UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Hrvoje Kalinić

# REGISTRATION AND MODEL-BASED ANALYSIS OF TRANSAORTIC VALVULAR FLOW ULTRASOUND IMAGES

DOCTORAL THESIS

Zagreb, 2012.

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Hrvoje Kalinić

# REGISTRATION AND MODEL-BASED ANALYSIS OF TRANSAORTIC VALVULAR FLOW ULTRASOUND IMAGES

DOCTORAL THESIS

Supervisors
Professor Sven Lončarić, Ph.D.
Professor Davor Miličić, Ph.D.

Zagreb, 2012.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Hrvoje Kalinić

# REGISTRACIJA I ANALIZA ULTRAZVUČNIH SLIKA PROTOKA PREKO AORTNOG ZALISKA UTEMELJENA NA MODELU

DOKTORSKI RAD

Mentori
Prof. dr. sc. Sven Lončarić
Akademik Davor Miličić

Zagreb, 2012.

This doctoral's thesis has been done at the University of Zagreb, Faculty of Electrical Engineering and Computing, Department of Electronic Systems and Information Processing.

Thesis Advisors: Professor Sven Lončarić, PhD
Professor Davor Miličić, PhD

The doctoral's thesis has 149 pages.

Thesis no.:

The dissertation evaluation and defense comittee:

1. Assistant Professor Marko Subašić, Ph.D.
   University of Zagreb, Faculty of Electrical Engineering and Computing
2. Professor Mislav Grgić, Ph.D.
   University of Zagreb, Faculty of Electrical Engineering and Computing
3. Professor Bart Bijnens, Ph.D.
   Pompeu Fabra University, Barcelona, Spain

Date of dissertation defence: $17^{th}$ October 2012

**Sven Lončarić** was born in Zagreb in 1961. He received Diploma of Engineering and Master of Science degrees in electrical engineering from University of Zagreb Faculty of Electrical Engineering and Computing (FER) in 1985 and 1989, respectively. He received Doctor of Philosophy (Ph.D.) degree in electrical engineering from University of Cincinnati, USA, in 1994.

In 1996, he became an assistant professor of electrical engineering, in 2001 an associate professor, in 2006 full professor, and in 2011 tenured full professor in electrical engineering and computer science at FER. He was researcher or project leader on a number of research projects in image processing and analysis. From 2001-2003, he was an assistant professor at New Jersey Institute of Technology, USA. He is the head of the Image Processing Group and was the Chair of the Department of Electronic Systems and Information Processing at FER. He founded the Center for Computer Vision at University of Zagreb. He authored or co-authored more than 180 scientific publications.

Prof. Lončarić is a senior member of IEEE and a member of Croatian Academy of Technical Sciences. He was the Chair of the IEEE Croatia Section. He founded Int'l Symposium on Image and Signal Processing and Analysis. He is a program committee member and reviewer for a number of scientific conferences and journals. He is an editor of EURASIP Journal of Image and Video Processing. He is a recipient of the Fulbright stipend and IEEE Croatia Section Outstanding Volunteer Award.

**Sven Lončarić** je rođen 1961. godine u Zagrebu. Diplomirao je i magistrirao u polju elektrotehnike na Sveučilištu u Zagrebu Fakultetu elektrotehnike i računarstva (FER), 1985. i 1989. godine. Doktorirao je u polju elektrotehnike na Sveučilištu u Cincinnatiju, SAD, 1994. godine.

Od 1996. je docent u polju elektrotehnike, od 2001. izvanredni profesor, od 2006. redoviti profesor, te od 2011. redoviti profesor u trajnom zvanju u polju elektrotehnike i polju računarstva na FER-u. Bio je suradnik ili voditelj na brojnim istraživačkim projektima u području razvoja metoda i sustava za obradu slika i signala. Od 2001. do 2003. bio je Assistant Professor na Sveučilištu New Jersey Institute of Technology, SAD. Voditelj je Grupe za obradu slike na FER-u, bio je predstojnik Zavoda za elektroničke sustave i obradbu informacija na FER-u. Osnivač je Centra izvrsnosti za računalni vid na Sveučilištu u Zagrebu. Publicirao je više od 180 znanstvenih i stručnih radova.

Prof. Lončarić je član stručne udruge IEEE i Akademije tehničkih znanosti Hrvatske. Bio je predsjednik Hrvatske sekcije IEEE. Osnivač je međunarodnog znanstvenog simpozija ”Int’l Symposium on Image and Signal Processing and Analysis”. Član je programskih odbora i recenzent više međunarodnih znanstvenih skupova i časopisa. Urednik je časopisa EURASIP Journal of Image and Video Processing. Dobitnik je Fulbrajtove stipendije i Nagrade Hrvatske sekcije IEEE za iznimni doprinos Sekciji.

**Davor Miličić** was born in Zagreb in 1962. He got his M.D. and Ph.D. degree from the University of Zagreb School of Medicine in 1986 and 1992, respectively.

From 1998, he is working at the University of Zagreb School of Medicine, starting from a position of a research fellow, scientific assistant, assistant professor, associate professor and, from 2008 as a full professor. In parallel, he is working at the University Hospital Centre Zagreb, starting from 1989. He is a specialist in Internal Medicine and subspecialist in cardiology. Today he is a Chairman of the Department of Cardiovascular Diseases at the University Hospital Centre Zagreb and a Dean, University of Zagreb School of Medicine, in his second mandate. He was a DAAD-research fellow at the Universiy of Hamburg, Germany in 1994-95, and trained in transplantation cardiology at the University of Udine and Padova, Italy, in 1995-96. He has published cca 200 papers in internationally peer reviewed journals and authored 14 articles in textbooks and monographies.

Prof. Miličić has been elected in 2012. as a Full Member (Fellow) of the Croatian Academy of Sciences and Arts. He is also a Fellow of the European Society of Cardiology and the American College of Cardiology. In 2012. he was elected as a Councillor of the European Society of Cardiology. He is a President of the Croatian Cardiac Society, as well as the Founder and Director of the Croatian Heart Foundation. He is a Board Member of the Croatian Medical Association and the Croatian Medical Chamber. He was a President of five national and four international cardiology congresses/meetings. He was a Conrgress Programme Commitee member of 4 congresses of the European Society of Cardiology, and cca 20 international or domestic scientific cardiology meetings. He is an Editorial Board Member of the Journal of Cardiovascular Medicine, Cor et Vasa and Liječnički vjesnik, and reviewer in many cardiology journals. He got the Rector's Award in 1985, the Ladislav Rakovac diploma in 2004, the International League of Humanists award in 2008, and the Croatian Academy of Sciences and Arts award in 2010.

**Davor Miličić** rođen je u Zagrebu 1962. godine. Diplomirao je i doktorirao na Medicinskom fakultetu Sveučilišta u Zagrebu Sveučilištu u Zagrebu, 1986. odnosno 1991. godine.

Od 1998. godine radi na Medicinskom fakultetu Seučilišta u Zagrebu, gdje je napredovao od znanstvenoga novaka preko višeg asistenta, docenta, izvanrednog profesora do redovitoga profesora 2008. g. Usporedno s fakultetskim obvezama radi i na KBC-u Zagreb od 1989., gdje je obavio specijalizaciju iz interne medicine a potom i subspecijalizaciju iz kardiologije. Danas je predstojnik Klinike za bolesti srca i krvih žila KBC-a Zagreb i dekan Medicinskog fakulteta Sveučičišta u Zagrebu u drugom mandatu. Bio je gostujući istraživač na Sveučilištu u Hamburgu kao stipendist DAAD-a 1994/95, a tijekom 1995/96. usavršavao se u transplantacijskoj kardiologiji na sveučilištima u Udinama i Padovi. Objavio je oko 200 radova u časopisima s međunarodnom citiranošću. Autor je 14 tekstova u udžbenicima i monografijama.

Prof. Miličić izabran je za redovitog člana HAZU 2012. g. Nosi počasnu titulu Fellow Europskoga i Američkoga kardiološkog društva. Izabran je za kancelara Europskoga kardiološkog društva 2012. Predsjednik je Hrvatskoga kardiološkog društva te utemeljitelj je i upravitelj Hrvatske kardiološke zaklade. Član je Glavnog odbora Hrvatskoga liječničkoga zbora i Hrvatske liječničke komore. Bio je predsjednik pet nacionalnih kardioloških kongresa, te 4 međunarodna kardiološka kongresa/skupa. Član je Kongresnog odbora Europskih kardioloških kongresa i dvadesetak međunarodnih ili domaćih znanstvenih skupova iz kardiologije. Član je uredničkog odbora časopoisa Journal of Cardiovascular Medicine, Cor et vasa te Liječnički vjesnik, te sudjeluje kao recenzent u većem broju časopisa. Dobitnik je Rektorove nagrade (1985.), plakete Ladislav Rakovac (2004.), nagrade International Leauge of Humanists (2008.), te godišnje nagrade za znanost HAZU (2010.).

*To my ancestors and for my children.*

[illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] ...

Μῆνιν ἄειδε, θεά, Πηληιάδεω, ᾿Αχιλῆος - begins the Homer's Illiad. However, it was not the wrath of Achilles that is remembered, nor its greatness – but the love of Hector. In ancient times, Achilles had some status as a sort of demigod – but he disappears altogether in late times. Hector grows greater as the ages pass, and it is his name that is the name of a Knight of the Round Table and his sword that legend puts into the hand of Roland, laying about him with the weapon of the defeated Hector in the last ruin and splendour of his own defeat. A legend speaks of stragglers from Troy that sailed across the sea and founded a republic on the Italian shore. The legend roots the republican virtue not in the Babylon or the Egyptian pride but in the honour of Hector and it's sacrifice, and on the shield of Hector a modern civilization is born, defying great civilizations of both Asia and Africa.

The history wanted that the tale of the end of Troy has no ending, for it has been lifted up on the shoulders of Hector, as it echoes throughout centuries and carves its verses in ours subconscious. Indeed, the name of Hector anticipates all the defeats through which our race were to pass – that survival of a hundred defeats that is its triumph.

I dedicate this work to all the Hectors of my people.

*...they stand unnaturally still while the arena and the world seem to revolve round them. And there shone on them in that dark hour a light that has never been darkened; a white fire clinging to that group like an unearthly phosphorescence, blazing its track through the twilights of history and confounding every effort to confound it with the mists of mythology and theory; that shaft of light or lightning by which the world itself has struck and isolated and crowned it; by which its own enemies have made it more illustrious and its own critics have made it more inexplicable; the halo of hatred...*

*Gilbert Keith Chesterton*
*The everlasting man*

*Navik on živi ki zgine pošteno!*

*Fran Krsto Frankopan*
*Pozvanje na vojsku*

# Summary

The dissertation describes a method for registration and model-based analysis of transaortic valvular flow ultrasound images. The image analysis is done from segmented images, where the segmentation is obtained by segmentation propagation from the atlas. The results showed that variability of the automated segmentation relative to the manual is comparable to the intra-observer variability. From the segmentation morphological feature are extracted and it was shown that the symmetry of aortic outflow profile relates to postoperative functional recovery of patients with aortic stenosis, and that it is is better predictor of functional recovery than other routinely used measurements. The accuracy of the segmentation depends on the registration, which in turn depends on the atlas and correct alignment of atlas and image. Thus, methods for atlas formation and definitions of image similarity measure are investigated. We show that the atlas constructed from multiple template images and multiple expert segmentation is fast method for atlas formation which produce atlas whose properties are comparable to atlas obtained by more complex atlas formation methods. It is also discussed under which circumstances the proposed method produces the same result as the classical atlas fusion method. Also, a novel image similarity measure, named absolute joint moments, is proposed. It was shown, that absolute joint moments can be observed as a generalization of the correlation, but also, via cumulant expansion of the probability density function as the approximation of the mutual information. Experimental results also showed that absolute joint moments combine good properties of both correlation coefficient and mutual information.

**Keywords:** medical image analysis, Doppler ultrasound imaging, cardiac outflow velocity profile, image registration, template matching, model-based segmentation, segmentation propagation, atlas formation/construction, similarity measure, absolute joint moments

# Sažetak

## Registracija i analiza ultrazvučnih slika protoka preko aortnog zaliska utemeljena na modelu

U ovom radu prikazana je metoda za registraciju i analizu ultrazvučnih slika protoka preko aortnog zaliska utemeljena na modelu. Uporaba ultrazvučnih snimaka protoka je česta u kliničkoj dijagnostici i predstavlja dio svakodnevnog posla za kardiologe. Automatiziranje tog posla može olakšati i ubrzati postupak donošenja kliničke dijagnoze, te povećati pouzdanost i ponovljivost mjerenja. Ultrazvučne slike protoka dobivene su na temelju Dopplerovog efekta, te prikazuju kretanje objekta u ravnini ultrazvučnog vala. Za dijagnostiku je od posebne važnosti mjeriti brzine protoka, a mjereći kretanje preko aortnog zaliska Dopplerovom metodom u zapisu, uz protok, često zabilježimo otvaranje i zatvaranje zaliska. Zahvaljujući svom iskustvu kardiolozi su u stanju razlikovati koji dio ultrazvučne snimke predstavlja klinički važnu informaciju, a koji nevažnu, te su u stanju izmjeriti valjane vrijednosti čak i kada otvaranje i zatvaranje zalistaka zakrije dio informacije o protoku krvi. Prvi korak u obradbi i razumijevanju sadržaja slike, pa time i mjerenja objekta u slici, je segmentacija slike. Metode segmentacije slike temeljene na modelima predstavljaju jednostavan način da se potrebno predznanje ugradi u proces segmentacije slike. Osnovna ideja je da se algoritmu za segmentaciju dodijeli primjer na temelju kojega će on segmentirati sve ostale slučajeve. Kao primjer može poslužiti jedan element skupa (npr. prototip) ili više njih (uzorci). Kada govorimo o modelu, podrazumijevat ćemo primjer oblikovan na temelju raspoloživih uzoraka ili odabran iz njih. Ipak, važno je napomenuti kako metode segmentacije modelom obično ne podrazumijevaju i ne zahtijevaju treniranje prilikom izgradnje modela, tipično za druge algoritme što pokušavaju implementirati računalnu inteligenciju. Tipičan način ugradnje znanja u proces segmentacije je definiranje para sastavljenog od slike i pripadajuće segmentacije koju je predložio stručnjak, u našem slučaju kardiolog. Par definiran slikom i pripadajućom segmentacijom obično nazivamo atlasom. Koristeći atlas (odnosno bilo koju segmentiranu sliku) možemo segmentirati neku drugu sliku koristeći metodu prenošenja (eng. propagation) segmentacije. Prenošenje segmentacije sa slike koja ima poznatu segmentaciju na sliku koja nije segmentirana vrši se tako da se dvije slike upare u procesu registracije (uparivanja) slika nakon čega se segmentacija s atlasa preslika na sliku koju želimo segmentirati. Pretpostavka ovog procesa je da je moguće upariti slike na način da segmentacija pridružena atlasu odgovara željenoj segmentaciji slike. U osnovi to ovisi o definiciji sličnosti između dviju slika i dopuštenoj geometrijskoj transformaciji, odnosno o sposobnosti optimizacijskog algoritma da pronađe najveću sličnost među slikama u okvirima dopuštene transformacije.

Nakon uvoda, pregleda nužnih matematičkih elemenata i pojmova, pregleda dosadašnjih istraživanja i opisa dohvata podataka ovaj rad opisuje četiri glavna dijela istraživanja, te njihove rezultate. Prvi dio se bavi segmentacijom ultrazvučnih slika protoka preko aortnog zaliska te izlučivanjem značajki iz dobivene segmentacije. Drugi dio ispituje metode za oblikovanje i/ili odabir modela, dok se treći dio bavi definicijom mjere sličnosti. Posljednji, četvrti dio pokazuje uporabu i korisnost razvijene metode za segmentaciju slika protoka preko aortnog zaliska. Prvi i posljednji dio predlažu rješenje zadanog problema koristeći varijacije postojećih metoda, dok središnji djelovi opisuju nov metodološki pristup, odnosno predstavljaju teoretsko unaprjeđenje i/ili poopćenje postojećih metoda. Od rezultata bismo istaknuli da predložena automatska segmentacija utemeljena na modelu daje segemntaciju čija je varijabilonost u odnosu na segmentaciju koju je načinio kardiolog tek neznatno veća od varijabilosti dviju segmentacija načinjenih od istog kardiologa. Također, morfološke značajke izlučene iz segmentacije su pokazale da je simetričnost protoka preko aortnog zaliska u svezi s oporavkom pacijenta nakon zamjene zalistaka i da mjera simetričnosti bolje predviđa oporavak pacijenta no druge, rutinske mjere aortnog protoka. U teoretskom dijelu rada raspravljali smo kako točnost segementacije ovisi o atlasu i mjeri sličnosti između dviju slika. Pokazali smo da atlas izgrađen na temelju više slika i više segmentacija dobivenih od kardiologa postiže svojstva slična atlasima izgrađenim mnogo složenijim metodama. Također samo pokazali uz koje uvjete ovaj način konstrukcije atlasa rezultira atlasom koji postiže istu točnost kao segmentacija nastala kombiniranjem više atlasa - poznatija i kao fuzija atlasa. Nova mjera sličnosti koju predlažemo nazvali smo međumomentna mjera jer je temeljena na apsolutnoj sumi međusobnih momenata. Pokazali smo da ovakva mjera može biti promatrana kao generalizacija korelacije, ali i da raspisom funkcije gustoće vjerojatnosti preko kumulanata, može biti gledana kao aproksimacija međusobne informacije. Eksperimentalni rezultati su također pokazali da međumomentna mjera kombinira dobra svojstva i korelacije i međusobne informacije.

**Ključne riječi:** dohvat i analiza medicinskih slika, Dopplerov ultrazvuk, aortni protoci, registracija slike, poravnavanje s prototipom, segmentacija modelom, propagacija segmentacije, odabir/izgradnja atlasa, mjera sličnosti, međusobni momenti

# Contents

# Chapter 1

# Introduction

Main focus of this thesis is analysis of transaortic valvular flow ultrasound images. Morphological features of the aortic outflow ultrasound images are often used in clinical practice for diagnosis of cardiovascular diseases. In current clinical practice, the features are extracted from expert's manual segmentation. To relieve the clinicians from the increasing load of work due to the increase of cardiovascular diseases, the goal is to develop an automated method. Due to the intrinsic properties of the problem the model-based approach is used. In order to implement a model-based segmentation, first a model needs to be costructed. Second, a correspondence between model and target image needs to be defined in order to propagate the segmentation from the model to a target. For this reason the thesis focuses on investigating model formation techniques and the very definition of correspondence, as well as the transaortic valvular flow segmentation and analysis.

The thesis is structured as follows. The rest of this chapter explains the motivation and describes the problem. Afterwards, a preliminary definitions and notation are given in Chapter 2. Following preliminaries, the overview of the prior art in the model-based segmentation and accompanied methods for model construction, and definition of correspondence between image and a model, are given in Chapter 3. In Chapters 4–8 data acquisition is presented, followed by presentation of the results to each challenge described in the problem statement (Section 1.3). Chapters 5 and 8 describe a solution of a novel problem using the problem specific variants of the existing methods. On the other hand, Chapters 6 and 7 describe a novel methodological approach and gives a different theoretical framework in order to generalize and/or increase properties of current methods. At the end appendices are listed and the conclusion is given.

## 1.1 Clinical motivation

To highlight the idea that motivated the research let's observe the heart as a muscular pump. The muscle develops force to pump blood through the blood vessels. Any change in muscle (e.g. myocardial deformation like ischemia) or in the diameter of the blood vessel (e.g. narrowing of the valve like aortic stenosis) would lead to change in force and pressure, and subsequently - flow. Thus, by observing the time-change of the velocities, the pathologies can be noticed. This indeed is the often clinical routine, where the cardiac patients are examined using Doppler ultrasound to get a useful information about blood velocities through the cardiac valves [20]. The Doppler signal used for this is a continuous wave Doppler signal that represents time-change of velocities along a scan line. Diagnostically useful informations are in the shape of this signal, or in it's morphological feature. To acquire them, the signal segmentation and quantification is necessary.

In current clinical practice, basic segmentation and measurements of aortic outflow Doppler traces are routinely obtained by manual tracking of Doppler traces. Manual tracking of the traces is often cumbersome, time-consuming and dependent on the expertise of the cardiologist/sonographer. Moreover, a detailed analysis of Doppler echocardiography traces is often limited by a high frequency workflow in the echocardiographic laboratory.

Several studies highlight the problem of cardiovascular diseases in a modern world, in th meantime indirectly showing the clinicians' increasing amount of work in the field [66, 7, 247]. Here are some statistical details extracted from these studies:

> At the beginning of the 20th century, cardiovascular disease was responsible for fewer than 10% of all deaths worldwide, while today that figure is about 30%, with 80% of the burden now occurring in developing countries [66]. In 2001, cardiovascular disease was the leading cause of death worldwide [66]. In United States, coronary heart disease caused 1 of every 5 deaths in 2004 [7]. Croatia is not an exception from this case, as according to the statistics of the Croatian National Institute of Public Health from 2004, cardiovascular diseases were responsible for 53% of all deaths in Croatia, which is roughly twice more than deaths caused by neoplasms or approximately three times more than all other causes of death [247]. There is also studies showing a tremendous increase of more than 60% in coronary heart disease death rates between 1988 and 1998 in Croatia [66].

Therefore, it is expected, that the automatic segmentation of images could relieve clinicians from the labor intensive aspects of their work while increasing the accuracy, consistency, and reproducibility of the interpretations, in the meantime allowing them to focus more on other aspect of their work. The reason why ultrasonic imaging is routinely used in clinical practice is its non-invasive property [201], however, ultrasonic imaging suffers from several issues, such as low spatial (or temporal) resolution, ill-defined boundary, poor contrast, acquisition artifact or other noise place additional demands on segmentation.

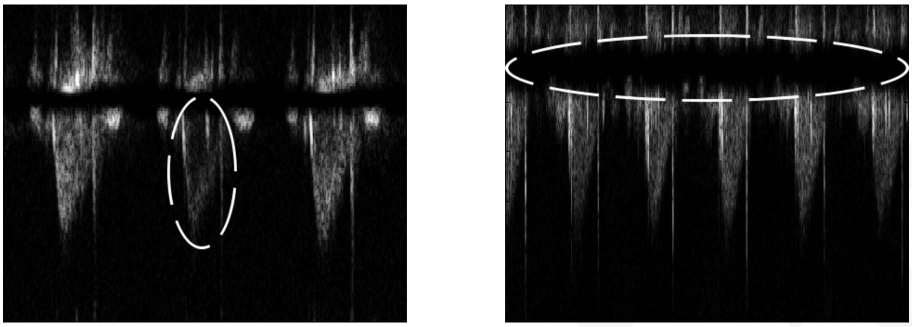
## 1.2 Technical motivation

Clinically obtained aortic outflow velocity images sometimes differ significantly, resulting in low (local) correlation and different resolutions with differing texture. Additionally, Doppler ultrasound images inherently contain a lot of (speckle) noise, various acquisition artifacts, and poor contrast, making the segmentation even more challenging. Some of these challenges are depicted in Figure 1.1, e.g. notice how the brightness and contrast vary across images. An example of the object that needs to be segmented is the object within the white ellipsoid in the Figure 1.1a. This object represents the velocities of the objects along the scanning line when the heart valves are open i.e. when the blood flows form the heart into the aorta. Although the region of interest is usually the largest bright region within image, this is not always the case, as we can see in the Figure 1.1d. Additional problems are low velocity rejection region which depends on the machine settings (clutter filter), and can significantly vary in size (see Figure 1.1b) depending on the user's settings, and aliasing and valve clicks (see Figure 1.1c and Figure 5.1).

To reduce the influence of some of the artefacts, the preprocessing step described in Chapter 4, will be applied to images, but this will not solve everything. For example, valve clicks represent a portion of the signal concatenated to the end or the beginning of the region of interest which arise at valve opening and closing. This obviously distorts the information about blood outflow velocities, since in many cases the maximal velocity recorded by Doppler ultrasound does not represent the maximal blood outflow velocity but rather the maximal valve velocity. This is a typical example of the ill-defined boundary between valve clicks and blood outflow since it is expected of the segmentation procedure to be able to differentiate between valve clicks and blood outflow.

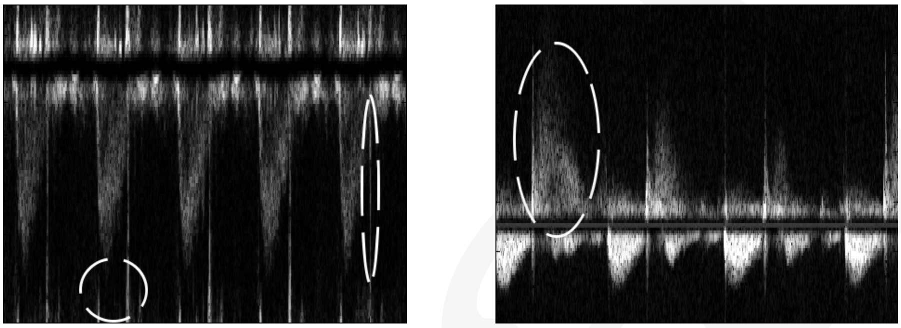
It is illusory to believe that segmentation facing all these problems can be achieved using pixels' intensities information only. As a solution to this, usually a prior knowledge is utilized. One way to do this is to incorporate the knowledge within the segmentation process in the form of the model that will be used as a prototype or template for segmentation of desired object. Comparative advantage of the model-based segmentation with respect to the other segmentation methods is the ability to segment the image with no well defined relation between regions and pixels' intensities. This is usually the case when the objects of the same structure need to be segmented (i.e. have the same texture), and the information about difference between these object is incorporated in spatial relationship between them, other objects, or within their morphometric characteristics.

Following this, the model-based approach seemed a natural choice for this type of problem. In addition, the use of models for image segmentation is one of most active and successful research area in image segmentation community (as we will see in Chapter 3).



(a) Dashed ellipsoid indicates problem with the poor contrast

(b) Dashed ellipsoid indicates the low velocity region



(c) Dashed ellipsoid indicates artifacts (referred by cardiologist as aliasing and valve clicks)

(d) Dashed ellipsoid indicates largest bright region

Figure 1.1: The figure depicts the issues with the region-based approach of the aortic profile detection and segmentation, such as poor contrast (a), large low velocity region (b), various artifacts (c), and other bright region apart from region of interest (d).

## 1.3 Problem statement

Challenges that analysis and quantization of any image impose are challenges to accurately segment the image and subsequently extract relevant (morphological) features. A decision is made to use the model-based segmentation for ultrasound aortic outflow profile segmentation. In order to perform a model-based segmentation, firstly we have to have a model, and secondly we have to have a definition of the correspondence between model image and the image. Thus, apart from segmentation, we have to solve the problem of model construction/selection, and the problem of defining the similarity between images. Finally, once the model-based segmentation is done, we are interested in answering to the challenge of practical use of the developed method in clinical practice, apart from reliability and reducing amount of manual work which is expected from the automated method.

Therefore the challenges imposed to this thesis are the following:

- Transaortic profile segmentation and feature extraction
- Model construction
- Image similarity measure definition
- Practical use of automated method for transaortic profile segmentation

The challenges are addressed in Chapters 5–8 in this exact order. Chapter 5 describes a model-based segmentation of the transaortic profile. The segmentation results and extracted features are compared to manual gold standard, showing the reliability, accuracy and the potential of the developed method. The chapter does not only solve the problem of aortic outflow velocity profile segmentation, but also presents a more general approach for segmentation of other (cardiac) images. Furthermore, it sets a framework for atlas construction that is used further on. Chapter 6 discuss several methods for model construction, and introduces several ideas to increase the accuracy of the segmentation, by careful model formation. In Chapter 7 image similarity measure (ISM) is presented, based on the absolute joint moments. In the chapter we show that the novel ISM relates to both correlation coefficient and mutual information, and combines their properties in a favourable way. Finally, Chapter 8 discuss the relationship between functional recovery of the AS patients after aortic valve replacement, and the shape of the aortic outflow velocity profile. Apart form reliability, reduced amount of manual work, and accuracy comparable to intra-observer accuracy of the method described in Chapter 5, Chapter 8 shows that the developed method is useful in clinical practice.

[illegible]...

*Missale Romanum Glagolitice - 1483.*
*First Croatian printed book*

# Chapter 2

# Preliminaries

## 2.1 Probability theory basics

**Definition 2.11 (Real-valued random variable)** *Let $\Omega$ be the set of all possible events and $\omega \in \Omega$ one (elementary) event. A real-valued random variable is a mapping from set $\Omega$ to real line $\mathbb{R}$:*

$$X : \Omega \to \mathbb{R}. \tag{2.1}$$

Definition 2.11 states that for each realization of elementary event $\omega$ a random variable $X(\omega)$ is a real value.

**Definition 2.12 (Relative frequency and probability)** *If event $\omega \in \Omega$ occurred $n_\omega$ times in a $n$ observation, then $n_\omega$ is called the frequency and the $n_\omega/n = p_\omega$ is called relative frequency, or experimental probability. The relative frequency satisfies certain properties which can be used to built up an axiomatic definition of the notion of the probability $P(\omega)$:*

- $0 \leq P(\omega) \leq 1, \forall \omega \in \Omega$
- $P(O) = 0, \quad P(I) = 1$, *where $I$ is certain event, and $O$ impossible event*
- $P(\omega_1 + \omega_2 + \dots) = P(\omega_1) + P(\omega_2) + \dots$ *if $\omega_i \in \Omega, (i = 1, 2, \dots)$ are finite or countable many mutually exclusive events (i.e. $\omega_i \omega_j = O \, \forall i \neq j$)*

**Definition 2.13 (Cumulative distribution function)** *The cumulative distribution function of a random variable $X$ is defined by:*

$$F_X(x) = \mathrm{P}(X < x), \tag{2.2}$$

*where the right side represents the probability that the random variable $X$ has value less than $x$.*

**Definition 2.14 (Expected value)** *The general definition expected value (expectation) of the random variable $X$ is given by:*

$$\mathrm{E}[X] = \int_{\Omega} X \,\mathrm{d}\, F = \int_{\Omega} X(\omega)\, F(\mathrm{d}\,\omega) \tag{2.3}$$

*The expectation is also know as mean (sometimes denoted by $\mu_X$).*

If $\Omega$ is a set of real numbers the expected value of a continuous random variable can be calculated as:

$$\begin{aligned} \mathrm{E}[X] &= \int_{-\infty}^{\infty} x \,\mathrm{d}\, F(x) && (2.4) \\ &= \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}\, x && (2.5) \end{aligned}$$

This follows from the relation[1] between cumulative distribution function and probability density function:

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x} F(x). \tag{2.6}$$

The Equation 2.5 corresponds well with its discrete case counterpart, where for discrete random variables the expectation is defined as:

$$\mathrm{E}[X] = \sum_{i=1}^{\infty} x_i\, p_i, \tag{2.7}$$

where $p_i$ stands for a relative frequency of the event $x_i$. The latter equation has also a physical interpretation where expectation represents the center of mass (gravity), and $p_i$ is also referred as probability mass function. The summation in the Equation 2.7 can be also interpreted as the vector product. More generally, for continuous random variable, the expected value of an arbitrary function of $x$, e.g. $g(x)$, with respect to the probability density function $f(x)$ is given by the inner product of $f$ and $g$:

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,\mathrm{d}\, x. \tag{2.8}$$

Notice the similarity with cross-correlation and convolution.

**Definition 2.15 (Properties of expectation operator value)** *The expected value operator (or expectation operator) E is linear in the sense that*

$$\begin{aligned} \mathrm{E}[X + c] &= \mathrm{E}[X] + c && (2.9) \\ \mathrm{E}[X + Y] &= \mathrm{E}[X] + \mathrm{E}[Y] && (2.10) \\ \mathrm{E}[cX] &= c\,\mathrm{E}[X] && (2.11) \end{aligned}$$

*where $c$ is a real number (constant).*

---

[1] With assumption that $f$ is continuous at $x$.

Combining the results from previous three equations, we can see that

$$\mathrm{E}[aX + b] = a\,\mathrm{E}[X] + b \tag{2.12}$$

$$\mathrm{E}[aX + bY] = a\,\mathrm{E}[X] + b\,\mathrm{E}[Y] \tag{2.13}$$

for any[2] two random variables $X$ and $Y$ and any real numbers $a$ and $b$. The Equation 2.13, generally holds for any number of random variables:

$$\mathrm{E}[\sum_{i=1}^{n} c_i X_i] = \sum_{i=1}^{n} c_i\,\mathrm{E}[X_i], \tag{2.14}$$

therefore we can say that expectation preserves linear combinations. Moreover, we can also write:

$$\mathrm{E}[\int_{\mathbb{R}} g(t)X(t)\,\mathrm{d}\,t] = \int_{\mathbb{R}} g(\mathrm{E}[X(t)])\,\mathrm{d}\,t \tag{2.15}$$

$$since$$

$$\int_{\Omega}\int_{\mathbb{R}} g(t)X(t)\,\mathrm{d}\,t\,\mathrm{d}\,F = \int_{\mathbb{R}} g(t)\int_{\Omega} X(t)\,\mathrm{d}\,F\,\mathrm{d}\,t \tag{2.16}$$

However, the expectation operator and functions of random variables generally do not commute:

$$\mathrm{E}[g(X)] \neq g(\mathrm{E}[X]) \equiv \tag{2.17}$$

$$\int_{\Omega} g(X)\,dF \neq g(\int_{\Omega} X\,dF) \tag{2.18}$$

but, for a convex function $g$, the relationship is more accurately described by Jensen's inequality [86, 42]:

$$\mathrm{E}[g(X)] \geq g(\mathrm{E}[X]) \tag{2.19}$$

Notice that the expected value operator is not multiplicative, i.e. $E[XY] = E[X] \cdot E[Y]$ holds if $X$ and $Y$ are independent. In case the variables $X$ and $Y$ are not independent, the error made by assuming the opposite is given by covarinace:

$$\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]. \tag{2.20}$$

**Definition 2.16 (Variance)** *The variance of a random variable $X$ is defined as:*

$$\mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2] = \mathrm{E}[X^2] - \mu^2\,. \tag{2.21}$$

*Standard deviation ($\sigma_X$) is square root of the variance, whereas the notation $\sigma_X^2$ is sometimes also used for variance.*

---

[2] $X$ and $Y$ do not need to be statistically independent or even defined on the same probability space.

As perhaps noticed, the variance is expected value of the variable $Y = (X - \mathrm{E}[X])^2$. We can generalize this by denoting the expected values of the powers of $X$ as the moments, and the expected values of powers of $X - \mathrm{E}[X]$ as central moments. Here are formal definitions:

**Definition 2.17 (Moments, central moments)** *The $k^{th}$ moment of a random variable $X$ is defined as:*

$$\mu_k = \mathrm{E}[X^k]. \tag{2.22}$$

*Similarly, the $k^{th}$ central moment of a random variable $X$ is defined as:*

$$\mu'_k = \mathrm{E}[(X - \mu_X)^k], \tag{2.23}$$

*where $\mu_X$ denotes mean, defined as in 2.14.*

**Definition 2.18 (Properties of central moments)** *The expected value operator (or expectation operator) E is linear in the sense that*

$$\mu'_k(X + c) = \mu'_k(X) \tag{2.24}$$

$$\mu'_k(cX) = c^k \mu'_k(X) \tag{2.25}$$

$$\mu'_k(X + Y) = \mu'_k(X) + \mu'_k(Y) \text{ if } k \leq 3 \text{ and } X \text{ and } Y \text{ independent.} \tag{2.26}$$

*where $c$ is a real number (constant).*

## 2.2 Generating functions

**Definition 2.21 (Moment-generating function)** *The moment-generating function of a random variable $X$ is defined as:*

$$m_X(t) := E\left[e^{tX}\right], \tag{2.27}$$

*where $t \in \mathbb{R}$.*

Notice that $m_X$ is equal to $G_p$ if $z = e^t$.

**Definition 2.22 (Characteristic function)** *The characteristic function of a random variable $X$ is defined as:*

$$m_X(it) := E\left[e^{itX}\right], \tag{2.28}$$

*where $t \in \mathbb{R}$.*

According to the usual convention of the continuous Fourier transform, the characteristic function is actually a complex conjugate of the continuous Fourier transform of $p(x)$.

