Improvements to Quantification Algorithms for Myocardial Infarction in CMR Images

- Validation in Human and Animal Studies

Felicia Seemann

Faculty of Engineering, Lund University, October 24, 2013
Supervisors: Jane Tufvesson and Einar Heiberg

Abstract

Cardiac magnetic resonance (CMR) images are used to investigate the heart for medical and research purposes. By injecting a contrast substance into the patient, myocardial infarctions (heart attacks) can be visualized in CMR image sets consisting of a number of image slices at different levels of the heart. Analysis of these images can detect an infarction, delineate it and estimate its size. This information is then processed by physicians in order to make a diagnosis and decide the course of treatment. Manual delineations are time consuming and observer dependent, why an automated algorithm is desired. Previous work presents a validated automatic segmentation algorithm that calculates a threshold used to separate the healthy tissue pixels from the infarction pixels, based on a fixed number of standard deviations. Theoretically, it is known that algorithms based on standard deviations are likely to be influenced by noise. Therefore, the aim of this thesis was to investigate if other techniques could be used to compute a threshold that is less noise sensitive in both humans and animals.

The study included 40 humans and 18 pigs. Two different techniques based on an Expectation-Maximization algorithm for threshold calculation was developed and implemented into the previous presented method. One implementation analyses each image slice separately (the slice method), and one takes all slices into account at once (the set method). The algorithms were evaluated by comparing computed infarction volume to volumes computed from manual delineations. Both algorithms show good agreement and low bias with the reference standard. The slice method yielded the best results on animal data with a high resolution. The set method yielded the best results in human CMR images, and it show an improved robustness for increasing noise levels. Both implementations show potential for fully automatic quantification of myocardial infarction.

Populärvetenskaplig sammanfattning

Denna masteruppsats tar fram en metod för att uppskatta volymen på hjärtinfarkter utifrån MR-bilder. Hjärtat är ett vitalt organ för att hålla oss vid liv. Dess uppgift är att pumpa blod för att förse alla organ med syre. Även hjärtmuskeln själv ska syresättas med blod från hjärtat och detta görs genom kranskärl. En förträngning i ett kranskärl kan orsaka syrebrist i hjärtmuskeln vilket leder till vävnadsdöd, en hjärtinfarkt. Det kan vara ett livsfarligt tillstånd då det drabbade hjärtat får svårt att pumpa. Hjärt- och kärlsjukdomar är i dagsläget den vanligaste dödsorsaken i världen. Det ställs stora krav på att korrekt diagnosticering kan ske i ett tidigt skede.

MR-kameran möjliggör visualisering av interna strukturer i kroppen genom att ta till vara på olika vävnaders magnetiska egenskaper, samtidigt som det är ofarligt för patienten. Genom att injicera en substans i kroppen som absorberas av muskel som innehåller infarkt åskådliggörs infarkten i MR-bilderna. Frisk muskelvävnad framställs som grå medan infarkten har en ljusare nyans i de svartvita bilderna. Genom att analysera bildserier som visar hjärtat på olika nivåer kan frisk vävnad och infarkter klassificeras.

Att rita ut infarkter för hand kräver en gedigen erfarenhet. Det är tidskrävande och subjektivt beroende på vem som hanterar bilderna. Därför är det av intresse att analysera bilderna i datorprogram som kan spara tid och ge objektiva resultat. Infarkten kan då markeras i bilderna och infarktsvolymen beräknas, vilket bistår läkare i deras arbete.

En metod som används idag för att detektera infarker i MR-bilder utgår från att mängden brus i bilderna ligger på ungefär samma nivå, oavsett patient och vilken MR-kamera. Syftet med uppsatsen är att modifiera metoden så att den klarar av att behandla bilder oberoende av brusnivå. Den nya metoden baseras på en algoritm som delar upp pixlarna i frisk och skadad vävnad utan antaganden kring bildkvaliteten.

Den utvecklade metoden har utvärderats på bilder på människor och djur genom att den beräknade infarktvolymen har jämförts med manuella volymsuppskattningar. Resultaten stämmer väl överens med varandra. Slutsatsen är att den nya metoden avsevärt förbättrat detekteringen av infarkter i högupplösta bildserier på djur. Resultaten i levande patienter är likvärdiga med den tidigare metoden under normala brusförhållanden, men har påvisat ett bättre resultat i bilder med hög brusnivå.

Acknowledgements

This is a master thesis in engineering mathematics at the Faculty of Engineering at Lund University. It has been carried out in the year of 2013 from June to October. The project has been executed at the Cardiac MR group, a Swedish research group represented by Lund University, the Department for Clinical Physiology at Skåne University Hospital in Lund, and the company Medviso. I wish to thank everybody in the Cardiac MR group for their welcoming and open atmosphere, their support and all the help that has been given to me. A special thanks to my supervisors Jane Tufvesson and Einar Heiberg, and to the fellow student Lisa Eneroth who has written her master thesis in engineering physics at the Cardiac MR group during the same period of time. I would also like to thank my family and friends for their infinite support in everything I do.

Abbreviations and Terminology

CMR	cardiac magnetic resonance
EM	expectation-maximization
Ex-vivo	outside the living organism
In-vivo	within the living organism
LV	left ventricular
LVM	left ventricular mass
MO	microvascular obstruction
MR	magnetic resonance
Myocardium	heart muscular tissue
SD	standard deviation

Contents

Chapter 1

Introduction

Heart attacks are one of the most common causes of death in the world and heart disease is one of the most expensive treatments in medical care [1]. When a heart attack (medically known as a myocardial infarction) has occurred, it is of great importance to study the extent of the scarred tissue in order to opt for a suitable treatment.

Magnetic resonance (MR) imaging is a tool that enables physicians to observe internal body structures non-invasivly. It is the gold standard method for determining the size of the heart and myocardial infarctions in medicine. Manual segmentation is time consuming and observer dependent, hence automatic segmentation methods can be of great use to extract objective information from the images [2].

In the Lund Cardiac MR group two different approaches are used for infarction delineation, one for patients (in-vivo studies) and another for explanted hearts in experimental animal studies (ex-vivo studies). Both approaches were presented by Heiberg et al [3] in 2008. The methods computes the mean and standard deviation of an area assumed to consist of normal healthy tissue (remote). The pixels are classified as infarction or remote tissue by calculating a threshold at a fixed number of standard deviations from the remote mean value. In 2011 Sjögren et al [4] presented an algorithm for segmentation in images with poorer contrast and showed that an Expectation-Maximization (EM) algorithm yielded better results than standard deviations from remote. Therefore the purpose of this study is to investigate if an EM-algorithm can improve the delineation of myocardial infarctions in both humans and animals.

1.1 Thesis overview

Chapter 2 in this thesis provides a short medical background to the heart and cardiac magnetic resonance (CMR) imaging. In Chapter 3 the mathematical theory is explained. Chapter 4 presents the aim of this thesis. Chapter 5 presents the methods used and Chapter 6 the validation strategies. The results are presented in Chapter 7 and Chapter 8 consists of conclusions and discussions.

Chapter 2

Medical Background

The algorithm developed in this thesis is designed to quantify myocardial infarction, the size of the heart attack, in the left ventricle as imaged by magnetic resonance.

2.1 Heart anatomy

The heart is a vital organ in the body in order to keep us alive. It is the organ responsible for pumping blood through the circulatory system in order to supply the body's organs with oxygenated blood [5].

The cardiac wall is called myocardium and consists mainly of muscular cells, it gets its blood supply from the coronary arteries. The myocardium chambers four cavities; two atria and two ventricles. The inside surface of the myocardium is called the endocardium and the outer boundary is called epicardium. The left ventricle (LV) pumps oxygenated blood to the whole body, causing the LV myocardium to be subject to the highest load and pressure within the heart. It has thicker walls and demands a higher blood supply than the right ventricle (RV), right atria (RA) and left atria (LA) [2][5]. A schematic illustration of the heart is shown in Figure 2.1 in the left panel, and an MR image of the four chambers is shown in the right panel.

2.2 Myocardial infarction

A myocardial infarction is the event of necrosis in the myocardium due to lack of oxygen, and may lead to serious symptoms and death. The lack of oxygen occurs due to partially or completely clogged coronary arteries [5]. Myocardial infarctions often occur in the LV myocardium due to the structure of the coronary arteries.

2.3 Magnetic resonance imaging

The magnetic resonance (MR) camera supplies a magnetic field on the body and detects the different structures due to the different magnetic properties of the tissues. It is a method used for visualizing internal images of cross sections inside the body with a high resolution, without causing any harm to the patients. Cardiac magnetic resonance (CMR) is the gold standard for measuring the size of myocardial infarction and of the heart. The right panel in Figure 2.1 shows a long-axis CMR projection of a human heart.

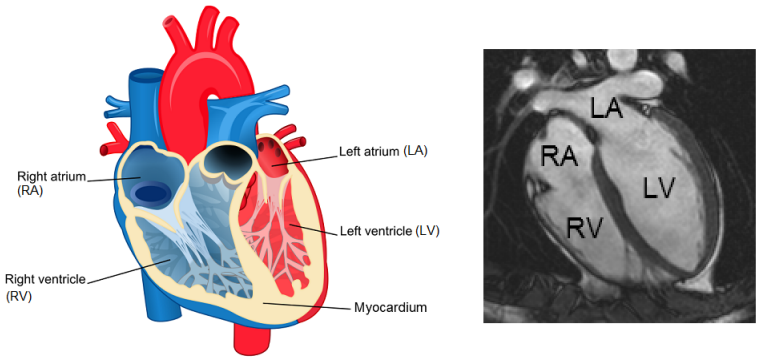

Figure 2.1: *Left panel illustrates the human heart, adopted from [6]. Right panel shows an MR image of a human heart in long-axis projection, adopted from [2]. The lighter area is blood and the surrounding darker area is the myocardium. LA is the left atrium, RA the right atrium. RV is the right ventricle and LV is the left ventricle.*

In order to visualize and quantify the myocardial infarction, contrast enhanced CMR images in the short-axis projection are used. They are acquired by injecting the patient with a contrast substance that is absorbed by the scarred tissue. This results in infarcted myocardium with lighter intensities compared to the darker healthy myocardium.

