Segmentation of C-Spine MRI Images Using the Watershed Transform

By

Jacobus Johannes Botha

A dissertation submitted to the Faculty of Engineering and the Built Environment in partial fulfilment of the requirements for the degree of

MAGISTER INGENERIAE

in

ELECTRICAL AND ELECTRONIC ENGINEERING

in the

FACULTY OF ENGINEERING AND THE BUILT ENVIRONMENT

at the

UNIVERSITY OF JOHANNESBURG

Supervisor: Prof A. Nel
November 2006
CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CHAPTER 1: INTRODUCTION</td>
<td>6</td>
</tr>
<tr>
<td>1.1</td>
<td>INTRODUCTION</td>
<td>6</td>
</tr>
<tr>
<td>1.2</td>
<td>BACKGROUND TO STUDY</td>
<td>6</td>
</tr>
<tr>
<td>1.3</td>
<td>RESEARCH GOALS</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>HYPOTHESIS</td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>STRUCTURE OF DISSERTATION</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>CHAPTER 2: BACKGROUND</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>IMAGE PROCESSING</td>
<td>9</td>
</tr>
<tr>
<td>2.2</td>
<td>MRI AS SUB AREA IN IMAGE PROCESSING</td>
<td>9</td>
</tr>
<tr>
<td>2.3</td>
<td>SEGMENTATION AS PROCESS</td>
<td>11</td>
</tr>
<tr>
<td>2.4</td>
<td>SEGMENTATION IN MRI</td>
<td>12</td>
</tr>
<tr>
<td>2.5</td>
<td>SUMMARY</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>CHAPTER 3: IMPLEMENTATION</td>
<td>14</td>
</tr>
<tr>
<td>3.1</td>
<td>METHOD RESEARCHED</td>
<td>14</td>
</tr>
<tr>
<td>3.2</td>
<td>THE WATERSHED TRANSFORM</td>
<td>21</td>
</tr>
<tr>
<td>3.3</td>
<td>MATHEMATICAL BACKGROUND</td>
<td>23</td>
</tr>
<tr>
<td>3.4</td>
<td>SOFTWARE IMPLEMENTATION NOTES</td>
<td>27</td>
</tr>
<tr>
<td>3.5</td>
<td>MRI SEGMENTATION PROCESSES</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>CHAPTER 4: RESULTS</td>
<td>33</td>
</tr>
<tr>
<td>4.1</td>
<td>RESULTS OF SEGMENTATION SYSTEM I</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>RESULTS OF SEGMENTATION SYSTEM II</td>
<td>36</td>
</tr>
<tr>
<td>4.3</td>
<td>COMPARISON BETWEEN TWO SYSTEMS</td>
<td>38</td>
</tr>
<tr>
<td>4.4</td>
<td>EVALUATION OF SEGMENTATION SYSTEM I ON DIFFERENT C-SPINE IMAGES</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>EVALUATION OF SEGMENTATION SYSTEM II ON DIFFERENT C-SPINE IMAGES</td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td>CHAPTER 5: INTERPRETATION OF RESULTS</td>
<td>49</td>
</tr>
<tr>
<td>5.1</td>
<td>RESULTS OF APPLICATION TO STANDARD TEST IMAGES</td>
<td>49</td>
</tr>
<tr>
<td>5.2</td>
<td>RESULTS OF APPLICATION TO MRI C-SPINE TEST IMAGES</td>
<td>50</td>
</tr>
<tr>
<td>5.3</td>
<td>COMPARISON TO OTHER KNOWN SEGMENTATION PROCESSES</td>
<td>54</td>
</tr>
<tr>
<td>5.4</td>
<td>NUMERICAL BENCHMARKS OF PERFORMANCE</td>
<td>55</td>
</tr>
<tr>
<td>5.5</td>
<td>SUBJECTIVE PERFORMANCE COMPARISON</td>
<td>61</td>
</tr>
<tr>
<td>5.6</td>
<td>SUMMARY</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>CHAPTER 6: SUMMARY</td>
<td>62</td>
</tr>
<tr>
<td>6.1</td>
<td>GOALS REVISITED</td>
<td>62</td>
</tr>
<tr>
<td>6.2</td>
<td>CRITICAL ANALYSIS OF THE WORK PERFORMED IN THE DISSERTATION</td>
<td>62</td>
</tr>
<tr>
<td>6.3</td>
<td>POTENTIAL FUTURE WORK</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>APPENDIX A: PRASA 2006 PAPER</td>
<td>67</td>
</tr>
<tr>
<td>FIGURE</td>
<td>DESCRIPTION</td>
<td>PAGE</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>SEGMENTATION CLASSIFICATION (1)</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>SEGMENTATION CLASSIFICATION (2)</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>IMAGE PROCESSING FLOW DIAGRAM</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>EROSION USING A 2X2 SQUARE STRUCTURING ELEMENT</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>HISTOGRAM EQUALIZATION</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>ONE-DIMENSIONAL WATERSHED SEGMENTATION, (A) GRAY INTENSITY IMAGE, (B) WATERSHED SEGMENTATION</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>(A) INPUT GRAYSCALE IMAGE, (B) WATERSHED IMAGE</td>
<td>22</td>
</tr>
<tr>
<td>8</td>
<td>(A) GRADIENT IMAGE, (B) WATERSHED OF THE GRADIENT IMAGE</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>THE GEODESIC DISTANCE BETWEEN X AND Y</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>GEODESIC INFLUENCE ZONE OF CONNECTED COMPONENT B1 INSIDE A</td>
<td>24</td>
</tr>
<tr>
<td>11</td>
<td>THREE POSSIBLE INCLUSION RELATIONS BETWEEN Y AND Y AND XH_MIN</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>FLOW DIAGRAMS OF SYSTEM I AND SYSTEM II</td>
<td>30</td>
</tr>
<tr>
<td>13</td>
<td>ORIGINAL MRI IMAGE</td>
<td>33</td>
</tr>
<tr>
<td>14</td>
<td>SEGMENTATION RESULTS OF SYSTEM I</td>
<td>35</td>
</tr>
<tr>
<td>15</td>
<td>SEGMENTED AREAS SUPERIMPOSED ON INPUT IMAGE FOR SYSTEM I</td>
<td>36</td>
</tr>
<tr>
<td>16</td>
<td>ORIGINAL C-SPINE MRI IMAGE</td>
<td>36</td>
</tr>
<tr>
<td>17</td>
<td>RESULTS OF SEGMENTATION SYSTEM II</td>
<td>37</td>
</tr>
<tr>
<td>18</td>
<td>SEGMENTED AREAS SUPERIMPOSED ON GRAYSCALE IMAGE (SYSTEM II)</td>
<td>38</td>
</tr>
<tr>
<td>19</td>
<td>SYSTEM I WATERSHED RESULTS</td>
<td>38</td>
</tr>
<tr>
<td>20</td>
<td>SYSTEM II WATERSHED RESULTS</td>
<td>39</td>
</tr>
<tr>
<td>21</td>
<td>C-SPINE IMAGE 1 SEGMENTATION</td>
<td>40</td>
</tr>
<tr>
<td>22</td>
<td>C-SPINE IMAGE 2 SEGMENTATION</td>
<td>40</td>
</tr>
<tr>
<td>23</td>
<td>C-SPINE IMAGE 3 SEGMENTATION</td>
<td>41</td>
</tr>
<tr>
<td>24</td>
<td>C-SPINE IMAGE 4 SEGMENTATION</td>
<td>41</td>
</tr>
<tr>
<td>25</td>
<td>C-SPINE IMAGE 5 SEGMENTATION</td>
<td>42</td>
</tr>
<tr>
<td>26</td>
<td>C-SPINE IMAGE 6 SEGMENTATION</td>
<td>42</td>
</tr>
<tr>
<td>27</td>
<td>C-SPINE IMAGE 7 SEGMENTATION</td>
<td>43</td>
</tr>
<tr>
<td>28</td>
<td>C-SPINE IMAGE 8 SEGMENTATION</td>
<td>43</td>
</tr>
<tr>
<td>29</td>
<td>C-SPINE IMAGE 9 SEGMENTATION</td>
<td>44</td>
</tr>
<tr>
<td>30</td>
<td>SEGMENTATION OF C-SPINE IMAGE 1</td>
<td>44</td>
</tr>
<tr>
<td>31</td>
<td>SEGMENTATION OF C-SPINE IMAGE 2</td>
<td>45</td>
</tr>
<tr>
<td>32</td>
<td>SEGMENTATION OF C-SPINE IMAGE 3</td>
<td>45</td>
</tr>
<tr>
<td>33</td>
<td>SEGMENTATION OF C-SPINE IMAGE 4</td>
<td>46</td>
</tr>
<tr>
<td>34</td>
<td>SEGMENTATION OF C-SPINE IMAGE 5</td>
<td>46</td>
</tr>
<tr>
<td>35</td>
<td>SEGMENTATION OF C-SPINE IMAGE 6</td>
<td>47</td>
</tr>
<tr>
<td>36</td>
<td>SEGMENTATION OF C-SPINE IMAGE 7</td>
<td>47</td>
</tr>
</tbody>
</table>

Segmentation Of C-Spine MRI Images Using The Watershed Transform
Segmentation Of C-Spine MRI Images Using The Watershed Transform
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Cervical</td>
</tr>
<tr>
<td>CT</td>
<td>Computer Tomography</td>
</tr>
<tr>
<td>IP</td>
<td>Image Processing</td>
</tr>
<tr>
<td>IZ</td>
<td>Influence Zone</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MS</td>
<td>Multiple Sclerosis</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>OOI</td>
<td>Objects of Interest</td>
</tr>
<tr>
<td>ROI</td>
<td>Region-of-interest</td>
</tr>
<tr>
<td>SKIZ</td>
<td>Skeleton by Influence Zone</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machines</td>
</tr>
</tbody>
</table>
1. CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Automatic classification of images has always been an important part of pattern recognition. The segmentation and classification of MRI images has always been a challenge. A segmented image is often a very important input to the classification process. Many classification techniques use segmented images as input to the classification process. Certain segments or areas of an image serve as important features that will be used for classification. Important information can be derived from the features that are present in the segmented image. Sometimes there might be a need to extract a certain object from an image to do classification on the object.