**Definition 2.23 (Cumulant-generating function)** *The cumulant-generating function of a random variable $X$ is defined as:*

$$k_X(t) := \ln E\left[e^{tX}\right], \tag{2.29}$$

*where $t \in \mathbb{R}$.*

The relationship between cumulant- and moment-generating functions is given by identity:

$$k_X(t) := \ln m_X(t). \tag{2.30}$$

An important property of the moment-generating function is that:

$$m_X(t) = m_Y(t) \Rightarrow F_X(x) = F_Y(x) \tag{2.31}$$

This means that if two distributions have the same moment-generating function, then they are identical at all points. This is not equivalent to "if two distributions have the same moments, then they are identical at all points", because in some cases the moments does not exist even when the moment-generating function exists, or vice versa. Log-normal or Cauchy distribution are examples of these cases (for details consult [54], p.224).

**Definition 2.24 (The moment problem)** *If there exists a sequence of moments $\mu_n (n = 1, 2, \dots)$ given by:*

$$\mu_n = \int x^n \, dF(x), \tag{2.32}$$

*can $F$ be reconstructed from $\mu_n$?*

To address the problem defined in 2.24 one have to answer whether the inverse mapping given by the Equation 2.32 exists and whether it is unique. There are three classical moment problems which differ by different support of $F$ (different integral boundaries in the problem definition), namely: Hamburger (support is $\mathbb{R}$), Stieltjes (support is $[0, \infty]$), and Hausdorff moment problem (support is $[0, 1]$).

The problem can be generalized to the from:

$$g(n) = \int h(x, n) \, dF(x), \tag{2.33}$$

for an arbitrary functions $g(n)$ and $h(x, n)$. Notice that this formulation of the problem corresponds to the problem of reconstructing $F$ from moment-generating function (Equation 2.37).

## 2.3 Moments and cumulants

A generating function is a clever way to write down a sequence of numbers [238]. For example, the moment-generating function is defined in Equation 2.27 can be used to find all the moments of the distribution. This becomes more clear if we expand $e^{tX}$:

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \frac{t^3X^3}{3!} + \cdots \tag{2.34}$$

which leads to:

$$m_X(t) = E(e^{tX}) = 1 + t\mu_1 + \frac{t^2\mu_2}{2!} + \frac{t^3\mu_3}{3!} + \cdots, \tag{2.35}$$

where $\mu_k$ stands for $k^{th}$ moment, which can be calculated (generated) from the $k^{th}$ derivative of the moment-generating function (hence the name):

$$E\left(X^k\right) = m_X^{(k)}(0) = \frac{d^k m_X}{dt^k}(0). \tag{2.36}$$

Therefore, we can say that the moment-generating function is a clever way of organizing the moments in one mathematical object. Notice also, that the moment-generating function is the function of the form:

$$\begin{aligned} m_X(t) &= \int_{-\infty}^{\infty} e^{tx}\, dF(x) && (2.37)\\ &= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx. && (2.38) \end{aligned}$$

Thus, the moment-generating function of a random variable is an alternative definition of its probability distribution.

Similarly, the cumulant-generating function generates cumulants, and we can define cumulants via cumulant-generating functions:

**Definition 2.31 (Cumulant)** *Cumulants are the coefficients of Taylor (Maclaurin) expansion of the cumulant-generating function about the origin:*

$$\kappa_n = k_X^{(n)}(0) = \frac{d^n k_X}{dt^n}(0). \tag{2.39}$$

The relationship between moments and cumulants can be established via their generating function, e.g.:

$$\begin{aligned} k_X(t) = \ln(\mathrm{E}(e^{tX})) &= -\sum_{n=1}^{\infty} \frac{1}{n}\left(1 - \mathrm{E}(e^{tX})\right)^n = -\sum_{n=1}^{\infty} \frac{1}{n}\left(-\sum_{m=1}^{\infty} \mu'_m \frac{t^m}{m!}\right)^n \\ &= \mu'_1 t + \left(\mu'_2 - {\mu'_1}^2\right)\frac{t^2}{2!} + \left(\mu'_3 - 3\mu'_2\mu'_1 + 2{\mu'_1}^3\right)\frac{t^3}{3!} + \cdots. \end{aligned} \tag{2.40}$$

## 2.4 Image registration basics

**Definition 2.41 (Image)** *Image is a function*

$$I : X \to Y \tag{2.41}$$

*where* $X \in \mathbb{R}^m$ *and* $Y \in \mathbb{R}$*.*

**Definition 2.42 (Source, target and reference image)** *Source image* $S(\mathbf{x})$ *is input image of the registration process. The input/source image is geometrically deformed* $(\mathcal{G}(\mathbf{x}))$ *to match the target image* $T(\mathbf{x})$*. The reference image* $R(\mathbf{x})$ *is the image against which the resulting image of the registration process* $(S(\mathcal{G}(\mathbf{x})))$ *is evaluated.*

Note that some authors do not distinguish between the target and reference image. However this distinction is helpful if we have training and testing (or target and reference) image sets, and it was suggested in one of the earliest papers dealing with atlas registrations [141].

**Definition 2.43 (Image registration)** *Image registration is a process of finding the optimal geometric transformation* $\mathcal{G}$ *for which the similarity measure* $SM$ *between source image* $S(x)$ *and target image* $T(x)$ *becomes maximal:*

$$\mathcal{G}_{optim} = arg\max_{\mathcal{G}} SM(S(\mathcal{G}(\mathbf{x})), T(\mathbf{x})) \tag{2.42}$$

**Definition 2.44 (Segmentation and labels)** *Image segmentation is the process of partitioning an image into multiple segments. Image segment is part of the image delineated by the border (segmentation). Label is identification object assigned to the part of an image. Thus, labels can be presented as an image:*

$$L : X \to [0, l] \tag{2.43}$$

*where* $X \in \mathbb{R}^m$ *and* $l \in \mathbb{N}$*.*

**Definition 2.45 (Template, prototype, model)** *Template is an object from a population used as a target or a reference. Prototype is an object selected as the most representative object of the population. Model is an object constructed with intention to be used as a population representative.*

**Definition 2.46 (Atlas)** *Atlas is an ordered pair of intensity image* $I(\mathbf{x})$ *and label image* $L(\mathbf{x})$*:*

$$A(\mathbf{x}) = (I(\mathbf{x}), L(\mathbf{x})) \tag{2.44}$$

## 2.5 Geometric transformation

The goal of image registration is to determine parameters of the geometric transformation (deformation field), that maps a source image into a reference image.

The notation used throughout the paper denotes $S(\mathbf{x})$ and $T(\mathbf{x})$ as the source and target image, respectively, where $\mathbf{x}$ stands for the coordinates vector of the image. The deformed source image, or the output image (of the registration process) is denoted as $O(\mathbf{x}) = S(\mathcal{G}(\mathbf{x}))$. Sometimes, $S'(\mathbf{x})$ will be used to denote the geometric transformation of the image $S(\mathbf{x})$ (i.e. $S'(\mathbf{x}) = S(\mathcal{G}(\mathbf{x}))$) obtained by the successive estimate of the registration transformation $\mathcal{G}$. The $\mathbf{x}$ denotes vector of the underlying image space. In our case $\mathbf{x}$ is the ordered pair in Cartesian coordinate system $(t, v)$, since Doppler ultrasound images represent the instantaneous blood velocity ($v$) within the sample volume (pulsed Doppler) or scan line (continuous wave Doppler) as a function of time ($t$).

Perhaps the most elegant way to describe the geometric transformation are homogeneous coordinates. The homogeneous coordinates were introduced by August Ferdinand Möbius in his 1827 work *Der barycentrische Calcül*, to ease claculus by representing affine (and projective) transformation in matrix form. Here, we give definitions for few geometric transformations in 2D case, since we will primarily use 2D images. The generalization for 3D case is straightforward and for further reference please consult [201, 78].

**Definition 2.51 (Translation matrix)** *Translation matrix is any matrix of the form:*

$$\mathcal{T} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2.45}$$

*where $t_x$ and $t_y$ stand for translation in $x$ and $y$ direction, respectively.*

**Definition 2.52 (Rotation matrix)** *Rotation matrix is any matrix of the form:*

$$\mathcal{R} = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.46}$$

*$\theta_z$ for where $\theta_z$ is the angle of rotation around z-axis.*

**Definition 2.53 (Scaling matrix)** *Scaling matrix is any matrix of the form:*

$$\mathcal{S} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.47}$$

*where $s_x$ and $s_y$ stand for scaling factor in $x$ and $y$ direction, respectively.*

The rigid transformation is defined as transformation that preserve all distances. This means that rigid transformation allows image only to rotate and translate. Using homogeneous coordinates this can be written in matrix notation as:

$$\mathbf{x}' = \mathcal{R}\mathcal{T}\mathbf{x} \tag{2.48}$$

where **x** represents column vector with the coordinates of the image with appended one in the last row (see Eq. 2.49)

**Definition 2.54 (Rigid transformation)** *Rigid transformation of the space* $(x, y)$ *to space* $(u, v)$ *can in homogeneous coordinates be written as:*

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{2.49}$$

All transformations mentioned and their combinations are special cases of more general affine transformation which preserves straightness of lines and parallelism, and whose homogeneous matrix is given by Definition 2.55.

**Definition 2.55 (Affine transformation)** *Affine matrix is any matrix of the form:*

$$\mathcal{A} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \forall a_{ij}, t_x, t_y \in \mathbb{R} \tag{2.50}$$

Curved transformations are the most flexible transformations. The simplest functional form in which we can write this transformation is a polynomial in the components of **x**:

$$\mathbf{x}' = \sum_{ij}^{IJ} \mathbf{c}_{ij} x^i y^j \tag{2.51}$$

## 2.6 Similarity measures

The purpose of an SM is to quantify the similarity between two images, usually referred to as source and target image. Considering the elementary problem of measuring the (dis)similarity between two images $O$ and $T$ , the simplest idea is to use a distance measure, e.g. euclidean distance.

**Definition 2.61 (Euclidean distance)** *Euclidean distance between* $O(x)$ *and* $T(x)$ *can be calculated as:*

$$d = \sqrt{\sum_{i=1}^{n} (O(x_i) - T(x_i))^2} \tag{2.52}$$

If images $O$ and $T$ are observed as random vectors we can define distance as:

$$
\begin{aligned}
D &= E[(O(x) - T(x))^2] && (2.53)\\
&= E[O(x)^2 - 2 \cdot O(x)T(x) + T(x)^2] && (2.54)\\
&= E[O(x)^2] - 2 \cdot E[O(x)T(x)] + E[T(x)^2] && (2.55)
\end{aligned}
$$

where $E[.]$ denotes the expectation operator. If the $E[O(x)^2]$ and $E[T(x)^2]$ from the equation 2.55 are constant, the negative of the distance measure $D$ will have qualitatively the same behavior as the correlation between $O$ and $T$:

$$C = E[O(x)T(x)] \tag{2.56}$$

Correlation has a simple geometrical interpretation if $O$ and $T$ are vectors. If $O$ and $T$ are (unit) vectors, their correlation is directly proportional (equivalent) to angle between them.

The correlation between centered random vectors $O$ and $T$ (i.e $E[(O(x)-\mu_O)(T(x)-\mu_T)]$) has interesting property to achieve maximum not only if $O$ and $T$ are identical but also if they are related by the function of the form: $T(x) = O(x) + k$. This means that, in case when $O$ and $T$ are images, this similarity measure is unaffected by the intensity offset which may happen due to the homogeneous change in lightning conditions. The correlation coefficient, given by Definition 2.62, represents a further enhancement in the robustness of similarity measure, due to its ability to detect similarity between images whose pixel values are in affine relationship ($T(x) = a{\cdot}O(x)+k$). This can be valuable e.g. when images of the same object but with different contrast are compared.

**Definition 2.62 (Correlation coefficient)** *Correlation coefficient between $O(x)$ and $T(x)$ is defined as:*

$$CC(T, O) = \frac{E[(T(x) - \mu_T)(O(x) - \mu_O)]}{\sigma_T \cdot \sigma_O} \tag{2.57}$$

*where $\mu_O$ and $\mu_T$ denotes the image average, and $\sigma_T$ and $\sigma_O$ the image standard deviation.*

Moreover, it was also shown that the correlation performs well if the images differ only by simple noise like Gaussian [231], or in some case Rician [201].

Another approach is to characterize the relationship between $T$ and $O$ in an information theoretic way, by utilizing the joint probability density function (PDF) $p_{TO} = p(T, O)$. The joint PDF allows us to describe any type of functional (or only statistical) relationship between images and any noise that may exists in the image acquisition process. One way to measure the amount of dispersion of the joint PDF is to use joint entropy [82]:

$$H(T, O) = -E[\ln p_{TO}(x, y)] \tag{2.58}$$

But alternatives are also mutual information (MI) and normalized mutual information (NMI).

**Definition 2.63 (Mutual information)** *Mutual information between $O(x)$ and $T(x)$ is defined as:*

$$MI(T,O) = E[\ln \frac{p_{TO}(x,y)}{p_T(x)p_O(y)}] \tag{2.59}$$
$$= H(T) + H(O) - H(T,O) \tag{2.60}$$

Mutual information was proposed as a more robust solution than joint entropy, since joint entropy is defined only on the overlapping region of images $T$ and $O$ (i.e. $T \cap O$), and was affected by the change in overlap. Later it was shown that the normalization of the form:

$$NMI(T,O) = \frac{H(T) + H(O)}{H(T,O)} \tag{2.61}$$

produces even better results.

For our further reference, it is worthwhile to notice that MI can be related to relative entropy [114] or Kullback-Leibler (KL) divergence [15], as a measure of the difference between two probability distributions $p$ and $q$.

**Definition 2.64 (KL divergence)** *KL divergence or relative entropy between two probability distributions $p$ and $q$ is given by:*

$$D_{KL}(p||q) = -\int_{-\infty}^{\infty} p(x)\ln q(x)dx - (-\int_{-\infty}^{\infty} p(x)\ln p(x)dx) \tag{2.62}$$
$$= \int_{-\infty}^{\infty} p(x)\ln \frac{p(x)}{q(x)}\,dx \tag{2.63}$$

*in case $p$ and $q$ are continuous random variables, analogously, for a discrete random variables $p$ and $q$:*

$$D_{KL}(p||q) = -\sum_x p(x)\ln q(x) + \sum_x p(x)\ln p(x) \tag{2.64}$$
$$= \sum_x p(x)\ln \frac{p(x)}{q(x)}. \tag{2.65}$$

## 2.7 Taylor series and Hermite polynomials

If a function $f(x)$ has all derivatives at $x = a$, then it can be represented with the Taylor expansion as a power series [18]:

**Definition 2.71 (Taylor series)** *The Taylor series of a real or complex function $f(x)$ that is infinitely differentiable in a neighborhood of a real or complex number $a$ is the power series:*

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots . \end{aligned} \tag{2.66}$$

Maclaurin series are special case of Taylor series with $a = 0$.

For example, the expansion for natural logarithm is given by Equation 2.67, and it is valid $\forall x \in \langle 0, 2]$, while the (Maclaurin) expansion of the exponential function is given by Equation 2.68 and it is valid $\forall x \in \mathbb{R}$.

$$\begin{aligned} \ln(x) &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{(x-1)^{n+1}}{n+1} = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(x-1)^n}{n} \end{aligned} \tag{2.67}$$

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned} \tag{2.68}$$

**Definition 2.72 (Hermite polynomials)** *Hermite polynomials are polynomials orthogonal with respect to measure $e^{-x^2}$ (normal distribution), and are given by:*

$$h_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2} \tag{2.69}$$

*In each field we must be careful to distinguish three aspects of the theory: (a) the formal logical content, (b) the intuitive background, and (c) applications. The character, and the charm, of the whole structure cannot be appreciated without considering all three aspects in their proper relation.*

*William (Srećko) Feller*

*An Introduction to Probability Theory and Its Applications - Volume I*

## Chapter 3

# Prior Art

Before we try to solve the problem specified in Chapter 1, in this chapter, we will first review the existing solutions of similar problems covered by the literature. This chapter also gives an overview of the existing techniques used in medical image analysis and image segmentation, with main accent on the methods proposed for segmentation of arotic outflow velocity profiles. Thus, the model based segmentation techniques used in medical applications are discussed. Since our main focus are atlas based segmentation techniques, an overview of atlas construction methods is given. This is followed by the state of the art in the field of image registration. Papers in the field of image registration can be classified in various ways, depending on the image acqusition technique or depending on the building blocks used, namley: geometrical transofrmation, optimization and similarity measure. A stronger emphasis is on similarity measures (SM) than other registration components since Chapter 7 presents a novel image similarity measure.

## 3.1 Medical image analysis

Medical image analysis deals with problem of extraction of meaningful information from images acquired by medical imaging techniques like: magnetic resonance imaging (MRI), nuclear medicine tomography (like PET or SPECT), computed tomography (CT), ultrasound, etc. Many works and textbooks cover the filed of medical image analysis (see [201, 217] and their references), however, our main interest is in the analysis of the ultrasound images, or to be more specific, ultrasound blood flow velocity profiles analysis. Often, the fist step in image analysis is image segmentation, whose goal is to simplify or change the representation of an image into something more meaningful or easier to analyze. Many ultrasound image segmentation methods are covered by survey paper of Noble et al. [150], however, our particular interest lies in the model-based segmentation since they provide an elegant way to incorporate a prior knowledge to the image segmentation process.

### 3.1.1 Blood flow velocity profile analysis

To the best of our knowledge there are no studies on the analysis of blood flow velocity profiles obtained by Doppler ultrasound published in literature, apart from the works of Tschirren et al. [221] and Bermejo et al. [10]. Tschirren et al. presented an automated cardiac cycle and envelope extraction of brachial artery flow profile based on image processing operations such as thresholding and correlation. However, this approach is not suitable for the cardiac outflow profiles mainly because it also segments the valve clicks (see Figure 5.1), not just the blood outflow. The work of Bermejo et al. analysed outflow profiles that are averaged and manually segmented, with a goal to analyse the valvular dynamics, so this work uses both a different methodological approach and a different hypothesis.

On the other hand, the published research on image segmentation and registration techniques is rather extensive (see [201, 251] and their references).

### 3.1.2 Image segmentation

Image segmentation, defined as the separation of the image into regions, is one of the first steps leading to image analysis, interpretation, and object quantification [47]. From segmented image the desired objects can be separated from the background, measured, counted or in other means quantified. Its goal is to simplify or change the representation of an image into something more meaningful or easier to analyze. It is used in many practical applications in machine vision, biometric measurements, medical imaging etc. for the purpose of detecting, recognition or tracking of an object. The image segmentation techniques can be classified with respect to the object, sensor (modality) or application, where different surveys already exist (see e.g. [160, 150] or [239]). If the image segmentation is observed as a pattern classification problem we can (similar to [237]) divide the field in the two schools: one that uses features and one that uses templates for segmentation. The approach of the first school can be observed as an indirect method where features are compared, rather than objects (or patterns [237]) as in the second school. However, the first approach is faster, due to the preprocessing step that reduced the amount (dimensionality) of the information. The feature-based and template-based techniques roughly corresponds to low-level and high-level segmentation techniques as classified in [47]. Here adjectives low-level and high-level do not describe the level of complexity of the problem, but rather the level of abstraction, same as in the low-level and high-level computer vision. Thus, in the template-based image segmentation the segmentation problem is moved away from the pixel intensity properties to more abstract formulation of the problem, where the higher level of a priori information is used in the segmentation process.

## 3.2 Model-based segmentation

Model-based segmentation is essentially a template or prototype matching technique. The problem of matching a template to an object is actually quite old, and has been approached in a number of different way, and addressed by numerous names. Here we give an overview on the existing techniques.

### 3.2.1 Template matching

The correlation can be observed as the most common solution to the problem of template matching. Matched or North filters [223, 151] address this problem via Fourier transform of the image and the template, where after multiplication the peak is sought. This kind of a template matching used a fixed template with a measure of "goodness" of matching [56], where the peak of the "goodness" function is sought by the exhaustive search. The Fischler and Elschlager's [56] spring-loaded templates and Widrow's rubber mask [237], came as a natural extension to the problem, since they introduce a deformable template.

### 3.2.2 Deformable models

Th approach proposed by Fischler and Widrow was popularized more than a decade later, by the works of Terzeopoulos et al. [103, 213, 214, 215], in which the term model superseded the term template, and the term energy/cost function superseded the term "goodness" of match. The formulation of the energy function via differential equations enabled easier implementation of the optimization algorithm, and faster search for minimal energy/cost. The deformable models grown to be one of most active and successful research area in image segmentation under various names such as deformable surfaces or contours [117, 67, 109, 250], active contours or surfaces [71, 11], snakes or balloons [103, 121, 222], active shape [41, 39] or appearance [38, 40], etc., and are usually divided to parametric deformable models [103] and geometric deformable models [23, 137]. Geometric deformable models utilize either curve evolution theory [190, 191, 24, 249] or level set methods [125, 154, 33, 187, 188]. The level set formalism [152] is interesting as it introduces several new possibilities by allowing topological changes. In addition, efficient numerical schemes are available, and the extension of the method to higher dimensions is easy. A comprehensive survey on the deformable models and their use in the medical image analysis can be found in [143], or [242], while the work of Gibson et al. [68] provides a survey on different techniques used for deformable models in computer graphics.

Compared to other methods of segmentation, model-based segmentation has the ability to segment images without well defined relation between objects and pixels' intensities. This turns out to be very valuable when the objects of the same texture need to be segmented, or when the border is occluded by excessive noise. To understand the possible shortcoming of the model based methods, it has to be understood that the successful mapping of one dataset into the other does not guarantee that it also makes sense as an anatomical mapping, even if the alignment looks perfect. For example, to map a leg into a nose is perfectly possible, but this does not allow any reasonable physiological interpretation. To cope with this problem some authors propose the use of statistical models [210], which use statistical information about possible anatomical variations. The parametrization and dimensionality reduction can usually be effectively achieved by PCA [93]. Active shape [41, 39] and active appearance models [37, 38, 40] fall into the category of statistical models.

### 3.2.3 Model image

If we model solely the shape of the object, or it's color and texture (appearance) and the physical property of the object (like elasticity), we build either active shape or active appearance model. While different classification and various extensions to these classes are possible (see [242]), we are interested in the case when the whole scene is interpreted as one object. This is particularly important when one wants to model not just one object and its morphological features, but multiple objects and geometrical relationship between them.

An image represents a scene captured by the imaging device (sensor). If we are to find the objects in the images that do not depend only on the pixel property of the object or the morphological characteristics of the objects itself, but on the relationship between objects as well, it is natural to model the whole scene, not just the objects. The model of the scene can be obtained in several ways, e.g. from computational (numerical) phantom (e.g [35, 185, 153, 120, 107, 95], or [88] p.439.), or from physical (imaging) phantom (e.g. [65, 203, 147, 204, 95]). Similarly, a sample image can be used as a template [173], or several images can be used to construct a model image [173, 240, 230]. A further discussion about various methods for model construction is given in Section 3.3.

If we are to determine the correspondence between target image/scene and the model, we utilize geometrical transformation to maximize similarity (or minimize distance) between them. This actually describes the process of image registration, whose background is described in more details in Section 3.4. When the correspondence between model and target has been determined, the segmentation of the target image is determined via segmentation propagation. Segmentation propagation is a process in which the segmentation/labels from the model is/are transformed by the same geometric transformation for which the maximal similarity between the model and the target is achieved. Thus, the model consists not only of image model but also of the segmentation model. An pair image plus segmentation (labels), we refer as the atlas.

### 3.2.4 Atlas-based Segmentation

The advantage of atlas-based segmentation when compared to other image segmentation is a relatively simple way to introduce a prior knowledge about shape and distribution of the segmented structures. This simplicity is due to the fact that only a presegmented image is necessary as a reference that guides the segmentation, e.g. no training is necessary to store the expert information.

On the other hand the main disadvantage of an atlas-based segmentation is the computational complexity, i.e. the time necessary for atlas construction and registration wherever iterative procedure (e.g. [138, 252, 119]), or a complex non-rigid registration (e.g. [45, 202]) is utilized. Since the atlas-based segmentation is usually used when the information from the gray level intensities are not sufficient, it is difficult to produce objective validation. Whenever the manual or semi-manual segmentation is used as a golden standard one cannot observe this as objective validation since it is user dependent. As a solution for this problem, the unsupervised evaluation which is automatic and user independent by definition and thus considered (more) objective, can be proposed (as in [248]). However, if we do have automatic (unsupervised) evaluation why not use it for further enhancement of the segmentation procedure? Naturally, after we have used this evaluation within a segmentation process we cannot (re-) use it to objectively evaluate the segmentation process, and by looking for the different evaluation method we again end up with (semi-) manual gold standard. Other evaluation methods for registration or segmentation can be fund in [225, 111, 234].

### 3.2.5 Atlas and its use in medical applications

Models with a common anatomical substrate are in medical applications often known as atlases. Atlas incorporates useful prior information for segmentation and registration tasks, so variation within population can be described with fewer (transformation) parameters. Atlases have broad application in medical image segmentation and registration and are often used in computer aided diagnosis to measure the shape of an object or detect morphological differences between patient groups. Various atlas-based techniques are developed for different human organs, like multiple abdominal organs (liver, kidneys etc.), [122, 156, 229, 198], prostate [113, 49], lungs [106, 199, 200], head and neck [71, 166, 70], heart [63, 129, 158, 128], and especially the brain [87, 2, 55, 94, 13, 127, 126, 8, 178, 52, 53, 206, 218].

## 3.3 Atlas construction

From a set of images, any image can be used as a template, where the most representative instance is referred as the prototype. Sometimes, the most appropriate image to represent the set will not belong to the set. Different methods to select or construct the representative instance from the set are presented in this section. To set the nomenclature, we distinguish between prototype, template and model. An instance from the population is referred as template, while the most representative one (under whichever criterion) is referred as prototype. If the information from several templates is used to construct the population representative we refer to it as model. Thus, the prototype is carefully selected from the population, while the model is constructed. It is important to notice that in an atlas formation procedure it is possible to have template, prototype or model for both image and labels. As we will see, most of the authors only discussed how to select or construct atlas image, since it is a common belief that insufficient similarity between the atlas and the target image often results in local mismatches, which in turn leads to segmentation errors [230, 43].

The most trivial approach for atlas image selection is just to pick a random image from a population sample, but by doing this, we have a great chance to pick an image that does not represent a typical instance from the population. To produce better result, one could try to use several instances from the sample simultaneously (templates). As an alternative, one could find the instance closest to the sample mean (prototype), hoping that it would be sufficiently close to the population mean. Another alternative is to try to construct the image (or atlas) that reflects the population mean. The latter approach actually constructs model from several templates. One might expect that by introducing more information in the atlas formation process would lead to better atlas, so that algorithms for prototype selections are of lesser importance, however, it turns out that the easy way to build a model is to root the model on the prototype. Moreover (as we will see later on), the assumption that more information lead to better atlas will turn out wrong.

So, the easy way to build a model is to pick one individual from a sample (a root image on which the model image is going to be built), and transform other instance onto that target, to assure that all instances have the same spatial frame for further processing [253, 155]. The prototype is most suitable as the root image, since model is always biased toward root image, which is especially visible if the root image is picked far from the population mean, as noticed by several authors (see [175, 16, 17]). Some authors proposed how to select the least biased root image, while others tried to find the way how to estimate (or converge to) the true mean of the population. Therefore, prototype selection is referred by many groups that actually sought another way to construct the least biased model. For example, Marsland et al. in [138] proposed a method for least biased selection of root image, using iterative algorithm that minimizes the distance and maximizes the mutual information. Park et al. in [155] proposed an alternative in which the selection of the root image is based solely on distance. Furthermore, Park argues that the least biased atlas should be done in this way because it is less affected by inherent noise in the images, and, since it uses estimation technique instead of iterative algorithm, is faster. Bathia et al. in [13] proposed the approach where one arbitrary image is used just as an intensity reference (but not shape), after which the similarity between images is maximized using non-rigid transformation. To assure that the image calculated in this fashion is actually the mean (with respect to transformation) they put the constrain that the sum of all transformation is equal to zero. In [94], Joshi et al. proposed the method which is invariant to target image selection since after the construction of the atlas in the space frame of the target image, the target image is transformed to the space frame of the mean transformation. The mayor improvement of this work is that this was done for large deformations which was not the case with [13]. As the dissimilarity measure the squared error distance was used and it was shown that the optimal atlas (for the selected dissimilarity measure) is an average intensity atlas. A similar work using Kullback-Liebler divergence is described in [127] by Lorenzen et al.. Warfield et al. [235], proposed a group-wise registration hoping that this would lead to average and anatomically valid atlas. Guimond et al. [73] propose an iterative averaging algorithm to reduce the bias, by calculating average image and average deformation in order to find average shape and intensity. In [218] the iterative technique whereby the atlas converges to the unknown population mean is suggested by Toga and Thompson. Both algorithms suggest independent shape and intensity averaging. Zöllei et al. in [253] investigated the performance of four approaches to atlas formation. Two approaches used root image while the other two try to converge towards the population average, where one of the algorithm is congealing algorithm proposed by Zöllei et al. and Learned-Miller in [252] and [119].

The prototype selection is easier, usually computationally less intensive, and gives the atlas with the smallest possible bias but nevertheless it is still biased. The estimation of the root image reduces bias but may give the result which, in some case, falls out from the space of possible samples. This is somewhat similar to calculating the mean value of the bits in a computer. The expected result is around 0.5, but it falls out from the sample of possible events which is {0,1}, since a bit with value 0.5 is an impossible event. In the case of medical atlas this could result with the atlas which is anatomically impossible. However, this can be solved with two separate atlases for each class (in the bit example above, this would be classes 1 and 0). Therefore, the critical underling assumption is that images used for model image construction and the allowed deformations are carefully selected so that they lead to anatomically valid atlas. Moreover, since the estimation of one mean image from the population is essentially a kind of an averaging procedure, the averaging of the image may lead to fuzzy borders and blurred anatomical features. To cope with the lack of detailed anatomical features on the average anatomical image some authors suggested to overlay the average anatomical image with high-resolution single-subject image [224].

The previously mentioned papers discussed various methods for prototype/root image selection and model construction. However, after the paper of Rohlfing et al. [173], multi-atlas based approach became popular. Rohlfing et al. in [173] (and in more details in [174]) compared atlas constructed from prototype image, model image and one or more template images. First, one random individual from the set was selected (a template), second, the average shape atlas was constructed (a model), third, the most similar instance from the set was selected as atlas (a prototype) and fourth, several individual images were used as atlases and multi-classifier approach (as in [90]) was introduced before final segmentation (several templates). In these works Rohlfing modelled only image (not segmentation) and showed that the multi-atlas based approach leads to most accurate segmentation. After this finding, more studies using multi-atlas segmentation emerged [112, 80, 240, 1, 113, 3, 230, 87, 245, 76, 70, 108]. Some of them also showed that multi-atlas segmentation outperforms methods that use a single atlas, like [112, 80], but it was recently noticed that the adding more atlases does not always improve result, on contrary, after some point adding more atlases can deteriorate the result, albeit usually very slowly [1]. This can be explained in a similar way as the feature selection phenomenon known from pattern recognition [89]. This approach to image segmentation became known as atlas fusion, by analogy to classifier fusion techniques. Among these works several variations exist, such as different stopping criterion for selection of atlases that are to be fused to provide the most accurate classification (segmentation) (e.g. [113, 230]), or local instead of global use of atlas (e.g. [240, 230]).

## 3.4 Image Registration

Due to the development of acquisition devices the diversity of applications for image registration have grown significantly. For example, Zitova et al. in [251] states that, according to the database of the Institute of Scientific Information (ISI), within the period 1993–2003 more than 1000 papers were published on the topic of image registration. More recent numbers, from the same database, show that in just last two years (2011 and 2012) more than 4000 papers were published on the topic of image registration. Therefore, we will not aim to cover all the publications within this fields, but rather just survey the methods, with the accent on the publications within field of medical imaging.

A comprehensive survey of image registration methods can be found in [19], and the studies [51, 140, 134, 83, 177] offer a review of the methods in the field of medical imaging, while paper by Toga and Thompson [219] discuss the role of image registration in brain segmentation. More recent surveys on image registration are the work of Mäkelä et al. [136], which covers the field of cardiac image registration, the work of Pluim et al. [161], which covers mutual information based studies in medical imaging and the work of Zitova and Flusser [251] which discuss various use of of image registration in computer vision. Various other disciplines utilized image registration methods such as: computer aided diagnosis, atlas construction, computer vision, remote sensing, cartography etc. ([19, 251, 134, 75]).

Noticing the variety of use of image registration, and having in mind that in just the last two years more than 4000 papers on topic of image registration were published, one can conclude that the definition that aims to cover all these disciplines can only state that image registration is the process of overlaying two or more images to achieve maximum correspondence. If we want to make the definition useful we have to specify the "overlaying", "correspondence" and how to "achieve [its] maximum". The interrelationship between these processes, and how they form the process of image registration is depicted in Figure 3.1.

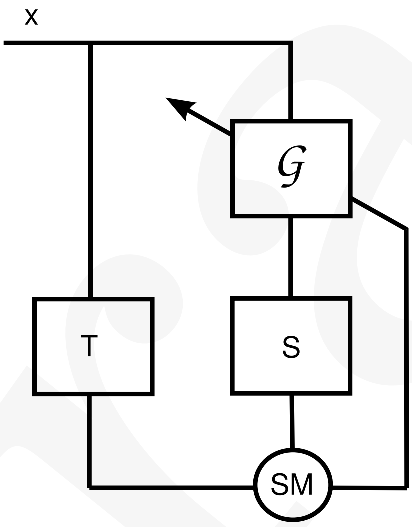


Figure 3.1: The diagram shows the process of registering two images $T$ and $S$. The correspondence between them is defined by similarity measure (denoted $SM$), whose output is used to control the geometric transformation ($\mathcal{G}$). The geometric transformation warps the underlying space of images $S$ (denoted with **x**) so that $S(\mathcal{G}(x))$ match the image $T$.

In Figure 3.1 we can notice the three basic building blocks of image registration:

1. Transformation
2. Similarity measure
3. Optimization algorithm

The transformation is a class of geometric transformation (or parameters of freedom) which explains how we actually overlay one image onto the other (denoted $T$ in the Figure 3.1). The similarity measure is also known as alignment measure, registration function etc., or more general, energy function, cost function or score. It quantifies the similarity (correspondence) between two images. The optimization algorithm drives the registration to find the similarity measure maximum, and in Figure 3.1 is shown as a feedback loop.

The registration algorithms can be classified with respect to expected variation in the scene. The variations can be due to the subject variation, or due to the sensor variation. Thus we distinguish between registration of two images of the same subject (intra-subject registration, e.g. [207, 102]), two images from different subject (inter-subject registration, e.g. [211, 189]), and registration pf an image to a model (model-based registration, e.g. [8]). Alternatively, we can have the change of the sensor type (multi-modal registration, e.g. [236, 132, 45, 126]), or change of the sensor position (multi-view registration, e.g. [27]). If we register two images which also differ by the time of image acquisition we can also speak of multi-temporal registration. The latter is especially important for cases like motion tracking, or segmentation of the growing tumor. Naturally, cross-classes image registration methods exist, so it is natural to have multi-modal intra-subject registration, or mono-modal inter-subject registration.

Alternatively, we can classify image registration with respect to algorithms used for its building blocks. This is partially done in the following sections.

### 3.4.1 Transformation

Geometric transformation used in image registration process can be roughly divided to global and local (as in e.g. [134]). However, global transformations are usually used for initial alignment which is later improved by more complex local transformation (e.g. [179, 159]). The transformations can also be divided with respect to their complexity, thus we distinguish between rigid and nonrigid transformations (formulae for different type of transformations are given in Chapter 2). The nonrigid transformations are usually divided into affine (e.g. [207, 6, 12, 189]) and curved transformations (e.g. [129, 168, 116, 202]), sometimes also referred as free form deformations (FFD, e.g. [179, 63, 45, 146]), and typically implemented [198, 178] using splines [227, 228, 226, 184]. More ideas for implementation of nonrigid geometric transformations can be found in e.g. [84, 50].