2.3.1 Ex-vivo and In-vivo imaging

Ex-vivo images are obtained by imaging explanted hearts, while in-vivo images are captured within a living patient. The main difference between ex-vivo and in-vivo images is the resolution. Ex-vivo captions allow clear images with a high resolution (about $0.5 \times 0.5 \times 0.5$ mm), a typical image set consists of over 150 short-axis image slices. In-vivo sets usually has a resolution of $1.5 \times 1.5 \times 8$ mm, consists of 12 slices, and they are noisier than the ex-vivo images. The patient's breathing and movements from heartbeats is the source of this extra noise. Despite fewer

slices than in ex-vivo data sets, manual delineation in in-vivo data sets is still time consuming and observer dependant. Figure 2.2 shows an example of the short-axis projection to the left, an ex-vivo animal infarction segmentation in the middle and an in-vivo human segmentation to the right.

2.3.2 Imaging artifacts

The presence of noise in the CMR images is a challenge that has to be overcome in order to implement an automatic segmentation algorithm. Noise is captured in all measurements, but it can be more or less present depending on facts such as camera settings and movements during the image acquisition.

The contrast substance is injected into the patient so that the infarction can be visualized in the CMR images. In some cases the contrast is not completely distributed within the infarction due to obstruction on a microvascular level. This leads to a dark core surrounded by light intensities. The phenomenon is called microvascular obstruction (MO) and occurs in both humans and animals. MO should be classified as infarction despite its dark color in the images.

Another event that may occur due to the contrast substance is the visualisation of artifacts. The contrast is not exclusively absorbed by the scarred myocardium, which may result in lighter areas that should not be classified as infarctions. Examples of sources of artifacts are blood vessels, movements while capturing an image or camera settings. Artifacts that are detected by automatic segmentation algorithms must be manually removed by experienced observers afterwards. The right panel in Figure 2.2 show a human in-vivo image with both MO and detected artifacts present.

2.4 Software

Segment is a validated software for cardiovascular image analysis that is freely available for research and educational purposes. It is developed in Matlab by the Lund Cardiac MR group at Lund University, and a commercial version is sold by the spin-off company Medviso *(http://segment.heiberg.se)*. The software is especially developed for MR images, but has expanded to also handle other cardiac imaging techniques. All cardiac image analysis in this thesis are carried out in Segment v. 1.9 [7] and the developed algorithms are based on a framework from previous methods in the software.

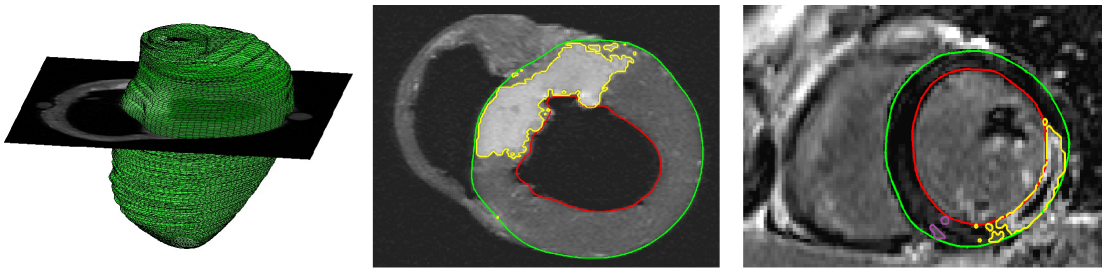

Figure 2.2: *Left panel shows a 3D model of the LV epicardium of a pig with an illustration of the position of a short-axis projection. Middle panel shows an example of an ex-vivo CMR image of a pig heart in short-axis projection with a segmented myocardial infarction in yellow. Right panel shows a short-axis projection of a human heart in-vivo, with MO viewed as a dark core inside the infarction area. Green indicates the epicardium and red lines the endocardium. Yellow lines illustrates infarction segmentation. Purple lines shows areas that are detected by the segmentation algorithm but should be considered artifacts rather than infarction.*

Chapter 3

Mathematical Theory

In image analysis, segmentation is the task of distinguishing different objects or patterns from each other within a picture [8]. With our eyes, we easily categorize different structures and objects in images. It may be the presence of a face, a pair of shoes, a red dot or an area of bright myocardium in a CMR image. For computers to automatically execute the same task, different segmentation techniques are required. The technique used in this thesis is segmentation based on clustering.

Clustering is the task of determining which components of a data set that naturally belong together [8]. Applying this to CMR imaging, the components are pixels and the data set consists of one or more images. The task is to partition the pixels into two groups; normal myocardium and scarred myocardium in the CMR images. All images are gray scaled and represented as a matrix where each element represents a pixel. All pixel intensities in the images are normalized between 0 and 1, where 0 corresponds to black pixels and 1 corresponds to white.

This chapter presents the mathematical theory used in this thesis. The outline is as follows:

1) The CMR images are assumed to consist of a mixture of Gaussian distributed pixel intensities.

2) A threshold that separates the healthy tissue from the scarred tissue is computed by implementing the k-means algorithm and the EM-algorithm into the weighted method, in order to delineate the infarction and compute its volume.

3) Otsu's method is used for distinguishing image slices that are completely infarcted from slices consisting of only healthy tissue.

4) The SD from remote method is noise sensitive and, when applied on ex-vivo images, time consuming. It would preferably be replaced by an automatic segmentation method.

3.1 Gaussian mixture densities

When data from two or more Gaussian distributions are mixed into one single data set, the set is said to consist of a Gaussian mixture density.

Definition 3.1.1. A Gaussian mixture density $p(x|\mu,\sigma)$ is the weighted sum of $K \geq 2$ Gaussian probability densities

$$g(x|\mu_i,\sigma_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \cdot exp\Big\{ - \frac{(x-\mu_i)^2}{2\sigma_i^2} \Big\},\ i=1,...,K$$

with mean μ_i and standard deviation σ_i,

$$p(x|\mu,\sigma) = \sum_{i=1}^{K} \alpha_i \cdot g(x|\mu_i,\sigma_i)$$

where x is a data vector and α_i is the components weights with $\sum_{i=1}^{K} \alpha_i = 1$ [8].

Figure 3.1 illustrates an example of a Gaussian mixture density $p(x|\mu,\sigma)$ with $K = 3$ components $g(x|\mu_i,\sigma_i)$.

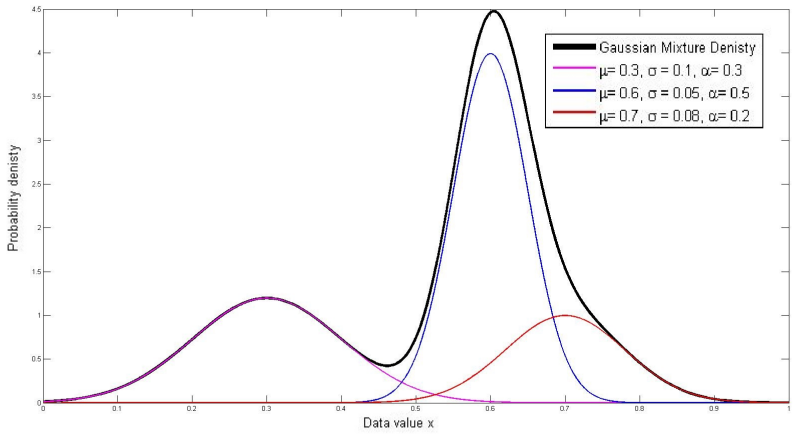

Figure 3.1: *The black line shows an example of a Gaussian mixture density with K=3 components. The purple, blue and red lines shows the corresponding weighted Gaussian probability densities.*

3.2 The k-means clustering algorithm

The k-means algorithm is an iterative method that partitions N data components into K clusters so that each data point belongs to the cluster with the closest cluster center [9]. The pseudocode for the k-means method with the input data vector x is given in Algorithm 1. If no a priori information for the initial guess of cluster centres is available, the Forgy initialization [10] that chooses K random centres from the data vector x can be applied.

Data: Vector x containing N components to be partitioned.
Result: A data vector classifying each component in x to a cluster $i = 1, ..., K$ and K different cluster centers $c_1, ..., c_K$.
Initialization: determine K initial cluster centres $c_1^{(0)}, ..., c_K^{(0)}$.
Convergence criteria: When the assignments does not change any more.
while *no convergence* **do**
  Assignment step;
  Find the cluster centre c_i with the least euclidean distance to each component. Assign every component to cluster $i = 1, ..., K$ accordingly.

  Update step;
  Find the mean of all components assigned to each cluster and update the cluster centres $c_1, ..., c_K$ to the corresponding mean value.
end

Algorithm 1: *K-means algorithm [9].*

3.3 Expectation-maximization algorithm

The Expectation-Maximization (EM) algorithm is an iterative method that estimates the parameters θ in statistical models, where some data the model depends on are unknown. It alternates between conducting two steps until convergence is reached; the expectation step (E-step) and the maximization step (M-step) [11]. In this section the EM-algorithm for Gaussian mixture densities will be described.

Assume that the data $x = [x_1, ..., x_N]$ is given. The unknown data is the vector $z = [z_1, ..., z_N]$ that indicates which class $k = \{1, ..., K\}$ each element in x belongs to.

$$x_i|(z_i = k) \sim N(\mu_k, \sigma_k),\ i=1,...,N$$

$$P(z_i = k) = \alpha_k,\ where\ \sum_{k=1}^{K} \alpha_k = 1.$$

The aim is to estimate the parameters $\theta = (\alpha, \mu, \sigma)$, where $\alpha = \{\alpha_1, ..., \alpha_K\}$, $\mu = \{\mu_1, ..., \mu_K\}$ and $\sigma = \{\sigma_1, ..., \sigma_K\}$.