In the case of MRI images, certain structures of the human body like organs and tissue can be isolated by the segmentation process. These objects of interest (OOI) can give vital information for the identification of medical abnormalities (anomalies) and diseases. Segmented objects can play an important role to assist medical practitioners in the diagnosis and treatment of medical problems.

I would like to test the performance of the watershed segmentation algorithm on MRI images of the cervical (C) spine.

1.2 BACKGROUND TO STUDY

Much work has been done on the segmentation and classification of MRI images. Various techniques have been generated and tested over the past decades. Segmentation techniques like thresholding, convolution, pyramid segmentation and morphological segmentation have been utilised. All these techniques have their advantages and disadvantages. The pre-processing of an image plays a very important role in the success of the segmentation process. Histogram manipulation, filtering, thresholding and edge detection are important pre-processing techniques to yield good segmentation results.

Many segmentation and classification techniques have been implemented on MRI images. The latest techniques include support vector machines (SVMs), neural networks (NNs), statistical methods, threshold techniques and normalised cuts. Segmentation of bony structures plays an important role in image guided surgery of the spine [1]. Physicians have commonly relied on computed tomography (CT) images to support their decisions in the diagnosis, treatment, and surgery of different pathologies of the spine due to the high resolution and good visualization of bone offered by this medical imaging modality. CT relies on the use of ionizing radiation, and does not depict soft tissue pathology, unlike magnetic resonance imaging (MRI) [1]. While the segmentation of vertebral bodies from CT images
of the spine has commonly been accomplished with seed growing segmentation techniques [1], this task is more difficult in MRI, with variations in soft tissue contrast, and with the RF inhomogeneities, which increase the level of complexity.

1.3 RESEARCH GOALS

The primary goal of this project is to develop segmentation techniques for C-spine MRI images. This method will also be compared against other methods like pyramid segmentation and morphological segmentation. The watershed segmentation will be implemented and tested as the final step of the segmentation process.

This project will try to use a combination of techniques, rather than to implement and evaluate one single method. It has been learned from literature and also from experience that the pre-processing of the raw data plays a crucial role in the quality of the segmentation process. Therefore, some attention will be given to the pre-processing of the images as part of the segmentation process.

1.4 HYPOTHESIS

The watershed segmentation in combination with other pre-processing techniques can yield better results on C-spine MRI images than other segmentation techniques.

1.5 STRUCTURE OF DISSERTATION

Chapter 1 gives an introduction to this dissertation: background to the study, research goals and the hypothesis for this dissertation are discussed.

Chapter 2 gives background on image processing, segmentation, MRI images and processing and some existing techniques used on MRI image segmentation.

Chapter 3 gives detail on the implementation of processes that is used for image segmentation like morphological processing, histogram equalization, and the watershed algorithm. The mathematical calculation of the watershed transform is also given in this chapter. Pseudo code of the implementation of the watershed segmentation system is given. The two segmentation systems that have been implemented and evaluated are discussed in this chapter.

Chapter 4 gives the results of the implementation of the two segmentation systems that have been implemented. The two systems are compared against each other. The results of the two systems on 9 different cervical-spine images are shown.

In chapter 5 the interpretation of the results are discussed. The segmentation systems have been evaluated on the Berkeley segmentation dataset. The watershed segmentation systems have been evaluated on ground truth images (manually segmented images). The numerical results are shown in
this chapter. A subjective comparison is made against a pyramid segmentation algorithm. A radiologist gives his interpretation on the performance of the results.

In chapter 6 the goals of this dissertation are revisited. A critical analysis of the work performed in this dissertation is done in the next section. Lastly some goals are discussed for potential future work.
2. CHAPTER 2: BACKGROUND

2.1 IMAGE PROCESSING

Image processing is any form of information processing for which both the input and output are images, such as photographs or frames of video. Most image processing techniques involve treating the image as a two-dimensional signal and applying standard signal processing techniques to it [2].

We begin with certain basic definitions. An image defined in the "real world" is considered to be a function of two real variables, for example, \( A(x,y) \) with \( A \) as the amplitude (e.g. brightness) of the image at the real coordinate position \( (x,y) \).

\[
\text{PixelAmplitude} = A(x, y)
\]

The usual mathematical representation of an image is a function of two spatial variables: \( f(x, y) \). The value of the function at a particular location \( (x,y) \) represents the intensity of the image at that point. An image may be considered to contain sub-images sometimes referred to as regions-of-interest, ROIs, or simply regions. This concept reflects the fact that images frequently contain collections of objects each of which can be the basis for a region. In a sophisticated image processing system it should be possible to apply specific image processing operations to selected regions. Thus one part of an image (region) might be processed to suppress motion blur while another part might be processed to improve color rendition.

The amplitudes of a given image will almost always be either real numbers (floating point) or integer numbers. The latter is usually a result of a quantization process that converts a continuous range (say, between 0 and 100%) to a discrete number of levels. In certain image-forming processes, however, the signal may involve photon counting which implies that the amplitude would be inherently quantized. In other image forming procedures, such as magnetic resonance imaging (MRI), the direct physical measurement yields a complex number in the form of a real magnitude and a real phase. For the remainder of this dissertation we will consider amplitudes as real numbers or integers unless otherwise indicated.

2.2 MRI AS SUB AREA IN IMAGE PROCESSING

MRI is widely used in the diagnosis and treatment of diseases in the human body [3]. MRI is extensively used for scanning of the brain, C-spine, lumbar spine, abdomen and knees. In fact, any part of the body can be scanned to give the physician an in-depth view of the organs and tissues that need to be investigated. It gives the physician the ability to view in detail components and structures inside the human body. Small slices of the body are been made. Usually the slices are 1mm thick. The images are grayscale images that can also be mapped to a color-map image to aid in the interpretation.
of certain areas like tissue, bone, muscles etc. The slices can be superimposed to yield a 3-dimensional view of the part of the body that has been scanned, for example the brain.

As the high-energy nuclei relax and realign, they emit energy which is recorded to provide information about their environment. The realignment of the magnetic field is termed longitudinal relaxation and the time in milliseconds required for a certain percentage of the tissue nuclei to realign is called ‘Time 1’ or $T_1$. This is the basis of $T_1$-weighted imaging [3].

T2-weighted imaging relies upon local dephasing of spins following the application of the transverse energy pulse; the transverse relaxation time is termed ‘Time 2’ or $T_2$. Both $T_1$- and $T_2$-weighted images are acquired for most medical examinations. Often, a paramagnetic contrast agent is administered, and both $T_1$ and $T_2$ images are obtained [3].

In order to create the image, spatial information must be recorded along with the received tissue relaxation information. Magnetic fields with an intensity gradient are applied in addition to the strong alignment field to allow encoding of the position of the nuclei. A field with the gradient increasing in each of the three dimensional planes is applied in sequence. The received signals are recorded in a temporary memory named the K-space. This is the spatial frequency weighting in two or three dimensions of a real space object sampled by the MRI. The information is inverse Fourier transformed by a computer into real space to obtain the desired image. Typical medical resolution is about 1 mm$^3$, while research models can exceed 1 μm$^3$ [3].

Major applications of MRI medical imaging:

- Diagnosing multiple sclerosis (MS)
- Diagnosing tumours of the pituitary gland and brain
- Diagnosing infections in the brain, spine or joints
- Visualizing torn ligaments in the wrist, knee and ankle
- Visualizing shoulder injuries
- Diagnosing tendonitis
- Evaluating masses in the soft tissues of the body
- Evaluating bone tumours, cysts and bulging or herniated discs in the spine
- Diagnosing strokes in their earliest stages

The fields of image processing are very wide. MRI is a specialist area in the field of medical imaging. Sonar images, CT-scans and MRI images are among the fields that are most commonly used. While CT provides good spatial resolution (the ability to distinguish two structures an arbitrarily small distance from each other as separate), MRI provides comparable resolution with far better contrast.
resolution (the ability to distinguish the differences between two arbitrarily similar but not identical tissues) [3]. The basis of this ability is the complex library of pulse sequences that the modern medical MRI scanner includes, each of which is optimized to provide image contrast based on the chemical sensitivity of MRI [3]. MRI is a widely employed imaging technique that can be used for both anatomical and functional images [4].

2.3 SEGMENTATION AS PROCESS

Segmentation is the partitioning of images/volumes into meaningful pieces [1]. Segmentation involves the isolation on a specific region of interest. Segmentation subdivides an image into objects of interest. These objects are often referred to as regions of interest (ROI) or volumes of interest (VOI) [4]. In this dissertation I will only use the term ROI, whether they are 2D or 3D.

Segmentation plays a key role in the identification of objects and features and subsequent classification of images. Objects of interest can be isolated and extracted by the segmentation process. The segmented objects can be used as input features for the classification process. The segmented regions can also be used as inputs to intelligent processes like fuzzy logic systems and neural network systems.

The purpose of image segmentation is:

- Detection or recognition of an object or feature of an object
- Quantifying of the properties of an object, like the size and quantity of the object
- Identification of statistical, morphological and geometrical properties of the object.

Figure 1 Segmentation classification (1)
There are three types of image segmentation: thresholding, edge-based and region-based segmentation (Figure 1). Segmentation techniques can also be classified as statistical segmentation, morphological and region-based segmentation (Figure 2).

One way to formulate segmentation is with normalized-cuts method. Normalized-cuts formulate segmentation as a graph-partitioning problem: it maximizes both the total dissimilarity between the different groups and the total similarity within the groups [1].

2.4 SEGMENTATION IN MRI

Segmentation is a key process in the image processing of MRI images. Physical structures like organs, bones and tissues can be isolated in the segmentation process. These regions of interest (ROI) will play a crucial role in further analysis and diagnosis of abnormalities and diseases. Specific objects of interest like tumours, fractured bones, lesions, etc. can be isolated and identified in the segmentation process [6], [7].

