<tr>
<td>635L</td>
<td>-0.036</td>
<td>-0.03</td>
</tr>
<tr>
<td>635R</td>
<td>0.007</td>
<td>0.031</td>
</tr>
<tr>
<td>636L</td>
<td>0.014</td>
<td>-0.009</td>
</tr>
<tr>
<td>636R</td>
<td>0.015</td>
<td>0.001</td>
</tr>
<tr>
<td>637L</td>
<td>-0.073</td>
<td>-0.073</td>
</tr>
<tr>
<td>637R</td>
<td>0.021</td>
<td>0.031</td>
</tr>
<tr>
<td>638L</td>
<td>-0.008</td>
<td>-0.002</td>
</tr>
<tr>
<td>638R</td>
<td>0.058</td>
<td>0.068</td>
</tr>
<tr>
<td>639L</td>
<td>-0.063</td>
<td>-0.06</td>
</tr>
<tr>
<td>639R</td>
<td>-0.041</td>
<td>-0.031</td>
</tr>
<tr>
<td>640L</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>640R</td>
<td>-0.06</td>
<td>-0.035</td>
</tr>
<tr>
<td>642L</td>
<td>0.02</td>
<td>0.082</td>
</tr>
<tr>
<td>642R</td>
<td>0.079</td>
<td>0.098</td>
</tr>
<tr>
<td>646L</td>
<td>0.001</td>
<td>-0.023</td>
</tr>
<tr>
<td>646R</td>
<td>-0.022</td>
<td>-0.027</td>
</tr>
<tr>
<td>648L</td>
<td>-0.05</td>
<td>-0.012</td>
</tr>
<tr>
<td>648R</td>
<td>0.016</td>
<td>0.144</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>0.004</strong></td>
<td><strong>0.018</strong></td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td><strong>0.041</strong></td>
<td><strong>0.052</strong></td>
</tr>
</tbody>
</table>
The results from Tables 4.1, 4.2, 4.3, show that out of the 24 kidneys segmented by rater A, the segmentations performed using the geometric and intensity statistics derived in this study generated higher volume overlaps for 16 of the cases. Out of the 24 kidneys segmented by rater B, the geometric and intensity statistics in this study generated better results for 14 of the cases. The average increase in volume overlap over all the cases was 1.6%.

4.1 Comparison of Segmentation Results

All the segmentations depicted in this chapter that were derived during this study are depicted with a white contour, while those obtained during the previous study are depicted with a black contour.

The volume overlap for Case 642R (from Table 4.1) was 79.8% (A to C) and 76.9% (B to C) for segmentation performed using the current geometric statistics. The previous study (from Table 4.2) obtained a volume overlap of 71.9% (A to C) and 67.1% (B to C) for the same image. Visual inspection of the image indicated that the change in the training process for this study may have resulted in geometric and intensity information that likely enhanced the model fitting ability for this image.
Fig. 4.1. Case 642R – Segmentation derived using current model (white contour) versus previous model (black contour). Shown in (a) axial view and (b) coronal view. Circled ROI shows the area that was improved using the new training process.

Figure 4.1 (a) shows a comparison of the contours on an axial view. This particular image illustrates quite well how the previous segmentation cut a flat line across the image as opposed to a slightly convex shape as illustrated by the white line. During this study, all the training images were manually segmented using slightly convex lines on this particular area of the kidney, while the previous study generally utilized flat lines. The models that were fit to these new training images trained the geometric statistics and intensity profiles to recognize this change in shape, which highly increased the accuracy of the segmentation for this particular case.
The volume overlap for Case 634R (from Table 4.1) was 67.1% (A to C) and 71.2% (B to C) using the model derived in this study. The previous study (from Table 4.2) derived a model that had an overlap of 60.8% (A to C) and 64.2% (B to C). Visual inspection of the images shows that the new intensity templates developed for this study were more effective in determining the kidney boundary in an environment that included neighboring organs of similar intensities.

Fig. 4.2. Case 634R – Segmentation derived using current geometry and intensity statistics (white contour) versus previous model (black contour). Shown in (a) axial view and (b) coronal view.

Fig. 4.2 (a) shows the crowded kidney environment and the neighboring organs of similar intensities. The new geometric statistics and intensity templates together produced better results than Rao achieved. The top left corner of the segmentations
in Fig. 4.2 (a) and the top right corner in Fig. 4.2 (b) potentially indicate the ability of the current intensity templates to obtain more accurate boundary estimations.

Figure (Fig. 4.3) illustrates a case in which the results for Rao’s study were more favorable in comparison to the results obtained using the current method. The volume overlap for case 639R (from Table 4.1) was 82.9% (A to C) and 81.5% (B to C) using the current mean model for initialization. The previous study obtained a volume overlap of 87.0% (A to C) and 84.5% (B to C) for segmentations produced for the same case.

Fig. 4.3. Case 639R shows a better volume overlap using the previous method’s segmentation.
This case illustrates an example of when placing greater emphasis on the model regularity penalty may overly restrict model deformation. The current method used a penalty weight to enforce grid regularity, and prevented deformation leading to the best possible image match. A comparison between the current m-rep versus the previous m-rep is shown in Fig. 4.4.

![Fig. 4.4](image)

(a)  (b)

Fig. 4.4. Case 639R – Deformed m-rep for (a) current study which placed greater emphasis on grid regularity versus (b) previously defined model which ignored regularity.