The initialization of the algorithm is to guess the initial parameters $\theta^{(0)}$. The E-step is carried out by calculating the weighted normal probability density functions $\alpha_k \cdot g(x; \mu_k, \sigma_k)$ and classifying each data component x_i to the class k by computing the responsibilities

$$\gamma_{k,i} = \frac{\alpha_k \cdot g(x_i; \mu_k, \sigma_k)}{\sum_{l=1}^{K} \alpha_l \cdot g(x_i|\mu_l, \sigma_l)} = \frac{p(x_i|\theta_k^{(t)})}{\sum_{l=1}^{K} p(x_i|\theta_l^{(t)})}$$

for each class k (Figure 3.2).

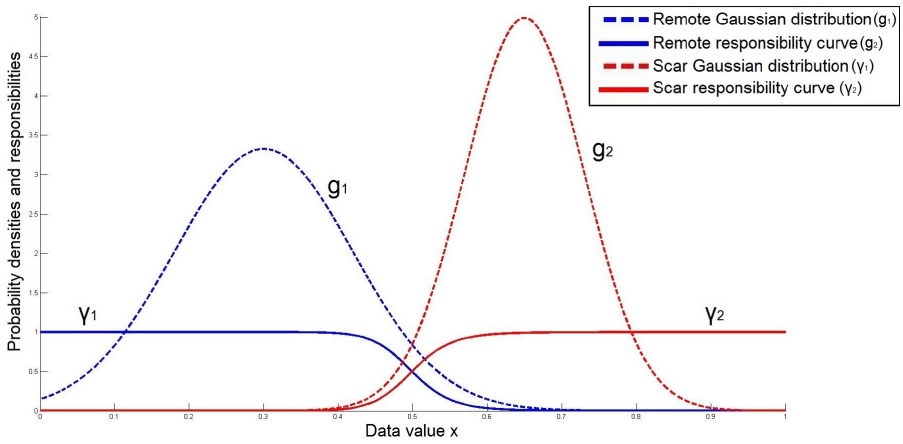

Figure 3.2: *Dashed lines represent the remote (g_1) and scar (g_2) probability densities and their corresponding responsibility curves γ_1 and γ_2 are presented as solid lines.*

In the M-step the parameters $\theta^{(t+1)}$ are updated for $k = 1, ..., K$ as

$$\alpha_k^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}{N},$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^{N} (\gamma_{k,i}^{(t)} \cdot x)}{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}$$

and

$$\sigma_k^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{k,i}^{(t)} \cdot \left(x - \mu_k^{(t+1)}\right)^T \left(x - \mu_k^{(t+1)}\right)}{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}.$$

Convergence is reached when the log-likelihood function

$$logL(\theta; x, z) = logP(x, z|\theta) = \sum_{i=1}^{N} log\left[\sum_{k=1}^{K} \gamma_{k,i}\right]$$

is constant [11]. Algorithm 2 illustrates the procedure.

Data: Vector x containing N components to be partitioned.
Result: The parameters $\theta = (\alpha, \mu, \sigma)$ for each class.
Initialization: Determine an initial guess $\theta^{(0)} = (\alpha^{(0)}, \mu^{(0)}, \sigma^{(0)})$.
Convergence criteria: When the log-likelihood function has converged.
while *no convergence* **do**
  Expectation step;
  Estimate the probability densities $g(x|\theta^{(t)})$ and compute the responsibilities;
  for *all components* $i = 1, ..., N$ **do**
    for *all classes* $k = 1, ..., K$ **do**

$$\gamma_{k,i} = \frac{p(x_i|\theta_k^{(t)})}{\sum_{l=1}^{K} p(x_i|\theta_l^{(t)})}$$

    end
  end
  Maximization step;
  Update the parameters θ for the new classification.
  for *all classes* $k = 1, ..., K$ **do**

$$\alpha_k^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}{N}$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{k,i}^{(t)} \cdot x}{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}$$

$$\sigma_k^{(t+1)} = \frac{\sum_{i=1}^{N} \gamma_{k,i}^{(t)} \cdot \left(x - \mu_k^{(t+1)}\right)^T \left(x - \mu_k^{(t+1)}\right)}{\sum_{i=1}^{N} \gamma_{k,i}^{(t)}}$$

  end
end

Algorithm 2: *EM-algorithm for Gaussian mixture densities [8].*

3.4 Otsu's method

Otsu's method is an automatic method for clustering gray scaled pixels in an image into two classes by computing a threshold that divides the two classes. The threshold is calculated by exhaustive search, so that the variance within the classes is minimized [12]. Algorithm 3 presents the procedure.

Data: Gray scale image with N pixels to be partitioned.
Result: The threshold k^* that separates the two classes.
Initialization: Compute the probabilities $p_i = \frac{n_i}{N}$ for each pixel intensity (n_i is the number of pixels with intensity i) and $\mu_T = \sum_i (i \cdot p_i)$.
for *all possible thresholds k (the range of all pixel intensities)* **do**

$$\omega(k) = \sum_{i=0}^{k} p_i$$
$$\mu(k) = \sum_{i=0}^{k} (i \cdot p_i)$$
$$\sigma_B^2(k) = \frac{(\mu_T \cdot \omega(k) - \mu(k))^2}{\omega(k)(1 - \omega(k))}$$

end
The optimal threshold $k^*$is the k that corresponds to the maximum $\sigma_B^2(k)$.

Algorithm 3: *Otsu's Method [12].*

3.5 Standard deviations from remote method

The standard deviations (SDs) from remote method is a method used for the quantification of myocardial infarction in CMR images [3]. The method requires that an area assumed to consist of healthy remote myocardium is defined. The method detects infarctions in the LV-myocardium by computing a threshold as the mean of the intensities in the remote region plus a fixed number of k standard deviations (SD). All pixel intensities with a larger value than the threshold is considered to be infarcted myocadrium.

When the SD from remote method is applied to ex-vivo images, the method requires that an observer manually delineates two regions in each image slice; a remote region and a myocardium of interest region [13]. The myocardium of interest region defines the area of myocardium where an infarction is suspected to be present (Figure 3.3). The fixed number k is empirically defined to $k = 8$ [3]. All pixels within the myocardium of interest region with a higher intensity than the threshold are defined as scarred myocardium. Algorithm 4 summarise the method.

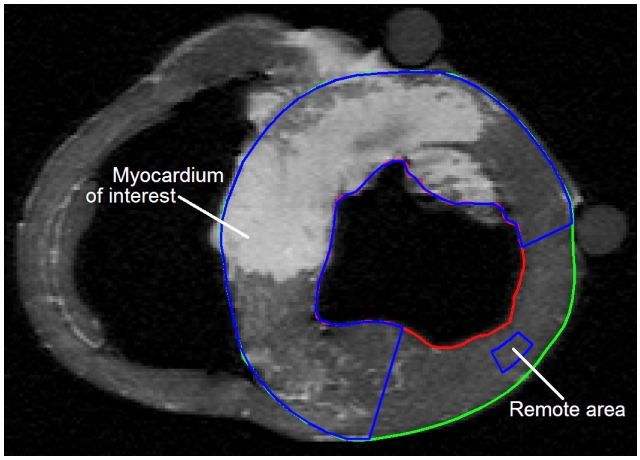

Figure 3.3: *Short-axis ex-vivo CMR image slice with manually drawn regions of interest marked with blue lines.*

Since the SD from remote method applied on ex-vivo CMR images is semi-automatic and requires manually drawn areas of interest in over 150 image slices, it is desired to implement an automatic segmentation method for ex-vivo images to replace it. An implementation of the SD from remote method is also used in the weighted method for in-vivo infarction segmentations. It is desired to use another method for in-vivo data as well, since it can not be concluded that the same fixed number k suits for every image set.

Data: Data set of CMR images
Result: The volume of scarred myocardium in % of left ventricular mass.
Initialization: Contours that defines the LV myocardium and manually drawn areas of interest.
for *each slice in the data set* **do**
  1. Compute the mean μ and SD σ of the pixel intensities in the remote region.
  2. Calculate the $threshold = \mu + k \cdot \sigma$ and define all pixels within the myocardium of interest region with a higher intensity than the threshold as scarred pixels.
  3. Calculate the infarction volume.
end

Algorithm 4: *SD from remote method for ex-vivo images [3] [13].*

3.6 Weighted method

The weighted method is a method for automated quantification of myocardial infarctions in in-vivo CMR images that takes partial volume effects into account [3]. Partial volume effects occur due to variability in the shape and spread of an infarction between two image slices, and must be taken into account in order to estimate an accurate infarction volume, see Figure 3.4. The method detects contrast enhanced infarctions in the left ventricular (LV) myocardium by defining a region of remote healthy myocardium and computing a threshold as the mean of the remote intensities plus a fixed number of k SDs. The fixed number k is empirically defined to $k = 1.8$ for in-vivo studies [3].

All pixel intensities above the threshold are defined as scarred myocardium and weighted according to their intensities in order to compensate for partial volume effects, then post-processing is performed in order to remove very small infarction areas and include microvascular obstruction. Algorithm 5 summarises the method and Figure 3.4 illustrates the idea behind partial volume effects and how the weighting of the pixels is applied.