Many segmentation techniques have been developed for MRI images. Statistical techniques like expectation/maximization, binary mathematical morphology and active contour models have been implemented. Nearest-neighbour techniques have also been implemented. Non-linear techniques like neural networks and Support Vector Machines (SVM) have also been widely used in the segmentation of MRI images [8].

Scale-invariant segmentation, a combination of expectation-maximization segmentation, binary mathematical morphology and active contour models [9],[10] are also used. A combination of texture based segmentation and neural network classification can also be implemented.

2.4.1 Histogram based segmentation in MRI
A histogram can be used as a feature or input for image segmentation. It gives the statistical distribution of the pixels in an image. The histogram gives a very strong indication of the contrast of the image. In this method the histogram of the image is used for segmentation. It is proportional to the probability density function of the image. Therefore, it gives an indication of the statistical distribution of the pixel intensities in the image. The threshold that is used for segmentation is selected from the histogram of the image. A localized threshold can also be calculated from the local histogram, giving an optimal threshold for the localized area.

Histogram based features that can be used, includes: windowed histograms of brightness and windowed histograms of textons with intensity and position [1].

2.4.2 Texture based segmentation in MRI

Texture is a fundamental feature, which provides significant information for scene interpretation and image classification. Textures can be defined as homogenous patterns or spatial arrangements of pixels that regional intensity or color alone does not sufficiently described. Most texture segmentation algorithms require the estimation of model parameters, which has proved to be difficult [11].

*Texture segmentation* is the identification of regions based on their texture [12]. Various methods can be used in texture based segmentation. Calculation of the entropy (measure of randomness) of the image can be used as an important parameter for texture based segmentation. Calculating the gray-level co-occurrence matrix of the image, by calculating how often a pixel with gray-level (grayscale intensity) value \(i\) occurs horizontally adjacent to a pixel with the value \(j\).

Region based methods like calculating the local range and statistical methods like the standard deviation can also be used as part of the segmentation process.

2.5 SUMMARY

Segmentation of bony structures plays an important role in image guided surgery of the spine. In this, a novel approach to the segmentation of vertebral bodies from 2D sagittal magnetic resonance images of the spine is discussed. I would like to show that the watershed segmentation in combination with morphological processing can be used as a powerful segmentation tool in the segmentation of C-spine MRI images. The structure, shape and morphology of the bony structures in spine images lend itself to the watershed segmentation.
3. CHAPTER 3: IMPLEMENTATION

3.1 METHOD RESEARCHED

A typical image processing system is described in Figure 3. It usually consists of a number of stages, i.e. pre-processing, thresholding, morphological processing, segmentation and classification. The focus in this dissertation will be on thresholding, morphological processing and segmentation.

The method that has been used includes a combination of processes. Thresholding, morphological processing, edge-detection, filtering and the watershed segmentation have been used to yield a segmented image. Some pre-processing need to be applied on the MRI images to improve segmentation of the raw data.

The input processing includes morphological operations. Morphological processing is the processing of images based on structures with a certain shape. The following morphological operations can be used as powerful pre-processing in segmentation of images:

<table>
<thead>
<tr>
<th>Morphological Operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosion</td>
</tr>
<tr>
<td>Dilation</td>
</tr>
<tr>
<td>Closing</td>
</tr>
</tbody>
</table>

Figure 3 Image Processing Flow Diagram
### Morphological Operations

- Opening
- Morphological reconstruction
- Extended-minima transform
- Extended-minima transform
- Morphological bottom hat operation

- Erosion removes pixels from object boundaries [13],[14]. It assigns the minimum value of all the neighbouring pixels in the structuring element to the pixel that is currently evaluated. In a binary image, the pixel will be set to zero, if any pixel in the structuring element is zero. The equation for erosion is as follows:

\[
(f \ominus g)(x) = \bigvee_{y \in E} f(y) + g(x - y)
\]

where \( \ominus \) refers to the Minkowski addition [15].

- Dilation adds pixels to boundaries of an object. It assigns the maximum value of all the neighbouring pixels in the structuring element to the pixel under test [13],[14]. In a binary image, the pixel will be set to one, if any pixel in the structuring element is one. In algebra, any transformation which is increasing, anti-extensive and idempotent is called an (algebraic) opening, increasing, extensive and idempotent is called a (algebraic) closing.

\[
(f \oplus g)(x) \equiv \bigwedge_{y \in E} f(y) - g(x - y)
\]

where \( \oplus \) refers to Minkowski subtraction [15].

- Closing is a dilation operation followed by erosion with the same structuring element. Morphological closing can be used to join shapes and structures and to fill up gaps. It can

---

*Figure 4: Erosion using a 2x2 square structuring element*
also be used to fill in basins and valleys in a topological image [13],[14]. The resulting structures can be used as markers.

\[ f \mapsto (f \oplus g) \ominus g \quad (4) \]

- Opening is erosion followed by dilation with the same structuring element. Morphological opening can be used to remove small objects from an image while preserving the larger objects in an image [13],[14]. An opening removes some peaks and crest lines.

\[ f \mapsto (f \ominus g) \oplus g \quad (5) \]

- Morphological reconstruction processes one image, called the marker, based on the characteristics of another image, called the mask. Morphological reconstruction can be used to reconstruct the topology of an image [13].

- The regional extrema of a raw image mark relevant as well as irrelevant image features. The h-maxima transformation suppresses all maxima whose depth is lower or equal to a given threshold level \( h \). This is achieved by performing the reconstruction by dilation of \( f \) from \( f - h \) [16]:

\[ HMAX_h(f) = R^\delta_f(f - h) \quad (6) \]

- The extended-minima transform calculates the regional minima of the corresponding h-minima transform. Regional minima are connected components of pixels with the same intensity value, whose external boundary pixels all have a higher value [16]. The extended minima is defined by:

\[ EMIN_h(f) = RMIN[HMIN_h(f)] \quad (7) \]

- The extended-maxima transform calculates the regional maxima of the k-maxima transform. Regional maxima are connected components of pixels with the same intensity value, whose external boundary pixels all have a lower value [16]. The extended maxima EMAX is defined by:

\[ EMAX_h(f) = RMAX[HMAX_h(f)] \quad (8) \]

- An important input to segmentation and image classification is the gradient image. The gradient image gives an indication of the level of the gradients (peaks and valleys) in an image. The gradient image can be calculated with the derivative of a Gaussian filter. Gradient images can be used to calculate edges in images. It can also be used as an input to the watershed transforms [16]. The gradient of a function of two variables, \( F(x,y) \) is defined as [18]:

Segmentation Of C-Spine MRI Images Using The Watershed Transform
A distance transform can be used to calculate the gradient of an image [13]. The distance transform provides a measure of the separation of points in an image. The distance transform calculates the distance of the current pixel to the nearest non-zero pixel of the input binary image. The distance function $dist_X$ of a set of $X \subset \mathbb{Z}^2$ associates the distance of each pixel to the background [14]:

$$\text{dist}_X \left( p \mapsto \min \{d_c(p, g) \mid q \in X \} \right) \quad (10)$$

The distance function $d_I$ of a binary image $I$ is equivalent to that of its set of feature pixels, i.e. pixels with value 1. In addition we put conveniently:

$$\forall p \in D_I, I(p) = 0 \Rightarrow \text{dist}_I(p) = 0 \quad (11)$$

The distance transform can be calculated with several distance metrics, for example the Euclidean distance, city block, chessboard and the quasi-Euclidean distance. In the figure below is an example of the distance transform with the chessboard distance metric.

The input is shown in the top matrix and the output in the bottom matrix.

<table>
<thead>
<tr>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

An image can be enhanced or modified by applying an image filter. A filter can be applied to emphasize or remove certain features. Image processing operations implemented with filtering include smoothing, sharpening, and edge enhancement [13].
Filtering is a neighbourhood operation, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighbourhood of the input pixel. A pixel's neighbourhood is some set of pixels, defined by their locations relative to that pixel.

The most common type of filters use convolution, correlation and finite impulse response (FIR) filters.

Noise in an image can be removed with a median filter. The median filter calculates the median of the pixels in the neighbourhood and assigns that value to the input pixel. Median filtering is a special case of ordered-statistics filtering, also known as ranked-filtering. The median filter is much less sensitive to outlier values than an average filter. A median filter will remove the outliers without sacrificing the sharpness of an image [13].

The contrast of an image can be enhanced by applying a histogram equalizer or by applying histogram manipulation. This is done by changing the intensity values of the image according to a mapping function that map the intensity values according to a specific transfer function based on the pixel input values.

Histogram equalization enhances the contrast of an image by transforming the intensity values in an image so that the histogram of the output image approximately matches an equal distributed (flat) or predefined histogram [13].
Figure 5 Histogram equalization:
Input image (a). Histogram equalized image (b). Histogram of original image (c). Histogram of histogram equalized image (d).

An example of histogram equalization can be seen in. The original image can be seen in the top left image (Figure 5). The histogram of this figure can be seen in the bottom left figure. This is a bar chart displaying the distribution of the gray scale levels of the image. The image after histogram equalization can be seen in the top right image. The histogram after histogram equalization can be seen in the bottom right image. It can clearly be seen that the contrast has been stretched in this image.