### 3.4.2 Image similarity measure

The purpose of image similarity measure (SM) is to quantify similarity between two images usually referred as source and target image. Various information from image data is used for calculating the SM, thus we can classify SM according to the information content used as landmark-based or (pixel property) intensity-based SMs. Landmarks can be extrinsic [192, 60] and intrinsic [165], where intrinsic landmarks can be sub-classified as anatomical [12] or geometrical landmarks [208, 243, 130]. If we are to compare the images based on their landmarks, the selection of Euclidean distance (or SSD) seems natural, with different variations such as incorporating weighting factor (e.g. Mahalanobis disatnce [133, 220]) or leaving some sample out (e.g. RANSAC algorithm [57, 64, 254]). An alternative to this is to use some kind of correlation or an entropy function. The alternative becomes clear if the intensity information is exploited to calculate the similarity between images. Actually, intensity-based SMs are increasingly used in medical image registration [170]. Probably the most prominent correlation and entropy functions are correlation coefficient (CC) and mutual information (MI).

It is well known [201, 231] that the selection of CC, defined as in Equation 2.57 is an optimal choice if the pixel values between images can be related by an affine function even if a reasonable amount of noise is added. As most mono-modal image registration techniques assume that type of a functional relationship between images, CC was considered a dedicated measure for this type of problems, and often the first choice. Similarly, MI defined in Equation 2.60 is established as a valuable SM for a multi-modal image registration [34, 232].

All good properties, as well as the shortcomings, of MI and CC come directly from their definition. For example, CC is fast since it uses only summary statistics. The calculation of only mean and standard deviation, which are sufficient to describe an affine functional relationship between pixel values, makes CC a good and fast measure for registration of images with affine relationship between their pixel values. However, many image registration problems violate the assumption of an affine relationship between image pixel values, thus making CC a suboptimal choice for images with more complex functional relationship between pixel values[171, 201]. On the other hand, the definition of MI includes much more statistical information about the relationship of the images $S(\mathbf{x})$ and $T(\mathbf{x})$. This property comes directly from the joint probability density function $p_{TS} = p(T, S)$, which fully characterized the relationship between images $T$ and $S$, without any assumptions on a type of functional relationship between images, and including any noise that may exist in the image acquisition process. Many of the disadvantages of the MI come from the very same thing - joint probability density function (PDF). As we are dealing with digital images (joint) PDF can be only estimated or approximated, which naturally leads to slower implementation, sensitivity to interpolation [171, 162], sensitivity to number of bins (in case of estimation) or sample selection (in case of approximation). Similarly, MI suffers from sensitivity to the change in image overlap region. These come directly from the inherent property of the joint entropy [82], and some methods to reduce these sensitivities are already proposed [205, 162].

In this perspective we can say that the main difference between CC and MI is whether SM incorporates a (few) summary statistic values, or all the statistical information in the form of a PDF. If we embrace this perspective on SM we can see that many measures proposed earlier in the literature are derivatives of two most prominent SM. Roughly we can classify them as either CC-based (e.g. [171, 12, 168, 5, 58, 62, 102, 104, 157, 244, 81]), or MI-based (e.g. [135, 232, 132, 235, 129, 181, 169, 131, 233, 21, 124, 216]), which are sometimes also referred as (information-theory) IT-based similarity measures (see [44, 246, 25, 123]).

It is worthwhile to mention several other papers which organize various similarity measure, or compare their performance, such as: [26, 185, 195, 196, 172].

### 3.4.3 Optimization algorithm

Optimization algorithm drives the image registration towards similarity measure maximum. Generally we classify optimization methods as constrained [13] and unconstrained, or as deterministic and nondeterministic [46, 91]. Most of the papers listed herein belong to classes of unconstrained and deterministic optimization algorithms. However, for use of evolutionary algorithms in image registration (as a special case of nondeterministic algorithms) we suggest to consult the paper by Damas et al. [46] and its reference. There are many existing algorithms for optimisation and for details we suggest to consult the in-depth and comprehensive book by Fletcher [61] or Numerical recipes by Press et al. [164].

## 3.5 Conclusion

In this chapter we presented the most important ideas and papers from the filed of model-based segmentation and image registration, where main accent was on the field of medical image registration. Hopefully, this overview will be helpful in understanding the ideas that motivated our work, research and contributions described in the following chapters.

We would also like to mention that there are several toolkits freely available on the Internet that can be used for image registration, e.g. [180, 112, 77], where *elastix* (described in [180]) is used in one of our experiments. Additionally, it is interesting to mention the works of Christensen et al., Skrinjar et al. and Lorenzen et al. (see [126, 31, 197, 30]), where image registration process is observed through prism of inverse consistency and transitivity. Although they do not guarantee the accuracy of the registration [126, 31], they are sometimes used as measure of quality of the registration, and one of our experiment (see Chapter 6) suggests that this would be a preferable in order to improve some of the registration (and segmentation) results.

> 尽信书不如无书。
>
> *Asian proverb*

## Chapter 4

# Data Acquisition

## 4.1 Ultrasound imaging and Doppler effect

Ultrasound imaging is a noninvasive medical imaging technique, which utilizes the ultrasonic wave to measure distance and velocity. In order to acquire data, ultrasonic signal is transmitted into a medium and both specular and scatter reflections are on the receiver recorded as a function of time. The time period between transmitting the signal and receiving an echo is known as "time of flight". The time of flight is proportional to the distance between the transducer and the object causing the reflection. If the velocity of the wave in the medium is known, the exact distance can be calculated. Velocity of a moving object can also be calculated from the ultrasound data by exploitation of the Doppler effect. An every-day example of this phenomenon is that of an object (e.g. car or motorcycle) passing by, when the sound emitted by the object is observed differently when the object approaches to the observer than when it moves away from her. The difference between the observed and the transmitted frequency is the Doppler frequency or Doppler shift. It is used to to estimate the relative velocity between the observer and sound source, or between the receiver and the scattering object. The relationship between the transmitted ($f_T$) and the received frequency ($f_R$) is given by the Doppler equation[209, 79]:

$$f_D = f_R - f_T = -\frac{2|\vec{v}|cos\theta}{c} f_T, \tag{4.1}$$

with $\vec{v}$ denoting the velocity vector, $\theta$ the angle between the velocity vector and the sound beam and $c$ the velocity of sound. It is worthwhile to notice that only the axial component of the motion of the object can be measured as any motion perpendicular to the direction of wave does not induce a Doppler shift.

Usually, the received signal is subdivided into sequential time interval. From each segment the spectral information (i.e. amplitude, frequency and phase) can be calculated. If the amplitude is encoded as a gray value and frequency as the amplitude of a time changing signal the resulting image is called a spectrogram (or sonogram). An example of such image can be seen in the middle part of the Figure 4.1. The image shows the distribution of the detected Doppler frequencies as a function of time and represents the differing velocities which occur simultaneously within the sound beam. In practice, because of technical limitations, a compromise has to be made between the velocity and temporal resolution of the spectrogram. In other words, when an accurate velocity estimate is required, temporal resolution of the spectrogram has to be reduced and vice versa. For this reason two different image acquisition systems using Doppler ultrasound have been developed: continuous wave (CW) Doppler and pulsed wave (PW) Doppler systems. As its name states CW-Doppler systems transmit continuous signal and therefore do not contain spatial information. Within the thesis, ultrasound images acquired by the CW-Doppler method are analyzed.

## 4.2 Data acquisition

As described in previous chapter CW-Doppler signal represents the time-change of velocities along a scan line in a 2-D ultrasound imaging plane and is known as the spectrogram. If the signal is generated by the blood outflow from the heart into aorta (acquired by continuous wave Doppler) it is referred as the aortic outflow profile (see Figure 4.1). The x-axis of the aortic outflow profile represents time and y-axis velocities. In the thesis we will mainly work with such signals.

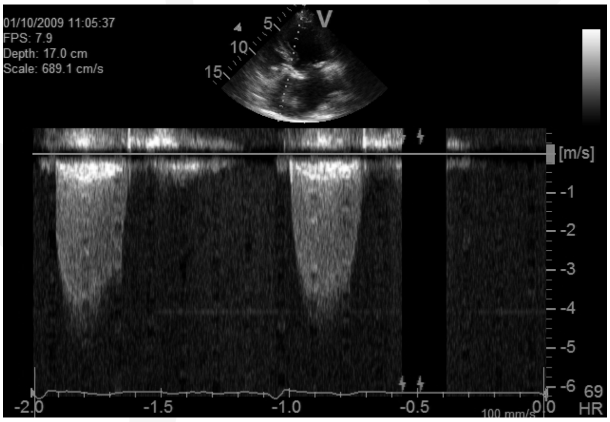


Figure 4.1: CW Doppler outflow velocity profile of the blood flow through a heart valve, as seen by the cardiologist at the workstation.

The Continuous Wave Doppler images of the aortic outflow used herein were acquired with a clinical echocardiographic scanner (Vivid 7, GE Healthcare) using an apical 5-chamber view. Images were digitally stored in 'raw' Dicom format, containing the spectral Doppler information in proprietary tags. These 'raw' Dicom images were converted into Hierarchical Data Format (HDF) using an Echopac workstation (GE Healthcare). From the HDF file, the image containing the aortic velocities was extracted (see Figure 4.2).

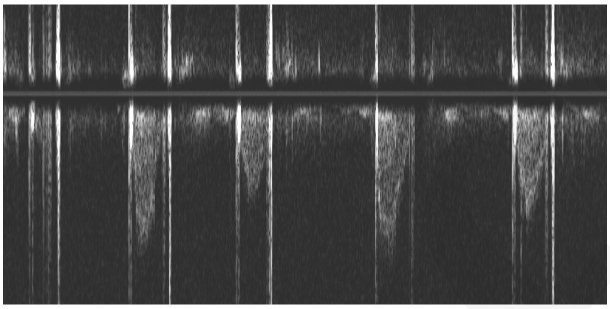


Figure 4.2: CW Doppler outflow velocity profile from the HDF file. For better visibility, color and brightness enhancement is done.

The traces were acquired in several hospitals and constitute of patients with diagnosed coronary artery disease (CAD) and aortic stenosis, and healthy volunteers. The total number of images used in each experiment varied around 140 images, and the exact numbers are given for each experiment separately.

## 4.3 Image preprocessing

After conversion of the image from the Dicom data format to HDF the aortic outflow profile needs to be extracted. This was done in semi-automatic way. First lower and upper signal boundary is detected, and secondly the ejection period was detected by an expert cardiologist. Finally, the speckle noise is reduced using median filter. The outline of preprocessing is shown below:

1. determine zero-velocity line (lower signal boundary)
2. determine maximal-velocity line (upper signal boundary)
3. extract ejection period
4. apply median filter for image noise reduction

### 4.3.1 Lower and upper signal boundary

The low velocity rejection region (see Figure 5.1) depends on the machine settings (clutter filter), and can significantly vary in size (see Figure 1.1(b)) depending on the user's settings. This velocity rejection region may present a problem for a region based approach to image segmentation since it may arbitrarily vary around zero and occlude the low velocities of the object of interest.

To detect the lower and upper signal boundary, the image is projected onto the y-axis. Afterwards, the image projection is smoothed, divided by 2, and plotted together with the image projection, as shown in Figure 4.3. Now, the low velocity region is easily recognized as the part of the signal for which the smoothed projection is higher than the original signal. The zero-velocity line is defined as the medial part of this region. The maximal-velocity line is detected as the rise of the smoothed projection by 10% from the end of the signal.

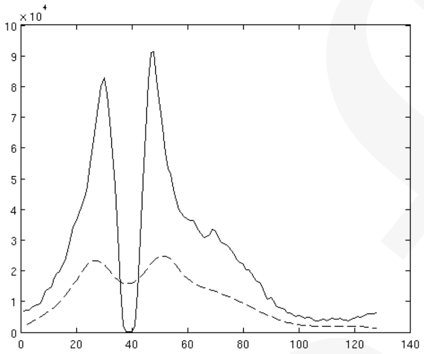


Figure 4.3: The projection of the image onto the *y*-axis (solid line) and smoothed signal divided by 2 (dotted line). *x*-axis is in pixels, *y*-axis shows the sum of intensities.

The smoothing of the projection of the image onto the y-axis is done using the convolution between the original signal and the Gaussian filter with $\sigma = 10px$.

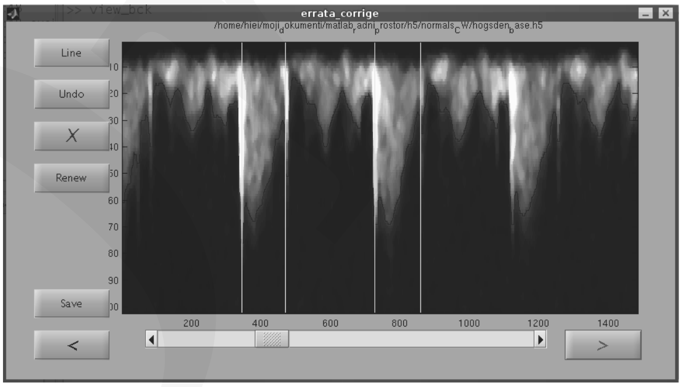


Figure 4.4: Illustration of graphical user interface used for indicating opening an closuring of the valve.

### 4.3.2 Ejection period extraction

To time the cardiac cycle, a single ejection period (from opening until closure of the valve) was manually indicated by an expert cardiologist. To ease this work, a software application was implemented with a graphical user interface as illustrated in Figure 4.4. From the individuals included in the study one or two heart beats were acquired, and the resulting outflow profiles are used as data for further processing.

### 4.3.3 Noise reduction

Since we primarily expect speckle noise in the images, two median filters were introduced. This is done just after the zero-velocity line detection, to avoid the low velocity region blurring in cases when this region is rather thin. The mask of the first median filter is 3-by-3, while the second mask dependents on the size of the original image. In the case the image size is M-by-N, the mask will have the size (M-by-N)/100. The result is rounded to integer.

γνῶϑι σεαυτόν

*The inscription in the pronaos of the Temple of Apollo at Delphi*

## Chapter 5

# Atlas-based Segmentation of Cardiac Outflow Velocity Profiles

## 5.1 Introduction

The chapter presents a method for segmentation of aortic outflow velocity profiles from cardiac Doppler ultrasound images. The core of this chapter is the research from one of our papers (namely [100]). Alternative approaches were also investigated by our team with varying success and some of them are described in papers [97, 101, 99]. In [101] slightly different geometric transformation and genetic algorithm were utilized in the image registration process. In [99] we investigated a less computationally intensive approach based on boundary modeling in the form of a (truncated) harmonic decomposition [163, 29, 186] and coupled with random sample consensus [57].

The method proposed herein is based on the statistical image atlas derived from ultrasound images of healthy volunteers. The chapter describes atlas formation, and atlas based segmentation of aortic blood velocity profiles followed by the experimental results. The ultrasound image segmentation is done by registration of the input image to the atlas, followed by a propagation of the segmentation result from the atlas onto the input image. In the registration process, the normalized mutual information is used as an image similarity measure, while optimization is preformed using a slightly modified multiresolution gradient ascent method. The geometric transformation used is specially formulated for outflow profiles. Since the inherent problem of the segmentation validation is the difficulty of obtaining the reliable reference, several validation techniques are used. The experiments can broadly be divided in two major groups of registration and segmentation validation.

The registration is evaluated using an in-silico phantom image, 30 deformed images from healthy volunteers, and inverse consistency test as proposed by Christensen et al. [31] and Lorenzen et al. [126].

The segmentation is evaluated on 59 images from healthy volunteers and 89 images from patients, using manual segmentation by an expert cardiologist. In this way the segmentation is tested on the set of images which are anatomically far from the atlas, since only the healthy volunteers were used for atlas construction. To check the usability of the proposed segmentation in clinical practice, several cardiac parameters with diagnostic potential are extracted from atlas-based segmentation and ground-truth segmentation. When compared to intra-observer variability these parameters also show the segmentation accuracy.

Experimental validation conducted shows excellent results. Cardiac parameter segmentation evaluation showed that the variability of the automated segmentation relative to the manual is comparable to the intra-observer variability. Therefore, the proposed method is useful for computed aided diagnosis and extraction of cardiac parameters.

## 5.2 Method

This section presents the proposed method for atlas-based segmentation. In atlas-based segmentation, the input image is registered to the pre-segmented atlas image. The registration result returns the parameters of geometric mapping from the input image onto the atlas image. With the inverse geometrical mapping the segmentation from the atlas is propagated on to the image.

An example of the outflow velocity profile from a healthy volunteer is given in Figure 5.1.

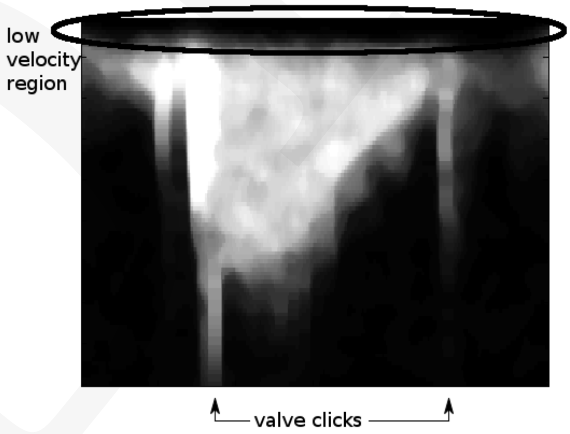


Figure 5.1: The outflow velocity profile from a healthy volunteer. The low velocity region is marked with the black ellipse, and the valve clicks which define the relevant part of the phase cycle are marked with arrows at the bottom of the figure.

In the following sections a new registration algorithm, composed of a geometric transformation, similarity measure and optimization algorithm is described. Next, the method for creation of an atlas image is proposed. The atlas image is manually segmented by an expert cardiologist and the result is mapped to the source image to provide the segmentation result. At the end of Section the atlas-based segmentation procedure is described.

### 5.2.1 Registration

The registration method consists of transformation and optimization with respect to the defined similarity measure. After detection of region of interest by detecting the ejection time and low and high velocity region as described in Chapter 4, the images are initially aligned. The initial alignment matched phases of all outflow velocity profiles and resized all images to have the same resolution.

The rest of the registration procedure stretches the image along the velocity axis, in several bands, and is described in more details below. NMI is used as a similarity measure. The similarity measure is maximized using a modification of the gradient ascent optimization algorithm. This section is divided into three subsections dedicated to the major parts of the registration procedure: transformation, similarity measure and optimization algorithm.

#### Transformation

After the relevant phase of the cardiac cycle is extracted as described in Chapter 4 and all images are aligned, by the geometric transformation $\mathcal{G}$ given by:

$$\mathcal{G}(t, v) = \begin{bmatrix} e(t, v) & 0 \\ 0 & f(t, v) \end{bmatrix} \cdot \begin{bmatrix} t \\ v \end{bmatrix} \tag{5.1}$$

where $e(t, v)$ and $f(t, v)$ are arbitrarily function of time and the velocity. Since the initial transformation resized all images to have the same resolution no transformation in $t$ dimension is required, therefore we can now set $e(t, v) = 1$. All the possible inter-individual changes in the profiles can now be governed only by the variable $f(t, v)$ from the Equation 5.1, which we call the stretching function. It is important to notice that the stretching function is a function of time, i.e. $f(t, v) = f(t)$, so the stretching function itself can be used to quantify the instantaneous blood velocity change for different outflow profiles.

For practical reasons, a parametrized stretching function is used. The function is parametrized by selecting N equidistant points, which are sorted in a row vector. The vector is denoted as $\mathbf{f}$ and will be addressed as the transformation vector. This can be written as follows:

$$\begin{aligned} t_i &= \frac{(i-1)\cdot P}{N-1}, \forall i = 1, .., N \\ f_i &= f(t_i) \\ \mathbf{f} &= [f_1 ... f_N] \end{aligned} \tag{5.2}$$

where $P$ stands for phase cycle of outflow velocity profile. Now, the transformation of an image is described and quantified with the transformation vector components. The reconstruction of a stretching function from the transformation vector is done using linear interpolation. If one selects $N = 11$, as we did in this study, the image transformation and the transformation vector components can be visualized as depicted in the Figure 5.2, where white circles represents the transformation vector components, and the curve interpolated between them represents the interpolated stretching function ($f(t, v)$).

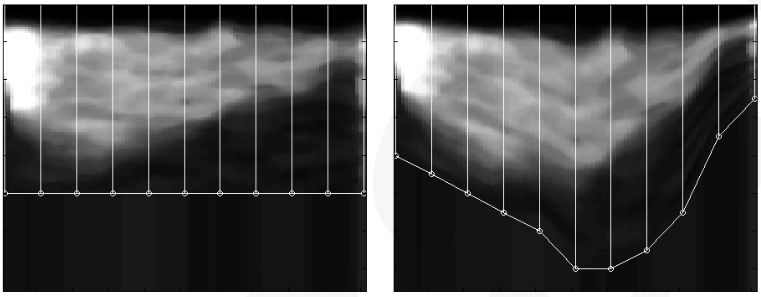


Figure 5.2: The original image (left) and the transformed image with transformation vector components (right).

It is also important to notice that since the transformation function is parametrized and has N degrees of freedom, the optimization space is N-dimensional. The details of the optimization algorithm are described further in the section 5.2.1. In the next section, we will first discuss the similarity measure.

### Similarity measure

As a similarity measure, normalized mutual information (NMI) is used. However, it is important to notice that the similarity measure is not calculated for the set of pixels in the overlapping region of $T(\mathbf{x})$ and $S'(\mathbf{x}) = S(\mathcal{G}(\mathbf{x}))$, i.e. within $C = T \cap S'$, as assumed in [59]. Instead, it is calculated for all pixels in the target image except for the low velocity region (see Figure 5.1). The region over which the NMI is calculated may be written as

$$D = R \setminus (\{S'(\mathbf{i})\} \cup \{T(\mathbf{i})\}), \forall \mathbf{i} \in L \tag{5.3}$$

where $L$ is the set of pixels from the low velocity region (both in image $S$ and $T$). Low velocity region is decided after projection of the image onto the y-axis as shown in Figure 5.3, as set of pixels having the projection lower than 10% of the projection maximum.

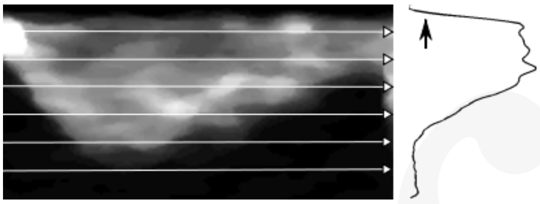


Figure 5.3: Region L is decided after projection of the image onto the y-axis. Black arrow indicates end of low velocity region.

The reason not to calculate NMI over the region $C$ is because $C$ is a function of $\mathcal{G}$, i.e. $C = C(\mathcal{G})$, so to avoid influence of the transformation function on the similarity measure. The problem of non-existent values for the source image is solved as suggested by Roche et al. in [170]. In short, these values are artificially generated during the transformation, using the pixels from the image border.

### Optimization Algorithm

The gradient ascent numerical optimization method is used to find the global maximum of the energy function. The pseudocode for this algorithm is given below, where $E$ stands for the energy function. The energy function $E$ is calculated as the Normalized mutual information between two images $S(\mathcal{G}(\mathbf{x}))$ and $T(\mathbf{x})$ (see Equation 2.61), over the region $D$, as defined in Equation 5.3 and described in section 5.2.1. Same as above, $f$ and $N$ denote the deformation vector and its dimension.

```
Function gradient_ascent($starting\_point, E$)
define:  $\mu, \gamma, \delta, tolerance$
$f = starting\_point$
do
  for $i = 1$ to $N$
    sample E around $f_i$ with $\mu$
    approximate $dE/df_i$ from sampled points
    if $dE/df_i > \gamma_i \cdot \delta$
      $\gamma_i = 0.95 \cdot \gamma_i$
    end if
  end for
  $f = f + 3 \cdot \mu \cdot \gamma \cdot (dE/df)/norm(dE/df)$
while $norm(dE/df) > tolerance$
return $f$
```

In this algorithm, $\delta$ is an estimation of the optimization function gradient at the starting point (according to Fletcher [61]). The gradient has to be smaller for every next step to assure that the algorithm converges. This is done using $\gamma$, which modifies the convergence rate, forcing the change of $f$ to be smaller for every next step.

Since Doppler ultrasound images contain a lot of noise, registration of these images is very sensitive to the initial conditions and the convergence step, and may easily end up in a local (instead of global) optimum. To assure the accuracy and robustness of the proposed method, a two-step multiresolution optimization approach is used. This approach is described in the pseudocode below.

```
Define $S(\overrightarrow{x}), T(\overrightarrow{x}), starting\_point$
Resize $S(\overrightarrow{x})$ and $T(\overrightarrow{x})$ to 100x100px
Define $filt =$ gaussian filter with $\sigma = 9px$
$S_b =$ convolution($S(\overrightarrow{x}), filt$)
$T_b =$ convolution($T(\overrightarrow{x}), filt$)
Define $E(\overrightarrow{f}) = NMI(S_b(\mathcal{G}((\overrightarrow{x})), T_b(\overrightarrow{x}))$
$\overrightarrow{f_1}$ = gradient_ascent($E(\overrightarrow{f}), starting\_point$)
Define $E(\overrightarrow{f}) = NMI(S(\mathcal{G}(\overrightarrow{x})), T(\overrightarrow{x}))$
$\overrightarrow{f_2}$ = gradient_ascent($E(\overrightarrow{f}), \overrightarrow{f_1}$)
```

With this implementation of the multiresolution approach a trade-off between speed and accuracy is made, since the images are not downsampled. The downsampling is avoided since it causes histogram changes, which in turn may cause some of the artefacts similar to the ones mentioned in [162].

### 5.2.2 Atlas Construction

The purpose of a statistical atlas is to combine many images into a single image, which represents a statistical average of all images. In this method, we have used the arithmetic image averaging operation to construct the atlas. After all aortic outflow velocity images are aligned and resized, the atlas is constructed as an average intensity atlas using the formula:

$$A(t, v) = \frac{1}{K} \sum_{i=1}^{K} S_i(t, v) \tag{5.4}$$

where $S_i$ are the images used to construct the atlas. Using this approach the atlas image from $K = 59$ images from 30 healthy volunteers is constructed for the purpose of this study. The resulting image is shown in Figure 5.5 (left).

### 5.2.3 Atlas-based Segmentation

The idea of atlas-based segmentation is based on the use of a representative reference (or atlas) image, where the desired structure is manually segmented by an expert cardiologist. In our case, the desired structure is the aortic outflow velocity profile. Expert segmentation is done only once, and is later automatically propagated to the other images of this type. When a new patient image is acquired, the segmentation is conducted in four steps:

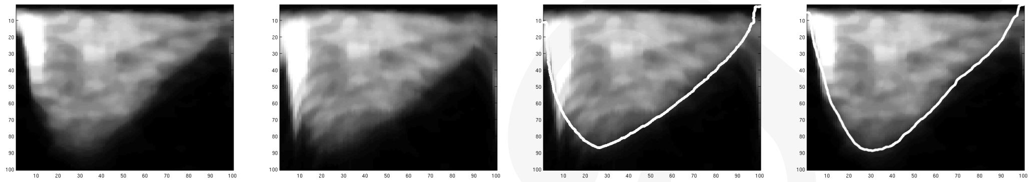
(a) The new image that needs to be segmented is declared as a source image.

(b) The source image is registered to the reference image, resulting in a set of parameters describing the geometric transformation.

(c) The segmentation of an aortic outflow profile from the reference image (the atlas) is propagated to the source image.

(d) The source image along with the propagated segmentation is backward transformed (using the inverse set of parameters) to its original form.

This procedure is depicted in Figure 5.4 where each step is represented with one image. As the reference image, the manually segmented atlas is used.



(a) Source image (b) $S(x) \rightarrow R(x)$ (c) Propagated seg. (d) Bck. transform

Figure 5.4: The segmentation procedure: Each image represents one step described in Section 5.2.3. All axis are in pixels.

## 5.3 Experiments and Results

This chapter is focused on evaluating the registration algorithm accuracy and the comparison of atlas-based segmentation with the segmentation done by an expert cardiologist. First, the registration validation using an in-silico phantom, along with the phantom construction, is presented. Second, the validation on real data is presented, where the exact geometric transformation between data sets is known. Third, the validation of the registration accuracy on a set of test data based on inverse consistency is presented. Fourth, the atlas-based segmentation is validated on 59 images from healthy volunteers manually segmented by an expert cardiologist. Fifth, the same segmentation validation is done on 89 images from patients with a diagnosis of either coronary artery disease or aortic stenosis. Sixth, the segmentation is validated by comparison of the cardiac parameters extracted from the manual and automated segmentation. Finally, intra-observer variation is studied and compared with the error between manual and automated segmentation.

### 5.3.1 Phantom study-based registration validation

The outflow velocity is modeled using a linear combination of sinusoidal functions. The attenuation is modeled by an inverse tangent function:

$$P(t,v) = c_1 - c_2 \cdot arctg(v - F(t) + c_3)) \tag{5.5}$$

where $c_1$, $c_2$ and $c_3$ are constants used for centering the image on the coordinate system, and $F(x)$ is constructed as:

$$F(t) = sin(\pi t) + \frac{sin(2\pi t)}{4} + \frac{sin(3\pi t)}{6} \tag{5.6}$$

The attenuation of low blood velocities is modeled similar to eq. 5.5 using the function:

$$A(v) = \frac{1}{\pi} \cdot arctg((v - a_1) \cdot c_4) + c_5 \tag{5.7}$$

The parameter $a_1$ can be used to set the percentage of the outflow velocities that will be attenuated, in our case it is set to 10%. The resulting image is shown in Figure 5.5.

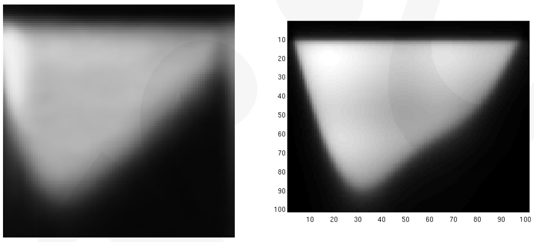


Figure 5.5: The average intensity atlas (left) and the phantom image (right)

To validate the registration accuracy on the phantom image, the source and target image have to be defined. Part of the registration error is due to the suboptimal performance of the optimization algorithm and the properties of the similarity measure. This error can be quantified if the desired transformation is known. For this reason the following experiments are constructed. An image with different velocity outflow profile shape is constructed from the phantom image using a random transformation. Random transformation is here defined by a transformation vector, whose eleven elements are picked from the uniform distribution on the interval [0.7,1.4]. Using the random transformation, 50 variations of the phantom image are constructed.

In the first experiment the original phantom is selected as a source image and the deformed phantoms as target images. Here, the goal of the registration algorithm is to reconstruct the deformation function. The registration error can now be calculated as difference between deformation and the transformation as found by optimization algorithm. Since the transformation is parametrized by the transformation vector this reduces to:

$$\mathbf{e}_1 = \|\mathbf{f}_1 - \mathbf{f}_2\|_2 \tag{5.8}$$

where $\mathbf{f}_1$ and $\mathbf{f}_2$ are row vectors which represent respectively random deformation of the phantom image and the deformation approximation as found by the registration algorithm. The average error vector components from 50 different phantom transformations equals to $\mu(\mathbf{e}_1) = 1.68\%$ with standard deviation of $\sigma(\mathbf{e}_1) = 0.92\%$.

In the second experiment, the original phantom is labeled as target image and the deformed phantoms as source images. Now the registration algorithm has to find the inverse transformation function. For each pair of images the registration error is calculated using the Equation:

$$\mathbf{e}_2 = \|\mathbf{1} - \mathbf{f}_1 \cdot \mathbf{f}_3{}^T\|_2 \tag{5.9}$$

where $\mathbf{f}_1$ and $\mathbf{f}_3$ are row vectors, which represent the random deformation of the phantom image and the inverse, as found by the registration algorithm. Same as before, the mean error vector components and their standard deviation are calculated. Mean error is $\mu(\mathbf{e}_2) = 2.15\%$ and standard deviation $\sigma(\mathbf{e}_2) = 1.92\%$. One may notice that the error and deviation is smaller in the first experiment, which is due to the direction of the registration algorithm. In the first experiment the registration algorithm searches for the transformation parameters in the same direction that is used for the deformation, while in the second experiment the opposite direction is used (i.e. in this experiment the deformation model is not the same as the transformation function).

### 5.3.2 Real image-based registration validation

The similar experiment, as the one explained above on the in-silico phantom image, is conducted on real images of cardiac aortic outflow velocities. This was done since the phantom image used in previous section does not have any (speckle) noise, does not model valve clicks and small deviations of the time frame which are possible to show up in the real images. In this experiment, from a set of 59 images, each was deformed with thirty random transformation vectors and the registration algorithm searched for vectors that will re-transform these images back to their original form. The vector elements used for the deformation are randomly picked from the uniform distribution on the interval $[0.7, 1.4]$, and the starting vector for the optimization algorithm is the unity vector. The error vector is calculated using the equivalent formula as in Equation 5.9 and denoted as $\mathbf{e}_3$. In this experiment the average error is $\mu(\mathbf{e}_3) = 2.93\%$ with standard deviation of $\sigma(\mathbf{e}_3) = 2.03\%$.

### 5.3.3 Inverse consistency-based registration validation

The final registration experiment is based on the inverse consistency test. Although inverse consistency does not guarantee the accuracy of the registration, it is often preferable or even used as measure of quality of the registration [31, 126]. This, along with the desire to quantify the bi-directional transformation error, are the main reasons for the additional validation using the inverse consistency test [30, 92]. Each of the images from the set is registered to all the others images from the set. In this way, the registration is done bi-directionally (i.e. the image $I_1$ s registered to image $I_2$ and vice-versa). Using the notation $.*$ for Hadamard product (where only the corresponding vector elements are multiplied) and $\mathbf{f}(I_2, I_1)$ for the transformation vector received after registration of image $I_1$ to image $I_2$, the mean error vector is calculated as:

$$\mathbf{e}_4 = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{1} - (\mathbf{f}(i)(I_2, I_1)\ .*\ \mathbf{f}(i)(I_1, I_2))| \tag{5.10}$$

where $N$ represents the total number of registration experiments. If the number of images is $n$ then the total number of registration experiments equals to $N = \frac{n^2+n}{2}$, since registration of an image onto itself is also taken into account. The average vector component error is equal to 2.89%.

### 5.3.4 Atlas-based segmentation validation: Healthy volunteers

The atlas-based segmentation validation is done on 59 outflow profiles from 30 healthy volunteers. The expert manually segmented the atlas image (constructed as described in section 5.2.2), and this atlas is used as a template for the segmentation.

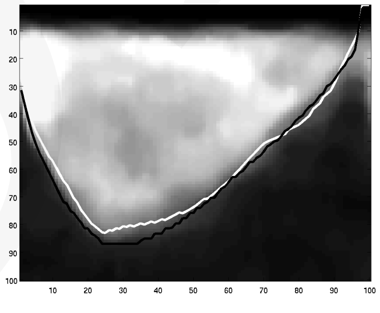


Figure 5.6: The comparison of manual (black) and propagated (white) segmentation. Both axes are in pixels.

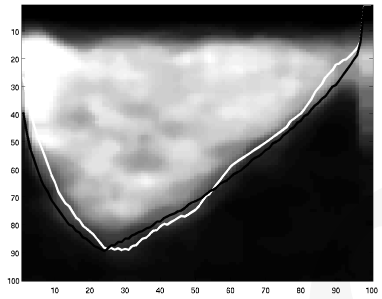


Figure 5.7: Propagated segmentation with small deviation of automated segmentation (white) from the manual segmentation (black). Both axes are in pixels.

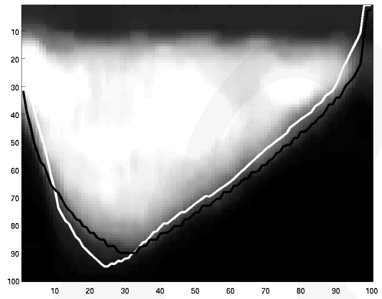


Figure 5.8: Propagated segmentation with larger deviation of automated segmentation (white) from the manual segmentation (black). Both axes are in pixels.