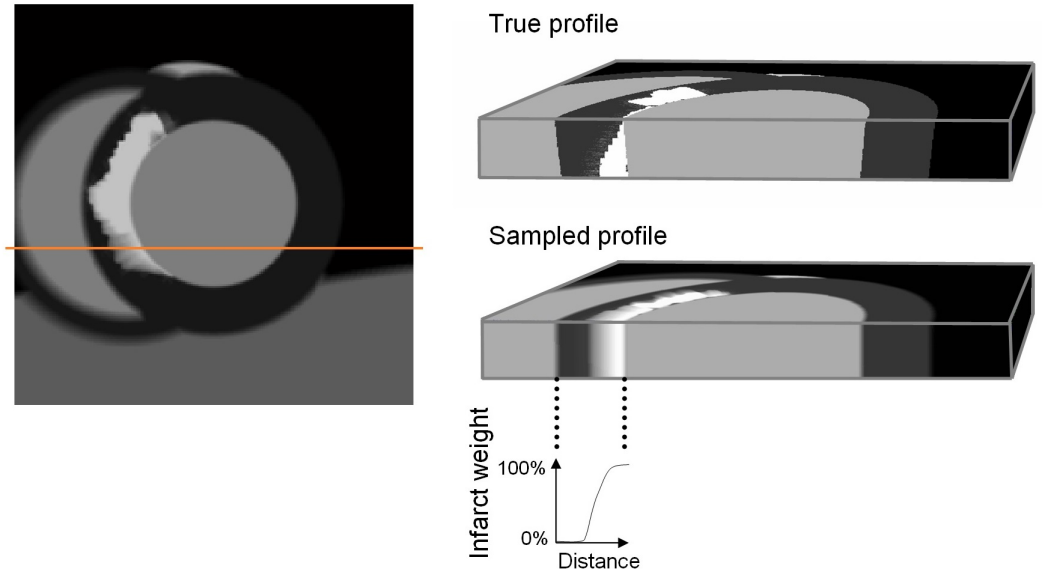

Figure 3.4: *Adopted from [3]. Description of partial volume weighting of an image from a computer-generated phantom. Left: Short-axis view of 8-mm-thick section with partial volume effects. Top right: True profile along the marked horizontal line in the short-axis image. Middle right: Sampled thick section with partial volume effects. Bottom right: Resulting infarction weight by using the weighted approach.*

Data: Data set of CMR images
Result: The volume of scarred myocardium in % of left ventricular mass (LVM)
Initialization: Contours that defines the LV myocardium.
for *each slice in the data set* **do**

1. Divide the LV myocardium into four equally sized sections. Compute the mean μ and SD σ for each section and define μ_{remote} as the minimum computed μ.
2. Calculate the $threshold = \mu_{remote} + k \cdot \sigma_{remote}$ and define all pixels with a higher intensity than the threshold as scarred pixels.
3. Apply a fast-level algorithm [14] with the speed term as a value calculated by subtracting the section-specific threshold level from the pixel intensites.
4. Remove infarction regions that are isolated and smaller than 1.5 cm^3, unless it is the only infarction region in that slice or if the volume consists of more than 1% of the LVM. If present, include MO into the scarred region.
5. Calculate the infarction volume by weighting each pixel linearly proportional to its intensity.

end

Algorithm 5: *Weighted method [3].*

Chapter 4

Aim

The aim of this thesis is to develop and validate a fully automated segmentation algorithm for quantification of myocardial infarction in contrast enhanced CMR images, based on an Expectation-Maximization (EM) algorithm. The purpose is to study whether an EM-algorithm can improve the delineation of infarctions in human and animal studies, compared to the previous suggested segmentation algorithms that use a fixed number of standard deviations from remote to determine the classification of pixels.

The specific aims are the following:

- Develop an EM-algorithm for classification of normal and scar myocardium in CMR images.
- Implement the developed EM-algorithm into the framework of the previous suggested method (the weighted method).
- Validate the performance of the new method in both humans and animals.
- Implement the new method into the software Segment.

Chapter 5

Method

The segmentation algorithms developed in this thesis are based upon the framework of the weighted method implementation in Segment. The modification of the method is the computation of the threshold that separates healthy and scar pixels, described as step 1 and 2 in Algorithm 5. An EM-algorithm provides the Gaussian probability densities for the healthy and scarred myocardium, based on a threshold which can be calculated in order to delineate the scar tissue. The computation of the infarction volume in the image set is executed by an already implemented function in Segment.

All pixel intensities in the images are normalized between 0 and 1, where 0 corresponds to black pixels and 1 corresponds to white. The algorithm only takes the pixels within the delineated LV-myocardium into account. Weighting is not used on ex-vivo CMR since there are so many image slices at such a short distance from each other that the partial volume effects are assumed to be negligible. Figure 5.1 illustrates the complete method.

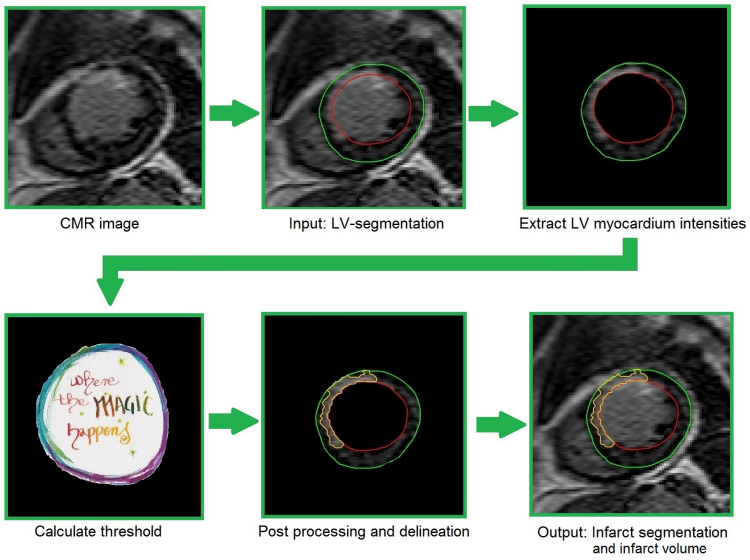

Figure 5.1: *Overview of the implemented algorithm.*

5.1 Algorithm requirements

Before implementing the segmentation algorithm, requirements were established. The goal was to implement new functionality and to fit the developed algorithm into the framework of the weighted method.

The algorithm must be able to:

- Cluster the myocardium intensities into two classes with the EM-algorithm; healthy and scarred regions.
- Detect if an image slice contains one or two classes.
- Decide whether an image slice with one class consists of 100% healthy or scarred myocardium.
- Be implemented into the framework of the weighted method.

The requirements where established based on the aim presented in Chapter 4, the imaging artifacts presented in Section 2.3.2 and the requirements of the EM-algorithm itself.

5.2 EM-algorithm implementation

The EM-algorithm was straight forward implemented according to Algorithm 2, but with a modified convergence criteria that is further described in the subsection below. The initial guess of the parameters $\theta^{(0)} = (\alpha^{(0)}, \mu^{(0)}, \sigma^{(0)})$ was obtained by the k-means algorithm. The output of the k-means algorithm, a vector classifying each pixel to a cluster, allows the EM-algorithm to compute $\theta^{(0)}$ by calculating the size of each class in percentage as $\alpha^{(0)}$, and taking the mean and standard deviation of each class as $\mu^{(0)}$ and $\sigma^{(0)}$. The initial guess for the k-means algorithm was the 5th and 95th percentile of the pixel intensities. The percentiles are calculated by sorting the vector with pixel intensities and dividing it in 100 equally sized spaces consisting of 1% of the data. The 5th percentile is the boundary point that divides the lowest 5% of the intensities from the 95% highest intensities. Accordingly, the 95th percentile is the point that divides 95% of the lowest intensities from the highest 5%.

5.2.1 Convergence criteria

The theoretical convergence criteria for the EM-algorithm is when the log-likelihood function $log\ L(\theta; x, z)$ no longer changes [11]. Implementing this criterion gives good

results on artificial phantom data. However, when analysing CMR image sets in the Segment software the convergence appears to be very slow and the computer often experience run time instabilities. An alternative convergence criterion was therefore implemented.

The EM-algorithm classifies all pixels intensities in the E-step in each iteration. When 0.05% or less of the pixels changed class during the last five iterations, the algorithm was considered to have reached convergence. This approach gives classifications in agreement with the log-likelihood convergence criteria and works well when implemented into Segment.

5.3 Detection of the number of clusters

Prior to the start, the EM-algorithm needs the information about the number of classes K it should cluster the data into. This is indicated to the EM-algorithm by the number of elements in $\theta^{(0)}$. In a CMR image set with myocardial infarction there are often image slices that contain scarred tissue and some that do not. There may also occur slices where all myocardium is scarred.

There are two ways to analyse the image data. If the LV segmentations in all slices are analysed as one data set one can assume it has two clusters. When analysing each slice separately the number of clusters must be detected before applying the EM-algorithm. If only one cluster is found it is important to classify the slice as completely healthy or scarred in order to get the correct infarction delineation and to compute an accurate infarction volume.

To detect whether two components are present in the current slice, the data is clustered into two classes by the k-means algorithm. The k-means clustering allows two Gaussian densities to be estimated. The distance d between the means of the densities, expressed in the number of standard deviations, is calculated for each slice and stored in the vector $\overline{d}$ (Figure 5.2). The vector $\overline{d}$ is then scaled into

$$\overline{D} = \frac{\overline{d} - d_{min}}{d_{max} - d_{min}}$$

where d_{min} is the minimum value and d_{max} is the maximum in $\overline{d}$. By analysing all scaled distances $\overline{D}$ in an image set, conclusions of the number of clusters in a slice can be drawn. If the scaled distance in an image slice is small compared to the rest of the elements in $\overline{D}$, it can be assumed to consist of one cluster since the two Gaussian probability densities are so close to each other that the data are better

represented by one density curve. If the scaled distance is larger than the average elements in $\overline{D}$ it can be assumed to consist of two clusters. A limit for when the distance is large enough to conclude that a specific slice contains two clusters can be computed by clustering the elements in $\overline{D}$ with Ostu's method [12], resulting in a threshold limit k^*. Each element in $\overline{D}$ is then compared to the threshold limit k^*. All slices with a scaled distance that is larger than k^* is considered to have two clusters, and one cluster otherwise.

In order to determine whether an image slice with one cluster is completely healthy or scarred, further analysis is required. This is done after the slices with two clusters are segmented, so that data of normal and scar pixel intensities are available. The mean of the pixel intensities is stored for each slice considered to consist of one class, and compared to the mean of all final thresholds that are computed with the EM-algorithm in the slices with two clusters. If the mean of the intensities is larger than the mean of all thresholds, that particular image slice is considered to be completely infarcted. If it is smaller than the mean of all thresholds it is considered to consist of only healthy myocardium.