The pseudo code to perform the histogram equalization can be seen in the following paragraph:

Input image: \( a \)
Output image: \textbf{output}

Scale input image to levels between 0 and 1
\[
    a = a / \text{max}(a(:));
\]

Choose number of histogram levels
\[
    n = 64;
\]
\[
    \text{hgram} = \text{ones}(1,n) \times (\text{numel}(a)/n);
\]
n = 256;

*Normalize histogram*

\[ hgram = hgram \times \left( \frac{\text{numel}(a)}{\text{sum}(hgram)} \right) \]

\[ m = \text{length}(hgram) \]

\[ \text{nn} = \text{imhist}(a,n) \]
\[ \text{cum} = \text{cumsum}(\text{nn}) \]

\[ \text{cumd} = \text{cumsum} \left( hgram \times \frac{\text{numel}(a)}{\text{sum}(hgram)} \right) \]

*Create transformation to an intensity image by minimizing the error between desired and actual cumulative histogram.*

\[ \text{tol} = \text{ones}(m,1) \times \text{min}([\text{nn}(1:n-1),0;0,\text{nn}(2:n)])/2; \]

\[ \text{err} = (\text{cumd}(:) \times \text{ones}(1,n) - \text{ones}(m,1) \times \text{cum}(:)) + \text{tol}; \]

\[ d = \text{find} (\text{err} < -\text{numel}(a) \times \text{sqrt} (\text{eps})) ; \]

\[ \text{if} \; \text{isempty}(d) \]
\[ \text{err}(d) = \text{numel}(a) \times \text{ones}(\text{size}(d)); \]
\[ \text{end} \]

\[ [\text{dum},T] = \text{min}(\text{err}); \]
\[ T = \frac{(T-1)}{(m-1)}; \]

\[ \text{max\_input} = \text{max}(a(:)); \]
\[ \text{max\_T\_index} = (\text{numel}(T)-1); \]
\[ \text{output} = \text{zeros}(\text{size}(a,1),\text{size}(a,2)); \]

\[ \text{for} \; i = 1: \text{size}(a,1), \]
\[ \text{for} \; j = 1: \text{size}(a,2) \]
\[ \quad \text{index} = \text{ceil}(\text{max\_T\_index} \times (a(i,j)/\text{max\_input})+0.5); \]
\[ \quad \text{output}(i,j) = T(\text{index}); \]
\[ \text{end} \]
\[ \text{end} \]

Another way to improve the contrast of an image is to do contrast stretching. Here is the pseudo code for the contrast stretching algorithm:

**Input image:** a  
**Output image:** out_image

\[ a = \text{round}(a \times 255/\text{max}(a(:))); \]
\[ a = a + 1; \]
\[ [m,n] = \text{size}(a); \]
\[ \text{out\_image} = \text{zeros}(m,n); \]

\[ \text{cmap} = \text{contrast}(a); \]
\[ \text{scale\_factor2} = (\text{length}(\text{cmap})-1)/\text{max}(a(:)); \]

\[ \text{scale\_factor} = 256; \]

\[ \text{for} \; i = 1:m, \]
\[ \text{for} \; j = 1:n, \]
\[ \quad \text{out\_image}(i,j) = \text{cmap}(\text{round}(a(i,j) \times \text{scale\_factor2}+1),1) \times \text{scale\_factor}; \]
Here follows the pseudo code for the contrast stretching algorithm:

function cmap = contrast(x,m);
    if nargin < 2, m = size(colormap,1);
    end
    xmin = min(x)
    xmax = max(x)
    x = round((m-1) x (x-xmin)/(xmax-xmin))
    f = find(diff(sort([x(:); (0:m)'])))
    f = f/sum(f)
    cmap = [f f f]

3.2 THE WATERSHED TRANSFORM

The Watershed transform can be used to separate touching objects in an image. This is often the case with MRI images, where brain structures touch on neighbouring objects. This can also be the case with spine images, where the vertebrae touch on adjacent mussels and tissue, making the segmentation of these objects extremely difficult. It is sometimes extremely difficult to separate touching or overlapping objects. The watershed transform provides a solution to this problem [16].

![Watershed segmentation](image)

Figure 6 One-dimensional watershed segmentation. (a) Gray intensity image. (b) Watershed segmentation.

The watershed transform can be used to separate regions (basins) in an image, by using the watershed ridge lines in the image (Figure 6) [16].

The watershed can be explained in the following way. Consider the input image to be a topology image. Each local minimum can be considered as a pierced hole in a topology surface. The surface is then slowly immersed into a lake. The water will progressively fill from the lowest minima until the catchment basin is flooded. At each vertex where the water coming from two catchment basins
merges, we built a dam. At the end of this flooding process, each minimum region is completely surrounded by dams, which delimit its associated catchment basin [16]. These dams correspond to the watershed lines of the input image.

Figure 7 (a) Input grayscale image. (b) Watershed image

Figure 8 (a) Gradient image. (b) Watershed of the gradient image

This transform is combined with other morphological processes; the watershed forms the basis of extremely useful and powerful segmentation processes.

When the watershed segmentation is applied on the raw input image, it usually results in over-segmentation (Figure 7). To overcome this problem, we first calculate the gradient image of the input image. The watershed transform is then applied to the gradient image, to yield a much better result for the final segmentation (Figure 8).
3.3 MATHEMATICAL BACKGROUND

A mathematical explanation of the watershed algorithm, as derived and explained by Luc Vincent and Pierre Soille [16], is given in the following paragraphs.

Let us denote the image under test as $I$. Name $h_{\text{min}}$, the smallest value of $I$ on its domain $D$. Also denote $h_{\text{max}}$ as the largest value of $I$ on its domain $D$. Denote $T_h(I)$ as the threshold of $I$ at level $h$:

$$T_h(I) = \{ p \in D, I(p) \leq h \}$$  \hspace{1cm} (12)

Denote $C(M)$ as the catchment basin associated with a minimum $M$ and $C_h(M)$ the subset of this catchment basin made of the points belonging to the minima less than or equal to altitude $h$:

$$C_h(M) = \{ p \in C(M), I(p) \leq h \} = C(M) \cap T_h(I)$$  \hspace{1cm} (13)

As concerns the minima of $I$, $\text{min}_h(I)$ refers to the set of points belonging to the minima at altitude $h$.

We need to review the geodesic distance to be used in this transform:

Let $A$ be a set which is supposed to be simply connected. The geodesic distance between the pair of pixels $x$ and $y$ in $A$ is the maximum of the paths which join $x$ and $y$ and are totally included in $A$:

$$d_A(x, y) = \inf \{ \text{length}(P), P \text{ is the path between } x \text{ and } y \text{ that is totally included in } A \}$$

This definition is illustrated in [16].

Let's say $A$ contains a set $B$ made of several connected components $B_1, B_2, ..., B_k$.

Definition: The geodesic influence zone $iz_A(B_i)$ of a connected component $B_i$ of $B$ in $A$ is the locus of points of $A$ whose geodesic distance to $B_i$ is smaller than their geodesic distance to any other component of $B$:

$$iz_A(B_i) = \{ p \in A, \forall j \in [1, k], i \neq j, d_A(p, B_i) < d_A(p, B_j) \}$$  \hspace{1cm} (14)

This concept is illustrated in Figure 10. Points of $A$ that do not belong to any geodesic influence zone constitute the skeleton by influence zone (SKIZ) of $B$ inside $A$, denoted SKIZ_A(B):
According to this digital definition, the geodesic SKIZ of B in A does not necessarily separate the different geodesic zones. The digital SKIZ may sometimes be a thick one, since the set of pixels from at an equal distance from two connected components may well be very thick.

\[ X_{\text{h}_{\text{min}}} = T_{\text{h}_{\text{min}}} (I) \cap X_{\text{h}_{\text{min}}} \]  

\( X_{\text{h}_{\text{min}}} \) consists of the points of I which belong to the minima of the lowest altitude. The threshold of I at level \( h_{\text{min}} + 1 \), i.e. \( T_{h_{\text{min}}} + 1(I) \), will now be considered. Obviously, \( X_{h_{\text{min}}} \subseteq T_{h_{\text{min}}} + 1(I) \). \( Y \), being one of the connected components of \( T_{h_{\text{min}}} + 1(I) \), there are three possible relations of inclusion between \( Y \) and \( Y \)

\[ Y \cap X_{h_{\text{min}}} = \phi \]  

\( Y \) is a new minimum of I. \( Y \) is a plateau at level \( h_{\text{min}} + 1 \), since
All the surrounding pixels do not belong to \( T_{h_{\text{min}+1}}(I) \) and have therefore a gray level strictly greater than \( h_{\text{min}+1} \). The catchment basin surrounding this minimum, that was discovered, will be progressively filled up with water.

\[ Y \cap X_{h_{\text{min}}} \neq \emptyset \quad \text{and is connected: } Y \text{ corresponds exactly to the pixels belonging to the catchment basin associated with the minimum } Y \cap X_{h_{\text{min}}} \text{ and having a gray level lower or equal to } h_{\text{min}+1}: \]

\[ Y = C_{h_{\text{min}+1}}(Y \cap X_{h_{\text{min}}}) \quad (19) \]

\[ Y \cap X_{h_{\text{min}}} \neq \emptyset \quad \text{and is not connected: } Y \text{ contains different minima of } I. \text{ Let's denote } Z_1, Z_2, \ldots, Z_k \text{ these minima, and let } Z_i \text{ be one of them. The best possible choice for } C_{h_{\text{min}+1}}(Z_i) \text{ is given by the geodesic influence zone of } Z_i \text{ inside } Y: \]

\[ C_{h_{\text{min}+1}}(Z_i) = iz_y(Z_i) \quad (20) \]

These inclusion relationships are illustrated in Figure 11. After all the possibilities have been discussed, take the second set of the recursion as the following one:

\[ X_{h_{\text{min}+1}} = \min_{h_{\text{min}+1}} \bigcup IZ_{T_{h_{\text{min}+1}}(I)}(X_{h_{\text{min}}}) \quad (21) \]

This relation holds for all levels of \( h \), and finally, we get the following definition:

**Definition (catchment basins and watersheds by immersion):**

The set of the catchment basins of the grayscale image \( I \) is equal to the set \( X_{h_{\text{max}}} \), and can be obtained after the following recursion:

\[ \begin{align*}
& \text{a) } X_{h_{\text{min}}} = T_{h_{\text{min}}}(I) \\
& \text{b) } \forall h \in [h_{\text{min}}, h_{\text{max}} - 1], X_{h+1} - \min_{h+1} \bigcup IZ_{T_{h+1}(I)}(X_h) \quad (23)
\end{align*} \]

The watersheds of \( I \) correspond to the complement of this set in \( D_I \), i.e., to the set of the points of \( D_I \) which do not belong to any catchment basin.
∀p ∈ D₁,
\[ I(p)_{OT} = \begin{cases} 
I(p) \oplus \tilde{T}_2 & \text{if } I(p) \oplus \tilde{T}_2 < I(p) \ominus \tilde{T}_1 \\
I(p) & \text{otherwise}
\end{cases} \tag{24} \]

In the above equation, \( \oplus \) and \( \ominus \) refer to the well-known Minkowski operation.