Although a greater volume overlap was achieved using the previous model, it is unlikely that the actual kidney has a structure similar to that indicated by the model in Fig. 4.4 (b). In this particular case, even though the segmentation may not match the human segmentation quite as well as the previous study was able to, improvement is made because a more regular kidney shape is defined for the model. Please view figures 5.2 and 5.3 for further discussion.
The majority of cases in this study used high penalties for the geometric and neighbor penalty weights. However, out of the 24 cases tested, 8 cases performed better when both penalties were relaxed. They were: 630L, 636R, 637L, 638L, 649L, 639R, 640L, and 642L.

One of the biggest variables in the ability to accurately segment CT images is the initialization of the starting model. Utilizing the same kidney model for initialization ensures consistency in model shape. However, the translational placement as well as the angular rotation of the starting model can have significant effects on the final segmentation.

Fig. 4.5 below shows two segmentations of case 642R that undergo the exact same process, used the same geometric statistics, intensity templates and optimization penalties. The only differences between these two segmentations were the translational and rotational placement of the starting mean model relative to the target kidney.
Fig.4.5. Case 642R – Segmentations indicated by the white contour represent the result of one initialization, while the black contour represents another initialization. Large difference can be seen when looking at both the (a) coronal and (b) axial views.

These results indicate that, currently, automatic segmentation is sensitive to the ability of the operator to manually initialize the starting model in a position that accurately approximates the target kidney.
CHAPTER 5 – Conclusions and Discussion

5.1 Analysis of the Results

The main objective of this study was to analyze the effectiveness of automatic segmentation of CT images using deformable models known as m-reps. This work tested a new way of obtaining a mean medial model with geometric and intensity statistics for a population of kidneys.

The results of this study suggest that the geometric statistics used in initialization for automatic segmentation produced results that were satisfactory but mixed. By one metric, volume overlap, the geometric statistics developed for use by this experiment generated superior segmentations on 62.5% of the cases studied compared to the previous study. By another metric, average surface distance, this study obtained worse segmentations than the prior study in 60% of the cases. And finally, a third metric, the Hausdorff distance, saw improvements in 60% of the cases in this experiment versus the previous one.
Fig. 5.1. Case 633L – Notice the high degree of regularity in the spacing between the atoms for the model obtained in the current study (a) versus the prior study (b). Both models achieved similar volume overlaps.

Using the current model, regularity and internal correspondence were enhanced in the final segmentations as compared to Rao’s study. As illustrated by Fig. 4.4, Rao focused on achieving the highest possible image-match while sacrificing grid regularity and thus internal correspondence. The majority of Rao’s segmented kidney models demonstrate similar grid irregularity. Her results achieved good agreement with human segmentations, while sacrificing an important method for achieving internal correspondence. In this study, the ability to acquire segmentations that were on par or even marginally better than previously available was accompanied by models that more closely preserved correspondence.
Fig. 5.2. Case 639R – A 3-D visual representation of a manual segmentation by human rater A. The actual kidney inside the patient is unlikely to resemble this.
Fig. 5.3. Case 639R – A 3-D solid visual representation of 639R from (a) the previous study (b) the current study. Although the previous study may have obtained a model that achieves a higher volume overlap when compared to Fig. 5.2, it is unlikely that the actual kidney contains the visible creases and bumps shown in Fig. 5.3 (a).

Penalty settings compared to the previous study were tweaked more frequently. While Rao used the same penalty parameters for each case, this study tested the use of two different settings. The most optimal segmentations achieved from these two settings were used as the final model for this study.

Manual initialization of the starting model has found to be one of the most important factors in determination affecting the final deformed model. In this study, several cases that resulted in poor segmentations using a particular hand initialization resulted in excellent segmentations using a different initialization. Case 642R which was examined in Chapter 4, showed the results of two different hand
placements. Both hand placements seemed reasonable upon visual inspection, but one resulted in volume overlaps of 76%-79% while the other resulted in overlaps of 62%-64%.

5.2 Future Direction

At this time, more work needs to be done before geometric and intensity statistics can be tested clinically. Enhanced correspondence and more consistent model initialization parameters are important factors in improving the automatic segmentation process. Placing the geometric statistics into a clinical setting where doctors could perform side by side real time comparisons between their own manual segmentations and those generated by the computer would be the first step towards clinical validation. The ultimate goal of this research is to eventually reach a point where radiation oncologists feel just as comfortable using the computer to perform kidney segmentation as they do contouring CT scans themselves.

One possible improvement to the approach used in this study is to replace the locally varying templates with a region based histogram technique [Broadhurst 2005]. The current local intensity templates have difficulty capturing the interrelations among pixels intensities in a region. Regional based intensity methods capture this information much more effectively.

Perhaps one of the more difficult but important tasks in the improvement of automatic segmentation methods involves the initialization procedure immediately prior to segmentation. Different placements can have significant effects on the final segmentation. Currently, optimal use of the computer requires the user to have a
large amount of knowledge of both kidney structure and the inner workings of the computer optimization process. Not only does the hand placement of models often require a lengthy trial and error process to obtain optimal segmentations, but it also requires a significant time commitment. Automatic initialization, therefore, could have great impact on the ability of the program to consistently achieve optimal computer segmentations.

Finally, to more accurately determine accurate geometric statistics, additional landmarks could be placed during the training stage on the binary images. One idea is to place a series of landmarks on the crest region of the kidney to ensure that the crest atoms correspond to the crest region of the image.


7. Yongwon Jeong (RPI), Dr. Richard J. Radke (RPI), Dr. Badri Roysam (RPI), Dr. Daniel Freedman (RPI), Dr. George TY Chen (MGH), Philip Yoon (RPI). Automatic segmentation of the prostate and surrounding structures for image-guided radiotherapy.