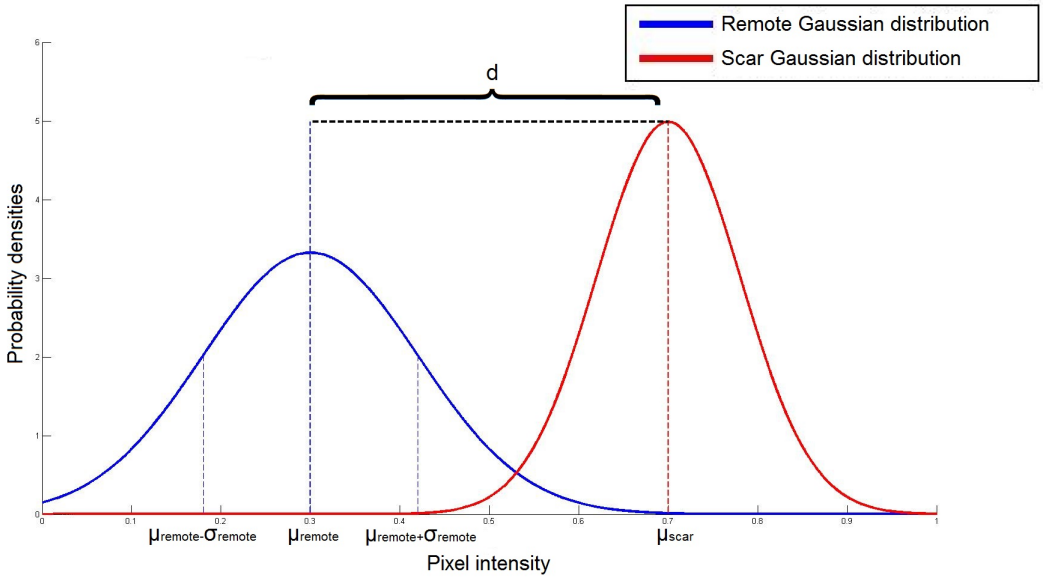

Figure 5.2: *Illustration of two probability density curves. The means are represented by red and blue dashed lines. The black dashed line show the distance between the means. The thin dashed lines show one SD from remote in both directions.*

5.4 Threshold calculation

In the slices where it is determined that two clusters are present, the EM-algorithm is applied in order to estimate two Gaussian distributions within the myocardium. The responsibility curve of the scar density is then calculated and the threshold is set to the pixel intensity where the scar responsibility curve has the value 0.5 (Figure 5.3).

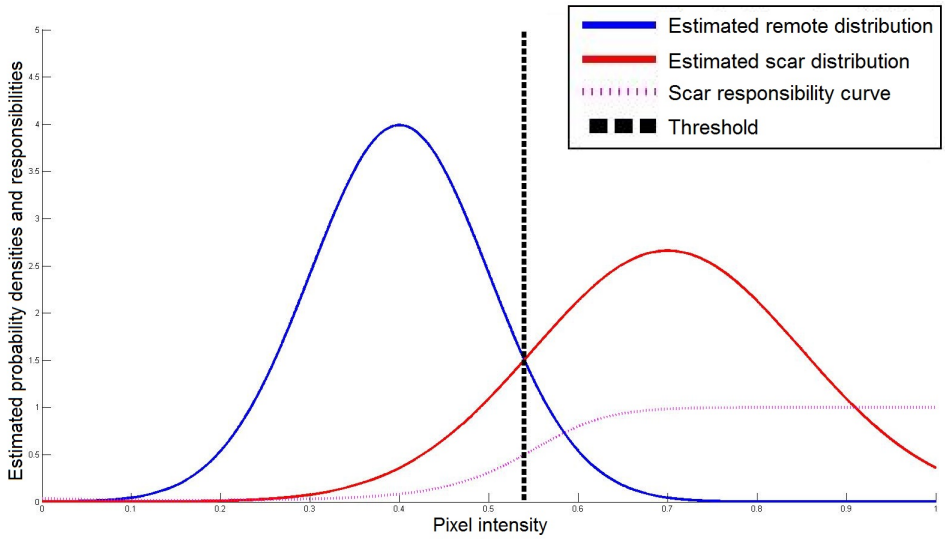

Figure 5.3: *Illustration of two estimated probability densities and the scar responsibility curve. The black dashed line shows the threshold. All pixels with larger intensity than the threshold are classified as scar tissue.*

5.5 Final implementation of algorithm

Two different algorithms were implemented in this thesis. One of the algorithms analyses each slice separately and is called the slice method, it is described in Algorithm 6. The second implementation is called the set method. The latter analyses all slices as one data set and is described in Algorithm 7. Both methods allows the observer to manually correct erroneous segmentations afterwards.

Data: Data set of CMR images.
Result: The volume of scarred myocardium [% LVM].
Initialization: Contours that defines the LV myocardium.
for *each slice in the data set* **do**
  1. Cluster the data into two classes with the k-means algorithm.
  2. Estimate the number of clusters K in the slice.
  if *K=2* **then**
    3. Estimate the probability densities of the data with the EM-algorithm. Use the result from step 1 as initialization.
    4. Calculate the threshold for this slice and store it.
  end
end
5. Analyse whether the slices with one cluster consists of scar or healthy tissue.
6. Segment each slice according to its threshold.
7. Remove infarction regions that are isolated and smaller than 1.5 cm^3, unless it is the only infarction region in that slice or if the volume consists of more than 1% of the LVM. If present, include MO into the scarred region.
8. Calculate the infarction volume.
if *In-vivo data* **then**
  9. Weight each pixel linearly proportional to its intensity.
end

Algorithm 6: *The slice method.*

Data: Data set of CMR images.
Result: The volume of scarred myocardium [% LVM].
Initialization: Contours that defines the LV myocardium.
1. Cluster the data into two classes with the k-means algorithm.
2. Estimate the probability densities of the data with the EM-algorithm. Use the result from step 1 as initialization.
3. Calculate the threshold for the data and segment the infarction.
4. Remove infarction regions that are isolated and smaller than 1.5 cm^3, unless it is the only infarction region in that slice or if the volume consists of more than 1% of the LVM. If present, include MO into the scarred region.
6. Calculate the infarction volume.
if *In-vivo data* **then**
  7. Weight each pixel linearly proportional to its intensity.
end

Algorithm 7: *The set method.*

5.6 Attempted approaches

Before reaching the final algorithms, some approaches that did not make it to the final implementations where studied and tested.

Different approaches for finding the number of Gaussian distributions contained within the image data were tried out during the development of the algorithms in this thesis without success. The performance of the jump method [15], a technique to determine the number of clusters in a data set by means of information theory, was implemented and tested on a number of CMR images unsuccessfully. Another unsuccessful approach was to analyse the histogram of the data and try to fit Gaussian mixtures with varying components onto a smoothed histogram function.

A full width half maximum (FWHM) approach [16] has also been implemented into the weighted method as an alternative threshold computation to the EM-algorithm. The performance of the FWHM with weighting did not differ enough from the weighted method based on SDs from remote for it to be studied further.

The possibility that the pixel intensities might consist of a Gaussian mixture with $K = 3$ components when MO is present in an image slice was investigated. Analysis of the pixel intensities in the MO area suggests that they had the same properties as the healthy myocardium, hence no implementation for Gaussian mixtures for three components was developed.

Chapter 6

Validation

Two kinds of segmentation algorithms were developed in this thesis; one that analyses each image slice in a data set separately (Algorithm 6: the slice method), and one which takes the data in all slices into account at once (Algorithm 7: the set method). Due to the differences between ex-vivo and in-vivo CMR images regarding resolution and the amounts of image slices in a data set, they need to be studied and validated separately. The performance of the methods proposed by Heiberg et al. [3]; 8 SDs from remote for ex-vivo data and the weighted method with 1.8 SDs from remote for in-vivo data are also studied for comparison.

The validation is performed on the same CMR images that were studied in [3]. The images originate from two MR cameras from different manufacturers. The results of the methods are compared to a reference standard. The algorithm performance was measured by linear regression and by computing the difference in scar volume between the methods and the reference standard (Bland-Altman analysis). Differences are presented as mean $\pm$ SD in percent of LVM.

6.1 Phantom data

The performance of the EM-algorithm was validated by measuring the percentage of erroneous classified pixels on artificially created images, so called phantom data, consisting of a Gaussian mixture density with two equally sized components and known representative means and SDs according to ex-vivo animal data. Figure 6.1 shows a phantom image and an illustration of its optimal classification.

6.2 Ex-vivo animal data

The ex-vivo study included 18 explanted pig hearts with experimentally induced infarction by occlusion of the left anterior descending artery. Segmentations using 8 SDs from remote with manual corrections made by an experienced medical professional were used as reference standard. The algorithms were evaluated as the

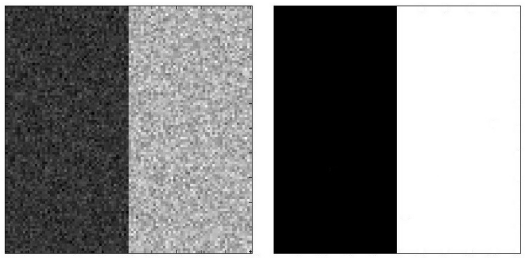

Figure 6.1: *Left panel show a phantom image used to validate the EM-algorithm. Right panel illustrates how each pixel in the phantom should be classified if the EM-algorithm performed flawlessly. Black corresponds to the first class and white to the second class.*

difference in scar volume between the new proposed methods based on the EM-algorithm and the reference standard.

6.3 In-vivo human data

The in-vivo study included 40 patients, 36 men and 4 women. The mean volume of manual delineations from three independent observers was used as reference standard. The observers had extensive experience in delineating CMR images and they used the same LV-segmentation, that is, the same pixels where investigated. Weighting was used in the algorithms in order to compensate for partial volume effects.