Figure 11 Three possible inclusion relations between Y and Y and X_{\text{h min}}
3.4 SOFTWARE IMPLEMENTATION NOTES

The software has been implemented in Matlab. Existing toolboxes and functions of Matlab have been used, however the watershed transform has been implemented from the Vincent and Soille modified algorithm [19]. It was decided to rather use existing functions that has been tested and proven, than to rewrite software. The low-level concepts have already been sorted out. The goal was to find a segmentation process that can be implemented easily and will yield good results on the test data. It should be maintainable and easily upgradeable. The power and flexibility of Matlab yield an ideal environment to implement and test the concepts to yield a system that will give good segmentation results.

The pseudo-code for a modified Vincent-Soille watershed algorithm is shown in the following paragraph [19].

```
Input: input_IM - digital grayscale image
Output: label - labeled watershed image

INIT = -1;
MASK = -2;
WSHED = 0;
FICTITIOUS = [-1;-1];
fifo = []; cur_label = 1;

NO_ROWS = size(input_IM,1);
NO_COLS = size(input_IM,2);

label = -1*ones(NO_ROWS,NO_COLS);
dist = zeros(NO_ROWS,NO_COLS);
neighbourhood = zeros(3,3);

[sorted_levels,index] = sort(input_IM(:));
h_min = sorted_levels(1)
sorted_length = length(sorted_levels);
h_max = sorted_levels(sorted_length)

indexc = zeros(1,sorted_length);
indexr = zeros(1,sorted_length);

% Sort all the pixels in the image
for i = 1:sorted_length,
    indexc(i) = ceil(index(i)/NO_ROWS);
    remr = rem(index(i),NO_ROWS);
    if (remr == 0)
        indexr(i) = NO_ROWS;
    else
        indexr(i) = remr;
    end
end
```
% Start the flooding procedure
index_st = 1;
index_end = 1;
while (index_st <= sorted_length)
    h = sorted_levels(index_st);
    if (index_end < sorted_length)
        while (sorted_levels(index_end+1) == h)
            index_end = index_end + 1;
            if (index_end == sorted_length)
                h = h_max+1;
                break;
        end
    end
    for I = index_st:index_end,
        image_indr = indexr(I);
        image_indc = indexc(I);
        label(image_indr,image_indc) = MASK;
        % Test all neighbours of masked pixel for watershed or labelled pixels
        for m = image_indr-1:image_indr+1,
            for n = image_indc-1:image_indc+1
                if (—((m == image_indr)&(n == image_indc))&(m >= 1)&(n >= 1)&(m <= NO_ROWS)&(n <= NO_COLS))
                    if ((label(m,n) > 0))
                        fifo = [fifo,[m;n]]; dist_image(m,n) = 1;
                    end
                end
            end
        end
    end
end
current_dist = 1;
fifo = [fifo,[FICTITIOUS]];
index_st = index_end + 1;
index_end = index_st;

%====== BIG LOOP: EXTEND BASINS ===============
while (~isempty(fifo))
    p = fifo(:,1);fifo(:,1) = [];
    if (p == FICTITIOUS)
        if (isempty(fifo))
            break;
        else
            fifo = [fifo,[FICTITIOUS]];
            current_dist = current_dist + 1;
            p = fifo(:,1);fifo(:,1) = [];
        end
    end
    cr = p(1);
    cc = p(2);
    for m = cr-1:cr+1,
        for n = cc-1:cc+1
            if (—((m == cr)&(n == cc))&(m >= 1)&(n >= 1)&(m <= NO_ROWS)&(n <= NO_COLS))
if ((dist(m,n) < current_dist)&(label(m,n) == WSHED))
    % pixel(m,n) belongs to an existing basin or watershed
    if label(m,n) > 0
        if ((label(cr,cc) == MASK)|(label(cr,cc) == WSHED))
            label(cr,cc) = label(m,n);
        elseif (label(cr,cc) == label(m,n))
            label(cr,cc) = WSHED;
        end
        elseif (label(cr,cc) == MASK)
            label(cr,cc) = WSHED;
        end
    elseif ((label(m,n) == MASK)&(dist(m,n) == 0)) % M,N is located on a plateau
        dist(m,n) = current_dist + 1;
        fifo = [fifo,[m;n]];
    end
    end
end

% END OF BIG LOOP

% Detect and process new minima at level h
for k = 1:NO_ROWS,
    for l = 1:NO_COLS
        if (input_IM(k,l) == h)
            dist(k,l) = 0;
            if (label(k,l) == MASK)
                cur_label = cur_label + 1;
                fifo = [fifo,[k;l]];
                label(k,l) = cur_label;
            end
            while (~isempty(fifo))
                q = fifo(:,1);fifo(:,1) = [];
                % Inspect neighbors of q
                image_indr = q(1);
                image_indc = q(2);
                for m = image_indr-1:image_indr+1,
                    for n = image_indc-1:image_indc+1,
                        if (~((m == image_indr)&(n == image_indc))&(m > 0)&(n > 0)&(m <= NO_ROWS)&(n <= NO_COLS))
                            if (label(m,n) == MASK)
                                fifo = [fifo,[m;n]];
                                label(m,n) = cur_label;
                            end
                        end
                    end
                end
            end
        end
    end
end

% SEGMENTATION OF C-SPINE MRI IMAGES USING THE WATERSHED TRANSFORM
Two systems have been implemented. The flow diagrams of these systems can be seen in Figure 12.

The first system (I) will be described in the following paragraphs:

1. The first step is to threshold the image with a fixed threshold. The threshold is calculated with Matlab's `graythresh` function. This function uses Otsu's method [20], calculating a threshold that minimizes the intraclass variance of the black and white pixels.

2. The next step is to implement a closing function. The `bwmorph` function in Matlab was used to perform this operation. This function performs a morphological dilation followed by erosion.
The operation is implemented 30 times, until the process settles and no more changes happen in the output image.

3. The next step is to perform a two-dimensional convolution. A uniform mask of 10-by-10 pixels is used. Matlab’s \texttt{conv2} function is used to perform this operation. This operation enhances block-shape structures in the input image.

4. An erosion operation is then performed on the convoluted image. A disk shape structure with a radius of 5 is used. The disk shape structure is approximated by a sequence of 4 periodic-line structuring elements.

5. The next step is to threshold the image with a regional minima transform of the H-minima transform. Regional minima are connected components of pixels with similar intensity values, and whose external boundary pixels all have a higher value. This transform is very useful as a pre-process for the segmentation process. The Matlab function \texttt{imextendedmin} is used to do this thresholding.

6. The next step is to implement a distance transform. This function calculates the Euclidean distance between each non-zero pixel and the nearest non-zero pixel. The Matlab function \texttt{bwdist} is used for this operation. This operation yields a gradient image giving an indication of the distance between objects and neighbouring objects in the image.

7. The sign of the values of the output image is inverted. Pixels belonging to the background of the image are set to minus infinity.

8. A median filter is implemented next to remove outliers and noise. This is a non-linear operation that reduces noise significantly, but preserves edges. Matlab’s \texttt{medfilt2} is used for this operation.

9. The final step is to calculate the watershed transformation. This function calculates the watershed regions, yielding the final segmented image. I use a modified version of the Vincent Soille \texttt{watershed} algorithm [19] to perform this operation.

The second system (II) is described in the following paragraphs. This system is much simpler and only consists of a small number of steps:

1. The first step is to equalize the histogram of the input image (3.1). This step enhances the contrast of the image. It is done with the function \texttt{jcontrast}.

2. The second step is to threshold the image to create a binary image. The Matlab function \texttt{graythresh} is used to calculate the threshold. The thresholding is done with the function \texttt{im2bw}. All pixels above the threshold is set to one and pixels below the threshold is set to zero. The output of this function is a binary (black-and-white) image.
3. The next step is to implement a distance transform. This function calculates the Euclidean distance between each non-zero pixel and the nearest non-zero pixel. The Matlab function `bwdist` is used for this operation.

4. The sign of the values of the output image is complemented. Pixels belonging to the background of the image are set to minus infinity.

5. A median filter is implemented next to get rid of outliers and noise. This is a non-linear operation that reduces noise significantly, but preserves edges. Matlab's `medfilt2` is used for this operation.

6. The final step is to calculate the watershed transformation. This function calculates the watershed regions, yielding the final segmented image. I used a modified version of the Vincent Soille `watershed` algorithm [19] to perform this operation.

7. Results of this process can be seen in the following section.
4. CHAPTER 4: RESULTS

4.1 RESULTS OF SEGMENTATION SYSTEM I

The C-spine input image can be seen in Figure 13. The images have kindly been supplied by Kloof Hospital in Erasmuskloof, Pretoria. This is a grayscale intensity image. The image has been threshold by the extended minima transform, described earlier. The threshold was calculated with the `graythresh` function, based on Otsu's method [20].

A number of morphological functions are then performed on the image. A morphological closing is done on the image (Figure 14, 2nd row, left), followed by a convolution with 10x10 square structured element (Figure 14, 2nd row, right). This is followed with a morphological dilation (Figure 14, 3rd row, left) and erosion (Figure 14, 3rd row right), also with a 10x10 structured element.

An extended minima transform is then applied on the image (Figure 14, 4th row left), followed by the distance transform (Figure 14, 4th row right). All pixels not belonging to any object is then moved to the background (minus infinity). The watershed transform is then applied on the image (Figure 14, 5th row left). The various segments are then labelled. The final result can be seen in (Figure 14, 5th row right).
Segmentation of C-Spine MRI Images Using The Watershed Transform
This system works well (see numerical calculation of results, section 5.4). However, it is dependent on the size and structure of the structured elements that's used for the morphological operators. Therefore, some pre-knowledge is needed of the characteristics of the C-spine data, to yield optimum segmentation results.