With the procedure as described in section 5.2.3 the manual segmentation is propagated from the atlas image to the rest of the 59 images. These images are compared with the ones that are segmented manually by the same cardiologist. For the brevity of the presentation, only some of the results, that are representative of all results, are presented. These images are shown in Figures 5.6, 5.7, 5.8 and 5.9. In Figure 5.6, we may see the manual and the automated segmentation results that correspond very well. In Figure 5.7 we want to point out the small bumps that exist in the automated segmentation, while there is no trace of them in the manual segmentation. Although, the automated segmentation corresponds well to the manual segmentation, the bumps may be explained as an inherent intensity change. If we take a look at the Figure 5.8 we may notice that, around the peak, the automated segmentation peaks over the manual segmentation. Nevertheless, this segmentation result explains well the shape of the signal despite the selection of a different

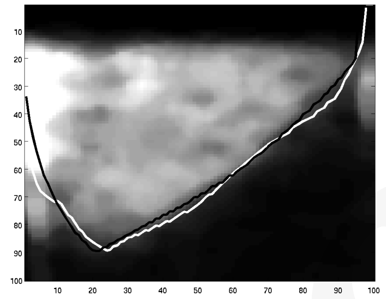


Figure 5.9: The comparison of manual (black) and propagated (white) segmentation on the outflow profiles with the starting valve click. Both axes are in pixels.

threshold. In the Figure 5.9 the outflow profiles with the starting valve click is shown. In clinical practice, the cardiologists try to distinguish between the blood flow and the valve click based on their experience, since only the blood flow bears significant information for diagnosis. It can be seen how the manual segmentation performs across different intensities, as if there is no valve click. When this is compared to the automated segmentation there is a difference, but automated segmentation also managed to ignore the valve click. This last results (Figure 5.9) demonstrate also the important improvement compared to the work of Tschirren et al. [221] since these results cannot be reproduced using just envelope detection. When the numerical results of the manual and automated segmentations are compared, this knowledge from the visual inspection should also be taken into account, since it is disputable whether some of these errors are errors indeed. If manual and propagated segmentations are observed as sampled function and denoted as $m_i[t]$ and $p_i[t]$, respectively, where $i$ stands for the instance of the Doppler outflow image, the error may be measured as average difference between $m_i[t]$ and $p_i[t]$ and written as:

$$d_e = \frac{1}{K \cdot M} \sum_{i=1}^{K} \sum_{t=1}^{M} |m_i[t] - p_i[t]| \tag{5.11}$$

where $K$ stands for the number of images, i.e. $K = 59$, and $M$ for the number of samples in the time (phase) frame, i.e. $M = 100$. Using this measure we may say that the propagated segmentation deviates in average by 4.6 pixels from the manual. Since all the images have been resized to 100-by-100 pixels, images have 100 samples in the velocity direction and so do the functions $m_i[t]$ and $p_i[t]$. Since the transformation is done along th $y$-axis this error corresponds to 4.6%.

The sample correlation coefficient between manual and propagated segmentation of all outflow profiles is also calculated. This is done using the

Equation:

$$r = \frac{1}{M-1} \sum_{t=1}^{M} \frac{m[t] - \mu_m}{\sigma_m} \cdot \frac{p[t] - \mu_p}{\sigma_p} \tag{5.12}$$

Here $r$ denotes the sample correlation coefficient for one instance of the outflow profile, and $M$, $m$ and $p$ are used as defined above. The average sample correlation coefficient of the population is $r = 0.98$ with the population standard deviation $\sigma_r = 0.024$. The minimal and maximal sample correlation coefficient between manual and propagated segmentation are $r_{min} = 0.86$ and $r_{max} = 0.99$, respectively, which shows excellent statistical correlation between manual results and the proposed method for atlas-based segmentation.

### 5.3.5 Atlas-based segmentation validation: Patients

In the previous subsections, the validation is done on the aortic outflow profile that is either artificially created or belongs to the data set that is used to create the atlas. To validate the segmentation procedure on the outflow profiles from different data sets 89 outflow profiles are selected. 36 of these outflow profiles belong to patients with coronary artery disease (CAD) and 53 of them belong to patients with aortic stenosis (AS). Again, the manual segmentation is propagated from the atlas to all the instances of the patients outflow profiles as described in section 5.2.3. These images are compared with the ones that are segmented manually by the same cardiologist. In Figure 5.10, representative images of the patients with diagnosed CAD and AS are presented, with both manual and automated segmentation of the outflow profile.

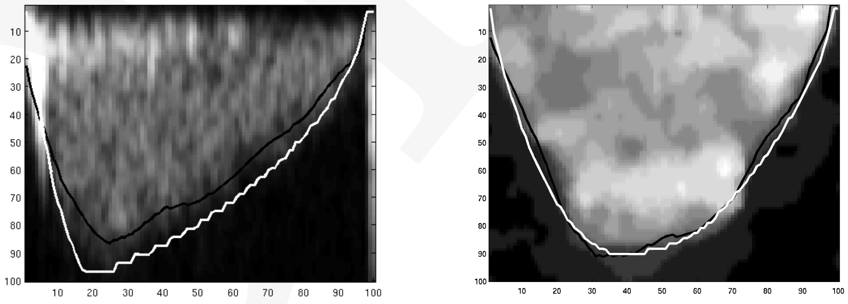


Figure 5.10: The comparison of manual (white) and propagated (black) segmentation for patients with diagnosed CAD (left) and AS (right)

If the same measurements as for normal patients are used (see Subsection 5.3.4), we can see that the average automated segmentation error with respect to the manual segmentation is $d_e = 5.08\%$ for the patients with diagnosed aortic stenosis, and $d_e = 8.70\%$ for the patients with diagnosed coronary artery disease. At the same time, the correlation coefficient between manual and automated segmentation is $r = 0.98$ both for the patients with AS and CAD. The maximum sample correlation coefficient is $r_{max} = 0.99$ for both set of patients, while the minimum sample correlation coefficient is $r_{min} = 0.96$, for patients with CAD, and $r_{min} = 0.92$, for patients with AS.

### 5.3.6 Cardiac parameter-based segmentation validation

In this subsection, we describe a segmentation validation procedure based on the comparison of the cardiac parameters extracted from two aortic outflow profiles. The first aortic outflow profile is obtained by the proposed automatic segmentation method, while the second aortic outflow profile is obtained by manual segmentation. Cardiac parameters that are measured are: time to peak, peak value and rise-fall time ratio. These parameters have shown to have potential for use in diagnosis of some of the cardiac disease (see [97] or [32]), however, they are not routinely used in clinical practice since their extraction is often subjective, being both dependent on computer display (brightness and resolution) as well user interpretation, as will be shown in the next section.

Let $ttp^m$ and $ttp^a$ denote time-to-peak values extracted by manual and automated procedures, respectively. Similarly, let the same notation be used for the maximum value and rise-to-fall-time-ratio parameters ($max^m$, $max^a$, $trf^m$, and $trf^a$). Since outflow velocity profiles belong to different patients, different pacing and different velocities are expected. Therefore, to exclude the variation due to different patient characteristics and to observe the segmentation variation only, relative parameter errors are calculated and given as percentages rather than absolute values.

In this experiment, we calculate the relative error between the automated and manual segmentation, which in the case of time-to-peak parameter is expressed as:

$$e_{ttp} = \frac{ttp^a - ttp^m}{ttp^m} \tag{5.13}$$

For comprehensive analysis of method accuracy we calculated three statistical error measures: mean error, standard deviation of error, and mean absolute error. If a systematic error (bias) is present, it will be evident from the mean error and from the mean absolute error. Standard deviation of error does not detect systematic error. If no systematic error is present, then the mean error will be equal to zero and hence is not useful for error evaluation. In this case, both standard deviation of error and mean absolute error can be used for accuracy evaluation.

In Table 5.1 mean error, standard deviation of error, and mean absolute error of the observed cardiac parameters (automated vs. manual) are presented. The results from patients with diagnosed coronary artery disease (CAD), patients with diagnosed aortic stenosis (AS), and volunteers with normal outflow profiles (N) are given in separate columns.

Table 5.1: Mean error, standard deviation of error, and mean absolute error between cardiac parameters obtained from manual and automated segmentation. Rows 1-3 show errors for time-to-peak parameter, rows 4-6 show errors for rise-to-fall-time-ration parameter, rows 7-9 show errors for peak-value parameter.

| | CAD | AS | N |
|---|---|---|---|
| $\mu(e_{ttp})$ | 2.71% | -4.15% | 8.91% |
| $\sigma(e_{ttp})$ | 17.62% | 14.77% | 18.16% |
| $\mu(\|e_{ttp}\|)$ | 12.65% | 11.75% | 13.38% |
| $\mu(e_{trf})$ | 4.94% | -5.74% | 11.88% |
| $\sigma(e_{trf})$ | 27.09% | 20.74% | 24.21% |
| $\mu(\|e_{trf}\|)$ | 18.34% | 16.67% | 18.47% |
| $\mu(e_{max})$ | -9.33% | 2.66% | 2.92% |
| $\sigma(e_{max})$ | 8.14% | 4.82% | 5.91% |
| $\mu(\|e_{max}\|)$ | 10.22% | 4.58% | 5.17% |

It is evident from Table 5.1 that certain amount of systematic error exists. Standard deviation of error and mean absolute error are measures that show the amount of error, other than systematic error. The table shows that standard deviation of error and mean absolute error are highly correlated. Therefore, we can conclude that both measures can be used for evaluation of error.

For the interpretation of the results, one should note that the time to peak falls somewhere around the first quarter of the ejection time frame. For the images presented here, that would be around 25 pixels. If e.g. time-to-peak parameter estimate is inaccurate by one pixel only this will result in 4% error. This can be observed on the Figure 5.10 (left) where the relative errors in terms of the cardiac parameters are: $e_{ttp} = 16.28\%$, $e_{trf} = 21.71\%$ and $e_{max} = 10.74\%$; which are the values that are comparable with the standard deviation of the relative error in Table 5.1.

In addition to error measures, we have calculated the correlation between cardiac parameters extracted from manual and automated segmentation. For example, for the time-to-peak parameter the correlation is defined as:

$$c(ttp^a, ttp^m) = \frac{\sum_{i=1}^{K} ttp^a \cdot ttp^m}{\sqrt{\sum_{i=1}^{K}(ttp^a)^2}\sqrt{\sum_{i=1}^{K}(ttp^m)^2}} \tag{5.14}$$

The results have shown a very high statistical correlation between the cardiac parameters extracted using our method and the cardiac parameters extracted by the expert cardiologist. For example when time-to-peak parameter is measured a correlation of $c(ttp^a, ttp^m) = 0.988$ is achieved, for the rise-fall time ratio the correlation is $c(trf^a, trf^m) = 0.974$, and for time to peak $c(max^a, max^m) = 0.997$.

### 5.3.7 Intra-observer variability

In the previous subsections, the proposed method is compared to an expert manual segmentation. However, it is well known that there can be a considerable intra-observer and inter-observer variability of the results of manual segmentation. The intra-observer error is the error between subsequent results of the segmentation of the same image performed by the same person. The inter-observer error is obtained when several different people segment the same image. Typically, the inter-observer error is larger than the intra-observer error. One must be aware of these errors when a manual segmentation by one or more expert cardiologists is used as a validation reference, as these errors limit the validation accuracy.

To quantify the intra-observer error the following experiment is conducted. An expert cardiologist segmented 21 images that she already segmented one week ago. If the segmentation results are observed as two sets of measurements, this gives a total of 2100 measurements (since images are resized to 100-by-100 pixels) for each set. If the measurements $m_1(i)$ from the first set are interpreted as realizations of the random variable $m_1$ and the measurements $m_2(i)$ from the second set are interpreted as realizations of the random variable $m_2$ then the random variable $d_m$ defined as:

$$d_m = m_1 - m_2 \tag{5.15}$$

describes the difference between the two measurements. Since we do not know which measurement is the reference one (which represents the correct segmentation) we calculate the standard deviation as an estimate of the variance $\sigma^2(d_m)$ of the random variable and the mean absolute error ($\mu(|d_m|)$). Similarly, let $d_a = a - m$ be the random variable representing the difference between the automated ($a$) and manual delineation ($m$). Since we had 59 images from volunteers, and 89 images from patients, this results in a total of 14800 random variables. The realizations of these two random variables are shown in Figure 5.11, with $\sigma^2(d_m) = 28.94$ and $\sigma^2(d_a) = 47.51$.

The mean absolute error between the automatic and the manual segmentation is equal to $\mu(|d_a|) = 5.57px$, while the standard deviation of the difference is $\sigma(d_a) = 6.89px$. The mean absolute error between two different segmentations of the same image made by the same cardiologist is equal to $\mu(|d_m|) = 3.62px$, while the standard deviation of error is $\sigma(d_m) = 5.38px$.

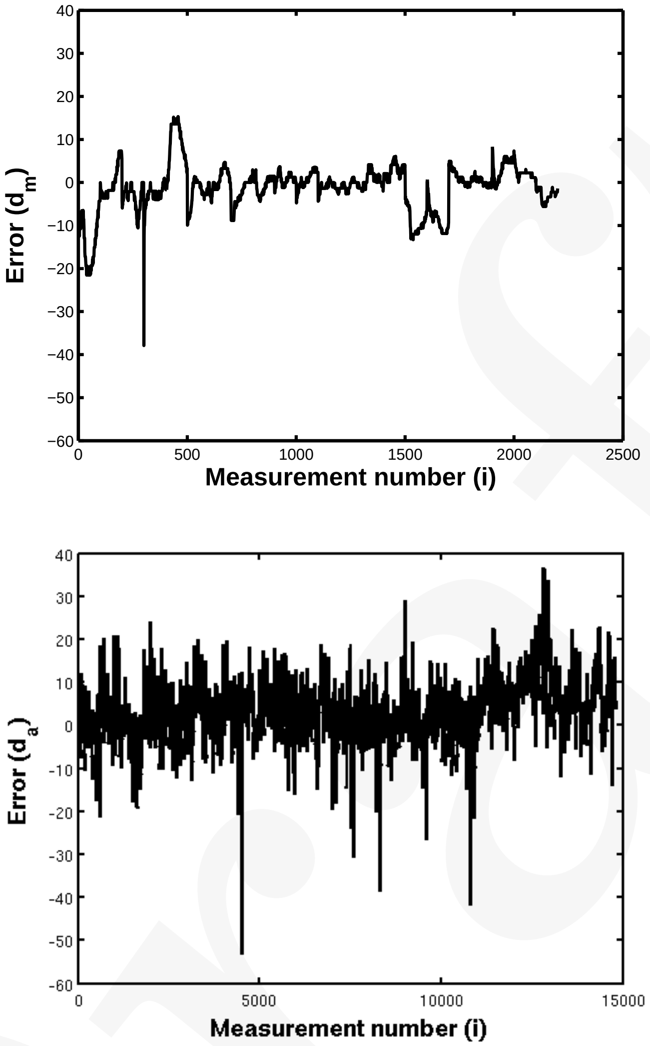


Figure 5.11: The upper graph shows the difference between two manual segmentations (intra-observer variability), while the lower graph shows the difference between manual and automated segmentation.

We conclude that the variability of the difference between the automatic and manual segmentation is comparable to the intra-observer variability (6.89px vs. 5.38px).

The intra-observer variability of the cardiac parameter extraction is also calculated. To compare it with the results from Section 5.3.6, the standard deviation of the relative time-to-peak error is calculated as in Table 5.1 and gives $\sigma(e_{ttp}) = 11.91\%$, while the standard deviation of relative error of rise-fall time ratio gives $\sigma(e_{trf}) = 16.28\%$. When we look at the standard deviation of relative peak value error, we can see that manual segmentation has a variability of $\sigma(e_{max})) = 5.15\%$. If these results are compared with the rest of the results in Table 5.1, we can see that the parameters from automated segmentation varies from the manual segmentation just slightly more than the manual segmentation from itself. The same is true if we observe the mean absolute error since $\mu(|e_{ttp}|) = 9.07\%$, $\mu(|e_{trf}|) = 12.54\%$, and $\mu(|e_{max}|) = 3.43\%$. While having in mind these results and the high correlation between manually and automatically extracted parameters we conclude that one may use the proposed atlas-based segmentation for cardiac parameters extraction.

### 5.3.8 Statistical analysis of manual and automatic parameter measurement

For statistical validation the automated and manual methods for parameter measurement, the t-test is used. Let $e^a_{param}$ and $e^m_{param}$ denote the automatic-to-manual error (error between automatic and manual parameter extraction) and manual error (human intra-observer error). The *param* in the subscript identifies which parameter is tested ($ttp$, $trf$ or $peak$).

The proposed null hypothesis is: The mean values of the errors $e^a_{param}$ and $e^m_{param}$ are equal i.e. the intra-observer parameter error is equal to the error between the automated and manual parameter extraction. The t-test allows a comparison of two datasets with different numbers of samples. In this experiment the first dataset has 21 and the second dataset has 148 elements (Section 5.3.7). The t-test is performed using Satterthwaite's approximation to calculate the number of degrees of freedom and without assumption of the same variability of both datasets (Behrens-Fisher problem). The p-values calculated from the t-test are given in Table 5.2.

Table 5.2: The p-values for tome-to-peak, rise-fall-time-ratio and peak cardiac parameter.

| | ttp | trf | peak |
|---|---|---|---|
| $p - value$ | 0.6843 | 0.7398 | 0.3908 |

The p-values for all three cardiac parameter errors (time-to-peak, rise-fall-time-ratio and peak value) are much above the traditionally used significance level ($\alpha$) of 0.05. One rejects the null hypothesis if the p-value is smaller than or equal to alpha. Since $\alpha = 5\%$ is much lower than the lowest p-value we may conclude that there is no statistically significant difference (at the 5% level) between the datasets or that there is no enough evidence to reject the null hypothesis that the intra-observer parameter error is equal to the error between the automated and manual parameter extraction. As we can see, this is true for all the cardiac parameters evaluated.

## 5.4 Discussion and Conclusion

A comprehensive validation of the registration method is conducted using an in-silico phantom (Section 5.3.1), 59 outflow profile ultrasound images from 30 healthy volunteers (Section 5.3.2), and the inverse consistency test (Section 5.3.3). The exhaustive validation of the atlas-based segmentation is done with respect to an expert manual segmentation. First, the 59 outflow profiles form the healthy volunteers are segmented using the atlas described in Section 5.2.2 and segmentation described in Section 5.2.3. The validation is described in Section 5.3.4. Second, 89 outflow profiles from the patients are segmented using the same atlas and the same segmentation procedure and validated in Section 5.3.5. In both experiments the difference and correlation between manual and propagated segmentation is calculated. Third, the segmentation is evaluated based on the cardiac parameters extracted from the automated segmentation (Section 5.3.6). Finally, the results are compared to the intra-observer variability of the manual segmentation (Section 5.3.7 and Section 5.3.8).

The phantom validation demonstrated that the registration is quite accurate, with an error of the transformation vector around 2% (see Section 5.3.1), at the same time the validation on real images gives an error of the transformation vector of around 3% (see Section 5.3.2). A portion of the errors is due to the asymmetry of the forward and backward transformation as explained in Section 5.3.3.

When the results of the automatic segmentation of healthy volunteers are compared to the manual segmentation by an expert cardiologist, the difference, as an error measure of the automated segmentation, is 4.6%, on average. The correlation between the manual and automatic segmentation is on average $r = 0.98$. Thus, we may conclude that the proposed method for the image registration may be used for the automatic segmentation of Doppler ultrasound images. Additionally, due to the intrinsic properties of the method, the method handles the valve click correctly and therefore is especially valuable in the automatic segmentation of the aortic outflow profiles.

The segmentation validation on the patients showed that the automatic segmentation with respect to the manual segmentation differs by 5.08% for the patients with the diagnosed aortic stenosis, and 8.70% for the patients with the diagnosed coronary artery disease. For both set of patients the correlation of automated and the manual segmentation is around $r = 0.98$. All of this shows us that the proposed atlas can be used for the patients as well as for the volunteers.

The registration and segmentation results are condensed in Table 5.3.

If the standard deviation of the difference between manual and automated segmentation is calculated over the whole set (volunteers and patients) and compared to the intra-observer variability we can see that both errors have

Table 5.3: Condensed experimental results. First two rows show the registration error (measured on synthesized examples), while the last three rows show the segmentation error (measured on empirical data).

| Validation type | Error |
|---|---|
| Phantom | 2.2% |
| Real images | 2.9% |
| Atlas/volunteers | 4.6% |
| Atlas/AS | 5.1% |
| Atlas/CAD | 8.7% |

the same order of magnitude (Section 5.3.7). The same conclusion holds for the average of absolute values, which is summarized in Table 5.4. In addition, Section 5.3.8 shows that there is no statistically significant difference between automatic-to-manual and manual (intra-observer) error. In this sense, we can conclude that the accuracy of the method is fundamentally limited by the (in)accuracy of the manual segmentation.

Table 5.4: Standard deviation and average of absolute values of percentage difference between cardiac parameters from manual and automated segmentation.

| | Comp-Human | INTRA |
|---|---|---|
| $\sigma(e_{ttp})$ | 17% | 12% |
| $\mu(\lvert e_{ttp}\rvert)$ | 12% | 9% |
| $\sigma(e_{trf})$ | 23% | 16% |
| $\mu(\lvert e_{trf}\rvert)$ | 18% | 13% |
| $\sigma(e_{max})$ | 6% | 5% |
| $\mu(\lvert e_{max}\rvert)$ | 7% | 3% |

As reported, the mean value of the absolute difference between the automated and the manual segmentation is equal to $\mu(|d_a|) = 5.57px$, while the standard deviation of the difference is $\sigma(d_a) = 6.89px$. When this is compared to the mean value of the absolute difference between two different segmentations of the same image made by the same cardiologist ($\mu(|d_m|) = 3.62px$) and the standard deviation of the difference ($\sigma(d_m) = 5.38px$) it is obvious that these two segmentations are relatively close to one another. This is even stronger emphasized when the correlation between cardiac parameter extracted from automated and manual segmentation is observed since correlations for time-to-peak, rise-fall time ratio, and peak parameter are $c(ttp) = 0.9875$,

$c(t_{r/f}) = 0.9741$, and $c(max) = 0.9966$, respectively. Therefore, we conclude that the proposed atlas-based segmentation has comparable accuracy and precision to a human expert.

*Atlas lifted up the heaven...*

*N.N.*

*A scene from Greek mythology*

*Qual è colui che forse di Croazia*
*viene a veder la Veronica nostra,*
*che per l'antica fame non sen sazia,*
*ma dice nel pensier, fin che si mostra:*
*'Segnor mio Iesù Cristo, Dio verace,*
*or fu sì fatta la sembianza vostra?';*

*Dante*

*Paradiso 31.103-108*

## Chapter 6

# Atlas Construction

## 6.1 Introduction

When comparing experimental data obtained from different subjects, a standard approach is to display results on an atlas, as a common anatomical substrate. Through time, different methodologies for atlas formation have been developed. Many authors discussed the problem of least biased atlas selection, while many others preferred multi atlas based segmentation, as previously already discussed (see Chapter 3). In the sections of this chapter we discuss several methods for atlas formation, and compare them based on their segmentation results, which we believe should be the criteria which atlas is the most appropriate atlas for the given task. First part of this chapter compares four methods for atlas construction and is closely related to our earlier work (see [98]). Second part of the chapter investigates the multi atlas based approach to the image segmentation problem and discuss the possible improvement regarding accuracy and speed.

The results from the first part of the chapter shows that there is no significant error reduction if different approaches are used for model image construction. However, further experiments shows that the proposed methodological improvement that is not based solely on image formation, but on the appropriate segmentation selection as well does significantly improves the atlas accuracy. We wanted to compare these results to other works that showed how multi atlas based approach provides better segmentation results. For this reason another experimental setup is constructed for the second part of the chapter. In it, we first discuss the similarities between our approach and multi atlas based approach, and follow with evaluation of the atlases based on the segmentation accuracy.

The image set used in this chapter consist out of 140 ultrasound images of cardiac outflow velocity profiles from both healthy volunteers and patients (see Chapter 4), and it was evaluated using image registration algorithm described in Chapter 5. Multi atlas based approach to image segmentation was additionally evaluated on 18 3D-MRI images of the human brain using third party software (elastix) for image registration.

## 6.2 Least biased atlas construction

In the further subsections an image registration method used for atlas evaluation is introduced, and next, four methods for atlas construction are described, along with the framework for atlas evaluation. In all methods for atlas construction, described below, it is assumed that all the images are already rigidly aligned and resized to the same resolution. The images are denoted with $I_i$, and the set of images can be denoted as $S = \{I_i; i = 1..N\}$, where $N$ stands for number of images. In this experimet, we use 140 images out of which 26 belongs to healthy volunteers, while the rest 114 belong to patients. An instance of an aortic outflow velocity profile of a patient and a healthy volunteer is presented in the Figure 6.1.

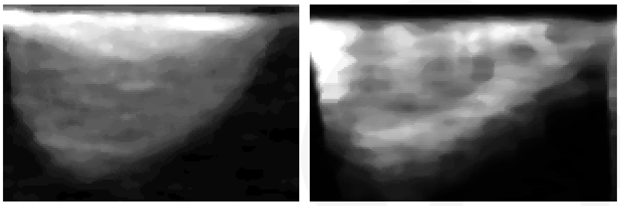


Figure 6.1: An example of the patient (left) and the healthy volunteer (right) aortic otuflow velocity profile.

### 6.2.1 Image registration

To register two images transformation function, similarity measure and optimization algorithm have to be defined. For registration of cardiac outflow profiles we have used the transformation function that non-linearly transforms image only along y-axis since this should be enough to describe all the possible physical changes. This can be expressed with the following formulae:

$$T(x, y) = t(x) \cdot y \tag{6.1}$$

where $t(x)$ denotes warping of image space frame along y-axis. This function was only estimated from N stripes selected from the image and calculating only the samples $t[k]$, after which the function $t(x)$ was reconstructed using linear interpolation. To measure the quality of the alignment between image, the similarity measure was defined in the form of normalized mutual information [59, 75, 132]. As an optimization algorithm a version of the gradient ascent algorithm [61] with multiresolution implementation was used.

### 6.2.2 Average intensity model

The average intensity image model is probably the easiest way of construction an statistical model. The construction of an model can be explained with formula:

$$A_{av}(x, y) = \frac{1}{N} \sum_{i=1}^{N} I_i(x, y) \tag{6.2}$$

### 6.2.3 Median intensity model

With median intensity image model we try to produce the atlas image only from the set of the pixels already existent in the image set $S$. The value of an each pixel at the position $(x, y)$ in the median intensity model was calculated as median of all pixels at the position $(x, y)$, selected from the set of images $I_i$:

$$A_m(x, y) = median(I_i(x, y)); i = 1..N \tag{6.3}$$

### 6.2.4 Construction of least biased model with respect to transformation function

Let's assume that all the images are already registered onto each other using the approach described in Section 6.2.1. Now we can define the distance measure to calculate how far is any image (let's say $I_i$) of the set from the rest of the images in the set (i.e. $I_j \forall i \neq j$). Let's define the distance measure as:

$$d_{ij} = \sum_{j=1, j\neq i}^{N} \sum_{k=1}^{n} |log(t_{ij}[k])| \tag{6.4}$$

Where $t_{ij}$ denote transfer vector that transforms the image $I_i$ onto the image $I_j$, and $k$ stands for $k$-th element of the transformation vector. The logarithm was used in distance measure since the scaling of the image is done by multiplication. In this way we assure the symmetry of the distance measure, i.e. that two images (one stretched by factor $\alpha$ and other squeezed by the same factor) have the same distance from theirs originals.

Defined in this way, the distance measure from the Equation 6.4 will lead to selection of the image on which we can root the model. Now, when the root image is selected, the rest of the images are mapped on this image and the average intensity image is calculated (similar to the formula from the Equation 6.2). Finally, the average image formed in this fashion is then deformed in the direction of the average deformation from all the images, that way assuring that atlas has the same distance from all the images from the set.

### 6.2.5 Average shape and intensity model

The atlas construction procedure described in this section is inspired with the idea of shape-based interpolation of multidimensional object described in the work of Raya and Udupa [167]. They presented the method that was used for shape interpolation between slices acquired from medical imaging scanner. The basic idea behind this approach is to convert the binary image (which represents the segmented object) to gray image where the gray value of the point represent the shortest distance (from the border of the binary image). The distance is defined as positive for the points within an object and negative for the points outside of an object. For an image this distance function can be observed in 3D space. For a circular object the distance transformation will have a shape of a cone, as shown in Figure 6.2.

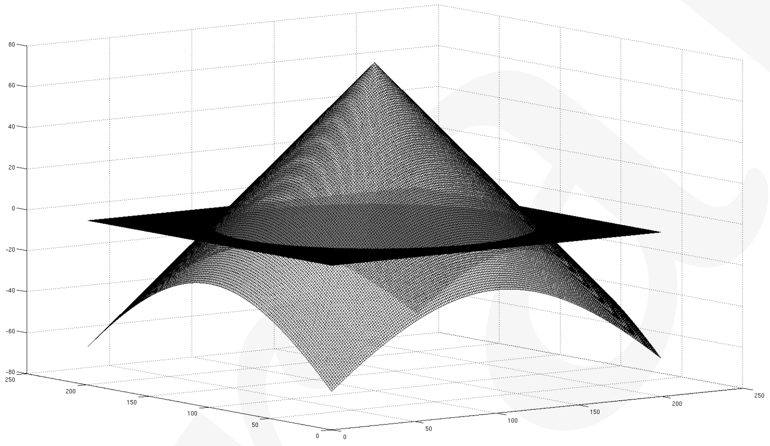


Figure 6.2: The distance transformation (cone) of an circular object on the plane.

Since the distance transformation of a binary image can again be represented as an (gray-level) image the process of averaging is by no means different. When the gray images from multiple objects are calculated and the images are averaged their mean shape can be extracted as a set of pixels with value zero in the average gray image. In other words, the border of the mean shape object is an isoline (contour line) with pixel value zero. We may think of the gray images from the Doppler ultrasound scanner as a set of 3-D object, if we represent the intensities as elevation (z-axis coordinate). This can help us to extend the idea of Raya and Udupa to a 3D space. Now, the shortest distance from the 3-D object is represented by isoplanes. The distance transformation will now give a function that exists in 4-D space and in its discrete form describes a set of isoplanes. After averaging this set we need to find the isoplane with value zero (i.e. the distance is zero). This isoplane is shape average of the 3-D object. Since the 3-D object contains the informations from both the intensity and shape of an object form an 2-D Doppler image we can say that this object is shape and intensity average of the images used for its construction. We convert this object back to gray (2-D) image and use it as a model.

### 6.2.6 Experiments and Results

In this section we evaluate four atlases type described in Sections 6.2.2 – 6.2.5. The proposed method for atlas evaluation is based on the segmentation accuracy of each atlas and the steps of this method are depicted in Figure 6.3.

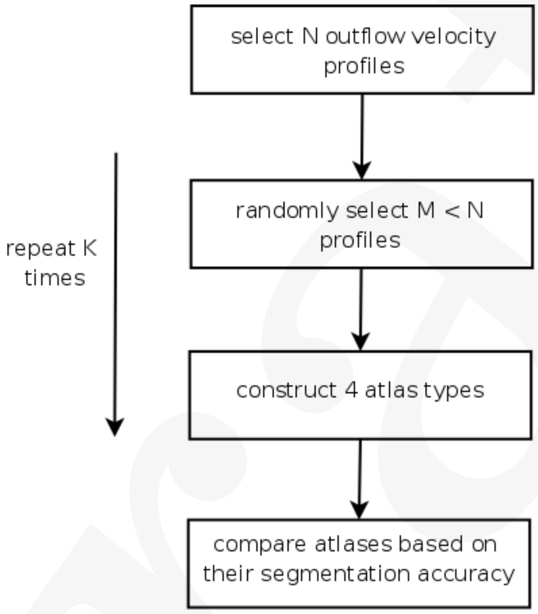


Figure 6.3: A flow chart of the proposed method for atlas evaluation.

The problem of segmentation evaluation lies in the fact that segmentation accuracy may vary based on error from manual segmentation of an atlas image (let's denote it with $e_m$), registration error ($e_r$), error from suboptimal choice of an atlas ($e_a$) and gold standard error ($e_g$). As a gold standard a manual segmentation of an image was used, and since the same image and same segmentation is used in all experiments this error is constant across experiments. In this way, we expect that only segmentation accuracy is affected. Similarly, one can use $\sigma(e_m + e_r)$ to denote the uncertainty of a method, since it depends on both the error from manual segmentation of an atlas image and the registration error. With this, only the precision of segmentation is affected, but since the same method is used in each experiment, the additional variation across experiments (due to $e_m$ and $e_r$) should be approximately the same.

When each instance of an atlas is constructed as shown in Figure 6.3 all the images that were not used for atlas construction, were transformed onto the atlas. The registration procedure for this is already described in Section 6.2.1. The transformation parameters were memorized and after the atlas is segmented, this segmentation was backward transformed onto the images where the segmentation evaluation is done. The segmentation accuracy and precision is used to evaluate the atlases performance. In these experiments, $M = 50$ images were used to construct an atlas image, and $K = 22$ times different images from the set were selected, to compare the variation across different atlas types, and across different image selection.

The preliminary results shows that segmentation error has a small variation across different atlas types. This is shown in Table 6.1. It has been widely proven in the pattern recogni- tion field that combining multiple classifiers can yield more ro- Index Terms—Atlas-based segmentation, classifier combination, combination of segmentations, majority voting, weighted voting. bust and accurate results than using single classifiers

Table 6.1: Atlas performance comparison based on segmentation accuracy. $A_1$ to $A_4$ denote atlases described in Sections 6.2.2 to 6.2.5, respectfully.

| | Err |
|---|---|
| $A_1$ | 5.2777% |
| $A_2$ | 5.2824% |
| $A_3$ | 4.5336% |
| $A_4$ | 4.7182% |

Additionally, the experiment was repeated $K = 22$ times for each method for atlas construction. The results can be depicted as shown in Figure 6.4, where vertical lines shows one standard deviation of the segmentation error for different atlas type. We can notice that the error distribution overlap significantly. From this we can conclude that there are cases when carefully selected images for atlas construction outperform the carefully selected method for atlas construction.

Knowing that segmentation also affects segmentation performance a concept of least biased segmentation selection (as compared with least biased image selection) is developed. All the images which contributed to atlas construction were segmented, and this segmentation was propagated along with image when atlas is constructed. The average segmentation is used as least biased segmentation of an atlas image. The segmentation evaluation was conducted in the same fashion as in the previous experiment. The results are shown in Figure 6.5.

From the results we can conclude that least biased segmentation selection improves performance of any atlas since it has lower mean error as well as the standard deviation.

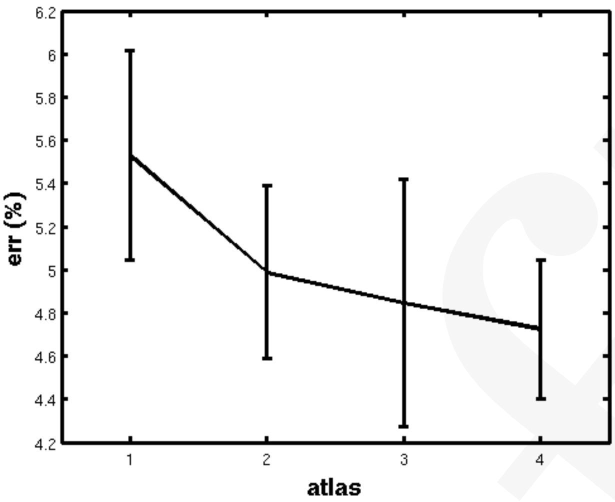


Figure 6.4: The average segmentation error along with denoted one standard deviation range. The 1 to 4 on y-axis denotes atlases described in Sections 6.2.2 to 6.2.5, respectfully.

### 6.2.7 Conclusion

In this section we proposed a novel method for least biased atlas construction (Section 6.2.5), and compared the results of this method with the performance of three different atlases (Figure 6.4). Although our method outperforms them, this improvement is not statistically significant. Additionally, we showed that the selection of a least biased atlas image did not lead to significantly better segmentation results, which may be contributed to the enough plastic transformation conducted within registration process. Finally, we presented the results that show how carefully selected atlas segmentation may have greater impact on segmentation accuracy than the atlas image selection.