Three levels of Gaussian noise was added to each subject in order to investigate the performance of the methods for increasing noise. Figure 6.2 shows an in-vivo CMR image with a manual delineation for three levels of noise.

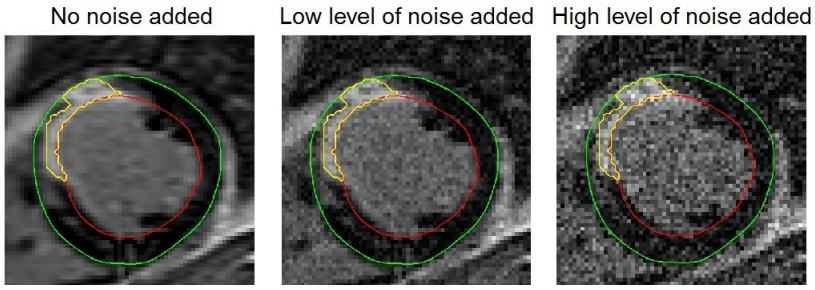

Figure 6.2: *Short-axis human in-vivo CMR image with the reference standard in yellow for three different levels of noise.*

Chapter 7

Results

7.1 Phantom segmentations

The performance of the EM-algorithm is measured by applying it to phantom image data, presented in Chapter 6.1. Three phantoms with 100×100 pixels are shown in Figure 7.1 together with an illustration of the classification results and the Gaussian distributions.

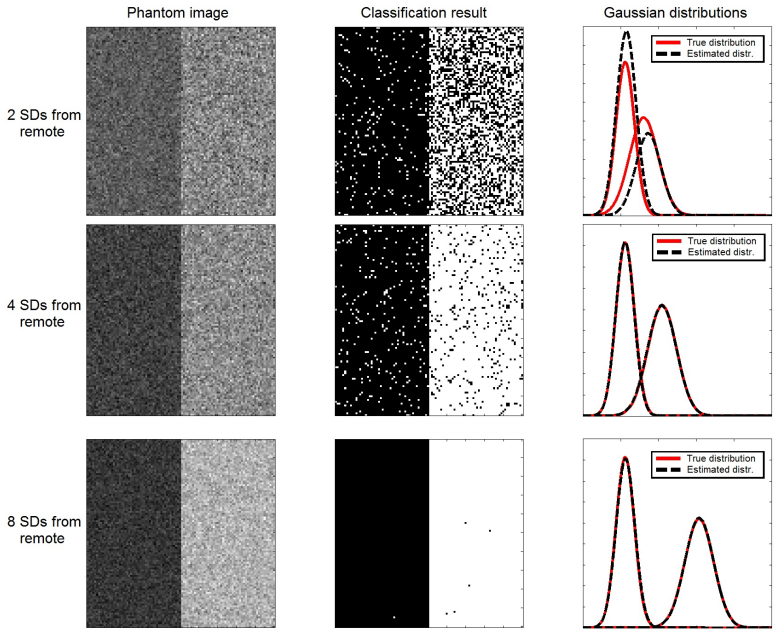

Figure 7.1: *Left column shows the generated phantoms with a fixed number of SDs from remote between the means of the classes. Middle column illustrates the estimated classification of each pixel after applying the EM-algorithm. Right column shows the true Gaussian distributions of the classes in red, while the estimated distributions are presented as black dashed lines. Top row shows consists of a Gaussian mixture with 2 SDs from remote, second row with 4 SDs from remote and third row with 8 SDs from remote.*

The segmentation results in the phantoms, measured in percentage of incorrectly classified pixels for 1 to 10 SDs from remote, are showed in Table 7.1 together with the percentage of area that overlaps in the true Gaussian probability densities, see Figure 7.2. The presented values are the average percentages after running the algorithm 100 times. The percentage of the erroneously classified pixels lies close to the percentage of the area that overlaps in the true Gaussian distributions.

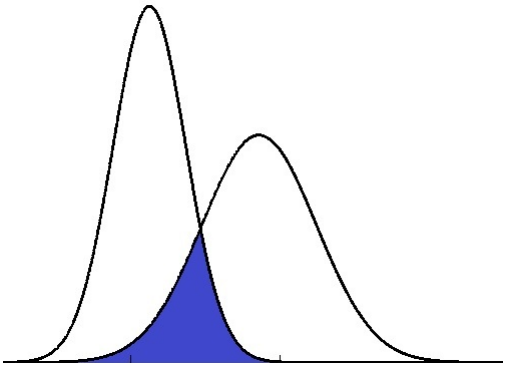

Figure 7.2: *Illustration of the overlap between two Gaussian probability densities, marked in blue.*

SDs from remote	Error [%]	Overlap [%]
1	35	33
2	23	21
3	12	12
4	5.8	5.8
5	2.5	2.5
6	1.0	0.95
7	0.3	0.3
8	0.09	0.09
9	0.02	0.02
10	0.005	0.005

Table 7.1: *Left column shows the number of standard deviations from remote of the components in the phantoms. Middle column shows the average percent of erroneously classified pixels. Right column shows the percentage of the area that overlaps in the true Gaussian distributions.*

7.2 Ex-vivo segmentations

The reference standard for ex-vivo segmentations is 8 SDs from remote with manual corrections. An example of the different segmentation results for the reference standard, 8 SDs from remote, slice method and set method is shown in Figure 7.3.

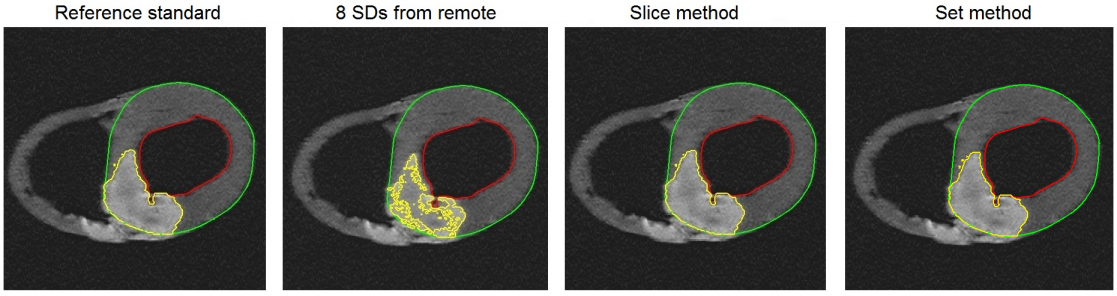

Figure 7.3: *An example of infarction segmentations by the reference standard, 8 SDs from remote and slice method.*

Figure 7.4 shows the linear regression curve and the Bland-Altman analysis for 8 SDs from remote, the slice method and set method applied on the 18 studied subjects.

The results of the linear regression and Bland-Altman analysis between the computed infarction volumes and the reference standard volumes are presented in Table 7.2. Differences are expressed as mean ± SD in percent of LVM. The slice method yields the best fit to the linear regression curve and the lowest bias.

Analisys method	$\boldsymbol{R}^2$	**Bias**
8 SDs from remote	0.87	-7.3 ± 5.0
The slice method	0.93	-1.4 ± 4.7
The set method	0.93	-4.8 ± 4.2

Table 7.2: *Ex-vivo results.*

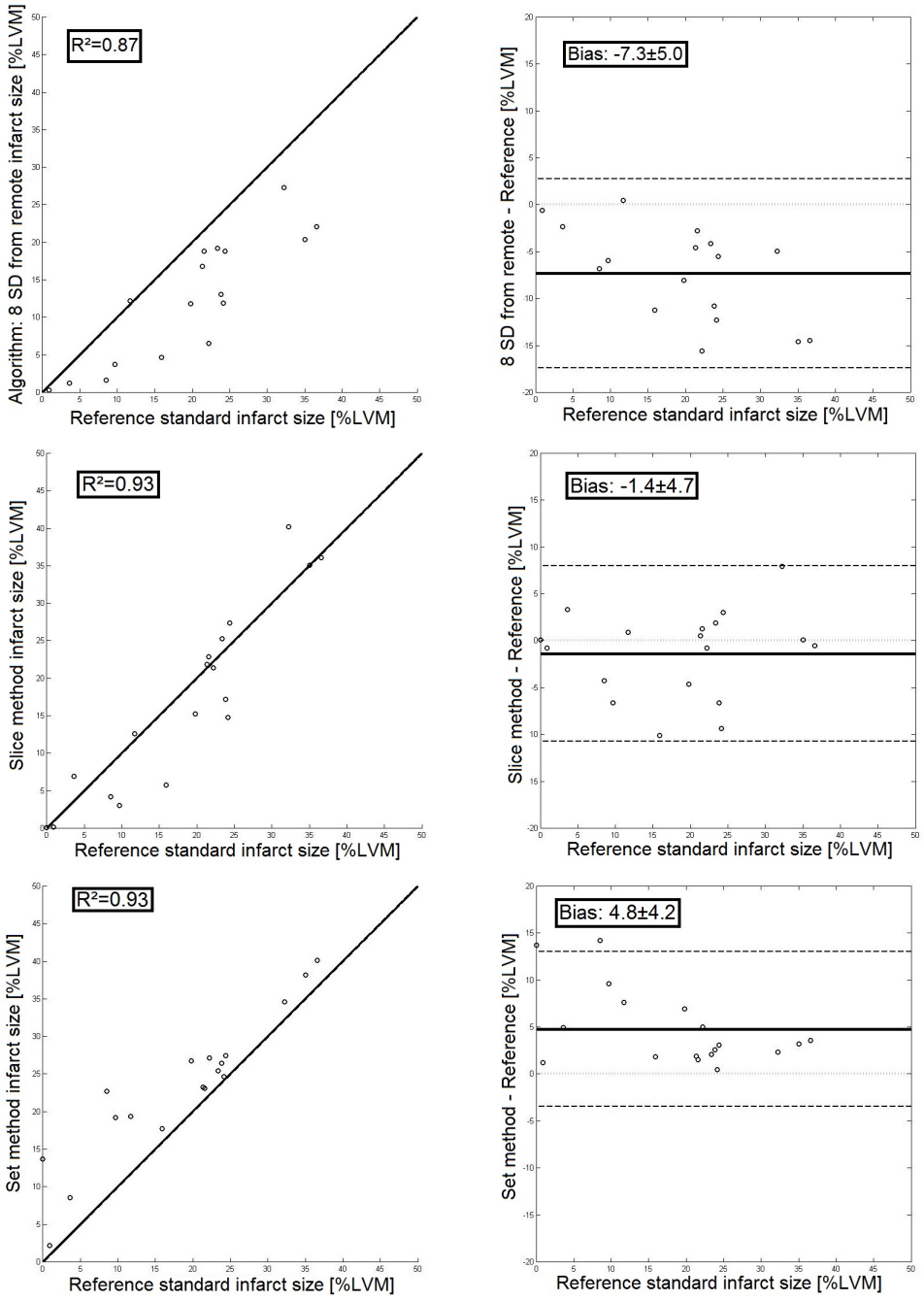

Figure 7.4: *Results for the ex-vivo data. Left column shows the linear regression curve, right column shows the Bland-Altman plot. Top row: 8 SDs from remote. Middle row: Slice method. Bottom row: Set method.*