The contours of the segmented areas have been superimposed on the input grayscale images (Figure 15 (a)). The labelled watershed image has been superimposed on the input grayscale image (Figure 15(b)). Not all the areas have been correctly segmented. Some areas (like the spinal cord in the neck) have been over-segmented. This can however be corrected with a merging algorithm.
Figure 15 Segmented areas superimposed on input image for system I

4.2 RESULTS OF SEGMENTATION SYSTEM II

The input image for this process can be seen in Figure 16. The segmentation results can be seen in Figure 17. The results of the histogram equalization can be seen in the top left image of Figure 17. The contrast has been enhanced by equalizing the histogram over all gray scale levels.

The threshold image can be seen in the top right image of Figure 17. Good results are obtained by first equalizing the threshold of the input image. The distance image is shown on the left of the middle row. The distance image is then filtered with a median filter to get rid of noise and outliers (middle row, right).

The watershed image is shown on the left of the bottom row. Very good results are obtained by applying the watershed algorithm on the distance image. The watershed image is labelled by
numbering each segmented area from top left to bottom right. The segmented areas are numbered in an increasing order.

This system yields similar results to system I, but is much simpler and quicker to execute. It's also not dependent on the size of the structured elements or masks used to do convolution.

The contours of the segmented areas have been superimposed on the histogram-equalized image. The results can be seen in Figure 18. The segmentation results are more accurate than compared...
to the results of System I (5.4) (Figure 15). The spinal cord has also been over-segmented, similar to segmentation system I.

Figure 18 Segmented areas superimposed on grayscale image (system II)

4.3 COMPARISON BETWEEN TWO SYSTEMS

The two systems have been applied to the same C-spine MRI image. The results can be seen in Figure 19 and Figure 20 respectively. The objects of system II are more accurately segmented. System I is also dependent on the relative size of objects. System II gives a simple and effective solution for the C-spine segmentation problem.
4.4 EVALUATION OF SEGMENTATION SYSTEM I ON DIFFERENT C-SPINE IMAGES

The segmentation system II has been applied on 9 different C-spine images (MRI-images from different patients). The results can be seen in the following sequence of figures. The top left image in each figure shows the original input MRI image. The top right image is the threshold binary image. The bottom left image is the distance image that is used as input for the watershed transform. The bottom right image is the result after the watershed transform.
Figure 21 C-spine image 1 segmentation

Figure 22 C-spine image 2 segmentation
Figure 23 C-spine image 3 segmentation

Figure 24 C-spine image 4 segmentation
Figure 25 C-spine image 5 segmentation

Figure 26 C-spine image 6 segmentation
Figure 27 C-spine image 7 segmentation

Figure 28 C-spine image 8 segmentation
4.5 EVALUATION OF SEGMENTATION SYSTEM II ON DIFFERENT C-SPINE IMAGES

Segmentation system II has also been applied on the same 9 different C-spine images (MRI-images from different patients). The results can be seen in the following sequence of figures.
The top left image in each figure shows the original input MRI image. The top right image is the threshold binary image. The bottom left image is the distance image that is used as input for the watershed transform. The bottom right image is the result after the watershed transform.

Figure 31 Segmentation of C-spine image 2

Figure 32 Segmentation of C-spine image 3
Figure 33 Segmentation of C-spine image 4

Figure 34 Segmentation of C-spine image 5
Figure 35 Segmentation of C-spine image 6

Figure 36 Segmentation of C-spine image 7
Figure 37 Segmentation of C-spine image 8

Figure 38 Segmentation of C-spine image 9
5. CHAPTER 5: INTERPRETATION OF RESULTS

5.1 RESULTS OF APPLICATION TO STANDARD TEST IMAGES

The segmentation system I has been tested on images from the Berkeley segmentation set [21], [22]. The original test image can be seen in the top left image of Figure 39. The threshold image can be seen in the top right image. The final watershed segmented images can be seen in the second row of Figure 39. The labels of the threshold image have been coloured randomly (grayscale) to increase the differentiation between segments (see right image of the middle row of Figure 39. Two manually segmented images (segmented by different people) can be seen in the third row of Figure 39. The various segments can be seen in different shades of gray.

Figure 39 Watershed Segmentation implemented on Berkely test image
The accuracy of the segmentation can be measured with sensitivity analysis by determining the receiver operating characteristic (ROC) or precision-recall plots of the segmentation results.

5.2 RESULTS OF APPLICATION TO MRI C-SPINE TEST IMAGES

9 Test images have been selected from C-spine MRI images from 9 different patients obtained from Radiology at Kloom Hospital in Pretoria. The test images can be seen in Figure 40. Ground truth images have been created by manual segmentation of the C-spine MRI test images. Binary segmented images have been created. All the segments in each image have been labelled. The manually segmented images can be seen in Figure 42.

Figure 40 C-spine MRI test images
Figure 42 Manually segmented (ground-truth) images
The edges of the manually segmented images have been calculated to yield ground truth edge images that can also be used for benchmark comparison measures (Figure 43).

The test images have been segmented with Segmentation System I and II. The results have been edge detected with a Canny algorithm and the results have been superimposed on the manually segmented (ground truth) edge images. The results can be seen in Figure 44 and Figure 45.
Figure 44: Edge images of test images superimposed on manually segmented images for System I

Figure 45: Edge images of test images superimposed on manually segmented images for System II
5.3 COMPARISON TO OTHER KNOWN SEGMENTATION PROCESSES

A comparison was done to a pyramid segmentation (multi-scale) algorithm. The results can be seen in Figure 46. A 255 level segmentation (containing 255 segments) image can be seen in Figure 46 (a). A 64 level segmentation (containing 64 segments) image can be seen in Figure 46 (b). A 16 level segmentation (containing 16 segments) image can be seen in Figure 46 (c). A 4 level segmentation (containing 4 segments) can be seen in Figure 46 (d). It is very evident that these results cannot compare to the results of the watershed segmentation system.

This segmentation algorithm will yield better results for images containing clustered areas of interest, like an area containing grassland, mountains, rivers or some crops.

FIGURE 46 PYRAMID SEGMENTATION RESULTS
5.4 NUMERICAL BENCHMARKS OF PERFORMANCE

To do a validation analysis, a very important component in the validation framework is a binary gold standard. The gold standard is the classification truth in terms of two mutually exclusive classes, for example tumor vs. non-tumor [24].

A popular effort for accessing classification or segmentation performance is a receiver operating curve (ROC). It is a function of sensitivity versus (1-specificity) at all possible decision thresholds. The sensitivity is the true positive rate and (1-specificity) is the false positive rate [25].

Two other quantities that can be used to calculate the performance of the segmentation: precision and recall. A precision-recall curve is generated and plotted for the algorithm for each segmented image. Precision and recall are similar, but different to ROC curves. Precision is the probability that a computer-generated boundary pixel is a true boundary pixel. Recall is the probability that a true boundary pixel is detected. Precision is a measure of how much noise is in the output of the detector. Recall is a measure of how much of the ground truth is detected. The curve shows the trade-off between these two quantities – misses and false positives – as the detector threshold changes.

The spatial overlap can also be used to measure the performance of the automated (watershed segmentation) vs. the ground truth. One parameter is the dice similarity coefficient (DSC), first proposed by Dice, which is a spatial overlap index and a reproducibility validation metric [24], [25].

The Berkeley segmentation benchmark (http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsdss/) was used to calculate the performance of the watershed segmentation system. I've also used Berkeley performance measurement tools to measure the performance of the watershed system. The Berkeley benchmark operates on the boundary map of the segmented images. The boundary map is compared to the ground truth boundaries. Ground truth boundaries are boundaries that are generated by a human expert. The image is divided into a number of segments corresponding to objects of interest in the image. The objects should have equal importance and not be part of the background or foreground. It is the objects that we want to separate and isolate from the rest of the image. The ground truth forms the benchmark against which the computer generated segmented images is compared.

The performance of the algorithm is also calculated in a single number. When two precision-recall curves do not intersect, then the curve furthest from the origin dominates the other. The summary statistic is a measure of this distance. It is called the F-measure, which is the harmonic mean of precision and recall [21], [26]. The F-measure is defined at all points on the precision-recall curve. The maximum F-measure value across an algorithm's precision-recall curve is given as its summary statistic [21], [27].
The performance of the segmentation algorithm can be determined by measuring the error between the ground truth segmentation boundary map and the computer-generated segmentation. The error in overlap can be measured.

Another measurement parameter that can be used to measure the performance of a segmentation system is the confusion matrix [28], [1]. The confusion matrix contains information about actual and predicted classifications done by a classification system. A confusion matrix is of size LxL, where L is the number of different label values.

a is the number of correct predictions that an instance is negative

b is the number of incorrect predictions that an instance is positive

c is the number of incorrect predictions that an instance is negative

d is the number of correct predictions that an instance is positive

\[
\begin{array}{cc|cc}
& & \text{Predicted} & \\
& & \text{Negative} & \text{Positive} \\
\text{Actual} & \text{Negative} & a & b \\
& \text{Positive} & c & d \\
\end{array}
\]

The following terms are defined for a two by two confusion matrix:

Classification accuracy:

\[
AC = \frac{(a + d)}{(a + b + c + d)} \quad (25)
\]

True positive rate (Recall, Sensitivity):

\[
TP = \frac{d}{(c + d)} \quad (26)
\]

True negative rate (specificity):

\[
TN = \frac{a}{(a + b)} \quad (27)
\]

Precision:
\[ P = \frac{d}{(b+d)} \] (28)

False positive rate:

\[ FP = \frac{b}{(a+b)} \] (29)

False negative rate:

\[ FN = \frac{c}{(c+d)} \] (30)

Classification accuracy, or its complement error-rate, defined as error:

\[ error = \frac{FN + FN}{N} = \pi_p FN_r + \pi_a FP_r \] (31)

This estimates the overall probability of correctly labelling a test sample.

\[ Recall = TP_r \] (32)

This indicates the probability of correctly detecting a positive test sample and is independent of class priors. \( TP_r \) is often utilised in medical applications where it is referred to as test sensitivity. The compliment to sensitivity is also used, namely specificity \((TN_r)\).