## 6.3 Atlas fusion

In the previous section we discussed whether the atlas constructed from image model and label model can produce better results than the atlas constructed from the image model and its manual segmentation. The proposed approach showed significant improvement from the classical approach. In this section we challenge the idea that modeling both image and labels can produce results similar to the multi atlas based approach. The idea is motivated by the fact that both approaches incorporate the same amount of the information, and was boosted by the results presented earlier.

In the following subsection we articulate the problem mathematically, and discuss theoretical aspects of the problem. Finally, both approaches are

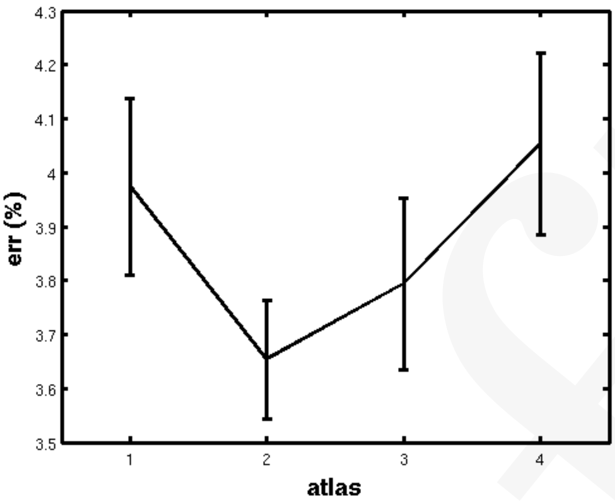


Figure 6.5: The average segmentation error along with denoted one standard deviation range. The 1 to 4 on y-axis denotes atlases described in Sections 6.2.2 to 6.2.5, respectfully. The results from least biased segmentation concept.

evaluated with respect to their segmentation accuracy.

In order to investigate the suitability of each atlas for image construction each atlas construction method will be applied to two sets of data, namely aortic outflow images, and MRI brain images. The registration itself is performed using our own algorithm when the aortic outflow images are used and a third-party software package named Elastix [180] is used for the registration of the MRI brain images.

The results show that following our approach for the model construction a time necessary to segment a set of images can be significantly improved, at the cost of negligible deterioration of the segmentation result.

### 6.3.1 Relation to other works

The idea to "average" both image and anatomical informations (e.g. labels) was used before. E.g. Chiavaras et al. [28] created the probability maps of sulci (depressions or fissures in the surface of the brain). Heckemann et al. [80] used label fusion similar to both methods proposed herein, where they referred to it as indirect method of atlas fusion, and to classical atlas fusion approach as a direct method of atlas fusion. However, in the indirect method they propagate a single atlas to a number of intermediate images, forming a set of atlases, which are in turn propagated to the subject's space. Thus, they do not average the manual segmentation in order to reduce error but rather average the different segmentation due to the image (subject) variability in order to reduce the segmentation error. This was probably done due to the difficulties in obtaining many multiple segmentations of different individuals. However, Heckemann et al. refer that the indirect method is similar to the probability mapping, which was used by many other research groups [211, 48, 69, 72, 74, 144, 193], so probably many variations on the theme do exist. The traditional approach of probabilistic atlas construction often still used [149, 148], constructs the probabilistic atlas by utilizing affine transformations only. The Heckemann et al. used nonrigid registration, but probably due to the difficulties in obtaining many multiple segmentations they used indirect method and ended up with different conclusion. It is also interesting to mention that Park et al. [156] constructs probability atlas with thin plate spline deformation, on 32 manually segmented CT images, and uses Bayesian framework to improve the results. From the paper, it seems that the core of the atlas formation method is the same as the method we propose herein. However, the paper does not compare the results to any other atlas formation alternative, and as seen in Chapter 3 the result was neglected in many other research papers (e.g [173, 80, 230] just to name a few) tracing the path towards atlas fusion as preferable multi-atlas based segmentation. We can assume that many research groups followed this direction due to the earlier insights from the pattern recognition field, where it was shown that combining multiple classifiers can yield more robust and accurate results than using single classifiers [110, 115]. Finally, the paper by Artaechevarria et al. [4] which discuss the efficient classifier combination on the example of atlas-based segmentation proposes an effective method for classifier combination which corresponds to the method that we named here as All-R-A-T. The paper compares the results with ideal classifier (oracle) in order to numerically show that there is no theoretical deterioration in the performance of the proposed lassifier. We, however, provide theoretical discussion on why this occurs, and were unaware of this paper at the time when we evaluated the results presented in this chapter and when we discussed the theoretical background behind this concept. We came across this paper when we wanted to additionally evaluate our results using independent software, since in the their paper Artaechevarria et al. also used Elastix.

### 6.3.2 Theoretical considerations

As discussed in Chapter 3 the atlas formation process is the process of finding the most representative object in the population, which we can observe as the process of modeling of the "average" object. Apart from simple averaging of the object intensities, one can try to "average" the labels utilizing some kind of voting algorithm or other classifier combination technique. Naturally, more abstract averaging is possible, such as averaging of the deformation field used to map the objects one to another, or averaging position of the objects in the manifold [241, 176, 212] of object etc. However, here the image and label averaging is implemented same as before, by averaging intensity and shape, and voting.

Let $\Xi$ denotes the averaging of either images or labels across image set $\{I_i\}$ or label set $\{L_i\}$. We write $\Xi I = \Xi_i I_i$ for image model, and $\Xi L = \Xi_i L_i$ for label model. Following this the atlas formed by manual segmentation of the image model we denote by $A_M = (\Xi I, L)$ . Similarly, we can write $A_A = (\Xi I, \Xi L)$ to denote image model paired with label model. The subscript $M$ indicates the manual segmentation in the model, while the subscript $A$ indicates the automatic formation of the segmentation from several instances of manual segmentation. Notice that $A_M$ and $A_A$ represent two approaches to atlas construction problem discussed in previous section. The multi atlas based approach usually referred in the literature as atlas fusion approach can be written as $\Xi A$, where each atlas $A$ is constructed from one manually segmented image, i.e. $A = (I, L)$.

After setting the notation it seems natural to wonder whether

$$(\Xi I, \Xi L) = \Xi(I, L)$$

or equivalently, whether

$$A_A = \Xi A$$

Since atlas is defined as pair of image and labels, averaging the atlas implies averaging both images and their labels. So, what is the difference between the two approaches? The difference lies in the root image on which the atlas is built. The multi-atlas based approach referred as atlas fusion maps all atlases onto the image space, or more precisely, onto the image that needs to be segmented. After mapping, all atlases are fused using some classifier fusion technique. This can be observed as if the image that needs to be segmented is used as the root image for atlas formation. On the other hand, the proposed approach maps all atlases within atlas space, by selecting the most appropriate image on which to root the atlas. After mapping, all atlases are blended into one atlas, therefore, for image segmentation one more mapping is necessary.

Following this reasoning, it is expected that the latter approach will contain some error. This error can be interpreted as error due to the inappropriate root image selection, since no root image is more suitable for atlas construction than the actual image that needs to be segmented. Alternatively, the error can be contributed to the mapping inaccuracy due to the parametrization of the deformation field. This will be discussed in more details further on and the error will be measured in the following subsection.

Actually, what interest us is whether we can first deform all atlases to an image, blend[1] all atlases into a single atlas and then deform that model to segment each image from the population, and would that produce the same segmentation results as to directly fuse the atlas on each image from the population. To express this in an analytical form we need to include the deformation

---

[1] via some kind of averaging procedure

filed in our equations. Let $D_{AB}$ be the deformation field necessary to map the object B to object A. Then the question is similar to the question of whether the following identity holds:

$$D_{AB}\Xi(D_{BC}I, D_{BC}L) = \Xi D_{AC}(I, L) \tag{6.5}$$

where each of the operators is written in order in which it is applied to the atlas, image or segmentation. However, since $D$ operates on underlying image space, and $\Xi$ operates on image intensities or labels, they are mutually independent operators, so the order of operations is irrelevant. This can be more obvious if we write:

$$\Xi_i I_i(D(\mathbf{x})) = \Xi_i I_i(D_i(\mathbf{x})) \iff D_i = D \forall i \tag{6.6}$$

where $\mathbf{x}$ stands for the coordinate vector of the underlying image space and subscript indicates the index of the image which spans across all images from the set (this was omitted in the Equation 6.5). If $D_i$ is constant with respect to $i$ (i.e. equals $D$), it is irrelevant whether we first apply the operator $D$ or operator $\Xi$ over the set of images $\{I_i\}$. This means that the $\Xi$ in the Equation 6.5 can switch places with $D_{AB}$, but not with $D_{BC}$, or $D_{AC}$, since both $D_{BC}$, and $D_{AC}$ are functions of $i$. If we rewrite Equation 6.5

$$\Xi D_{AB} D_{BC}(I, L) = \Xi D_{AC}(I, L) \tag{6.7}$$

we can see that the left and the right side of equations are the same if $D_{AB}D_{BC} = D_{AC}$ holds. This is known as the transitivity property, and generally does not hold for the deformation field calculated by the image registration of two discrete images. However, there are several works that discuss and aim to improve this (see [126, 31, 197, 30]). In following section we will not propose a registration process that result in a deformation field with a transitivity property, but we will make a tacit assumption that the transitivity holds for sufficiently small deformations, and that our deformations are sufficiently small.

If the left and the right side of the Equation 6.5 are the same, this means that both atlas formation approaches are the same. Both approaches reflect the atlas fusion, whereas our approach fuses atlases in the atlas space, and the classical approach fuses atlases in the image space. Our approach has one more deformation than the classical approach, however, this is the only deformation that needs to be done on-line, which means that by blending the atlases in an off-line manner we can speed up the registration procedure $N$ times, where $N$ is the number of atlases used. Whether this reduces the accuracy remains yet to be seen.

### 6.3.3 Experiments and results

In this section several atlases are compared with respect to their segmentation accuracy. Atlases were evaluated using leave-N-out method on two sets.

First set consisted from 140 images of aortic outflow profiles (same as in previous section), while the second set consisted of 18 3D-MRI brain images from IBSR2[2] database. In each case K times N randomly selected images (without repetition) were selected as the test set. The rest of the images were used for atlas formation, which was evaluated on the test set. For the aortic outflow profile images we selected $N = 100$ and $K = 22$, and for the MRI brain images we used $N = 11$ and $K = 50$. The registration procedure used for the aortic outflow images is the same as in previous Chapters, while the registration of the MRI brain images was done using Elastix [180] with B-spline interpolation [226, 179, 216] and Mattes MI implementation [139].

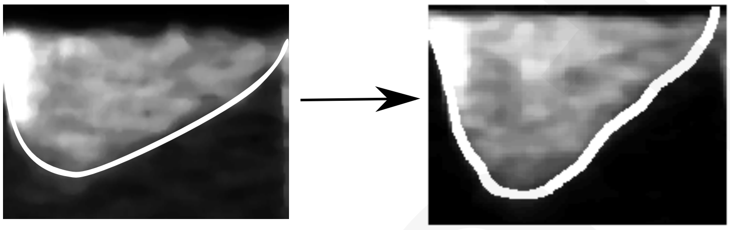


Figure 6.6: Atlas-based segmentation by propagation of manual segementation from prototype source image.

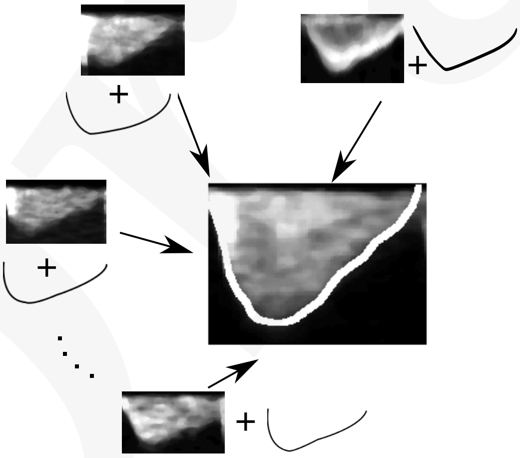


Figure 6.7: Atlas-based segmentation by propagation of manual segmentation from multiple template images.

Three alas formation methods schematically presented in Figures 6.6–6.8 and evaluated here are:

- A classical atlas based segmentation (Figure 6.6)

---

[2] `http://www.cma.mgh.harvard.edu/ibsr`

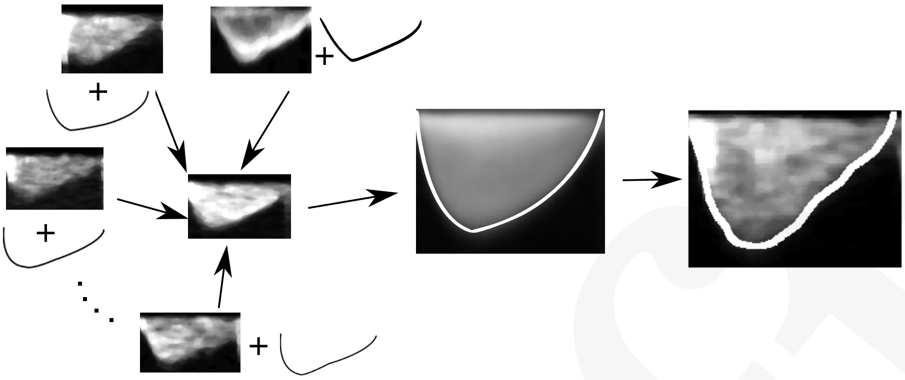


Figure 6.8: Atlas-based segmentation by segmentation propagation from one atlas whose image and segmentation is constructed from multiple template images and multiple manual segmentations.

- A classical atlas fusion approach (Figure 6.7)
- The proposed method of atlas fusion in the atlas space (Figure 6.8)

The first method (Figure 6.6) is actually a selection of prototype image paired with manual segmentation. Arrow in the Figure 6.6 represent the deformation from source to target image. The prototype and the root image was selected as the image that needs to be least deformed to all other images from the set. The method is further on referred as S-T method, since it maps the source image (used as an atlas) to target image. Second method (Figure 6.7) directly maps all images (atlases) to target image, and is therefore denoted as All-T method. The arrows in the Figure 6.7 represent the deformation field from template images to target image, and to be consistent with our previous notation (see Equation 6.5) we could each arrow denote by $D_{AC} = D_{AC}(i)$. Third method maps all images to root image and then (after atlas formation) mapping of the atlas to the target image. This method is referred as All-R-A-T method. The arrows in the left side of the Figure 6.8 represent the deformation field from template images to root image, selected as the most appropriate instance from the set, and represent the deformation denoted in Equation 6.5 by $D_{AB} = D_{AB}(i)$. Similarly, the arrows in the right side of Figure 6.8 represent the deformation field from atlas constructed on top of the root image to target image. This deformation is denoted in Equation 6.5 by $D_{BC}$ and is done only once. Note that this deformation is the only deformation for which is necessary to do the online registration, while in previous case to calculate the deformation $D_{AC}$ the registration between target image and each image form the set $\{I_i\}$ was necessary. Thus, the latter approach may be used to significantly improve the multi-atlas based segmentation. As always, this does not come at no cost. To see how this increase in speed reflects to the accuracy, the segmentation results for each method are presented in Figures 6.9 and 6.10.

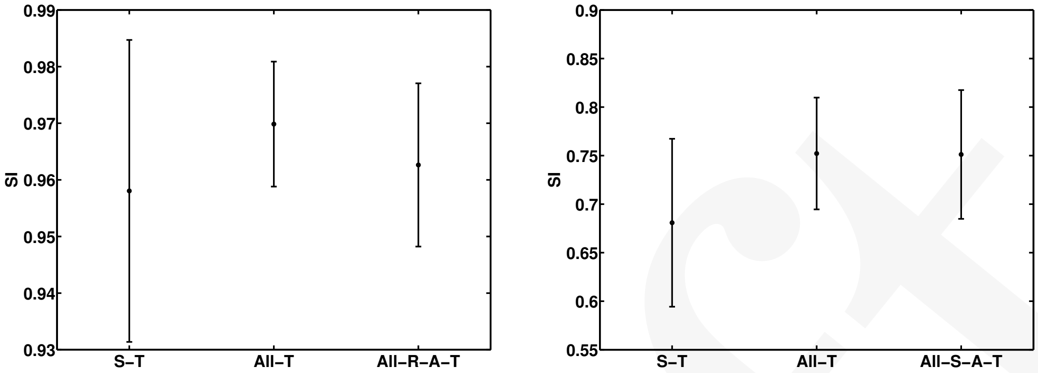


Figure 6.9: Similarity index between gold standard and atlas based segmentation calculated on aortic outflow image database and MRI brain database.

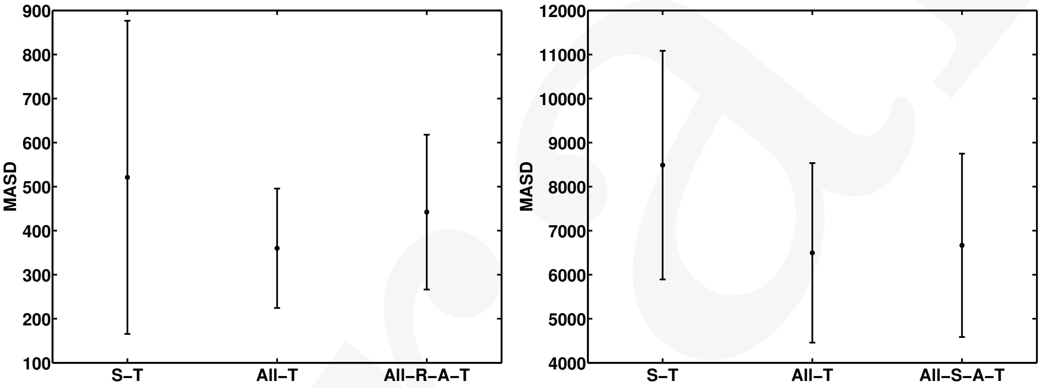


Figure 6.10: MASD between gold standard and atlas based segmentation calculated on aortic outflow image database and MRI brain database.

The accuracy of each method is evaluated using similarity index (SI) [3, 120, 182] and mean average surface distance (MASD) [199, 3]. The SI is given by:

$$SI = \frac{2|S_a \cap S_b|}{|S_a| + |S_b|} \tag{6.8}$$

While MASD is given by:

$$MASD = d(S_a, S_b) + d(S_a, S_b) \tag{6.9}$$

where $d(S_a, S_b)$ is approximated with $S_a \setminus S_b$ and $d(S_b, S_a)$ is approximated with $S_b \setminus S_a$.

### 6.3.4 Conclusion

From Figures 6.9 and 6.10 we can see that in both experiments All-R-A-T method performs almost as good as All-T method, and both method perform better than S-T method. Overall better results are achieved for aortic outflow database (larger SI) than for the MRI brain database, but this can be contributed to more complex registration procedure necessary for the registration of the 3D images. The results show that off-line atlas blending produces atlas with almost the same segmentation accuracy as the on-line atlas blending (fusion). The main advantage of the All-R-A-T method for atlas formation is its ability to break the segmentation process in two steps, thus allowing the atlas blending to be done only once. If this is done off-line, significant speed up of the segmentation algorithm can be achieved, directly proportional to the number of templates used for the atlas formation.

*Nihilominous tamen scio, homines ita esse affectos, ut mayor pars eorum; qvi has meas Machinas viderint, imo ij ipsi qvi antequam eas vidissent catalogum earu legerant, et mirificas esse censuerunt: postea spernent, et pro vulgaribus habebunt.*

*Faust Vrančić*
*Machinae novae*

# Chapter 7

# A Framework for Image Similarity Measure Construction

In this chapter we propose a novel approach for estimating image similarity. This measure is of importance in assessing image correspondence or image alignment and plays an important role in image registration. Currently, this problem is approached rather one-dimensionally, since most registration methods consider the problem as either mono- or multi-modal. This perspective leads to the selection of some form of either the correlation coefficient (CC) or mutual information (MI) as image similarity measure (ISM). We propose a more generic framework for ISM construction, based on absolute joint moments, which can be considered as a generalization of CC. Within this framework we propose a specific ISM that provides a different trade-off between MI and CC in terms of performance and computational cost for general registration problems. To illustrate this, we compare CC and MI with the proposed ISM and performed extensive experiments with regard to accuracy, robustness and speed. The evaluation demonstrated that the proposed ISM is a good combination of CC and MI, with respect to speed and performance. Therefore the proposed method is complementary to the existing CC and MI measures.

The idea to use the sum of absolute joint moments is actually an extension of the idea presented in one of our papers [96], and many results presented here overlap with one of our paper currently under review. In the following sections and related appendices, we will show not only the performance of the absolute joint moments, but also that absolute joint moments can be simplified to covariance (building block of CC) or converge to MI. While the first idea generates naturally from the very definition of the CC, the second idea was inspired by the Hausdorff moment problem [194, 54, 14, 145].

## 7.1 Introduction

Image registration is an optimization process that utilizes a similarity measure (SM) to find the optimal alignment of two images. The registration accuracy depends on the selection of the optimization algorithm and geometric transformation as well as on the definition of the similarity measure. Therefore, it is essential to use a suitable similarity measure for a given problem. The selection of an image similarity measure, especially in case of medical image registration, is usually reduced down to the question whether a multi or mono-modal registration is required. This black-and-white perspective leads to well known answers and results in the selection of correlation coefficient (CC) for mono-modal registration and mutual information (MI) for multi-modal registration.

However, the problem of registration can be approached from several perspectives, and often the registration of the same two images can be performed using different methodological approaches. For example, Zitova et al. in [251] distinguishes not only a multi-modal approach, but also a multi-view or multi-temporal. Similarly, Maintz et al. in [134], classify registration not only from a intra-/inter-modality perspective, but also from a intra-/inter-subject perspective. Perhaps for this reason many other ISM emerged over time (see Chapter 3).

In Chapter 3 we roughly classified ISMs as either CC-based or MI-based. It was also noticed that all good properties, as well as the shortcomings, of MI and CC come directly from their definition. For this reason we will primary focus on CC and MI[1] , and investigate their properties and compare them to the ISM that we propose. In the following section we will give some motivating examples to show that there is room for improvement besides the exiting ISMs. Next, we aim to propose a framework for ISM construction which will utilize a chosen amount of statistical values, instead of only a few (such as: standard deviation, skewness, kurtosis, etc.). Within this framework, we propose a specific ISM to show that when constructed in this way, it combines beneficial properties of the two most used ones: CC and MI.

Whatever similarity measure is used for image registration, it has to satisfy one basic condition - at the exact location of the correct alignment of two images, the similarity (measure) has to be maximal. To find the maximum, image registration incorporates an optimization algorithm, which iteratively calculates a similarity measure. The number of calculations may thus easily reach a number of a few thousand, especially if the geometric transformation is nonlinear. Therefore, the complexity of the similarity measure also plays an important role in registration methods. For this reason in our experiments we measure the performance of the similarity measure in terms of accuracy, robustness and speed.

---

[1] In [96] we compared accuracy and speed of more CC- and MI-based measures

## 7.2 Motivation

In previous sections we mentioned that CC is considered ideal for mono-modal image registration due to its elegance, swiftness and with the implicit assumption that in mono-modal image acquisition the relationship between pixel intensity values is affine and only corrupted by Gaussian noise [201, 231]. Due to its intrinsic properties, MI can be used in both mono-modal and multi-modal applications, the latter for which it is extensively used.

This section provides three simple examples where either CC or MI (or both) do not perform well (see Figure 7.1). There are other shortcomings of either CC or MI (or both), many of them well known from the literature, that you can find separately discussed in Appendix A.2.

As will be shown in the further sections, the performance of the MI may vary, depending on the number of bins selected to approximate the PDF. To distinguish one MI implementation from another, the number of bins is used as index (e.g. $MI_8$ and $MI_{256}$). The three examples shown in Figure 7.1 are given for each ISM implementation: $CC$, $MI_8$ and $MI_{256}$. The comparison between the images is done while applying three different geometric transformation: scaling, rotation and translation. The images which are to be aligned are constructed from the same template image selected from the test set (see Section 7.4.1) by simple noise addition, with $A_n/A_s = 0.5$ (for the details about noise degradation model see Section 7.4.2). Since both images are constructed from the same template, the correct alignment is already known, i.e. the ISM should have a maximal value for unity scaling, zero rotation and zero translation.

In the first example, the image is consecutively translated by one pixel in the range $[-30, 30]$. In the second example the image is consecutively rotated by one degree in the range $[-30, 30]$ degrees. In the third example the image is scaled by $s = 0.9 + 0.01 \cdot n$ where $n \in \{0, \dots, 20\}$. Figure 7.1 shows the values for $MI_8$, $MI_{256}$ and $CC$ calculated for the examples. Notice that $MI_{256}$ does not perform well in the first example, and neither of the ISMs has the maximum at the correct location for both rotation and scaling. Therefore, an ISM defined in a different way might be able to produce a better result in these examples, but then it obviously remains to be seen if its better performance would be related only to this particular image. Both question will be addressed further on (see Section 7.6).

## 7.3 Absolute Joint Moments

In this section we define source and target image as random variable and denote them as $S$ and $T$ instead $S(x)$ and $T(x)$. The following notation is used: $\mu_S = E[S]$, and $E[(S - \mu_S)^n]$ stands for the $n$-th central moment of the random variable $S$, or more generally the $E[(S - \mu_S)^n(T - \mu_T)^m]$ stands for

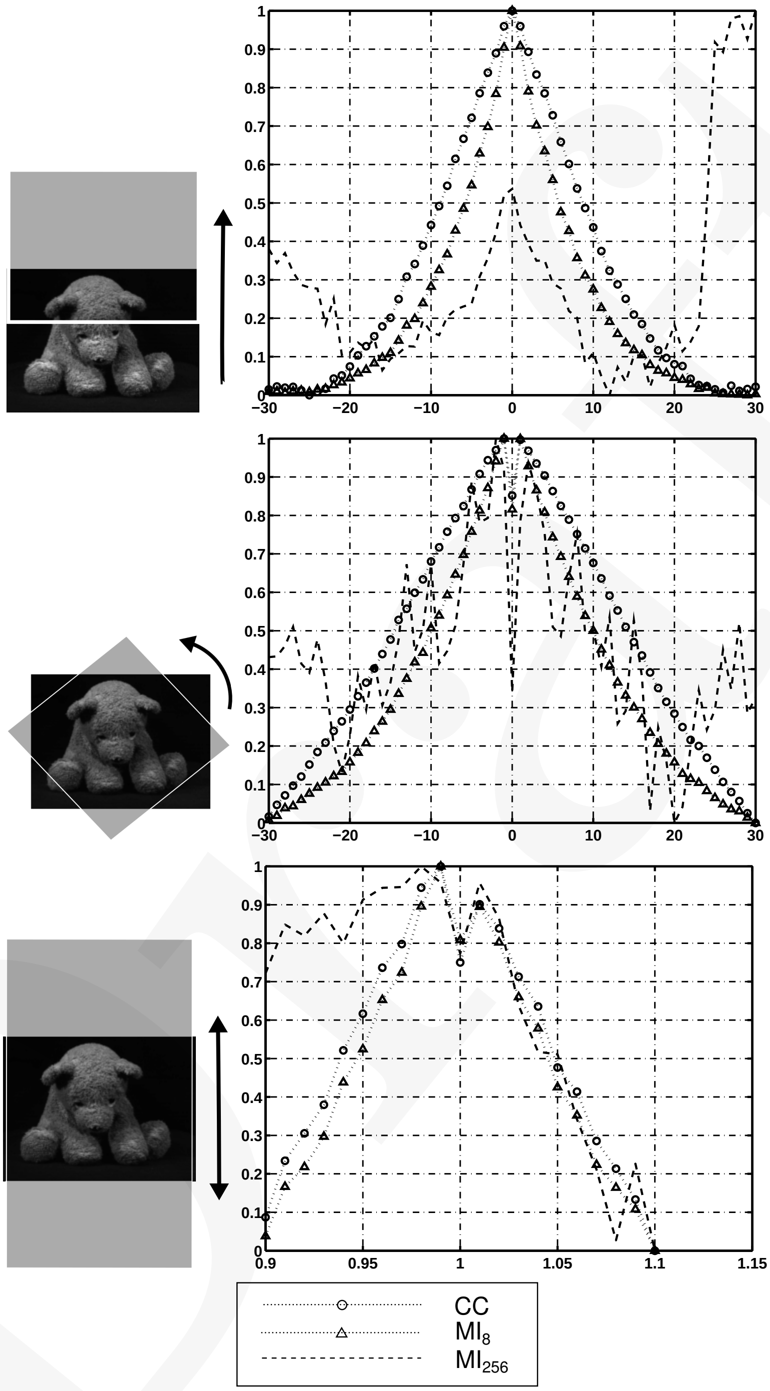


Figure 7.1: The graphs show the behaviour of $CC$, $MI_8$ and $MI_{256}$ for translation, scaling and rotation of an image (shown besides the graphs together with the geometric transformation used).

the joint central moment of the order $(m, n)$ of the random variables $S$ and $T$.

### 7.3.1 Framework for constructing image similarity measures

As mentioned in the previous section, the idea is to use more statistical information from the images (instead only a limited amount as in CC) for the ISM construction. The proposed information to use are the joint moments. The motivation behind this becomes clear if we write Equation 2.57 in the form:

$$CC(T, S) = \frac{E[(T(x) - \mu_T)(S(x) - \mu_S)]}{\sigma_T \cdot \sigma_S}, \tag{7.1}$$

were $\sigma_T$ and $\sigma_S$ denote standard deviations.

If we rewrite Equation 7.1 to the following form:

$$CC(T, S) \cdot \sigma_T \cdot \sigma_S = E[(T(x) - \mu_T)(S(x) - \mu_S)], \tag{7.2}$$

we can see that this is equal to the joint central moment of order (1,1). From this, we propose a generalization in the form:

$$AJM = |\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1}{\omega_n}\frac{1}{\omega_m} \cdot E[(T - \mu_T)^n (S - \mu_S)^m]| \tag{7.3}$$

that we refer to as the absolute joint moment (AJM) framework for ISM construction. Notice that, if we take only the first element of the sum and select $\omega_n = \omega_m = 1$, this reduces down to the absolute covariance (numerator of CC from the Equation 7.1). However, by writing the equation in this way we incorporate higher order joint moments which turn out to be valuable for establishing relationship between AJM and MI (see Appendix A.1 for details).

### 7.3.2 Proposed image similarity measure

For the proposed image similarity we use a specific selection of weights $\omega_n$ and $\omega_m$, which is computationally efficient and guarantees convergence of the sum. Using the selected weights, Equation 7.3 can be rearranged into the form of Taylor expansion of the exponential function:

$$\begin{aligned} AJM_{142} &= |\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1}{n!}\frac{1}{m!} E[(T - \mu_T)^n (S - \mu_S)^m]| = \\ &= |E[\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \frac{1}{n!}\frac{1}{m!} (T - \mu_T)^n (S - \mu_S)^m]| = \\ &= |E[\sum_{n=1}^{\infty} \frac{1}{n!}(T - \mu_T)^n \sum_{m=1}^{\infty} \frac{1}{m!}(S - \mu_S)^m]| = \\ &= |E[(e^{S-\mu_S} - 1)(e^{T-\mu_T} - 1)]| \end{aligned} \tag{7.4}$$

As the expectation is estimated on the overlapping region of images $S$ and $T$ we will have only the estimation of $AJM_{142}$:

$$\widetilde{AJM}_{142} = \frac{1}{N} \left| \sum_{x \in D_X} (e^{S-\mu_S} - 1)(e^{T-\mu_T} - 1) \right| \tag{7.5}$$

where we use $N$ to denote the number of pixel pairs. The index was used to differentiate the framework for ISM construction and the ISM itself. Index 142 is used to indicate that weights $\omega_n$ and $\omega_m$ are a sequence of factorial numbers, which is sequence $A000142$ from the OEIS[2] database. In further sections the $\widetilde{AJM}_{142}$ will be the only ISM from this framework used and tested, so we will denote it simply as $AJM$. For theoretical discussion on the existence of the expectation and AJM, as well as the numerical computation consideration please consult Appendices A.3 and A.4.

We will show that this ISM selected from the proposed framework, compared to MI and CC, will have a different trade-off between speed and performance as will result directly from the definition of AJM. In order to show this, in further sections we will investigate the properties of the AJM and compare them to MI and CC with regard to robustness, accuracy and speed.

## 7.4 Experimental data

### 7.4.1 Dataset

Figure 7.2: A representative collection of the images used in our experiments.

Images from publicly available databases were used to test the properties of the registration implemented using the AJM as similarity measure. The test set is constructed so as to have as much diversity as possible. First, we use all 44 miscellaneous images from the SIPI database[3]. Since this database does not have medical images, all 19 medical images from the VIS database[4], are

---

[2] `http://oeis.org/A000142`

[3] `http://sipi.usc.edu/database`

[4] `http://vis-www.cs.umass.edu/~vislib/Medical/InfarctScan/images.html`

added to the set, as well as 3 mammography images from the MIAS database[5]. From the MIAS set, only three images are selected since the variability of the images from this set is low. Finally, 34 images of different objects from the ALOI database[6] are added to the set which makes the total number of images in the testing set 100. The set constructed in this way contains images with different context, from natural to artificially constructed images. Both color and gray-scaled images are represented in this set, but all images are converted to gray-scaled images before the registration. All the images are coded with either 7 or eight bits per pixel and the resolution of images ranges from 128-by-128 to 1024-by-1024. For the purpose of the experiment all images are converted to double floating point precision and scaled to interval [0,1].

A few images from each database, forming a representative collection of images from the test set, are depicted in the first row of Figure 7.2. In the second row of the Figure 7.2 the same images after artificial degradation is shown (see Section 7.4.2 for details). The pairs of images constructed in this way will be used for the evaluation of the ISMs in most of the further experiments.

### 7.4.2 Image degradation model

To evaluate the performance of the ISMs we will need a pair of each image. Therefore we introduce several degradation models that will be applied on each image in order to simulate different effects that may happen during the image acquisition process such as: excessive noise, contrast changes or nonlinear intensity distortion. The image degradation models are described in the following subsection and, as can be noticed, are inspired by the paper of Maes et al. [132].

#### Contrast inhomogeneity

A linear spatially-variant intensity transformation is used to simulate image contrast inhomogeneity effects and is modelled by the following expression:

$$I'(x,y) = \alpha(x,y) I(x,y) \tag{7.6}$$

where $\alpha(x,y)$ is defined as:

$$\alpha(x,y) = \frac{1}{1 + k_1 \cdot ((x - x_c)^2 + (y - y_c)^2)} \tag{7.7}$$

with $(x_c, y_c)$ being the coordinates of the point around which the curve is positioned and $k_1$ the distortion parameter.

---

[5] `http://peipa.essex.ac.uk/ipa/pix/mias/`

[6] `http://staff.science.uva.nl/~mark/aloi//aloi_grey_red4_view.tar`

### Noise

Additive uniform noise from the interval [0, $k_2 A_s$] is superimposed on the original image. Here, $A_s$ stands for the amplitude of the signal, and $k_2$ for the amplitude ratios between noise and signal ($A_n/A_s$).

### Nonlinear intensity distortion

A non-linear intensity transformation is used to simulate pixel value distortion and is described by the following polynomial:

$$f(I_{xy}) = (I_{xy} - i_1)(I_{xy} - i_2)(I_{xy} - i_3)\dots(I_{xy} - i_n) \tag{7.8}$$

where $I_{xy}$ stands for the intensity level (at position $(x, y)$), and $i_1, i_2, i_3 \dots i_n$ are roots of the polynomial that simulates the intensity distortion.

After each distortion, the image is normalized to keep the original range of pixel values.

## 7.5 Experiments and Results

In the first two experiments (Sections 7.5.1 and 7.5.2) the image pairs, between which the correct alignment is to be determined, are the original and the degraded image. Therefore, the gold standard is well known, since the correct alignment is for unity scaling, and zero translation and rotation. To evaluate the performance of the ISM, an exhaustive search for the global maximum is done. This was done in order to assure that the suitability of the ISM, rather than the search strategy, is evaluated. Each ISM between image pairs is calculated for a progressive shift of one pixel in the interval [-100,100], for stepwise rotation of one degree in the interval [-180,179] and for scaling in steps of 0.01 to increase or decrease the scaling factor in the interval [0.5,2].