7.3 In-vivo segmentations

The reference standard for in-vivo segmentations is the mean volume from three manual delineations. An example of the different segmentation results for the reference standard, the weighted method with 1.8 SDs from remote, slice method and set method is shown in Figure 7.5. Figure 7.6 shows the linear regression curve and the Bland-Altman analysis for the weighted method with 1.8 SDs from remote, the slice method and set method applied on the 40 studied subjects.

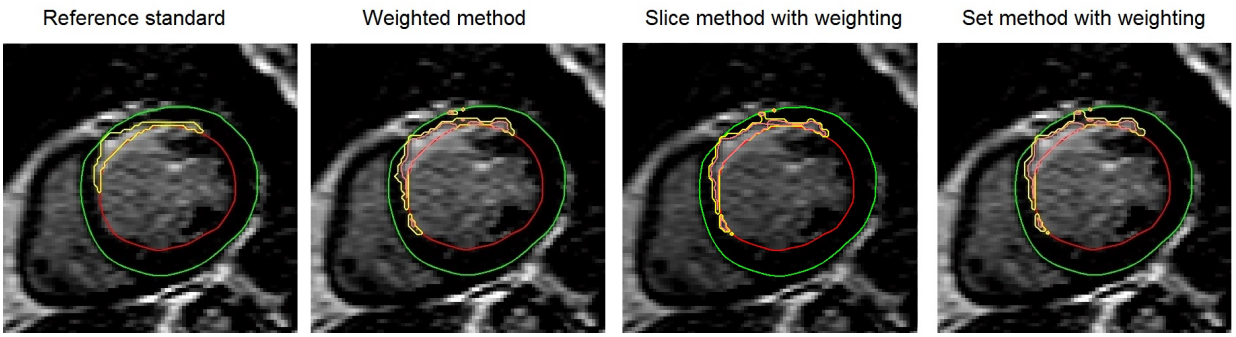

Figure 7.5: *An example of infarction segmentations by the weighted method, slice method and set method compared and the reference standard.*

The results of the linear regression and Bland-Altman analysis between the computed infarction volumes and the reference standard volumes are presented in Table 7.3. Differences are expressed as mean ± SD in percent of LVM. The weighted method with 1.8 SDs from remote yields the best fit to the linear regression curve and the lowest bias.

Analysis method	R^2	**Bias**
1.8 SDs from remote	0.95	0.4 ± 2.9
Slice method	0.85	-2.2 ± 5.1
Set method	0.93	-0.7 ± 3.5

Table 7.3: *In-vivo results.*

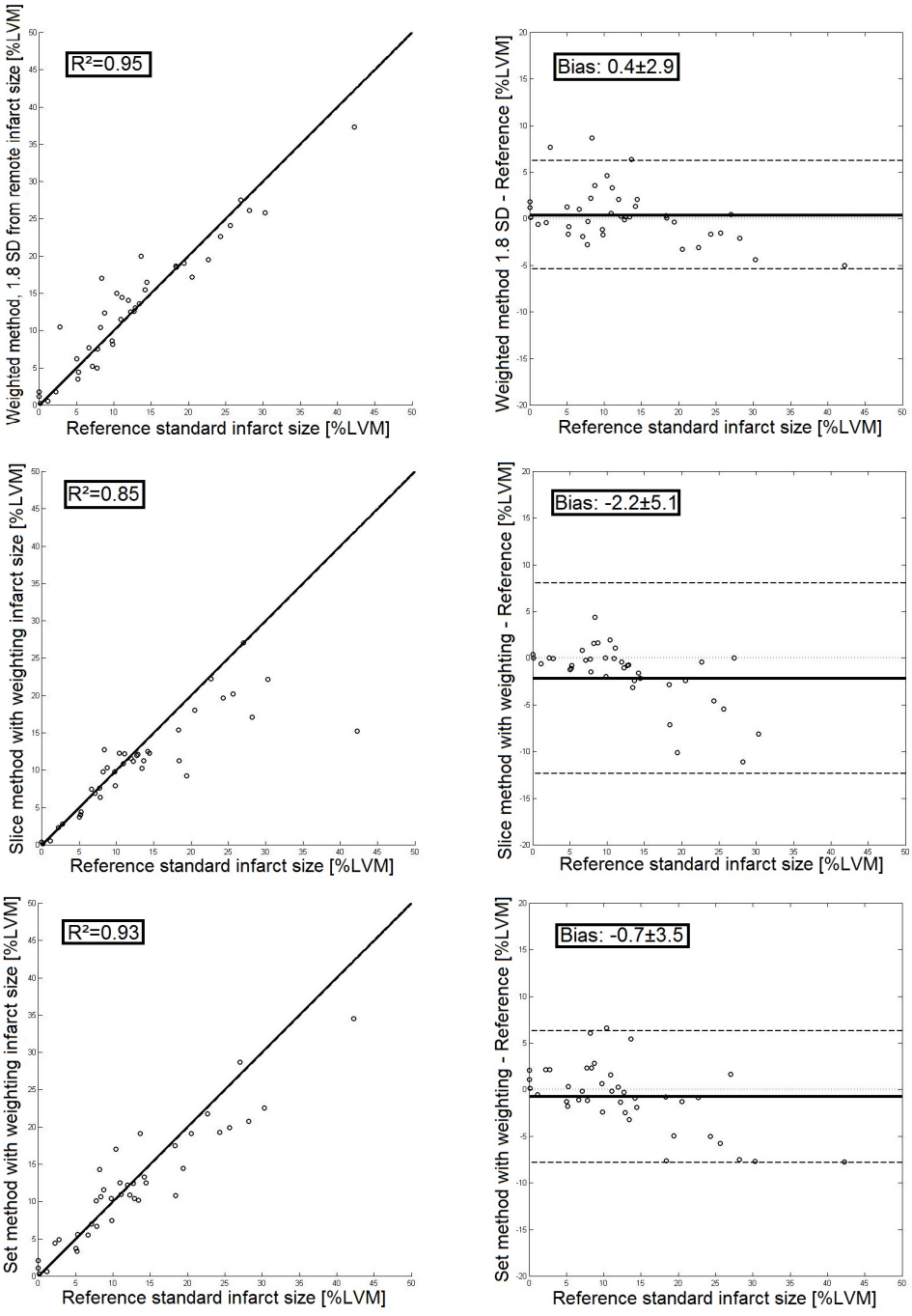

Figure 7.6: *Results for in-vivo data. Left column shows the linear regression, right column shows the Bland-Altman plot. Top row: Weighted method with 1.8 SDs from remote. Middle row: Slice method. Bottom row: Set method.*

7.4 Noise sensitivity

The performance of the weighted, slice and set method for increasing levels of added Gaussian noise is illustrated in Figure 7.7. Top row show an example of the noise levels, bottom row show the mean bias and variability in percent of LVM.

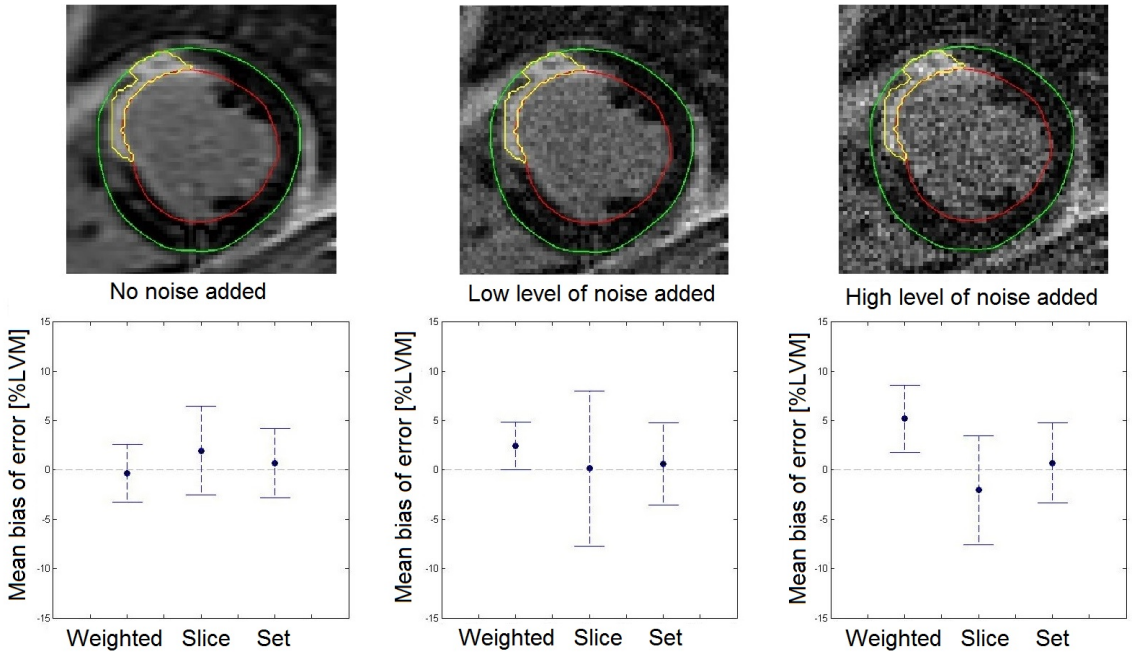

Figure 7.7: *Top row show an example of the reference standard for three different levels of added nois;, no added noise (left panel), low level of noise (middle panel) and high level of noise(right panel). Bottom row show the mean bias and variability for the weighted method, slice method and set method for increasing amount of noise.*

The weighted method presents the lowest variability and is the most sensitive method to the added noise. The slice method presents a slightly higher variability than the weighted method, and its mean bias is consistently low for all noise levels.