\[ Precision = \frac{TP}{TP + FP} \] (33)

This indicates the fraction of positives that are actually correct. Precision effectively estimates an overall posterior probability and is therefore a meaningful performance measure when detecting rare events 0.

\[ Posfrac = \frac{TP + FP}{N} \] (34)

Ground truth images have been created by manual segmentation of the C-spine MRI images. 9 ground truth images have been created. Binary segmented images have been created. All the segments in each image have been labelled. The edges of the segmented images have been created with the Canny edge detection method. The edges of the machine segmented images have also been calculated with the Canny edge detection algorithm. The machine segmented edge images have been compared to the manually segmented ('ground-truth') images. The number of matching pixels for the manually and machine segmented images have been calculated with a
Berkeley segmentation benchmark function. The precision, recall and F-measure parameters have been calculated for each ground truth image. The results can be seen in Table 1.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image No</td>
<td>System I</td>
<td>System II</td>
<td>System I</td>
</tr>
<tr>
<td>1</td>
<td>0.5915</td>
<td>0.5307</td>
<td>0.5392</td>
</tr>
<tr>
<td>2</td>
<td>0.5258</td>
<td>0.5167</td>
<td>0.5267</td>
</tr>
<tr>
<td>3</td>
<td>0.5304</td>
<td>0.5045</td>
<td>0.5284</td>
</tr>
<tr>
<td>4</td>
<td>0.4616</td>
<td>0.4279</td>
<td>0.4013</td>
</tr>
<tr>
<td>5</td>
<td>0.5381</td>
<td>0.495</td>
<td>0.5218</td>
</tr>
<tr>
<td>6</td>
<td>0.4924</td>
<td>0.4797</td>
<td>0.452</td>
</tr>
<tr>
<td>7</td>
<td>0.4605</td>
<td>0.495</td>
<td>0.3932</td>
</tr>
<tr>
<td>8</td>
<td>0.3762</td>
<td>0.3945</td>
<td>0.5011</td>
</tr>
<tr>
<td>9</td>
<td>0.4386</td>
<td>0.4314</td>
<td>0.4683</td>
</tr>
</tbody>
</table>

Table 1 Performance measurement results
The precision-recall graphs have been generated for the boundary maps of the MRI segmentations that have been generated by System I can be seen in Figure 47. The best F-measure has been calculated for each test figure and is indicated in red on each plot.

Figure 47 Precision-recall graphs for Segmentation System I
The precision-recall graphs for the segmentation that has been performed with System II can be seen in Figure 48. The best F-measure has been calculated for each test figure and is indicated in red on each plot.

**Figure 48 Precision-recall curves for Segmentation System II**
A comparison of the manually segmented images and the machine segmented watershed images segmented with System II can be seen in Figure 48. The difference in relative position of the edges of the images is shown in the figure below.

![Figure 48 Ground truth vs. machine segmented image comparison](image)

5.5 SUBJECTIVE PERFORMANCE COMPARISON

Feedback was received from two radiologists. Their general feedback is that the segmentation results look good in terms of detection, localization and single detection. The detection of edges of bony structures is clear and useful. The bony structures are accurately segmented and will aid in diagnosis and treatment of patients.

5.6 SUMMARY

The C-spine images have been successfully segmented. Some over segmentation has occurred in certain images.
6. CHAPTER 6: SUMMARY

6.1 GOALS REVISITED

Techniques have successfully been developed to segment C-spine MRI images. A combination of thresholding, morphological processing and the watershed transform has been implemented to yield a promising segmentation system. These techniques produce promising results. With some adjustments and refinement of the algorithm, a very useful system can be produced to support physicians in their decisions in the diagnosis, treatment, and surgery of different pathologies of the spine.

6.2 CRITICAL ANALYSIS OF THE WORK PERFORMED IN THE DISSERTATION

I did the performance analysis on the results of the segmentation algorithms on the boundary maps of the segmented images. There is a number of reasons for doing this rather than measuring the overlap of segmented areas:

- it is far more discriminative
  - it is more sensitive to boundary complexity
  - it does not match offset interiors
- it is an intuitive measure and unit
  - it is a detection-oriented framework
- this is both a low-level and mid-level benchmark
  - it is sensible for evaluating both edges and regions
  - you can quantify improvement between levels

As mentioned before precision-recall curves have been used a performance measure. Precision is the fraction of detections that are true positives rather than false positives. Recall is the fraction of true positives that are detected rather than missed. In probabilistic terms, precision is the probability that the detector's signal is valid, and recall is the probability that the ground truth data was detected.

Although ROC and PR curves qualitatively show the same trade off between misses and false positives, ROC curves are not appropriate for quantifying boundary detection. Fallout is not a meaningful quantity for a boundary detector since it depends on the size of the pixels.
I've found that by simply corresponding coincident boundary pixels and declare all unmatched pixels either false positives or misses, is too strict. This approach doesn't tolerate any localization errors and consequently over-penalize an algorithm that generates useful, though slightly mislocalized boundaries. From the previous chapter it is clear that the assignment of machine boundary pixel s to ground truth boundaries must allow localization errors since even the ground truth data contains localization errors. Allowing small localization errors yielded significantly better precision-recall curves. However, the bony structures have still been accurately segmented, although some localization errors occurred.

The intention with this approach is to calculate precision and recall matches as closely as possible the impression one would have if scoring the outputs visually. All three desirable properties of a segmentation system of boundary detector – detection, localization and single detection – are encouraged by the method and visible in the results.

Over segmentation occurred in the segmentation of some of the MRI images. This leads to an increase in false positives causing a drop in the precision performance indicator. This will also lead to a drop in the F-measure.

6.3 POTENTIAL FUTURE WORK

The watershed segmentation tends to over-segment the distance image. These segments need to be merged. Care need to be taken not to over-merge or under-merge these segments. The watershed algorithm might need to be updated to prevent over-segmentation of the distance image.

Possible future work on the segmentation of MRI images:

- Investigate the application of these techniques to CT images.
- Investigate the segmentation of lung tumours or lesions with morphological processing and the watershed transform.
- Investigate the segmentation of tumours on mammography images to determine the size, localization and in growth in volume of these tumours.
- Investigate the application of these techniques to 3-dimensional MRI C-spine images
- Investigate the application of the segmentation systems described in this dissertation on MRI brain and thorax images.
- Investigate the combination of these techniques with the normalised cuts method.
REFERENCES


Morphological and Watershed Segmentation of C-Spine MRI Images

Jacobus J Botha, André Nel

Department of Electrical and Electronic Engineering Science
University of Johannesburg, South Africa
aln@ing.rau.ac.za

Abstract
Segmentation of bony structures plays an important role in image-guided surgery of the spine. In this, a novel approach to the segmentation of vertebral bodies from 2D sagittal magnetic resonance images of the spine is discussed. We would like to show that morphological and watershed segmentation could be used as a powerful segmentation tool in the segmentation of cervical (C)-spine MRI images. The structure, shape and morphology of the bony structures in spine images lend itself to the watershed segmentation.

1 Introduction
Automatic classification of images has always been an important part of pattern recognition. The segmentation and classification of magnetic resonance imaging (MRI) images have always been a challenge. A segmented image is often a very important input to the classification process. Many classification techniques use segmented images as input to the classification process. Certain segments or areas of an image serve as important features that will be used for classification. Important information can be derived from the features that are present in the segmented image. Sometimes there might be a need to extract a certain object from an image to do classification on the object.

In the case of MRI images, the segmentation process can isolate certain structures of the human body like organs and tissue. These objects of interest (OOI) can give vital information for the identification of medical abnormalities (anomalies) and diseases. Segmented objects can play an important role to assist medical practitioners in the diagnosis and treatment of medical problems.

Many segmentation and classification techniques have been implemented on MRI images. The latest techniques include support vector machines (SVMs), neural networks (NNs), statistical methods and threshold techniques.

The goal of this work was to test the performance of the watershed segmentation algorithm on MRI images of the vertebral spine.

2 Segmentation Process
Segmentation is the partitioning of images/volumes into meaningful pieces [1]. Segmentation involves the isolation on a specific region of interest. Segmentation subdivides an image into objects of interest. These objects are often referred to as regions of interest (ROI) or volumes of interest (VOI). In this paper we will only use the term ROI, whether they are 2D or 3D.

Segmentation plays a key role in the identification of objects and features and subsequent classification of images. Objects of interest can be isolated and extracted by the segmentation process. The segmented objects can be used as input features for the classification process. The segmented regions can also be used as inputs to intelligent processes like fuzzy logic systems and neural network systems.

The purpose of image segmentation is:
- Detection or recognition of an object or feature of an object
- Quantifying of the properties of an object, like the size and quantity of the object
- Identification of statistical, morphological and geometrical properties of the object.

![Figure 1 Segmentation classification](image_url)

There are three types of image segmentation: thresholding, edge-based and region-based segmentation (Figure 1). Segmentation techniques...
can also be classified as statistical segmentation, morphological and region-based segmentation (Figure 2).

### Figure 2 Segmentation classification (2)

#### 3 Segmentation in MRI

Segmentation is a key process in the image processing of MRI images. Physical structures like organs, bones and tissues can be isolated in the segmentation process. These regions of interest (ROI) will play a crucial role in further analysis and diagnosis of abnormalities and diseases. Specific objects of interest like tumours, fractured bones, lesions, etc. can be isolated and identified in the segmentation process [3].

Many segmentation techniques have been developed for MRI images. Statistical techniques like expectation/maximization, binary mathematical morphology and active contour models have been implemented. Nearest-neighbour techniques have also been implemented.

Scale-invariant segmentation [4], a combination of expectation-maximization segmentation, binary mathematical morphology and active contour models are also used [5]. A combination of texture based segmentation and neural network classification can also be used.

#### 4 Morphological Processing Theory

A typical image processing system is described in Figure 3. It typically consists of a number of stages, i.e. pre-processing, thresholding, morphological processing, segmentation and classification. Sometimes the thresholding and classification stages are omitted. The focus in this paper will be on thresholding, morphological processing and segmentation.