### 7.5.1 Robustness test

Various image degradations may alter the intensities and may affect the performance of a similarity measure. To evaluate the robustness of a similarity measure with respect to additive noise, contrast change, or nonlinear intensity distortion, each degradation model (described in Section 7.4.2) was applied to each image from the set (see Section 7.4.1) and the similarity between the original and degraded image was calculated for different translation, rotation, and scaling factor.

For the degradation model the parameters listed in the Table 7.1 were used.

As can be assumed from the Table 7.1, only third order polynomials were used for the robustness test. However, this selection ensures nonlinear distortion, as can be seen in Figure 7.3.

Table 7.1: Parameters range for the robustness test

| | Range |
|---|---|
| $(x_c, y_c)$ | image center |
| $k_1$ | $[0.0001, 0.0004]$ |
| $k_2$ | $[0.1, 1]$ |
| $i_1$ | $[0.1, 0.5]$ |
| $i_2$ | $0.5$ |
| $i_3$ | $[0.5, 0.9]$ |
| $i_1 + i_3$ | $1$ |

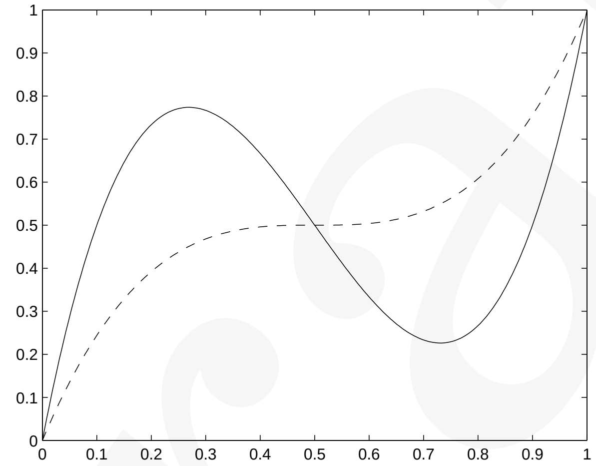


Figure 7.3: Polynomials used in the robustness test. The solid line represents the case $i_1 = 0.1$, while the dashed line represents the case $i_1 = 0.5$.

Here, we will focus on the robustness of the ISMs with respect to the amount of deformation. If the ISM is robust to the degradation effects, the maximum of the ISM will be achieved for approximately the same transformation for any amount and type of degradation. Therefore, in this test, we assess how the number of images with alignment error less than $\xi$, changes with respect to the amount of degradation. $\xi$ is defined as $5px$ for translation, 0.05 for scaling (see Eq. 7.9 for error definition), and $9^\circ$ for rotation - which represents 5% of the x-axis range (maximal deformation) in the Figure 7.7– 7.9. The experiment is done for each degradation model and the alignment error was measured for three different type of image transformation: translation, scaling, and rotation. Only one degree of freedom was allowed in each degradation model, namely: $k_1$, $k_2$, and $i_1$. Each parameter was changed in ten equidistant steps within the range allowed by the Table 7.1. The experimental results are given in Figures 7.4– 7.6. Figure 7.4 compares the performance of ISMs when only translation is allowed, Figure 7.5 shows the performance for rotation, and Figure 7.6 for scaling.

As anticipated, this experiment showed that $MI_{256}$ is not robust to noise, and that $CC$ is not robust to nonlinear intensity distortions. For the $AJM$, one can observe that it is affected by high nonlinear intensity distortion. However, for a moderate distortion, it still performs satisfactory. Therefore, $AJM$ seems robust to noise and to moderate amounts of nonlinear intensity distortion. $AJM$ is also fairly robust to contrast inhomogeneity, since the performance of $AJM$ is comparable to this of the other ISMs. All these conclusions hold for translation, rotation and scaling.

The results of the previous experiment could vary if a different error threshold $\xi$ is selected. To estimate how this results would change for different $\xi$ selected, we performed the following test. Again, all three degradation models were applied and the images were aligned after three different transformation. However, the fixed parameters in this case are $k_1$, $k_2$, and $i_1$. The parameters are set to represent the largest degradation. The results are presented in Figures 7.7– 7.9, where graphs of the ISMs for different distortion are plotted, and in each figure a different transformation is used to achieve the correct image alignment. In Figure 7.7 the alignment is achieved using translation only, in Figure 7.8 using rotation, and in Figure 7.9 scaling. The graphs show how many images are aligned with an error lower than a certain amount, and therefore give insight in the distribution of the ISM error for the image dataset.

Figure 7.7– 7.9 show that, no matter what error threshold ($\xi$) is selected, the order of the ISMs for their relative performance would remain approximately the same. For all experiments (with the exception of the combination of rotation and contrast inhomogeneities), $AJM$ is between $CC$ and $MI$ with regard to the overall number of correct alignments and sometimes even between the two different implementations of $MI$. This was as might be assumed from its theoretical properties.

As expected, the experiments showed that no ISM could compete with the performance of MI for large nonlinear intensity distortions. However, the use of higher order moments was helpful, since $AJM$ performs better than CC for nonlinear intensity distortion. From the graphs, it is also clear that in a noiseless environment $MI_{256}$ performs better than $MI_8$, but in a noisy environment $MI_{256}$ is not such a good choice.

### 7.5.2 Accuracy test

This test is used to evaluate the accuracy of localizing the exact position of the maximum of the ISM. The image pairs for the experiments are the original images from the test set and the degraded images, both of which are explained in Section 7.4. The degraded images are constructed by applying all three

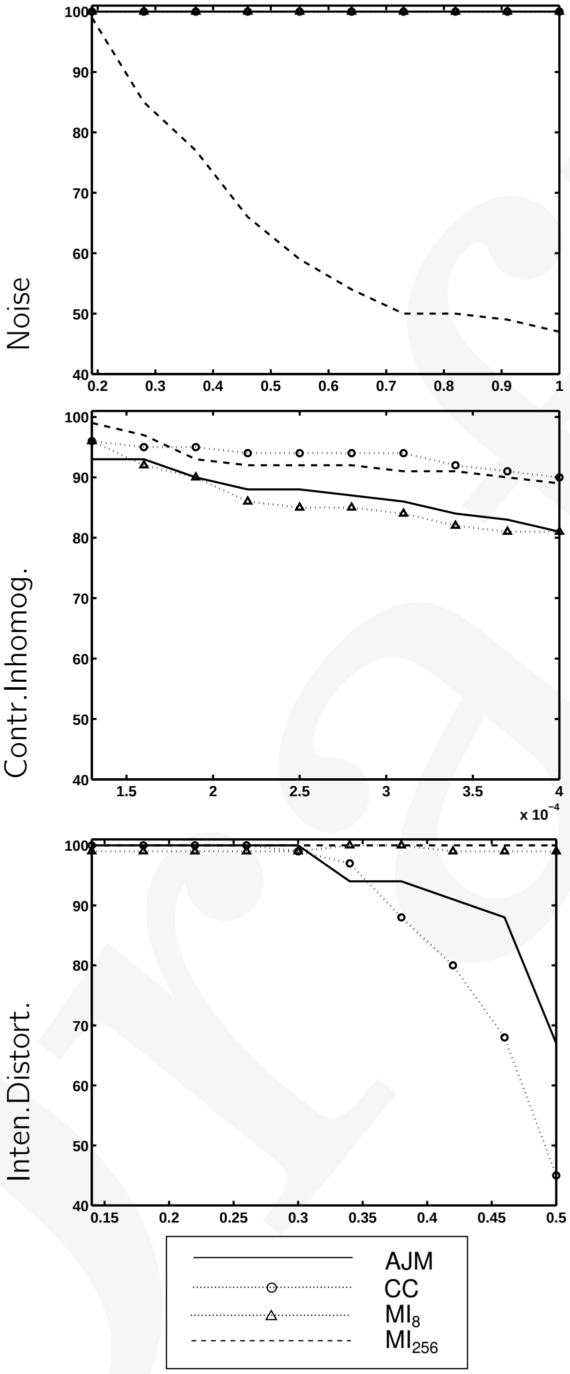


Figure 7.4: Robustness of ISMs when only translation is used for alignment. The graphs show the change in the performance of each ISM with regard to the distortion. For contrast inhomogeneities, this parameter is $k_1$, for noise it is $k_2$, and for nonlinear distortion the parameter is $i_1$. For each ISMs, and for each distortion, the y-axis shows the the number of images with alignment error less than $\xi$, and the x-axis shows amount of distortion (i.e. $k_1$, $k_2$, and $i_1$). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively..

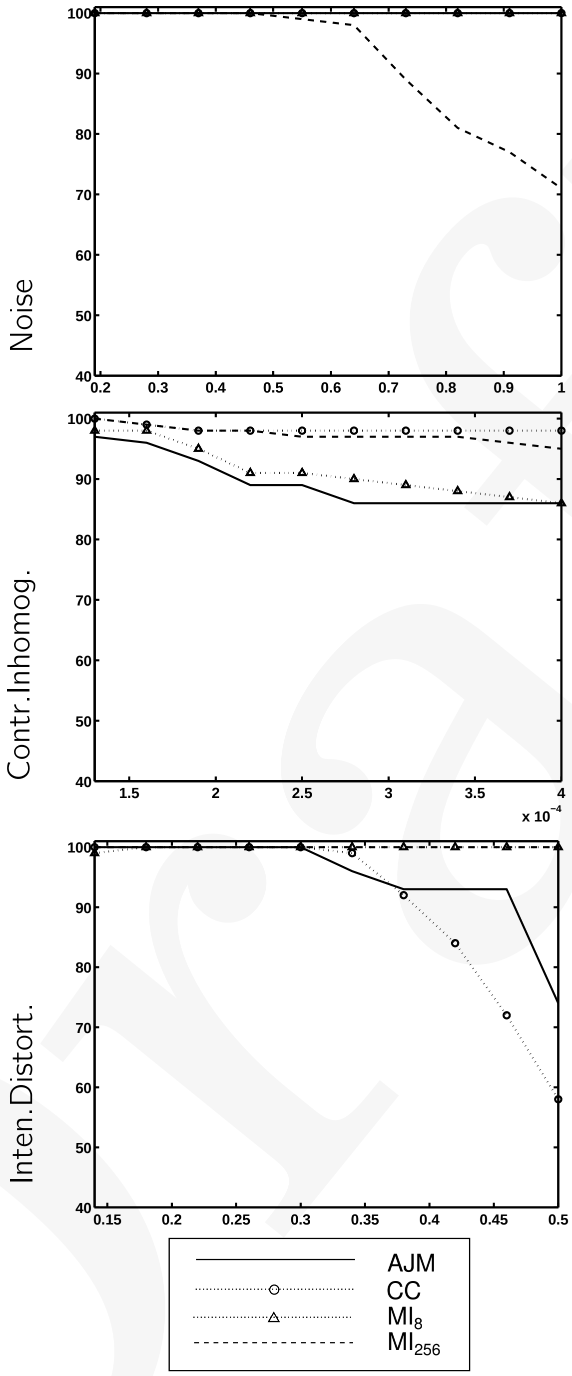


Figure 7.5: Robustness of ISMs when only rotation is used for alignment. The graphs show the change in the performance of each ISM with regard to the distortion. For contrast inhomogeneities, this parameter is $k_1$, for noise it is $k_2$, and for nonlinear distortion the parameter is $i_1$. For each ISMs, and for each distortion, the y-axis shows the the number of images with alignment error less than $\xi$, and the x-axis shows amount of distortion (i.e. $k_1$, $k_2$, and $i_1$). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively.

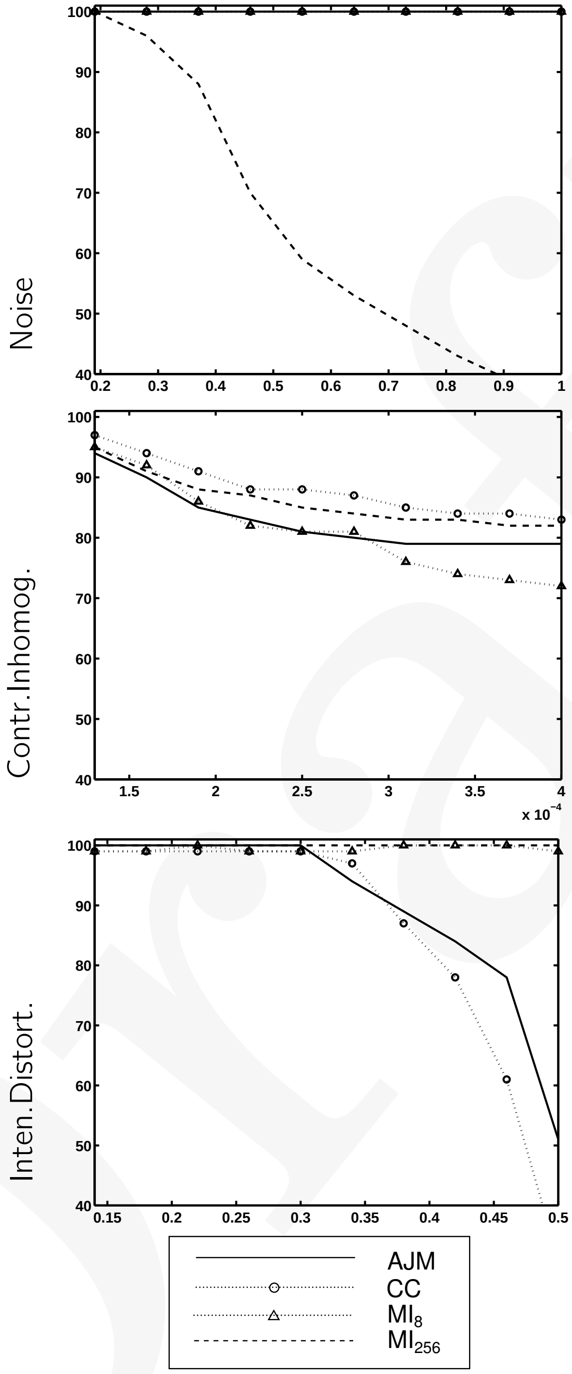


Figure 7.6: Robustness of ISMs when only scaling is used for alignment. The graphs show the change in the performance of each ISM with regard to the distortion. For contrast inhomogeneities, this parameter is $k_1$, for noise it is $k_2$, and for nonlinear distortion the parameter is $i_1$. For each ISMs, and for each distortion, the y-axis shows the the number of images with alignment error less than $\xi$, and the x-axis shows amount of distortion (i.e. $k_1$, $k_2$, and $i_1$). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively..

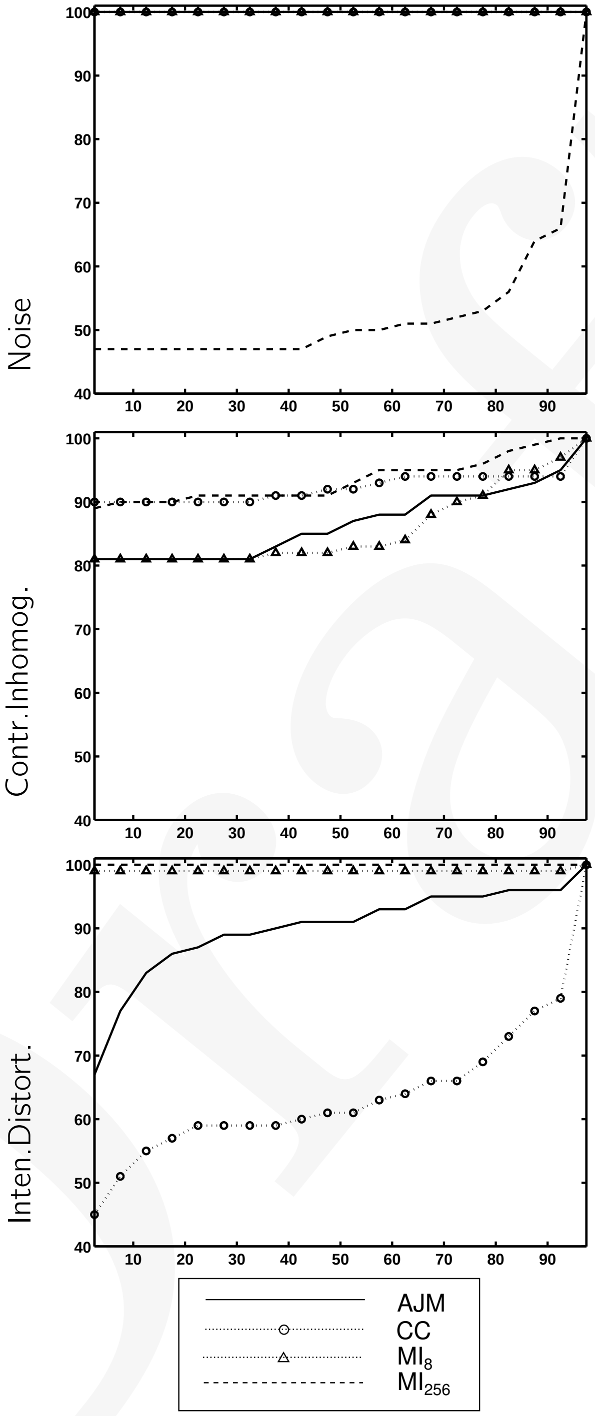


Figure 7.7: Distributions of the ISM errors when translation is used as geometric transformation. The graphs show the total number (y-axis) of aligned images with an error lower than x (value on the x-axis). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively.

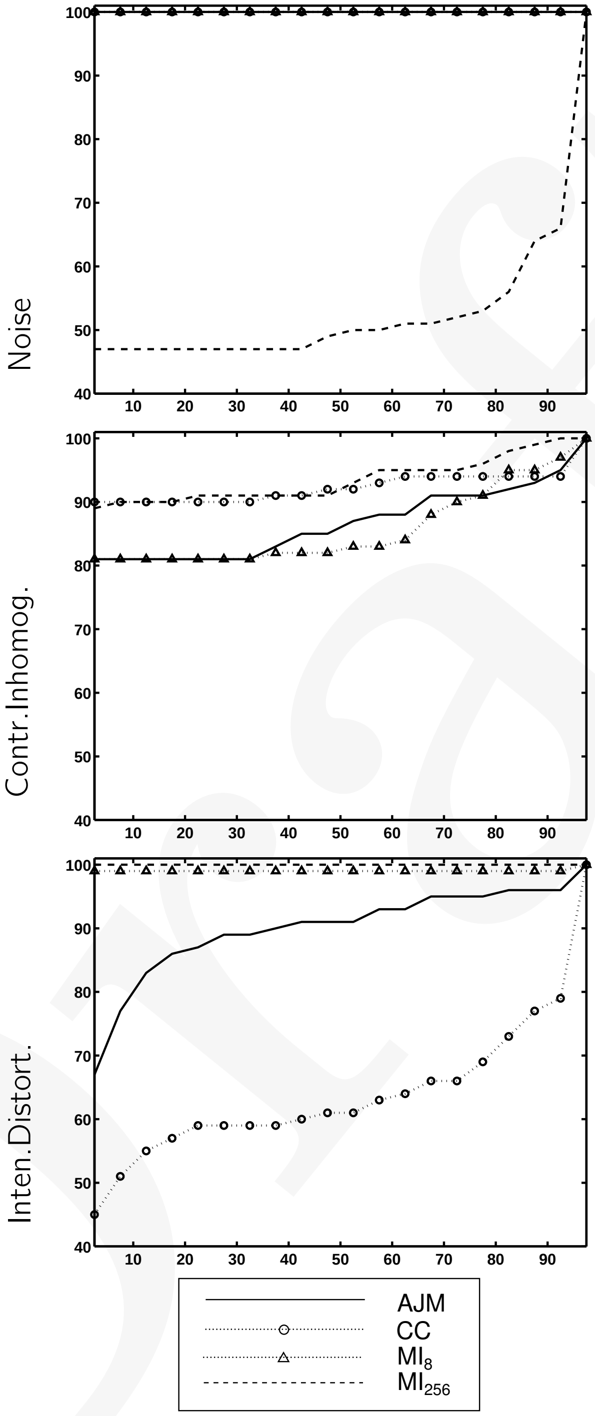


Figure 7.8: Distributions of the ISM errors when rotation is used as geometric transformation. The graphs show the total number (y-axis) of aligned images with an error lower than x (value on the x-axis). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively.

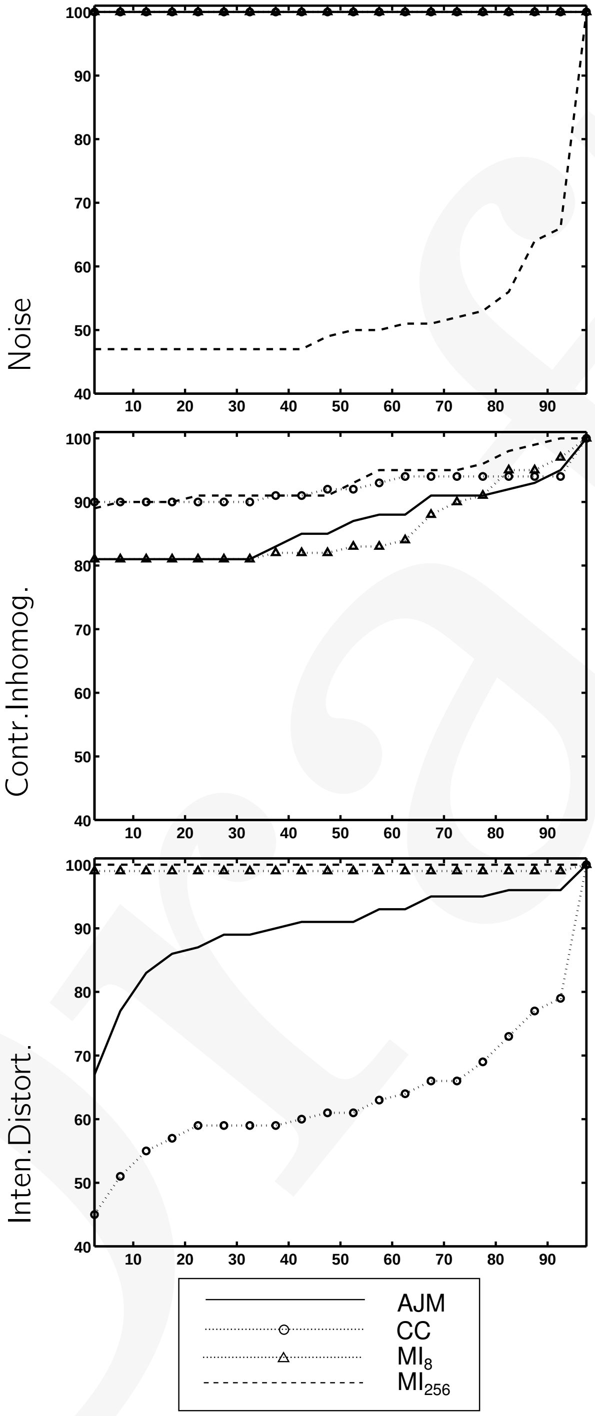


Figure 7.9: Distributions of the ISM errors when scaling is used as geometric transformation. The graphs show the total number (y-axis) of aligned images with an error lower than x (value on the x-axis). The performance of $CC$, $AJM$, $MI_8$, and $MI_{256}$ are plotted using circles, solid line, triangles and dashed line, respectively.

degradation model, where first additive noise is added, next contrast degradation is simulated and, finally, nonlinear intensity distortion is done on the pixel value. For degradation of each image the parameters are selected randomly within the value range listed in the Table 7.2.

Table 7.2: Parameter range for the accuracy test

| | Range |
|---|---|
| $(x_c, y_c)$ | within image |
| $k_1$ | $[0.00005, 0.0005]$ |
| $k_2$ | $[0, 1]$ |
| $i_1 \dots i_n$ | $[0, 1]$ |
| $n$ | $\{2, 3, 4, 5, 6\}$ |

As can be noticed from Table 7.2, the range of parameters is a bit larger in this test than in the previous one. So higher contrast inhomogeneity is allowed, as well as lower noise, similarly, nonlinear image distortion is implemented as an $n$-th order polynomial, where $n$ can range from two to six. Notice that, according to the robustness test, the increase of the range of these values will not go in favour of $AJM$.

The error in image alignment is measured separately for translation, scaling and rotation. This is depicted in Figures 7.10, 7.12, and 7.11. The circle represents the average error of each ISM, and the vertical line shows one standard deviation from the average. The error for translation and rotation is given in pixel and degrees, and the error for scaling is a unitless value calculated as:

$$\epsilon_S = log_2(D_i) \tag{7.9}$$

where $D_i$ is the scaling factor (deformation amount) for which the maximal ISM value is achieved. Index $i$ stands for the measurement (image) number and $N$ denotes the total number of measurements (images). This is done so that the scaling error is symmetrical, i.e. it gives the same error for squeezing and stretching the image by the same factor. Also, it gives no error if the images are scaled by the same factor.

Figures 7.10 and 7.12 show that $CC$, $MI_8$, and $AJM$ have approximately the same error, but $MI_8$ has the least and $CC$ the largest variance. It is also clear that $MI_{256}$ has both larger errors and variance than $AJM$. The Figure 7.11 shows the results for rotation. Notice that for rotation, $MI_8$, $MI_{256}$ and $AJM$ show similar results, all three being better than $CC$.

In Table 7.3, the overall averaged absolute registration error calculated for translation (first column), scaling (second column), and rotation (third column) is shown.

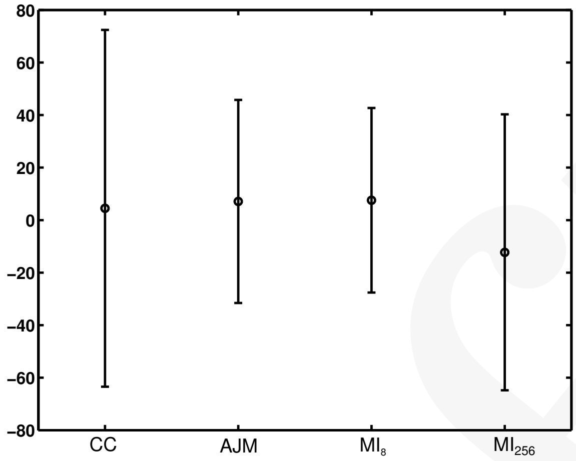


Figure 7.10: Alignment error and its standard deviation for different ISM using translation. y-axis is in pixels.

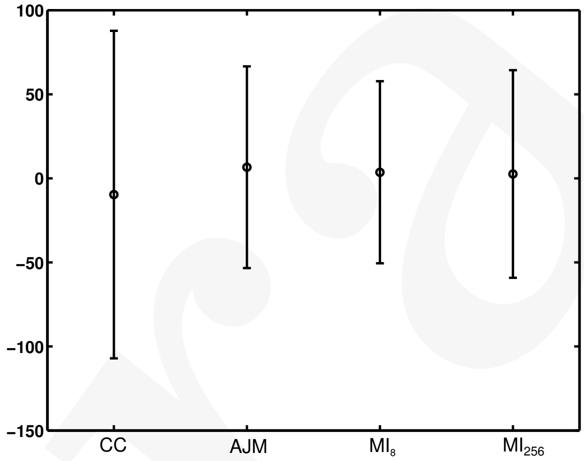


Figure 7.11: Alignment error and its standard deviation for different ISM using rotation. y-axis is in degrees.

### 7.5.3 Execution time test

Since CC, MI and AJM are implemented as defined in equations 2.57, 2.60 and 7.4 respectively, we can notice that the computational complexity of all three measures is $\mathcal{O}(N)$. However, it is expected that CC will work faster than AJM since it has to calculate only a product and a ratio instead of an exponential. Similarly, we can expect that AJM is faster than MI since it does not requires the histogram formation. To evaluate this, we measured the execution time of all three algorithms. For this purpose, the Matlab Profiler was used.

The algorithm was implemented on a standard quad-core PC without parallelization. The computation time required for one evaluation of the similarity measure varies linearly with the number of samples in the overlapping region between the two images. The results are shown in Table 7.4. The speed

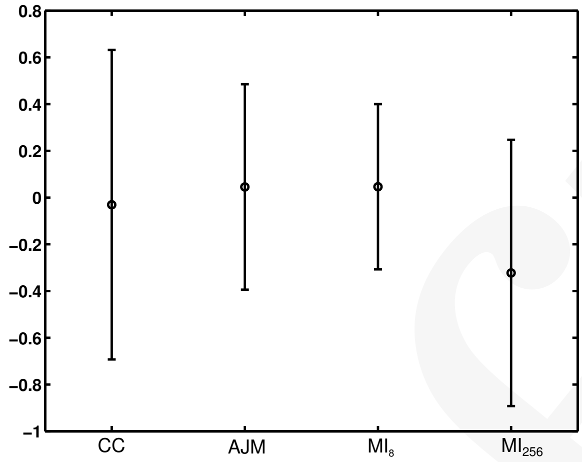


Figure 7.12: Alignment error and its standard deviation for different ISM using scaling. y-axis is unitless value.

Table 7.3: The average absolute error for translation, rotation and scaling.

| | $\overline{\epsilon_T}$ | $\overline{\epsilon_R}$ | $\overline{\epsilon_S}$ |
|---|---|---|---|
| $CC$ | 51.27 | 72.17 | 0.49 |
| $AJM$ | 19.57 | 31.04 | 0.24 |
| $MI_8$ | 15.48 | 23.70 | 0.15 |
| $MI_{256}$ | 33.01 | 37.73 | 0.45 |

performance evaluation is calculated as the average from 1000 function calls for an image size 512 × 512.

Table 7.4: Execution time measurements.

| | Time |
|---|---|
| $CC$ | 20.772 ms |
| $AJM$ | 24.609 ms |
| $MI_8$ | 50.010 ms |
| $MI_{256}$ | 78.627 ms |

### 7.5.4 Medical image registration test

The test investigates the suitability of the AJM for image registration. In the test, 26 medical images are used. The images are aortic outflow images acquired by Doppler ultrasound from healthy volunteers. In the registration the Nelder-Mead simplex algorithm is used as an optimization algorithm, and the same geometric transformation as in previous experiments (see Section 5.2.1) is also used here. Each image is registered onto itself, and the registration error is measured as:

$$\epsilon = \sum_i |ln(f_i)| \tag{7.10}$$

where $f_i$ stands for scaling factor of the transformation vector **f**. Ideally, each image registered to itself should result with the transformation vector whose components are ones. The registration error measured by Equation 7.10 is symmetrical, i.e. it returns same error for the same amount of squeezing and stretching, and it returns zero if images are correctly registered.

For AJM used as ISM, the average registration error is 6.24% and for $MI_{256}$ the average error is 7.48%. The average number of function evaluation necessary for AJM to converge was 231.54, while for $MI_{256}$ it was 264.12.

## 7.6 Overall comparison

At the beginning of the Chapter (Section 7.2) some examples are given, where $CC$, $MI_8$, and $MI_{256}$ are unable to find the exact alignment. To investigate how $AJM$ performs in a similar setting, the same experiments were repeated using $AJM$. The results are presented in Figures 7.13, showing the performance of $AJM$ alongside $CC$, $MI_8$, or $MI_{256}$.

As can be noticed, in all three cases, AJM outperforms the other ISMs. To check that this is due to the nature of this particular image, we ran this experiment on all images from the dataset. To recall, both images were created from the same original by simple noise addition, and we tried to align them using translation, rotation and scaling. In each experiment we counted the number of cases where AJM outperforms the other ISMs. The results are given in Table 7.5.

Table 7.5: The total number of aligned image pairs for which AJM outperforms other ISMs, with respect to given geometric transformation. Image pairs are altered using a noise degradation model.

| | $CC$ | $MI_8$ | $MI_{256}$ |
|---|---|---|---|
| Translation | 1 | 0 | 66 |
| Rotation | 69 | 60 | 67 |
| Scaling | 68 | 51 | 72 |

## 7.7 Discussion and Conclusion

The experiments show that $AJM$ is robust to noise, fairly robust to contrast inhomogeneities, and more robust than $CC$ and less robust than $MI$ for non-

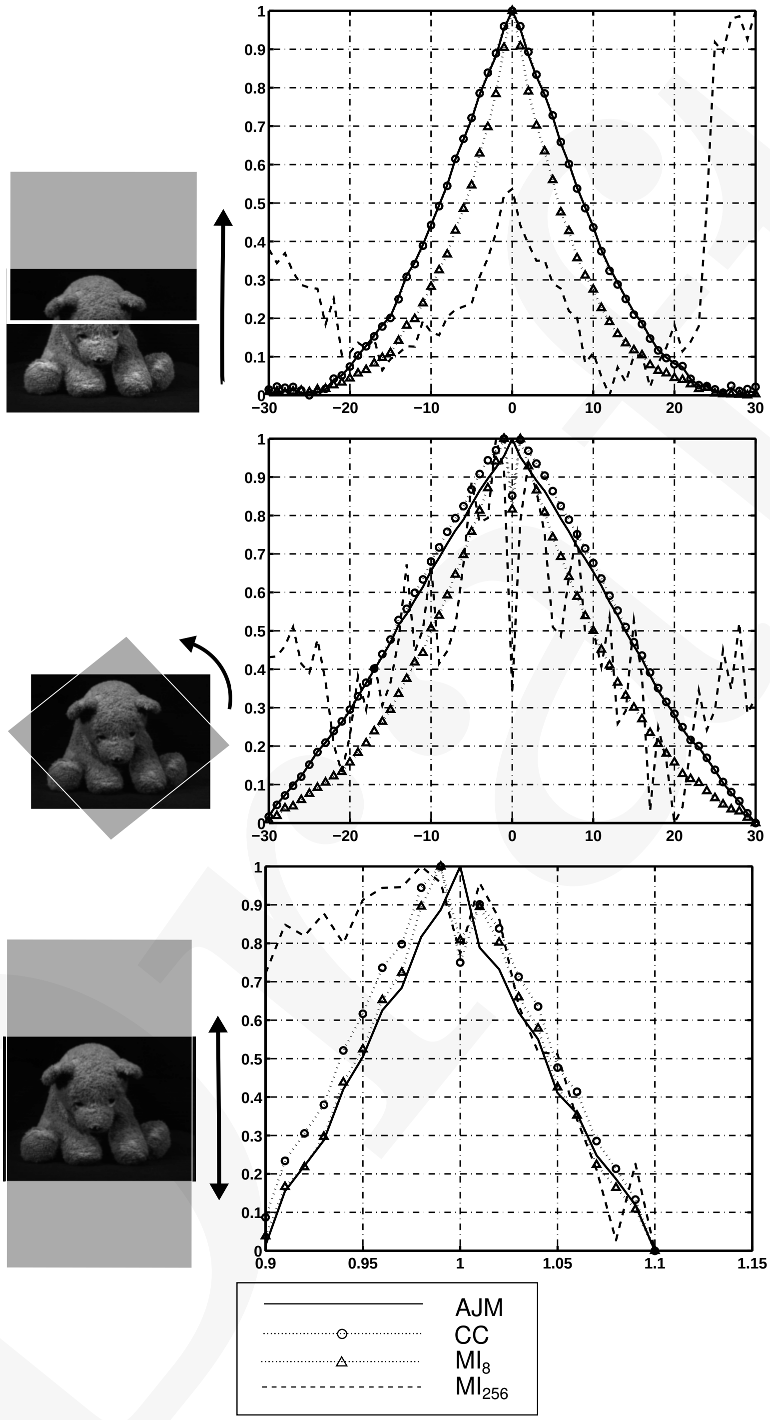


Figure 7.13: The graphs show the behavior of $AJM$ alongside $CC$, $MI_8$ and $MI_{256}$ for translation, scaling and rotation of an image (shown besides the graphs together with the geometric transformation used).

linear intensity distortion. The robustness test also showed that $MI_{256}$ in not very robust to noise. As a general conclusion, we can say that the robustness of the $AJM$ is a compromise between $CC$ and $MI$.

The accuracy test shows that $AJM$ is less accurate than $MI_8$ and overall better than $MI_{256}$. This difference between different MI implementation also emphasizes how $MI$ is affected by the number of bins of the histogram and the interpolation techniques, while $AJM$ is not. Both $MI$ and $AJM$ outperform $CC$ significantly, primarily due to the fact that $CC$ cannot cope with nonlinear intensity distortion. An additional strength of $AJM$ is that it does not require a statistically significant number of pixels for the calculation of entropy, so it can be calculated for a smaller region compared to $MI$.