Chapter 8

Conclusion and Discussion

This thesis presents two implemented methods for segmentation of myocardial infarction in CMR images, based on an EM-algorithm. The developed algorithms has been implemented into the framework of the weighted method and the Segment software, and the results have been validated on phantom data, ex-vivo CMR images and in-vivo CMR images. In this chapter the results are discussed and conclusions are drawn.

8.1 Discussion

An EM-algorithm has been implemented and validated on phantom data. The results suggest that the algorithm detects the components in the Gaussian mixtures with a high accuracy. The more overlap the estimated components have, the more pixels are classified incorrectly. By looking at the true and estimated probability density curves in Figure 7.1 and Table 7.1, the incorrect classification does not seem to be caused by poor probability density estimations, but rather because an overlap creates a theoretical area of uncertainty where a single threshold can not properly distinguish the two components.

The developed EM-algorithm has been implemented into two variants of the weighted method. The slice method analyses and computes a threshold for each image slice in a data set. The set method takes all slices into account at once, calculating a single threshold that is applied to all slices. Both methods have been implemented into the framework of the weighted method implementation in the software Segment, and validated on ex-vivo and in-vivo images.

The results suggest that the set method improves the segmentation of myocardial infarction in both ex-vivo and in-vivo images compared to the previous method based on a fixed number of standard deviations from remote. The slice method improved the segmentation in ex-vivo data even further, and no prior manual interaction with the images is required any more. Apart from the abolished requirement of time-consuming manually drawn areas of interest in each image slice, the improvement can also been concluded by studying the R^2 value from linear regression

and the bias from Bland-Altman analysis. Both new methods tends to underestimate the scar volumes for ex-vivo and in-vivo data.

In in-vivo images the detection of the number of clusters within a slice does not perform very well, resulting in a suboptimal infarction quantification. The detection works well for ex-vivo data, but tends to fail in very noisy images. A reason that the detection does not work as well for in-vivo CMR images might be due to the fact that Otsu's method calculates a global threshold based on only 12 slices, which might be too little data. A disadvantage of the set method is that it might experience difficulties in classifying the pixels correctly if the noise levels differs between the image slices within the same set.

The set method performs slightly worse than the original weighted method with 1.8 SDs from remote under normal noise levels. When adding Gaussian distributed noise to the images, the set method maintained its low bias and variability while the previous weighted method's bias increased accordingly to the noise levels. This suggests that a segmentation method based on an EM-algorithm is more robust to different signal to noise ratios.

The run time for both the slice and the set method are good enough to outperform the time needed for manual delineations, an image set is delineated in about 10-30 seconds depending on the number of image slices and noise levels.

8.1.1 Limitations

The data studied in this thesis is the same that was used in the paper that proposed the weighted method [3]. The reason that the implemented methods based on an EM-algorithm do not consistently outperform the previous method might be that the weighted method was developed based on these data.

The slice and the set method are both designed assuming that an infarction has occurred in the studied CMR image sets. This is a limitation since studies of CMR image sets where no myocardial infarctions have occurred are forced to detect an infarction anyway. From Figure 7.3 it can be concluded that the methods is able to handle cases where there is no or very little infarction. This is due to the post processing removing incoherent infarction areas, suggesting that this limitation can be overcome in subsequent processing steps.

8.1.2 Future work

The detection of the number of clusters within a slice does not work properly for in-vivo data and for very noisy ex-vivo data in the slice method. The development of a method that better analyses each image slice and detects the presence of a myocardial infarction correctly regardless of the noise level might improve the performance of the in-vivo quantification algorithm.

The next step for developing the segmentation algorithm further regarding the medical use would be to add a tool that detects which vessel that caused the myocardial infarction. A priori information about the heart and the cardiovascular system could then lead to the ruling out of artifacts as infarctions by studying the physiologically possible areas of infarction in the myocardium.

There is no guarantee that the data actually is distributed according to a Gaussian mixture density model. It would be an idea to study other density mixtures for comparison, such as a Rician-Gaussian mixture model since the noise in MR images are Rician distributed [17].

8.2 Conclusions

The overall results suggest that the implemented methods are equivalent to, or an improvement over thresholding with a fixed number of SDs from remote.

The implemented EM-algorithm outperforms previously used SD from remote method for ex-vivo data. Both the slice and the set method show good agreement and low bias with the reference standard. The slice method appears to have a slightly lower bias than the set method. Both methods tends to underestimate the infarction volume. The previously used SD from remote method was semi-automatic since it required manually drawn areas of interest in over 150 image slices. The new methods are automatic and do not require any manual interaction prior to the application of the algorithms, which significantly reduces the time spent by the observer when quantifying infarctions in ex-vivo data.

For in-vivo images, the slice and the set method with weighting shows good agreement and low bias with the reference standard. Both tend to underestimate the infarction volume. The set method outperforms the slice method.

The ordinary weighted method with 1.8 SDs from remote showed the lowest bias and the best fit in the linear regression, but it presented increasing bias with increasing amounts of noise. The set method showed low bias regardless of the noise level.

The set method is theoretically more robust to noise than the ordinary weighted method is, and they perform equally under normal noise levels.

In conclusion, two methods based on a developed EM-algorithm was presented and implemented into the framework of the weighted method. The automatic slice and the set method outperformed the previous used semi-automatic 8 SDs from remote method in ex-vivo CMR images, and the set method outperformed the weighted method with 1.8 SDs from remote in in-vivo CMR images regarding robustness to different noise levels. Both implemented methods in this thesis show potential for fully automatic quantification of myocardial infarction in high resolution contrast enhanced CMR images.

Bibliography

[1] Marcus Carlsson, "Aspects on Cardiac Pumping", Faculty of Medicine Doctoral Dissertation Series 2007:47, Doctoral Thesis at the Department of Clinical Physiology, Faculty of Medicine, Lund University, 2007, ch. 1, pp. 1-2.

[2] Helen Soneson, "Methods for Quantitative Analysis of Myocardial Perfusion SPECT, Validated with magnetic resonance imaging, phantom studies and expert readers", Doctoral Thesis at the Centre for Mathematical Sciences at Lund University, 2012, ch. 2, pp.3-5.

[3] Heiberg et al. "Automated Quantification of Myocardial Infarction from MR Images by Accounting for Partial Volume Effects: Animal, Phantom and Human Study". *Radiology* vol. 246, no. 2, pp. 581-588, February 2008.

[4] Sjögren et al. "Semi-automatic segmentation of myocardium at risk in T2-weighted cardiovascular magnetic resonance" in *Journal of Cardiovascular Magnetic Resonance* **14**:10, 2012.

[5] Eric P. Widmaier, Hershel Raff, Kevin T. Strang, "Cardiovascular Physiology. Section B: The Heart" in *Vander's Human Physiology: The Mechanisms of Body Function*, 11th edition, New York, McGraw-Hill, 2008, ch. 12, pp. 360-367 and 422-424.

[6] Collaborative. "The Human Heart", http://en.wikipedia.org

[7] Heiberg et al.: "Design and validation of Segment - freely avaliable software for cardiovascular image analysis" in *BMC Medical Imaging* **10**:1, 2010.

[8] Forsyth and Ponce, "Segmentation and Fitting Using Probabilistic Methods", in *Computer Vision: A Modern Approach*, International Edition, Upper Saddle River, Pearson Education Inc., 2003, ch. 16, pp. 301-325, 354-372.

[9] J. MacQueen. "Some Methods for Classification and Analysis of Multivariate Observations". *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1 , pp. 281-297, University of California Press, 1967.

[10] Greg Hamerly, Charles Elkan, "Alternatives to the k-means algorithm that find better clusterings". *Proceedings of the eleventh international conference on Information and knowledge management (CIKM)*, 2002.

[11] C. F. Jeff Wu, "On the Convergence Properties of the EM Algorithm". *The Annals of Statistics*, vol. 11, no. 1, pp 95-103, 1983.

[12] Nobuyuki Otsu, "A Threshold Selection Method from Gray-Level Histograms". *IEEE Transactions on systems, Mean, and Cybernetics*, vol smc-9, no. 1, 1979.

[13] Götberg et al. "A pilot of rapid Cooling by Cold Saline end Endovascular Cooling Before Reperfusion in Patients With ST-Elevation Myocardial Infarction". *Circulation: Cardiovascular Interventions*, 3:400-407, 2010.

[14] B. Nilsson, A. Heyden, "A fast algorithm for level set-like active contours". *Pattern Recognition Letters*, 24:1331-1337, 2003.

[15] Catherine A.Sugar and Gareth M. James. "Finding the Number of Clusters in a Dataset: An Information-Theoretic Approach". *Journal of the American Statistical Association* vol. 98, no. 463, pp. 750-763, September 2003.

[16] Hsu et al. "Quantitative Myocardial Infarction on Delayed Enhancement MRI. Part I: Animal Validation of an Automated Feature Analysis and Combined Thresholding Infarct Sizing Algorithm". *Journal of Magnetic Resonance Imaging*, 23:298-308, 2006.

[17] R. Mark Henkelman, "Measurement of signal intensities in the presence of noise in MR images", *Medical Physics*, 12:232, 1985