The method that has been used includes a combination of processes. Thresholding, morphological processing, edge-detection, filtering and the watershed segmentation have been used to yield a segmented image.

Some pre-processing like filtering or histogram equalization needs to be applied on the MRI images to improve segmentation of the raw data.

#### Morphological Operations

- Erosion
- Dilation
- Closing
- Opening
- Morphological reconstruction
- Extended-minima transform
- Extended-minima transform

The input processing includes morphological operations. Morphology is a broad set of image processing operations that process images based on shapes. Morphological operations that can be used as powerful pre-processing techniques are discussed in the following paragraphs.

Erosion removes pixels from object boundaries. It assigns the minimum value of all the neighbouring pixels in the structuring element to the pixel that is currently evaluated. It can be used on grayscale or binary images. In a binary image, the pixel will be set to zero, if any pixel in the structuring element is zero. The equation for erosion is as follows:

\[
(f \oplus g)(x) = \bigvee_{y \in E} f(y) + g(x - y)
\]

where \(\oplus\) refers to the Minkowski addition \(0\).

Dilation adds pixels to boundaries of an object. It assigns the maximum value of all the neighbouring pixels in the structuring element to the pixel under test. In a binary image, the pixel will be set to one, if any pixel in the structuring element is one. In algebra, any transformation which is increasing, anti-extensive and idempotent
is called an (algebraic) opening, increasing, extensive and idempotent is called a (algebraic) closing.

\[(f \ominus g)(x) = \bigwedge_{y \in E} f(y) - g(x - y) \quad (2)\]

where \(\ominus\) refers to Minkowski subtraction.

Closing is a dilation operation followed by erosion with the same structuring element. Morphological closing can be used to join shapes and structures and to fill up gaps. It can also be used to fill in basins and valleys in a topological image [1]. The resulting structures can be used as markers.

\[f \mapsto (f \ominus g) \ominus g \quad (3)\]

Opening is erosion followed by dilation with the same structuring element. Morphological opening can be used to remove small objects from an image while preserving the larger objects in an image. An opening removes some peaks and crest lines.

\[f \mapsto (f \ominus g) \oplus g \quad (4)\]

Morphological reconstruction processes one image, called the marker, based on the characteristics of another image, called the mask [1]. Morphological reconstruction can be used to reconstruct the topology of an image.

The regional extrema of a raw image mark relevant as well as irrelevant image features. The h-maxima transformation suppresses all maxima whose depth is lower or equal to a given threshold level \(h\). This is achieved by performing the reconstruction by dilation of \(f\) from \(f - h\):

\[\text{HMAX}_h(f) = R^h_f(f - h) \quad (5)\]

The extended-minima transform calculates the regional minima of the corresponding h-minima transform. Regional minima are connected components of pixels with the same intensity value, whose external boundary pixels all have a lower value. The extended minima are defined by:

\[\text{EMIN}_h(f) = \text{RMIN}[\text{HMIN}_h(f)] \quad (6)\]

The extended-maxima transform calculates the regional maxima of the k-maxima transform. Regional maxima are connected components of pixels with the same intensity value, whose external boundary pixels all have a lower value. The extended maxima \(\text{EMAX}\) is defined by:

\[\text{EMAX}_h(f) = \text{RMAX}[\text{HMAX}_h(f)] \quad (7)\]

An important input to segmentation and image classification is the gradient image. The gradient image gives an indication of the level of the gradients (peaks and valleys) in an image. The gradient image can be calculated with the derivative of a Gaussian filter. Gradient images can be used to calculate edges in images. It can also be used as an input to the watershed transforms 0. The gradient of a function of two variables, \(F(x,y)\) is defined as:

\[\nabla F = \frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} \quad (8)\]

A distance transform can be used to calculate the gradient of an image. The distance transform provides a measure of the separation of points in an image. The distance transform calculates the distance of the current pixel to the nearest non-zero pixel of the input binary image. The distance function \(d_{st}\) of a set of \(X \subset Z^2\) associates the distance of each pixel to the background [12]:

\[d_{st}(p) \mapsto \min\{d_{st}(p,q) | q \notin X\} \quad (9)\]

The distance function \(d_t\) of a binary image \(I\) is equivalent to that of its set of feature pixels, i.e. pixels with value 1. In addition we put conveniently:

\[\forall p \in D_1, I(p) = 0 \Rightarrow d_t(p) = 0 \quad (10)\]

The distance transform can be calculated with several distance metrics, for example the Euclidean distance, city block, chessboard and the quasi-Euclidean distance.

An image can be enhanced or modified by applying an image filter. A filter can be applied to emphasize or remove certain features. Image processing operations implemented with filtering include smoothing, sharpening, and edge enhancement. Many different kinds of filters can be applied to an image to enhance to image.

**5 MRI Segmentation System**

Two systems have been implemented. The flow diagrams of these systems can be seen in Figure 4.
The first system will be described in the next section.

5.1 Segmentation System I

The first step is to threshold the image with a fixed threshold. The threshold is calculated using Matlab's `graythresh` function. This function uses Otsu's method [13], calculating a threshold that minimizes the intraclass variance of the black and white pixels.

```
Figure 4 Flow diagrams of system I and II
```

The next step is to implement a closing function. The `bwmorph` function in Matlab was used to perform this operation. This function performs a morphological dilation followed by erosion. The operation is implemented 30 times, until the process settles and no more changes happen in the output image.

The next step is to perform a two-dimensional convolution. A Gaussian mask of 10-by-10 pixels is used. Matlab's `conv2` function is used to perform this operation. This operation enhances block-shape structures in the input image.

An erosion operation is then performed on the convoluted image. A disk shape structure with a radius of 5 is used. The disk shape structure is approximated by a sequence of 4 periodic-line structuring elements.

The next step is to threshold the image with a regional minima transform of the H-minima transform. Regional minima are connected components of pixels with similar intensity values, and whose external boundary pixels all have a higher value. This transform is very useful as a pre-process for the segmentation process. The Matlab function `imextendedmin` is used to do this thresholding.

The next step is to implement a distance transform. This function calculates the Euclidean distance between each non-zero pixel and the nearest non-zero pixel. The Matlab function `bwdist` is used for this operation. This operation yields a gradient image giving an indication of the distance between objects and neighbouring objects in the image.

The sign of the values of the output image is inverted. Pixels belonging to the background of the image are set to minus infinity.

A median filter is implemented next to remove outliers and noise. This is a non-linear operation that reduces noise significantly, but preserves edges. Matlab's `medfilt2` is used for this operation.

The final step is to calculate the watershed transformation. This function calculates the watershed regions, yielding the final segmented image. We used a modified version of the Vincent Soille watershed algorithm to perform this operation.

5.2 Segmentation System II

This system is much simpler and only consists of a small number of steps:

The first step is to equalize the histogram of the input image. This step enhances the contrast of the image. It is done with the function `jcontrast`.

The second step is to threshold the image to create a binary image. The Matlab function `graythresh` is used to calculate the threshold. The thresholding is done with the function `im2bw`. All pixels above the threshold are set to one and pixels
below the threshold is set to zero. The output of this function is a binary (black-and-white) image.

The next step is to implement a distance transform. This function calculates the Euclidean distance between each non-zero pixel and the nearest non-zero pixel. The Matlab function \texttt{bwdist} is used for this operation.

The sign of the values of the output image is complemented. Pixels belonging to the background of the image are set to minus infinity.

A median filter is implemented to get rid of outliers and noise. This is a non-linear operation that reduces noise significantly, but preserves edges. Matlab's \texttt{medfilt2} is used for this operation.

The final step is to calculate the watershed transformation. This function calculates the watershed regions, yielding the final segmented image. A modified version of the Vincent Soille watershed algorithm was used to perform this operation.

6 Results

The two systems have been applied to the same C-spine MRI images. The results can be seen in Figure 5 and Figure 6. The objects of system II have been more accurately segmented (see Table 1).

6.1 Results of Application to MRI Test Images

Test images have been selected from MRI C-spine images from 9 different patients obtained from Radiology at Kloof Hospital. Ground truth images have been created by manually segmenting the C-spine MRI test images. Binary-segmented images have been created. All the segments in each image have been labelled. Input test MRI images can be seen in the first row of Figure 7. The filled and boundary images of the manual segmentations are in the 2nd and 3rd row respectively of Figure 7.
Precision-recall curves were generated and plotted for the algorithm for each segmented image (Figure 8).

![Figure 8 Precision-recall curve for a test image](image)

Precision is the probability that a computer-generated boundary pixel is a true boundary pixel. Recall is the probability that a true boundary pixel is detected [14]. The F-measure, which is the harmonic mean of precision and recall, is also calculated.

Table 1 Performance numbers for test images

<table>
<thead>
<tr>
<th>Im No</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sys I</td>
<td>Sys II</td>
<td>Sys I</td>
</tr>
<tr>
<td>1</td>
<td>0.59</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td>2</td>
<td>0.53</td>
<td>0.52</td>
<td>0.53</td>
</tr>
<tr>
<td>3</td>
<td>0.53</td>
<td>0.50</td>
<td>0.53</td>
</tr>
<tr>
<td>4</td>
<td>0.46</td>
<td>0.43</td>
<td>0.40</td>
</tr>
<tr>
<td>5</td>
<td>0.54</td>
<td>0.50</td>
<td>0.52</td>
</tr>
<tr>
<td>6</td>
<td>0.49</td>
<td>0.48</td>
<td>0.45</td>
</tr>
<tr>
<td>7</td>
<td>0.46</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td>8</td>
<td>0.38</td>
<td>0.39</td>
<td>0.50</td>
</tr>
<tr>
<td>9</td>
<td>0.44</td>
<td>0.43</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Looking at the F-measures in Table 1, System II performed better than system I on all test images.

7 Conclusion

The results show that a combination of morphological functions and the watershed transform can be used to segment the C-spine from the rest of the MRI-image. System II performed better than System I and gives a simple and effective solution for the C-spine segmentation problem.

8 Acknowledgements

The authors would like to thank the Radiology department at Kloof Hospital in Pretoria for kindly assisting with the MRI images.

9 References