Although all three ISMs have similar complexity, $AJM$ is a faster method than $MI$ since it does not require histogram calculations (nor estimations), but it is still slower than $CC$. The comparison is qualitatively summarized in Table 7.6.

Table 7.6: Performance of similarity measures

| | CC | MI | AJM |
|---|---|---|---|
| noise | + | -/+ | ++ |
| contrast inhomogeneity | + | + | + |
| nonlin. intensity deform. | - | ++ | + |
| speed | ++ | - | + |
| overap invarinat | + | - | - |
| insensitive to no. of bins | + | - | + |
| insensitive to interpolation | + | - | + |

As a general conclusion, we can say that the experiments have shown that the proposed ISM is able to determine the correspondence among images with complex relationships between the pixel values and is computationally more efficient and does not have some of the inherent disadvantages of $MI$. It is also clear (Table 7.6) that $AJM_{142}$ from the $AJM$ framework for ISM construction provides a different trade-off between speed and performance and can be seen as a useful compromise between MI and CC.

*Birds of a feather flock together*
*A proverb*

... دآيْشه دآ اوَبْوَسَلِمْ ...

*Chirvat türkisi a.k.a. Croatian song, 1588*
*One of the oldest texts written on Croatian using Arabic script*

# Chapter 8

# The Use of Developed Models in Clinical Practice

This chapter describes an ongoing research that tries to relate the shape of the aortic outflow velocity profile with functional recovery of the patients with aortic stenosis (AS), after aortic valve replacement. The idea is to investigate whether the shape of the aortic outflow velocity profile predicts the functional recovery of patients with AS after aortic valve replacement. In order to do this from the aortic outflow segmentation few morphological features are extracted and their relation to three different population groups is studied preopearive patients (PRE group), postoperative patients (POST group) and population of healthy volunteers used as control group (CTRL group). Some features (rise time and fall time) have previously been proposed in aortic outflow analysis, while we also measure the asymmetry of aortic outflow profile and use it as a feature. After their mutual comparison, the fetures are related to mean pressure gradient ($PGmean$) and aortic valve area (AVA) as typical measures of stenosis severity, and to change of ejection fraction ($\Delta EF$) as measure of functional recovery. The results show that rise time, fall time and asymmetry index vary across different populations, however POST group is closer to the CTRL group than the PRE group, showing that selected features can meaningfully quantitize the shape of aortic outflow velocity profiles. Furthermore the results show that asymmetry index correlates to $\Delta EF$ and has better diagnostic accuracy than either $PGmean$ or AVA which are echocardiography measurements routinely used in aortic stenosis severity quantification.

## 8.1 Introduction

The profile of Doppler traces of aortic outflow provides information on blood flow velocities and pressure gradients which indicates the stenosis severity in patients with AS. However, apart from valve properties (valve area) blood flow velocities are also determined by other factors such as left ventricular (LV) contractility, or (systemic) afterload. Thus, the Doppler trace of aortic outflow profile in patients with AS should reflect not only stenosis severity, but potential LV functional (e.g. myocardial) damage as well.

It is a challenge to recognize patients with subclinical myocardial dysfunction in order to operate early enough to prevent irreversible myocardial damage, but not so early as to subject them to unnecessary risks. In other words, we would like to have a parameter with good diagnostic accuracy of the functional recovery of patients with AS if they undergone the aortic valve replacement. For this purpose we propose a novel parameter that we named asymmetry index ($asymm$), which is calculated as the relative ratio of the areas under the aortic flow velocity curves extracted from the model-based segmentation of the aortic outflow profile images (see Figure 8.1 and Equation 8.1).

First we compare the asymmetry index with other morphological features extracted from the aortic flow velocity curves. Since the functional recovery of the patients with AS after aortic valve replacement is measured by the change of the ejection fraction ($\Delta EF$), we also compare the asymmetry index with $\Delta EF$. To assess diagnostic accuracy of asymmetry index and other parameters usually used for the same purpose, such as mean systolic pressure gradient ($PGmean$) or aortic valve area (AVA), the comparison of the area under receiver operating characteristic (ROC) curves was performed. The ROC curves are calculated based on the functional recovery of the patients ($\Delta EF > 0$ measurements).

## 8.2 Method

The study is conducted on 31 patients undergoing aortic valve surgery for severe AS (as defined by the European Society of Cardiology guidelines [183]) preoperatively (PRE) and 9 months postoperatively (POST).

The analysis and feature extraction of morphological features is done on the aortic outflow curves acquired from the atlas-based segmentation of the aortic outflow images. The atlas used in atlas-based segmentation is constructed by All-R-A-T method described in Chapter 6 using 29 images from normal patients. Extracted morphological features are rise time ($t_r$), fall time ($t_f$) and asymmetry index ($asymm$), where $asymm$ is calculated as:

$$asymm = \frac{A_1 - A_2}{A} \tag{8.1}$$

where $A_1$ is area under first half of the aortic outflow curve, and the $A_2$ is area under second half of the curve. The $A$ denotes the total area ($A_1 + A_2$) under the curve. The morphological feature extraction is graphically described in Figure 8.1.

Additional echochardiographic measurements are done using dedicated software (Echopac, GE Horten, Norway). Ejection fraction (EF) was measured

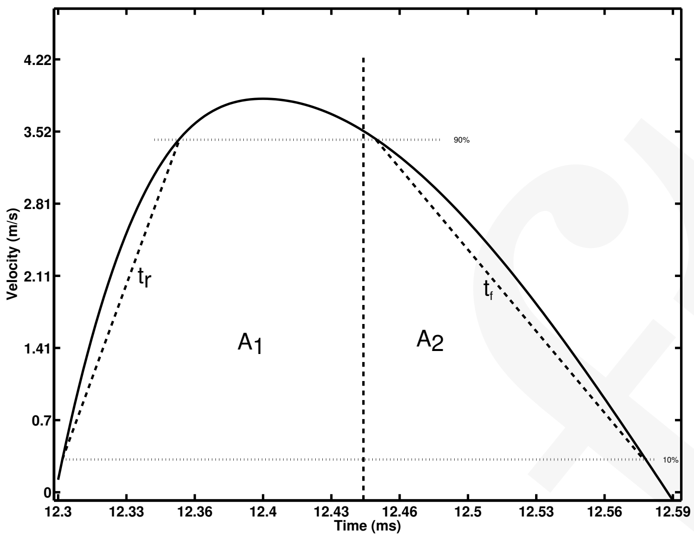


Figure 8.1: Image shows the aortic outflow curve with $t_r$ and $t_f$ measured from 10% to 90% and from 90% to 10 % respectively. $A_1$ and $A_2$ denote the left and right area under the curve, respectively.

by the Simpson's biplane method [118] using equation:

$$EF = \frac{EDV - ESV}{EDV} \tag{8.2}$$

where $EDV$ denotes end-diastolic volume and $ESV$ denotes end-sistolic volume. Continuous wave Doppler traces of the aortic outflow were analyzed by planimetry of the Doppler envelope in order to obtain the mean systolic transaortic pressure gradient ($PGmean$), which was calculated by the modified Bernoulli equation [9]:

$$\Delta P = 4 \cdot (V_2^2 - V_1^2) \tag{8.3}$$

Aortic valve area (AVA) was calculated from the continuity equation [9]:

$$A_1 \cdot v_1 = A_2 \cdot v_2 \tag{8.4}$$

which states that area-velocity product is constant. In Equation 8.4 the $A_1$ and $v_1$ are area and velocity measured at the left ventricular outflow tract, and $v_2$ is measured aortic outflow velocity, and $A_2$ is aortic valve area (AVA).

## 8.3 Experiments and Results

Table 8.1 shows the morphological features extracted from aortic outflow curves alongside manual echochardiographic measurements. Measurements from the patient group preoperatively (PRE) and nine months postoperatively

(POST), are compared to measurements obtained on the control group. For each population (PRE, POST and CTRL) the table lists mean values plus standard deviations of each parameter. In addition, for each population a percentage of population having the asymmetry index lower than 0.25 is calculated. Already from this we can notice that the postoperative measurements of asymmetry are more similar to the healthy population than the preoperative measurements.

Table 8.1: Morphological features extracted from aortic outflow curves and manual echochardiographic measurements from the same patient group preoperatively (PRE) and postoperatively (POST), and control group (CTRL).

| | PRE | POST | CTRL |
|---|---|---|---|
| $asymm < 0.25(\%)$ | 96.77 | 16.13 | 0.00 |
| $asymm$ | $0.16 \pm 0.04$ | $0.28 \pm 0.04$ | $0.33 \pm 0.06$ |
| $t_r\,(ms)$ | $58.94 \pm 12.44$ | $41.69 \pm 11.48$ | $32.00 \pm 8.00$ |
| $t_f\,(ms)$ | $125.56 \pm 22.46$ | $155.26 \pm 23.66$ | $153.00 \pm 13.00$ |
| $AVA\,(cm^2)$ | $0.80 \pm 0.18$ | $1.94 \pm 0.44$ | $3.54 \pm 0.66$ |
| $PGmean\,(mmHg)$ | $48.29 \pm 13.24$ | $13.26 \pm 5.04$ | $3.14 \pm 0.53$ |

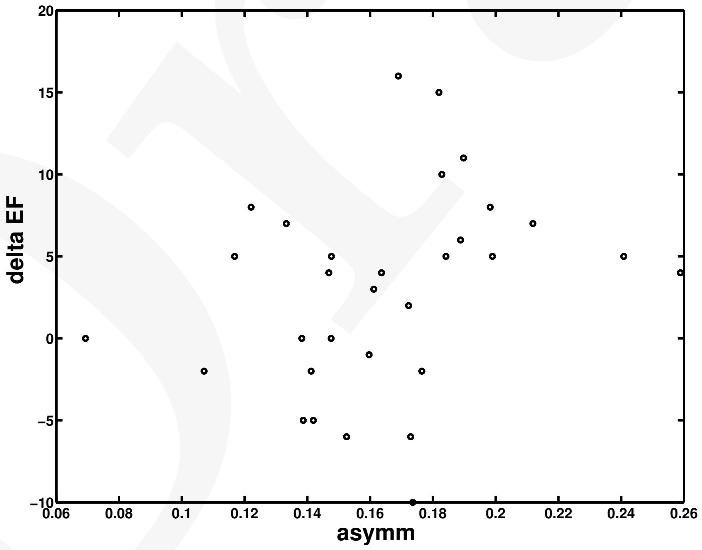


Figure 8.2: The correlation between asymmetry index of the PRE group and delta EF.

From Table 8.1 we can see that *asymm* was the lowest in the preoperative patient group indicating the most symmetrical traces, while its values rise significantly after operation and becomes more alike the ones in the control group. Notably, 96.77% of the AS patients had an abnormally symmetric trace ($asymm < 0.25$) before aortic valve surgery, while only 16.13% of the AS patients retained such symmetric traces after operation and none of the control group patients had such abnormal traces. Similarly, $t_r$ was the longest in the PRE group, with a significant shortening in the POST group, while it was the shortest in the control group. On the contrary, $t_f$ was the shortest in the PRE group, rising significantly in the POST group while it was the longest within the control group.

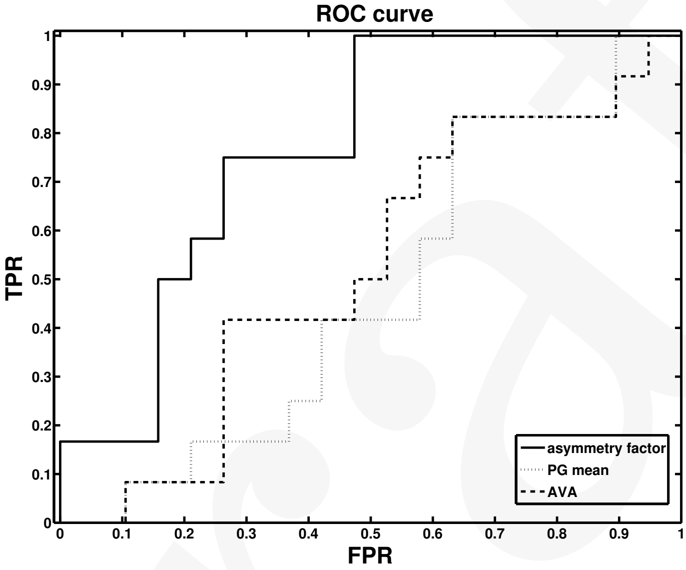


Figure 8.3: ROC curve for $asymm$ (solid line), $PGmean$ (dotted line) and AVA (dashed line).

Within the PRE group of patients, a correlation between $asymm$ and $\Delta EF$ is present (as demonstrated in Figure 8.2 ). To measure the diagnostic potential of $asymm$, the ROC curve is used. Figure 8.3 demonstrates the ROC curves for $asymm$, AVA and $PGmean$ in relation to $\Delta EF$.

Figure 8.2 shows that the asymmetry index ($asymm$) has better diagnostic accuracy than mean systolic pressure gradient ($PGmean$) or aortic valve area (AVA).

## 8.4 Conclusion

In this Chapter, we have shown some possible use of the aortic outflow segmentation and developed models in clinical practice. From the results it is clear that the aortic outflow profiles patients with AS vary in the degree of symmetry and that this can be quantified by several morphological features. Furthermore, it was shown that the symmetry of aortic outflow profile relates to postoperative functional recovery of patients. We have seen that the shape of the instantaneous peak velocity in patients with AS is markedly more symmetric than in the control group and that an increase of symmetry relates to poorer EF recovery 9 months after surgery. When compared to AVA or $PGmean$, the proposed morphological feature ($asymm$) is a better predictor of functional recovery than either AVA or $PGmean$.

*Vorhersagen sind schwierig, besonders wenn sie die Zukunft betreffen.*
*zugeschrieben Niels Bohr, Albert Einstein, Karl Valentin u.a.*

# Chapter 9

# Conclusion

In the thesis we have presented a method for registration, segmentation and model-based analysis of transaortic valvular flow ultrasound images.

In order to analyze transaortic valvular flow ultrasound images a model-based segmentation utilizing image registration was done. Chapter 5 presented one possible solution to transaortic valvular flow segmentation via model-based image registration. The registration was evaluated using an in silico phantom, and the segmentation was evaluated based on gold standard provided by the expert cardiologist. It was shown that variability of the automated segmentation relative to the manual is comparable to the intra-observer variability. The same holds for the morphological features extracted from the segmentation, showing that the automated method can be used in clinical practice. Atlas used in Chapter 5 was manually segmented model constructed from multiple template images. Alternative methods for atlas formation were discussed in Chapter 6 and we proposed an atlas constructed from multiple template images presegmented by expert cardiologist. We demonstrated that such atlas is less biased and more accurate than simple atlas formation techniques, and faster than more complex image segmentation techniques utilizing multi-atlas based segmentation. We also showed that if the registration has the transitivity property the proposed method for atlas formation produces the same result as the classical atlas fusion technique. Furthermore the experiments have shown that for our experimental setup (where the registration is not restricted to have the transitivity property) the error between atlas fusion and the proposed approach is not significant. In order to accurately propagate the segmentation from atlas to images the correspondence between image and atlas has to be defined. This problem is discussed in Chapter 7 where we have proposed a novel image similarity measure that we named absolute joint moments (AJM). There we have shown that AJM relates to both correlation coefficient and mutual information, and combine their good properties. Finally, we used the aortic outflow image segmentation to extract morphological features and asses the use of the developed method in clinical practice. The experimental results presented in Chapter 8 showed that the aortic outflow profiles patients with

AS vary in the degree of symmetry and that this can be quantified by several morphological features. Moreover, it was shown that the symmetry of aortic outflow profile relates to postoperative functional recovery of patients and that it is better predictor of functional recovery than some other routinely used measurements such as AVA or $PGmean$.

*Sed contrahenda iam vela sunt et consistendum.*

*Josip Ruđer Bošković*

*De continuitatis lege*

# Bibliography

[1] Paul Aljabar, R. Heckemann, Alexander Hammers, Joseph Hajnal, and Daniel Rueckert. Classifier selection strategies for label fusion using large atlas databases. In *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention*, MICCAI'07, pages 523–531, Berlin, Heidelberg, 2007. Springer-Verlag.

[2] Paul Aljabar, R. A. Heckemann, Alexander Hammers, J. V. Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726–738, July 2009.

[3] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz de Solorzano. Combination strategies in multi-atlas image segmentation: Application to brain mr data. *Medical Imaging, IEEE Transactions on*, 28(8):1266–1277, August 2009.

[4] Xabier Artaechevarria, Arrate Muñoz-Barrutia, and Carlos Ortiz de Solorzano. Efficient classifier generation and weighted voting for atlas-based segmentation: Two small steps faster and closer to the combination oracle. In *Proceedings of SPIE*, volume 6914, pages 1W–1–1W–9. SPIE, 2008.

[5] K. V. Arya, P. Gupta, P. K. Kalra, and P. Mitra. Image registration using robust m-estimators. *Pattern Recogn. Lett.*, 28:1957–1968, November 2007.

[6] J. Ashburner and K. Frinston. Multimodal Image Coregistration and Partitioning - A Unified Framework. *NeuroImage*, 6(3):209–217, October 1997.

[7] American Heart Association. Heart disease and stroke statistics - 2008 update, 2008.

[8] M. Bach Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, and J. Thiran. Atlas-based segmentation of pathological brain mr images. In *Proceedings of International Conference on Image Processing*

*2003 , ICIP'03, Barcelona, Spain*, volume 1, pages 14–17, Boston/Dordrecht/London, 2003. IEEE.

[9] Helmut Baumgartner, Judy Hung, Javier Bermejo, John B Chambers, Arturo Evangelista, Brian P Griffin, Bernard Iung, Catherine M Otto, Patricia A Pellikka, and Miguel Quiñones. Echocardiographic assessment of valve stenosis: Eae/ase recommendations for clinical practice. *Journal of the American Society of Echocardiography official publication of the American Society of Echocardiography*, 10(1):1–25, 2009.

[10] J. Bermejo, J. C. Antoranz, M. A. García-Fernández, M. M. Moreno, and J. L. Delcán. Flow dynamics of stenotic aortic valves assessed by signal processing of doppler spectrograms. *Am J Cardiol*, pages 611–617, 2000.

[11] M. Bertalmio, G. Sapiro, and G. Randall. Morphing active contours. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(7):733–737, 2000.

[12] Margrit Betke, Harrison Hong, and Jane P. Ko. Automatic 3d registration of lung surfaces in computed tomography scans. In *MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 725–733, London, UK, 2001. Springer-Verlag.

[13] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. In *Proceedings of the IEEE Symposium on Biomedical Imaging (ISBI)*, pages 908–911, 2004.

[14] Patrick Billingsley. *Probability and Measure*. John Wiley & Sons, New York, 1995.

[15] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, October 2006.

[16] Daniel J. Blezek and James V. Miller. Atlas stratification. In *MICCAI (1)*, pages 712–719, 2006.

[17] Daniel J. Blezek and James V. Miller. Atlas stratification. *Medical Image Analysis*, 11(5):443–457, 2007. Special Issue on the Ninth International Conference on Medical Image Computing and Computer-Assisted Interventions - MICCAI 2006.

[18] I.N. Bronshtein, K.A. Semendyayev, G. Musiol, and H. Muehlig. *Handbook of mathematics*. Springer, 2007.

[19] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[20] J. T. Bushberg, J. A. Seibert, E. M. jr. Leidholdt, and J. M. Boon. *The Essential physics of medical imaging*. Lippincott Williams and Wilkins, Philadelphia, 2nd ed edition, 2002.

[21] Nathan D. Cahill, Julia A. Schnabel, J. Alison Noble, and David J. Hawkes. Overlap invariance of cumulative residual entropy measures for multimodal image alignment. In Josien P. W. Pluim and Benoit M. Dawant, editors, *Medical Imaging 2009: Image Processing*, volume 7259, 2009.

[22] Jean-François Cardoso. Dependence, correlation and gaussianity in independent component analysis. *J. Mach. Learn. Res.*, 4:1177–1203, October 2004.

[23] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik*, 66:1–31, 1993.

[24] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic Active Contours. *International Journal of Computer Vision*, 22(1):61–79, February 1997.

[25] L. Cazzanti and M.R. Gupta. Information-theoretic and set-theoretic similarity. In *IEEE International Symposium on Information Theory, 2006*, pages 1836–1840, Seattle, WA, USA, 2006.

[26] Sung-Hyuk Cha. Comprehensive survey on distance / similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences*, 1(4):300–307, 2007.

[27] Chao-I Chen, Dusty Sargent, Chang-Ming Tsai, Yuan-Fang Wang, and Dan Koppel. Uniscale multi-view registration using double dog-leg method. In *Proceedings of SPIE*, volume 7261, pages 1F–1–1F–10. SPIE, 2009.

[28] Mary M. Chiavaras, Georges LeGoualher, Alan Evans, and Michael Petrides. Three-dimensional probabilistic atlas of the human orbitofrontal sulci in standardized stereotaxic space. *NeuroImage*, 13(3):479 – 496, 2001.

[29] Gregory Chirikjian and Alexander Kyatkin. *Engineering Applications of Noncommutative Harmonic Analysis*. CRC Press, 2001.

[30] G. Christensen and H. Johnson. Consistent image registration. *IEEE Transactions on Medical Imaging*, pages 568–582, 2001.

[31] Gary E. Christensen, Hans J. Johnson, and Michael W. Vannier. Synthesizing average 3d anatomical shapes. *Neuroimage*, pages 146–158, 2006.

[32] M. Cikes, H. Kalinic, A. Baltabaeva, S. Loncaric, C. Parsai, J. Separovic Hanzevacki, I. Cikes, G. Sutherland, and B. Bijnens. Does symmetry of the aortic outflow velocity profile reflect contractile function in coronary artery disease?: An automated analysis using mathematical modeling. *Circulation*, 177(19):196, 2008.

[33] Cybèle Ciofolo and Christian Barillot. Atlas-based segmentation of 3d cerebral structures with competitive level sets and fuzzy control. *Medical Image Analysis*, 13(3):456 – 470, 2009.

[34] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. pages 263–274, 1995.

[35] D. Louis Collins, Alex P. Zijdenbos, Vasken Kollokian, John G. Sled, N. J. Kabani, Colin J. Holmes, and Alan C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging*, 17(3):463–468, 1998.

[36] Pierre Comon. Independent component analysis, A new concept? *Signal Process.*, 36(3):287–314, April 1994.

[37] T. Cootes, C. Beeston, G. Edwards, and C. Taylor. A unified framework for atlas matching using active appearance models. *16th Conference on Information Processing in Medical Imaging*, 1999.

[38] T. F. Cootes and C. J. Taylor. Statistical Models of appearance for medical image analysis and computer vision. *Proceedings of SPIE - The International Society for Optical Engineering*, 4322(1):236–248, 2001.

[39] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models—their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59, 1995.

[40] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):681–685, 2001.

[41] Timothy F. Cootes, Andrew Hill, Christopher J. Taylor, and J. Haslam. Use of active shape models for locating structures in medical images. *Image Vision Comput.*, 12(6):355–365, 1994.

[42] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, 2006.

[43] W. R. Crum, L. D. Griffin, D. L. Hill, and D. J. Hawkes. Zen and the art of medical image registration: correspondence, homology, and quality. *NeuroImage*, 20(3):1425–1437, November 2003.

[44] William R. Crum, Derek L. G. Hill, and David J. Hawkes. Information theoretic similarity measures in non-rigid registration. In *IPMI, LNCS 2732*, pages 378–387, 2003.

[45] Emiliano D'Agostino, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*, 7(4):565–575, 2003. Medical Image Computing and Computer Assisted Intervention.

[46] Sergio Damas, Oscar Cordón, and Jose Santamaría. Medical image registration using evolutionary computation: An experimental survey. *IEEE Comp. Int. Mag.*, 6(4):26–42, 2011.

[47] Benoit M. Dawant and Alex P. Zijdenbos. *Handbook of Medical Imaging*, volume 2. Medical Image Processing and Analysis, chapter Image Registration, pages 71–128. SPIE Press, 2000.

[48] Jörn Diedrichsen, Joshua H. Balsters, Jonathan Flavell, Emma Cussans, and Narender Ramnani. A probabilistic mr atlas of the human cerebellum. *NeuroImage*, 46(1):39 – 46, 2009.

[49] Jason Dowling, J. Fripp, P. Greer, J. Patterson, S. Ourselin, and O. Salvado. Automatic atlas-based segmentation of the prostate: A MICCAI 2009 prostate segmentation challenge entry. In *Proceedings of 2009 MICCAI Prostate Segmentation Challenge (London, September 2009)*, 2009.

[50] Gershon Elber. Geometric deformation-displacement maps. In *Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*, PG '02, pages 156–, Washington, DC, USA, 2002. IEEE Computer Society.

[51] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever. Medical image matching - a review with classification. *IEEE Eng. in Medicine and Biol.*, pages 26–38, 1993.

[52] David C. Van Essen and Heather A. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Journal of Neuroscience*, 17(18):7079–7102, September 1997.

[53] A. C. Evans, D. L. Collins, S. R. Millst, E. D. Brown, R. L. Kelly, and T. M. Peters. 3d statistical neuroanatomical models from 305 mri volumes. In *IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.

[54] W. Feller. *An introduction to probability theory and its applications*. Number s. 2 in Wiley series in probability and mathematical statistics. Wiley, 1971.

[55] P. Fillard, X. Pennec, P. M. Thompson, and N. Ayache. Evaluating brain anatomical correlations via canonical correlation analysis of sulcal lines. In *Proc. of MICCAI'07 Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation*, Brisbane, Australia, 2007.

[56] M A Fischler and R A Elschlager. The representation and matching of pictorial structures. *IEEE Transactions on Computers*, C-22(1):67–92, 1973.

[57] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[58] A.J. Fitch, A. Kadyrov, W.J. Christmas, and J. Kittler. Fast robust correlation. *IEEE Transactions on Image Processing*, 14(8):1063 – 1073, Aug. 2005.

[59] J. M. Fitzpatrick, D. L. G. Hill, and C. R. Jr. Maurer. *Handbook of Medical Imaging*, volume 2. Medical Image Processing and Analysis, chapter Image Registration, pages 447–514. SPIE Press, 2000.

[60] J. Michael Fitzpatrick. Fiducial registration error and target registration error are uncorrelated. *Proceedings of SPIE*, 7261(1):726102–726102–12, 2009.

[61] R. Fletcher, editor. *Practical Methods of Optimization*. John Wiley and Sons, second edition, 2001.

[62] Hassan Foroosh, Josiane B. Zerubia, and Marc Berthod. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 11(3):188–200, March 2002.

[63] A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen. Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling. *Medical Imaging, IEEE Transactions on*, 21(9):1151–1166, 2002.

[64] Orazio Gallo, Roberto Manduchi, and Abbas Rafii. Cc-ransac: Fitting planes in the presence of multiple surfaces in range data. *Pattern Recognition Letters*, 32(3):403 – 410, 2011.

[65] M. A. Gavrielides, R. Zeng, L. M. Kinnard, K. J. Myers, and N. Petrick. A template-based approach for the analysis of lung nodules in a volumetric ct phantom study. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7260 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, February 2009.

[66] Thomas A. Gaziano. Cardiovascular disease in the developing world and its cost-effective management. *Circulation 2005*, pages 3547–3553, 2005.

[67] Davi Geiger, Alok Gupta, Luiz A. Costa, and John Vlontzos. Dynamic programming for detecting, tracking, and matching deformable contours. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(3):294–302, 1995.

[68] Sarah F. F. Gibson and Brian Mirtich. A survey of deformable modeling in computer graphics. Technical report, 1997.

[69] Ali Gooya, Kilian M. Pohl, Michel Bilello, George Biros, and Christos Davatzikos. Joint segmentation and deformable registration of brain scans guided by a tumor growth model. In *MICCAI '11*, pages 532–540. Springer-Verlag, 2011.

[70] Subrahmanyam Gorthi, Meritxell Bach Cuadra, Ulrike Schick, Pierre-Alain Tercier, Abdelkarim S. Allal, and Jean-Philippe Thiran. Multi-atlas based segmentation of head and neck ct images using active contour framework. In *13th International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI2010*, pages 313–321, 2010.

[71] Subrahmanyam Gorthi, Valérie Duay, Nawal Houhou, Meritxell Bach Cuadra, Ulrike Schick, Minerva Becker, Abdelkarim S. Allal, and Jean-Philippe Thiran. Segmentation of head and neck lymph node regions for radiotherapy planning, using active contour based atlas registration. *IEEE Journal of selected topics in signal processing*, 3(1):135–147, 2009.

[72] Sylvain Gouttard, Martin Styner, Sarang Joshi, Rachel G. Smith, Heather C. Hazlett, Guido Gerig, Josien P. W. Pluim, and Joseph M. Reinhardt. Subcortical structure segmentation using probabilistic atlas priors. volume 6512, pages 65122J1–65122J11, 2007.

[73] Alexandre Guimond, Jean Meunier, and Jean-Philippe Thirion. Average brain models: a convergence study. *Computer Vision and Image Understanding*, 77(9):192–210, February 2000.

[74] Piotr A. Habas, Kio Kim, James M. Corbett-Detig, Francois Rousseau, Orit A. Glenn, A. James Barkovich, and Colin Studholme. A spatiotemporal atlas of mr intensity, tissue probability and shape of the fetal brain with application to segmentation. *NeuroImage*, 53(2):460 – 470, 2010.

[75] J. V. Hajnal, L. G. Hill, and D. J. Hawkes, editors. *Medical Image Registration*. CRC Press, Cambridge, first edition, 2001.

[76] Xiao Han, Lyndon S. Hibbard, Nicolette P. O'Connell, and Virgil Willcut. Automatic segmentation of parotids in head and neck ct images using multi-atlas fusion. In *13th International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI2010*, pages 297–304, 2010.

[77] T. Hartkens, D. Rueckert, J. A. Schnabel, D. J. Hawkes, and D. L. G. Hill. Vtk cisg registration toolkit - an open source software package for affine and non-rigid registration of single- and multimodal 3d images. In *Bildverarbeitung für die Medizin*, 2002.

[78] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

[79] L. Hatle and B. Angelsen, editors. *Doppler ultrasound in cardiology - Physical principles and clinical applications*. Lea & Febiger, Philadelphia, second edition, 1982.

[80] Rolf A. Heckemann, Joseph V. Hajnal, Paul Aljabar, Daniel Rueckert, and Alexander Hammers. Automatic anatomical brain mri segmentation combining label propagation and decision fusion. *Neuroimage*, 33(1):115–26, 2006.

[81] A. J. H. Hii, C. E. Hann, J. G. Chase, and E. E. W. Van Houten. Fast normalized cross correlation for motion tracking using basis functions. *Computer Methods and Programs in Biomedicine*, 82:144–156, May 2006.

[82] Derek L. Hill, Colin Studholme, and David J. Hawkes. Voxel similarity measures for automated image registration. *Visualization in Biomedical Computing 1994*, 2359(1):205–216, 1994.

[83] Derek L. G. Hill, Philipp G. Batchelor, Mark Holden, and David J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1–R45, March 2001.

[84] Mark Holden. A review of geometric transformations for nonrigid body registration. *IEEE Trans. Med. Imaging*, 27(1):111–128, 2008.

[85] Hyvärinen and E Oja. Independent component analysis: algorithms and applications. *Neural Networks: The Official Journal of the International Neural Network Society*, 13(4-5):411–430, Jun 2000.

[86] Aapo Hyvärinen. Independent component analysis. *Neural Computing Surveys*, 2, 2001.

[87] Ivana Isgum, Marius Staring, Annemarieke Rutten, Mathias Prokop, Max A. Viergever, and Bram van Ginneken. Multi-atlas-based segmentation with local decision fusion–application to cardiac and aortic segmentation in ct scans. *IEEE Transactions on Medical Imaging*, 28(7):1000–1010, 2009.

[88] A. K. Jain. *Fundamentals of digital image processing*. Prentice-Hall information and system sciences series. Prentice Hall, 1989.

[89] Anil Jain and Douglas Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Transaction on Pattern Analisys Machine Intelligence*, 19(2):153–158, February 1997.

[90] Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4–37, January 2000.

[91] Zsolt Jankó, Dmitry Chetverikov, and Anikó Ekárt. Using genetic algorithms is computer vision: registering images to 3d surface model. *Acta Cybern.*, 18(2):193–212, 2007.

[92] H. Johnson, G. Christensen, H. J. Johnson, and G. E. Christensen. Consistent landmark and intensity based image registration. *IEEE Trans Med Imaging*, 21(5):450–461, 2002.

[93] I. T. Jolliffe. *Principal Component Analysis*. Springer, second edition, October 2002.

[94] S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *Neuroimage*, 23(Supplement 1):–151, 2004. Mathematics in Brain Imaging.

[95] Sven Kabus, Tobias Klinder, Jens von Berg, and Cristian Lorenz. Functional non-rigid registration validation: A ct phantom study. In Bernd Fischer, Benoît M. Dawant, and Cristian Lorenz, editors, *Biomedical Image Registration*, Lecture Notes in Computer Science, pages 116–127. Springer, 2010. 4th International Workshop, WBIR 2010 Lübeck, Germany, July 11-13, 2010 Proceedings.

[96] Hrvoje Kalinić, Sven Lončarić, , and Bart Bijenens. A novel image similarity measure for image registration. In D. Seršić S. Lončarić, G. Ramponi, editor, *Proceedings of the Seventh Int'l Symposium on Image and Signal Processing and Analysis*, pages 195–199. University of Zagreb, Croatia, 2011.

[97] Hrvoje Kalinić, Sven Lončarić, Maja Čikeš, A. Baltabaeva, C. Parsai, Jadranka Šeparović, Ivo Čikeš, G. R. Sutherland, and Bart Bijenens. Analysis of doppler ultrasound outflow profiles for the detection of changes in cardiac function. In *Proceedings of the Fifth Int'l Symposium on Image and Signal Processing and Analysis (ISPA2007)*, pages 326–331, 2007.

[98] Hrvoje Kalinić, Sven Lončarić, Maja Čikeš, Davor Miličić, and Bart Bijnens. Comparison of different methods for atlas construction. In *The 17th European Signal Processing Conference (EUSIPCO 2009)*, pages 1319–1323, Glasgow, Scotland, 2009.

[99] Hrvoje Kalinić, Sven Lončarić, Maja Čikeš, Davor Miličić, and Bart Bijnens. Model-based segmentation of aortic ultrasound images. In D. Seršić S. Lončarić, G. Ramponi, editor, *Proceedings of the Seventh Int'l Symposium on Image and Signal Processing and Analysis (ISPA2011)*, pages 739–743. University of Zagreb, Croatia, 2011.

[100] Hrvoje Kalinić, Sven Lončarić, Maja Čikeš, Davor Miličić, and Bart Bijnens. Image registration and atlas-based segmentation of cardiac outflow velocity profiles. *Computer Methods and Programs in Biomedicine*, 106:188–200, 2012.

[101] Hrvoje Kalinić, Sven Lončarić, Maja Čikeš, Davor Miličić, Davor, Ivo Čikeš, George Sutherland, and Bart Bijnens. A method for registration and model-based segmentation of doppler ultrasound images. *Medical Imaging 2009: Image Processing*, 7259(1):72590, 2009.

[102] S. Kaneko, Satohb Y., and Igarashia S. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36(5):1165–1173, 2003.

[103] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[104] Y. Keller and A. Averbuch. A projection-based extension to phase correlation image alignment. *Signal Process.*, 87:124–133, January 2007.

[105] M. G. Kendall. *The advanced theory of statistics*, volume 1. C. Griffin & company limited, second edition, 1945.

[106] Shoji Kido and Akinori Tsunomori. Automated extraction of pleural effusion of three-dimensional thoracic ct images. In 7260, editor, *Proceedings of SPIE*, pages 0D1–0D8, 2008.

[107] A. P. King, C. Buerger, and T. Schaeffter. Cardiac respiratory motion modelling by simultaneous registration and modelling from dynamic mri images. In Bernd Fischer, Benoît M. Dawant, and Cristian Lorenz, editors, *Biomedical Image Registration*, Lecture Notes in Computer Science, pages 222–233. Springer, 2010. 4th International Workshop, WBIR 2010 Lübeck, Germany, July 11-13, 2010 Proceedings.

[108] Hortense A. Kirisli, Michiel Schaap, Stefan Klein, Lisan A. Neefjes, Annick C. Weustink, Theo van Walsum, and Wiro J. Niessen. A template-based approach for the analysis of lung nodules in a volumetric ct phantom study. In *Fully automatic cardiac segmentation from 3D CTA data: a multi-atlas based approach*, volume 7623, pages 05–1–05–9, 2009.

[109] Yasuyo Kita. Elastic-model driven analysis of several views of a deformable cylindrical object. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(12):1150–1162, 1996.

[110] Josef Kittler, Mohamad Hatef, Robert P. W. Duin, and Jiri Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:226–239, 1998.

[111] Arno Klein, Satrajit S. Ghosh, Brian Avants, B.T.T. Yeo, Bruce Fischl, Babak Ardekani, James C. Gee, J.J. Mann, and Ramin V. Parsey. Evaluation of volume-based and surface-based brain image registration methods. *NeuroImage*, 51(1):214 – 220, 2010.

[112] Arno Klein, Brett Mensh, Satrajit Ghosh, Jason Tourville, and Joy Hirsch. Mindboggle: Automated brain labeling with multiple atlases. *BMC Medical Imaging*, 5(1):1–14, October 2005.

[113] S. Klein, U.A. van der Heide, I.M. Lips, M. van Vulpen, M. Staring, and J.P.W. Pluim. Automatic Segmentation of the Prostate in 3D MR Images by Atlas Matching using Localized Mutual Information. *Medical Physics*, 35(4):1407–1417, April 2008.

[114] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statistics*, 22:79–86, 1951.

[115] L. I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley, 2004.

[116] Jan Kybic and Michael Unser. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11):1427–1442, 2003.

[117] Kok F. Lai and Roland T. Chin. Deformable contours: Modeling and extraction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17:1084–1090, 1995.

[118] Roberto M. Lang, Michelle Bierig, Richard B. Devereux, Frank A. Flachskampf, Elyse Foster, Patricia A. Pellikka, Michael H. Picard, Mary J. Roman, James Seward, Jack S. Shanewise, Scott D. Solomon, Kirk T. Spencer, Martin St. John Sutton, and William J. Stewart. Recommendations for chamber quantification: a report from the american society of echocardiography's guidelines and standards committee and the chamber quantification writing group, developed in conjunction with the european association of echocardiography, a branch of the european society of cardiology. *Journal of the American Society of Echocardiography official publication of the American Society of Echocardiography*, 18(12):1440–1463, 2005.

[119] Erik G. Learned-Miller. Data driven image models through continuous joint alignment. *PAMI*, 28:2006, 2006.

[120] Jong-Min Lee, Uicheul Yoon, Sang Hee Nam, Jung-Hyun Kim, In-Young Kim, and Sun I. Kim. Evaluation of automated and semi-automated skull-stripping algorithms using similarity index and segmentation error. *Computers in Biology and Medicine*, 33(6):495 – 507, 2003.

[121] F. Leymarie and M. D. Levine. Tracking deformable objects in the plane using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):617–634, June 1993.

[122] Changyang Li, Xiuying Wang, Yong Xia, Stefan Eberl, Yong Yin, and David Dagan Feng. Automated PET-guided liver segmentation from low-contrast CT volumes using probabilistic atlas. *Computer Methods and Programs in Biomedicine*, 107:164–174, 2011.

[123] Dekang Lin. An information-theoretic definition of similarity. In *Proc. 15th International Conf. on Machine Learning*, pages 296–304. Morgan Kaufmann, San Francisco, CA, 1998.

[124] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens. Nonrigid image registration using conditional mutual information. *Medical Imaging, IEEE Transactions on*, 29(1):19–29, January 2010.

[125] Sven Loncaric, Marko Subašić, and Erich Sorantin. 3-d deformable model for abdominal aortic aneurysm segmentation from ct images. In *Proceedings of the 22nd Annual International Conference of the IEEE*, 2000.

[126] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi. Multi-modal image set registration and atlas formation. *Medical Image Analisys*, 10(3):440–451, June 2006.

[127] Peter Lorenzen, Brad Davis, Guido Gerig, Elizabeth Bullitt, and Sarang Joshi. Multi-class posterior atlas formation via unbiased kullback-leibler template estimation. In *In LNCS*, pages 95–102, 2004.

[128] Maria Lorenzo-Valdés, Gerardo I. Sanchez-Ortiz, Andrew Elkington, Raad Mohiaddin, and Daniel Rueckert. Segmentation of 4d cardiac mr images using a probabilistic atlas and the em algorithm. *Medical Image Analysis*, 8(3):255–265, 2004.

[129] Maria Lorenzo-Valdés, Gerardo I. Sanchez-Ortiz, Raad Mohiaddin, and Daniel Rueckert. Atlas-based segmentation and tracking of 3d cardiac mr images using non-rigid registration. In *MICCAI (1)*, pages 642–650, 2002.

[130] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.

[131] Hongxia Luan, Feihu Qi, Zhong Xue, Liya Chen, and Dinggang Shen. Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. *Pattern Recognition*, 41:285–298, January 2008.

[132] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transaction on Medical Imaging*, 16:187–198, April 1997.

[133] P. C. Mahalanobis. On the generalised distance in statistics. In *Proceedings National Institute of Science, India*, volume 2, pages 49–55, April 1936.

[134] J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[135] J. B. Antoine Maintz, Erik H. W. Meijering, and Max A. Viergever. General multimodal elastic registration based on mutual information. In *Image Processing*, pages 144–154. SPIE Press, 1998.

[136] Timo Mäkelä, Patrick Clarysse, Outi Sipilä, Nicoleta Pauna, Quoc-Cuong Pham, Toivo Katila, and Isabelle E. Magnin. A review of cardiac image registration methods. *IEEE Trans. Med. Imaging*, 21(9):1011–1021, 2002.

[137] Ravi Malladi, James A. Sethian, and Baba C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(2):158–175, 1995.

[138] S. R. Marsland, C. J. Twining, and C. J. Taylor. Groupwise non-rigid registration using polyharmonic clamped-plate splines. In *in Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 2879 (2), pages 771–779, June 2003.

[139] David Mattes, David R Haynor, Hubert Vesselle, Thomas K Lewellen, and William Eubank. Pet-ct image registration in the chest using free-form deformations. *IEEE Transactions on Medical Imaging*, 22(1):120–128, 2003.

[140] Calvin R. Maurer and J. Michael Fitzpatrick. A review of medical image registration. In *Interactive image-guided neurosurgery*, pages 17–44, 1993.

[141] John C. Mazziotta, Arthur W. Toga, Alan Evans, Peter Fox, and Jack Lancaster. A probabilistic atlas of the human brain: Theory and rationale for its development. *NeuroImage*, 2(2):89–101, June 1995.

[142] P. McCullagh. *Tensor Methods in Statistics (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman & Hall, 1 edition, August 1987.

[143] Tim McInerney and Demetri Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, June 1996.

[144] Donald G. McLaren, Kristopher J. Kosmatka, Terrance R. Oakes, Christopher D. Kroenke, Steven G. Kohama, John A. Matochik, Don K. Ingram, and Sterling C. Johnson. A population-average mri-based atlas collection of the rhesus macaque. *NeuroImage*, 45(1):52 – 59, 2009.

[145] Robert M. Mnatsakanov and Artak S. Hakobyan. Recovery of distributions via moments. *IMS Lecture Notes*, pages 252–265, 2009.

[146] Marc Modat, Gerard R. Ridgway, Zeike A. Taylor, Daves J. Hawkes, Nick C. Fox, and Sebastien Ourselin. A parallel-friendly normalised mutual information gradient for free-form registration. In *Proceedings of SPIE*, volume 7259, pages 0L–1–0L–8. SPIE, 2009.

[147] Diane M. Muratore, Jeannette Herring Russ, Benoıt M. Dawant, and Robert L. Galloway. Three-dimensional image registration of phantom vertebrae for image-guided surgery: A preliminary study. *Computer Aided Surgery*, 7:342–352, 2002.

[148] M. Murgasova, P. Aljabar, L. Srinivasan, S. J. Counsell, V. Doria, A. Serag, I. S. Gousias, J. P. Boardman, M. A. Rutherford, A. D. Edwards, J. V. Hajnal, and D. Rueckert. A dynamic 4d probabilistic atlas of the developing brain. *NeuroImage*, 54(4):2750 – 2763, 2011.

[149] M. Murgasova, L. Srinivasan, I. S. Gousias, P. Aljabar, J. V. Hajnal, A. D. Edwards, and D. Rueckert. Construction of a dynamic 4D probabilistic atlas for the developing brain. pages 952–955, 2010.

[150] J. A. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. *Medical Imaging, IEEE Transactions on*, 25(8):987–1010, 2006.

[151] Dwight O. North. An analysis of the factors which determine signal / noise discrimination in pulsed-carrier systems. *Proceedings of the IEEE*, 51:1016–1027, 1963.

[152] S. Osher and J. Sethian. Fronts Propagating With Curvature-dependent Speed - Algorithms Based On Hamilton-Jacobi Formulations. *Journal Of Computational Physics*, 79(1):12–49, NOV 1988.

[153] Danielle F. Pace, Andinet Enquobahrie, Hua Yang, Stephen R. Aylward, and Marc Niethammer. Deformable image registration of sliding organs using anisotropic diffusive regularization. In *ISBI*, pages 407–413, 2011.

[154] G. Papandreou and P. Maragos. Multigrid geometric active contour models. *IEEE Trans. on Image Process.*, 16(1):229–240, January 2007.

[155] Hyunjin Park, Peyton H. Bland, Alfred O. Hero, and Charles R. Meyer. Least biased target selection in probabilistic atlas construction. In *MICCAI (2)*, pages 419–426, 2005.

[156] Hyunjin Park, Peyton H. Bland, and Charles R. Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans. Med. Imaging*, 22(4):483–492, 2003.

[157] Senthil Periaswamy and Hany Farid. Medical image registration with partial data. *Medical Image Analysis*, 10(3):452–464, June 2006.

[158] D. Perperidis, R. Mohiaddin, and D. Rueckert. Construction of a 4d statistical atlas of the cardiac anatomy and its use in classification. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, 8(Pt 2):402–410, 2005.

[159] Dimitrios Perperidis, Raad Mohiaddin, and Daniel Rueckert. Fast spatio-temporal free-form registration of cardiac mr image sequences. In *FIMH*, pages 414–424, 2005.

[160] D. L. Pham, C. Xu, and J. L. Prince. A survey of current methods in medical image segmentation. In *Annual Review of Biomedical Engineering*, volume 2, pages 315–338. 2000.

[161] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever. Mutual information based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[162] Josien P. W. Pluim, J. B. Antoine Maintz, and Max A. Viergever. Interpolation artefacts in mutual information-based image registration. *Comput. Vis. Image Underst.*, 77(9):211–232, 2000.

[163] Alexander D. Poularikas, editor. *The Transforms and Applications Handbook*. The Electrical Engineering Handbook Series. CRC Press, second edition, 2000.

[164] W. H. Press. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2007.

[165] Keerthi Ram, Yogesh Babu, and Jayanthi Sivaswamy. Curvature orientation histograms for detection and matching of vascular landmarks in retinal images. In *Proceedings of SPIE*, volume 7259, pages 1J–1–1J–8. SPIE, 2009.

[166] Liliane Ramus and Grégoire Malandain. Multi-atlas based segmentation: Application to the head and neck region for radiotherapy planning. In *13th International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI2010*, pages 281–288, 2010.

[167] Sai Prasad Raya and Jaram K. Udupa. Shape-based interpolation of multidimensional object. *IEEE Transaction on Medical Imaging*, 9(1):32–42, 1990.

[168] Jan Rexilius, Simon K. Warfield, Charles R. G. Guttmann, X. Wei, R. Benson, L. Wolfson, Martha Elizabeth Shenton, Heinz Handels, and Ron Kikinis. A novel nonrigid registration algorithm and applications. In *MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 923–931, London, UK, 2001. Springer-Verlag.

[169] J. Rigau, M. Feixas, M. Sbert, A. Bardera, and I. Boada. Medical image segmentation based on mutual information maximization. In *In International Conference on Medical Image Computing and Computed Assisted Intervention (MICCAI 2004), Proceedings, Rennes-Saint*, pages 135–142. Springer, 2004.

[170] Alexis Roche, Gregoire Malandain, Nicholas Ayache, and Sylvain Prima. Towards a better comprehension of similarity measures used in medical image registration. In *MICCAI*, pages 555–566, 1999.

[171] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *MICCAI '98: Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1115–1124, London, UK, 1998. Springer-Verlag.

[172] Peter Rogelj and Stanislav Kovacic. Similarity measures for non-rigid registration. In *Medical Imaging 2001: Image Processing, volume 4322 of Proc. SPIE*, pages 569–578. SPIE Press.

[173] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *Neuroimage*, 21(4):1428–1442, April 2004.

[174] Torsten Rohlfing, Robert Brandt, Randolf Menzel, Daniel B. Russakoff, and . Calvin R. Maurer, Jr. *The Handbook of Medical Image Analysis – Volume III: Registration Models*, chapter Quo Vadis, Atlas-Based Segmentation?, pages 435–486. Kluwer Academic / Plenum Publishers, 2005.

[175] P. E. Roland and K. Zilles. Brain-atlases - a new research tool. *Trends In Neurosciences*, 17(11):458–467, 1994.

[176] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.

[177] D. Rueckert and P. Aljabar. Nonrigid registration of medical images: Theory, methods, and applications. *IEEE Signal Processing Magazine*, 27(4):113–119, July 2010.

[178] D. Rueckert, A. F. Frangi, and J. A. Schnabel. Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE Trans Med Imaging*, 22(8):1014–1025, August 2003.

[179] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: Application to breast mr images. *IEEE Transactions on Medical Imaging*, 18:712–721, 1999.

[180] S. Klein AND M. Staring AND K. Murphy AND M.A. Viergever AND J.P.W. Pluim. elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196 – 205, January 2010.

[181] Mert R. Sabuncu and Peter J. Ramadge. Spatial information in entropy-based image registration. In *Proceedings of 2nd Workshop in Biomedical Image Registration (WBIR'03)*, LNCS 2717, pages 132—-141, 2003.

[182] Mehul P. Sampat, Zhou Wang, Shalini Gupta, Alan Conrad Bovik, and Mia K. Markey. Complex wavelet structural similarity: A new image similarity index. *IEEE Transactions on Image Processing*, 18(11):2385–2401, 2009.

[183] N.B Schiller, PM Shah, M Crawford, A DeMaria, Devereux R, H Feigenbaum, and et al. Recommendations for quantitation of the left ventricle by two-dimensional echocardiography. *J Am Soc Echocardiogr*, 2:358–67., 1989.

[184] L. L. Schumaker. *Spline Functions: Basic Theory*. Cambridge Mathematical Library. Cambridge University Press, 2007.

[185] Daniel Schwarz, Tomas Kasparek, Ivo Provaznik, and J. Jarkovsky. A deformable registration method for automated morphometry of mri brain images in neuropsychiatric research. *IEEE Trans. Med. Imaging*, 26(4):452–461, 2007.

[186] John L. Semmlow, editor. *Biosignal and Biomedical Image Processing MATLAB based Applications*. Marcel Dekker, Inc., 2004.

[187] J. Sethian. Advancing interfaces: Level set and fast marching methods. In *Proceedings of the International Conference on Industrial and Applied Mathematics*. Cambridge University Press, 1999.

[188] J. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 2005.

[189] Ran Shadmi, Arnaldo Mayer, Nir A. Sochen, and Hayit Greenspan. Piecewise smooth affine registration of point-sets with application to dt-mri brain fiber-data. In *ISBI*, pages 528–531, 2010.

[190] Jayant Shah. A common framework for curve evolution, segmentation and anisotropic diffusio. In *CVPR*, pages 136–142, 1996.

[191] Jayant Shah. Curve evolution and segmentation functionals: application to color images. In *In Proceedings IEEE ICIP'96*, pages 461–464, 1996.

[192] Reuben R. Shamir, Leo Joskowicz, and Yigal Shoshan. Optimal landmarks selection and fiducial marker placement for minimal target registration error in image-guided neurosurgery. In *Proceedings of SPIE*, volume 7261, pages 2N–1–1N–8. SPIE, 2009.

[193] Feng Shi, Pew-Thian Yap, Yong Fan, John H. Gilmore, Weili Lin, and Dinggang Shen. Construction of multi-region-multi-reference atlases for neonatal brain mri segmentation. *NeuroImage*, 51(2):684 – 693, 2010.

[194] J. A. Shohat and J. D. Tamarkin. *The problem of moments*. Mathematical surveys and monographs. American Mathematical Society, 1943.

[195] Darko Skerl, Bostjan Likar, and Franjo Pernus. A protocol for evaluation of similarity measures for rigid registration. *IEEE Trans. Med. Imaging*, 25(6):779–791, 2006.

[196] Darko Skerl, Bostjan Likar, and Franjo Pernus. A protocol for evaluation of similarity measures for non-rigid registration. *Medical Image Analysis*, 12(1):42–54, 2008.

[197] Oskar M. Skrinjar and Hemant Tagare. Symmetric, transitive, geometric deformation and intensity variation invariant nonrigid image registration. In *ISBI*, pages 920–923, 2004.

[198] Pieter Slagmolen, An Elen, Dieter Seghers, Dirk Loeckx, Frederik Maes, and Karin Haustermans. Atlas based liver segmentation using nonrigid registration with a b-spline. In T. Heimann, M. Styner, and B. van Ginneken, editors, *3D Segmentation in The Clinic: A Grand Challenge*, MICCAI, pages 197–206, 2007.

[199] Ingrid Sluimer, Mathias Prokop, and Bram van Ginneken. Toward automated segmentation of the pathological lung in ct. *IEEE Trans. Med. Imaging*, 24(8):1025–1038, 2005.

[200] Ingrid Sluimer, Arnold M. R. Schilham, Mathias Prokop, and Bram van Ginneken. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans. Med. Imaging*, 25(4):385–405, 2006.

[201] M. Sonka and J. M. Fitzpatrick, editors. *Handbook of Medical Imaging*, volume 2. Medical Image Processing and Analysis. SPIE Press, 2000.

[202] C. O. S. Sorzano, P. Thevenaz, and M. Unser. Elastic registration of biological images using vector-spline regularization. *IEEE Transactions on Biomedical Engineering*, 52(4):652–663, 2005.

[203] Jasbir Sra, David Krum, Angela Malloy, Melissa Vass, Barry Belanger, Elisabeth Soubelet, Regis Vaillant, and Masood Akhtar. Registration of three-dimensional left atrial computed tomographic images with projection images obtained using fluoroscopy. *Circulation*, 112(24):3763–3768, 2005.

[204] James D. Stefansic, Alan J. Herline, Yu Shyr, William C. Chapman, J. Michael Fitzpatrick, and Robert L. Galloway. Registration of physical space to laparoscopic image space for use in minimally invasive hepatic surgery. *IEEE Trans. Med. Imaging*, 19(10):1012–1023, 2000.

[205] C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(1):71–86, January 1999.

[206] Colin Studholme and Valerie Cardenas. A template free approach to volumetric spatial normalization of brain anatomy. *Pattern Recognition Letters*, 25(10):1191–1202, July 2004.

[207] Colin Studholme, John A. Little, Graeme P. Penney, Derek L. G. Hill, and David J. Hawkes. Automated multimodality registration using the full affine transformation: Application to mr and ct guided skull base surgery. In *VBC*, pages 601–606, 1996.

[208] Gérard Subsol. *Brain warping*, chapter Crest lines for curved-based warping. Academic Press, 1999.

[209] G. R. Sutherland, L. Hatle, P. Claus, J. D'hooge, and B. H. Bijnens, editors. *Doppler Myocardial Imaging - A Textbook*. Hasselt-Belgium: BSWK, first edition, 2006.

[210] Gábor Székely and Guido Gerig. Model-based segmentation of radiological images. *KI*, 14(3):18–23, 2000.

[211] A. M. Tahmasebi, P. Abolmaesumi, C. Wild, and I. Johnsrude. Quantification of inter-subject variability in human brain: A validation framework for probabilistic maps. In *SPIE Medical Imaging*, pages 726218–1–8, Orlando, 2009.

[212] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–23, 2000.

[213] Demetri Terzopoulos. On matching deformable models to images. In *Opical Meeting on Machine Vision*, volume 12 of *Technical Digest Series*, 1987.

[214] Demetri Terzopoulos and Kurt W. Fleischer. Deformable models. *The Visual Computer*, 4(6):306–331, 1988.

[215] Demetri Terzopoulos, Andrew P. Witkin, and Michael Kass. Constraints on deformable models: Recovering 3d shape and nonrigid motion. *Artif. Intell.*, 36(1):91–123, 1988.

[216] P. Thevenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *Image Processing, IEEE Transactions on*, 9(12):2083–2099, 2000.

[217] A. W. Toga. *Brain Warping*. Academic Press, 1999.

[218] A. W. Toga and P. M. Thompson. The role of image registration in brain mapping. *Image and Vision Computing*, 19(1-2):3–24, January 2001.

[219] Arthur W. Toga and Paul M. Thompson. The role of image registration in brain mapping. *Image Vision Comput.*, 19(1-2):3–24, 2001.

[220] R. Toth, S. Doyle, M. Rosen, A. Kalyanpur, S. Pungavkar, B. N. Bloch, E. Genega, N. Rofsky, R. Lenkinski, and A. Madabhushi. WERITAS: weighted ensemble of regional image textures for ASM segmentation. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7259 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, February 2009.

[221] Juerg Tschirren, Ronald M. Lauer, and Milan Sonka. Automated analysis of doppler ultrasound velocity flow diagrams. *IEEE Trans. Med. Imaging*, 20(12):1422–1425, 2001.

[222] Gavrill Tsechpenakis, Konstantinos Rapantzikos, Nicolas Tsapatsoulis, and Stefanos D. Kollias. A snake model for object tracking in natural sequences. *Sig. Proc.: Image Comm.*, 19(3):219–238, 2004.

[223] G Turin. An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329, 1960.

[224] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *NeuroImage*, 15(1):273–289, 2002.

[225] Jayaram K. Udupa, Vicki R. LeBlanc, Ying Zhuge, Celina Imielinska, Hilary Schmidt, Leanne M. Currie, Bruce Elliot Hirsch, and James Woodburn. A framework for evaluating image segmentation algorithms. *Comp. Med. Imag. and Graph.*, 30(2):75–87, 2006.

[226] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, November 1999. IEEE Signal Processing Society's 2000 magazine award.

[227] Michael Unser, Akram Aldroubi, and Murray Eden. B-spline signal processing. i. theory. *IEEE Transactions on Signal Processing*, 41(2):821–833, 1993.

[228] Michael Unser, Akram Aldroubi, and Murray Eden. B-spline signal processing. ii. efficiency design and applications. *IEEE Transactions on Signal Processing*, 41(2):834–848, 1993.

[229] E. M. van Rikxoort, Y. Arzhaeva, and B. van Ginneken. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. In T. Heimann, M. Styner, and B. van Ginneken, editors, *3D Segmentation In The Clinic: A Grand Challenge*, pages 101–108, 2007.

[230] Eva M. van Rikxoort, Ivana Isgum, Yulia Arzhaeva, Marius Staring, Stefan Klein, Max A. Viergever, Josien P. W. Pluim, and Bram van Ginneken. Adaptive local multi-atlas segmentation: application to the heart and the caudate nucleus. *Med Image Anal*, 14(1):39–49, 2010.

[231] P. Viola. *Alignment by maximization of mutual information*. PhD thesis, 1995. Co-Supervisors: Lozano-Perez, Tomas and Atkeson, Christopher G.

[232] P. Viola and W. M. Wells. Alignment by maximization of mutual information. volume 24, pages 137–154, 1997.

[233] Yong Wang and Bao-Gang Hu. Derivations of normalized mutual information in binary classifications. *Fuzzy Systems and Knowledge Discovery, Fourth International Conference on*, 1:155–163, 2009.

[234] S. K. Warfield, K. H. Zou, and W. M. Wells III. Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. 23(7):903–921, 07 2004.

[235] Simon K. Warfield, Jan Rexilius, Petra S. Huppi, Terrie E. Inder, Erik G. Miller, William M. Wells III, Gary P. Zientara, Ferenc A. Jolesz, and Ron Kikinis. A binary entropy measure to assess nonrigid registration algorithms. In *MICCAI '01: Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 266–274, London, UK, 2001. Springer-Verlag.

[236] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, pages 35–51, 1996.

[237] Bernard Widrow. The "rubber-mask" technique - i. pattern measurement and analysis. *Pattern Recognition*, 5(3):175–176, 1973.

[238] H. S. Wilf. *Generatingfunctionology*. Ak Peters Series. Academic Press, Inc., 2 edition, 1994.

[239] O. Wirjadi. Survey of 3d image segmentation methods. Technical report, 2007.

[240] Minjie Wu, Caterina Rosano, Pilar Lopez-Garcia, Cameron S. Carter, and Howard J. Aizenstein. Optimum template selection for atlas-based segmentation. *NeuroImage*, 34(4):1612–8, 2007.

[241] Yuchen Xie, Jeffrey Ho, and Baba C. Vemuri. Image atlas construction via intrinsic averaging on the manifold of images. In *CVPR*, pages 2933–2939, 2010.

[242] Chenyang Xu, Dzung L. Pham, and Jerry L. Prince. *Handbook of Medical Imaging*, volume 2. Medical Image Processing and Analysis, chapter Image Segmentation Using Deformable Models, pages 447–514. SPIE Press, 2000.

[243] Hongkui Xu, Mingyan Jiang, and Mingqiang Yang. A new landmark selection method for non-rigid registration of medical brain images. In *ICSP2010 Proceedings*, 2010.

[244] Hongshi Yan and Jian Guo Liu. Robust phase correlation based motion estimation and its applications. In *British Machine Vision Conference*, 2008.

[245] Jinzhong Yang, Yongbin Zhang, Lifei Zhang, and Lei Dong. Automatic segmentation of parotids from ct scans using multiple atlases. In Bram van Ginneken, Keelin Murphy, Tobias Heimann, Vladimir Pekar, and Xiang Deng, editors, *Proc. MICCAI Workshop Medical Image Analysis for the Clinic: A Grand Challenge*, pages 323–330, 2010.

[246] John Zachary and S. S. Iyengar. Informaton theoretic similarity measures for content based image retrieval. *J. Am. Soc. Inf. Sci. Technol.*, 52:856–857, August 2001.

[247] Hrvatski zavod za javno zdravstvo. Kardiovaskularne bolesti u republici hrvatskoj, 2004.

[248] Hui Zhang, Jason E. Fritts, and Sally A. Goldman. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260–280, 2008.

[249] Ying Zheng, Guangyao Li, Xiehua Sun, and Xinmin Zhou. A geometric active contour model without re-initialization for color images. *Image Vision Comput.*, 27(9):1411–1417, 2009.

[250] Yu Zhong, Anil K. Jain, and Marie-Pierre Dubuisson-Jolly. Object tracking using deformable templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(5):544–549, 2000.

[251] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.

[252] Lilla Zöllei, Erik G. Learned-Miller, W. Eric L. Grimson, and William M. Wells III. Efficient population registration of 3d data. In *CVBIA*, pages 291–301, 2005.

[253] Lilla Zöllei, Martha Shenton, William Wells, and Kilian Pohl. The impact of atlas formation methods on atlas-guided brain segmentation. *Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation at the 10th International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 39–46, 2007.

[254] Marco Zuliani. Ransac for dummies, 2012.

*Nay, we have not lived in vain. Have they not built towers of our bones?*

*Khalil Gibran*
*Sand and Foam*

*Prah po našim koracima,*
*prah slavnih nam pređah to je,*
*prah Zrinjskoga i Hrvoje*
*i mnogijeh sličnih njima!*
*Bud'mo vredni čestitosti*
*U njem shranit svoje kosti!!*

*Petar Preradović*
*Naša zemlja*

# Appendix A

# A1

## A.1 Relationship between AJM and MI

By definition, MI is:

$$MI(X,Y) = -H(X,Y) + H(X) + H(Y) \tag{A.1}$$

Negentropy [85] or (non-) Gaussianity [22] is defined as an entropy difference:

$$J(\mathbf{X}) = H(\mathbf{X}^G) - H(\mathbf{X}) \tag{A.2}$$

where $\mathbf{X}^G$ stands for the Gaussian (multivariate) random variable of the same covariance matrix as $\mathbf{X}$. In case of two-dimensional random variable (our case), we can write:

$$J(X,Y) = H(X^G, Y^G) - H(X,Y) \tag{A.3}$$

Using the definition of entropy we can express this as:

$$J(X,Y) = -\int p_{XY}^G \ln p_{XY}^G + \int p_{XY} \ln p_{XY} \tag{A.4}$$

The $MI$ can be expressed via negentropy in the form:

$$MI(X,Y) = J(X,Y) - J(X) - J(Y) - \frac{1}{2} ln \frac{E[XX]E[YY]}{\begin{vmatrix} E[XX] & E[XY] \\ E[YX] & E[YY] \end{vmatrix}} \tag{A.5}$$

This is just a simplification of the equations given in [36] and [86] (e.q. 2.4 and 23, respectively). If $X$ and $Y$ are uncorrelated [86] or statistically independent centralized variables [36] the last term vanishes, therefore, we write:

$$MI(X,Y) = J(X,Y) - J(X) - J(Y) \tag{A.6}$$

If $p_{XY}^G \approx p_{XY}$ (assumption similar to one done in [22] eq.42) we can write:

$$J(X,Y) = -\int p_{XY}^G \ln \frac{p_{XY}^G}{p_{XY}} dxdy \tag{A.7}$$

$$J(X) = -\int p_X^G \ln \frac{p_X^G}{p_X} dx \tag{A.8}$$

If we denote Gaussian distribution $p_{XY}^G$ (or $p_X^G$) with $n$ and $p_{XY}$ (or $p_X$) with $p$, we can notice that the negentropy is expressed in the form of relative entropy (KLD):

$$D_{KL}(n||p) = -\int n \cdot \ln \frac{n}{p} dx \tag{A.9}$$

If we expand the the natural logarithm in Taylor series KLD can be written as:

$$D_{KL}(n||p) = -\int n[\frac{p}{n} - 1 - \frac{(\frac{p}{n}-1)^2}{2} + \ldots]dx \tag{A.10}$$

$$\approx \cancelto{0}{-\int n(\frac{p}{n} - 1)dx} + \frac{1}{2}\int n(\frac{p}{n} - 1)^2 dx \tag{A.11}$$

$$= \frac{1}{2}\int n(\frac{p}{n} - 1)^2 dx = \widehat{D_{KL}}(n||p) \tag{A.12}$$

The step from Equation A.10 to A.11 simply truncates the series and step from Equation A.11 to A.12 is valid since the first summand in the Equation A.11 vanishes by the definition of the PDF.

A (multivariate) Gram-Charlier expansion of $p(\mathbf{x})$ around a reference distribution $n(\mathbf{x})$ is an expansion of $p(\mathbf{x})$ in the form:

$$\frac{p}{n} - 1 = \sum_{i}^{2} \eta_i h^i(\mathbf{x}) + \frac{1}{2!}\sum_{ij}^{2} \eta_{ij} h^{ij}(\mathbf{x}) + \frac{1}{3!}\sum_{ijk}^{2} \eta_{ijk} h^{ijk}(\mathbf{x}) + \ldots \tag{A.13}$$

where $\sum_{ij\ldots k}^2 = \sum_i^2 \sum_j^2 \cdots \sum_k^2$ and the $\eta$'s are coefficients and the $h$'s are fixed functions of $\mathbf{x}$ which depend on $n(\mathbf{x})$ (see [22], eq. 43). Since $n$ is Gaussian distribution, $h$'s are Hermite polynomials, and $\eta$'s are cumulants, which gives us cumulant approximation of the KLD [22]:

$$\begin{aligned}\widehat{D_{KL}}(n||p) = \frac{1}{2}\Bigg(&\sum_i (\kappa_i^p - \kappa_i^n)^2 + \frac{1}{2!}\sum_{ij}(\kappa_{ij}^p - \kappa_{ij}^n)^2 \\ &+ \frac{1}{3!}\sum_{ijk}(\kappa_{ijk}^p - \kappa_{ijk}^n)^2 + \frac{1}{4!}\sum_{ijkl}(\kappa_{ijkl}^p - \kappa_{ijkl}^n)^2 \ldots\Bigg)\end{aligned} \tag{A.14}$$

where $\kappa^n$ and $\kappa^p$ denote cumulants of the distribution $n$ and $p$. If $n$ is Gaussian distribution centralized around mean, we have $\kappa_i^n = 0$, $\kappa_{ij}^n = \delta_{ij}$ and all higher-order cumulants vanish.

Each summation in the Gram-Charlier expansion (Equation A.13) sums across all dimensions of $\mathbf{X}$ which will be only one ($X$) or only two dimensions ($X$ and $Y$).

In the case of two-dimensional random vector $\mathbf{X}$ each summation from the Equation A.13 expands to:

$$\sum_{i}^{2} \kappa_i^2 = \kappa_1 + \kappa_2 \tag{A.15}$$

$$\sum_{ij}^{2} \kappa_{ij}^2 = \kappa_{11}^2 + \kappa_{12}^2 + \kappa_{21}^2 + \kappa_{22}^2 \tag{A.16}$$

$$\sum_{ijk}^{2} \kappa_{ijk}^2 = \kappa_{111}^2 + \kappa_{112}^2 + \kappa_{121}^2 + \kappa_{211}^2 + \kappa_{122}^2 + \kappa_{212}^2 + \kappa_{221}^2 + \kappa_{222}^2 \tag{A.17}$$

$$\begin{aligned}\sum_{ijkl}^{2} \kappa_{ijkl}^2 = {} & \kappa_{1111}^2 + \kappa_{1112}^2 + \kappa_{1121}^2 + \kappa_{1211}^2 + \kappa_{1122}^2 + \kappa_{1212}^2 + \kappa_{1221}^2 + \kappa_{1222}^2 \\ & + \kappa_{2111}^2 + \kappa_{2112}^2 + \kappa_{2121}^2 + \kappa_{2211}^2 + \kappa_{2122}^2 + \kappa_{2212}^2 + \kappa_{2221}^2 + \kappa_{2222}^2\end{aligned} \tag{A.18}$$

Now, we write negentropy via cumulant approximation of the relative entropy:

$$\begin{aligned} & J(X) \approx \widehat{D_{KL}}(n||p_X) = \\ & = \frac{1}{2}\left(\kappa_X^2 + \frac{1}{2!}(\kappa_{XX} - \delta_{XX})^2 + \frac{1}{3!}\kappa_{XXX}^2 + \frac{1}{4!}\kappa_{XXXX}^2 + \dots\right)\end{aligned} \tag{A.19}$$

$$\begin{aligned} & J(X,Y) \approx \widehat{D_{KL}}(n||p_{XY}) = \frac{1}{2}\left(\kappa_X^2 + \kappa_Y^2 + \right. \\ & + \frac{1}{2!}[(\kappa_{XX} - \delta_{XX})^2 + (\kappa_{XY} - \delta_{XY})^2 + (\kappa_{YX} - \delta_{YX})^2 + (\kappa_{YY} - \delta_{YY})^2] + \\ & + \frac{1}{3!}[\kappa_{XXX}^2 + \kappa_{XXY}^2 + \kappa_{XYX}^2 + \kappa_{YXX}^2 + \kappa_{XYY}^2 + \kappa_{YXY}^2 + \kappa_{YYX}^2 + \kappa_{YYY}^2] + \\ & \frac{1}{4!}[\kappa_{XXXX}^2 + \kappa_{XXXY}^2 + \kappa_{XXYX}^2 + \kappa_{XYXX}^2 + \kappa_{YXXX}^2 + \\ & + \kappa_{XXYY}^2 + \kappa_{XYXY}^2 + \kappa_{YXXY}^2 + \kappa_{XYYX}^2 + \kappa_{YXYX}^2 + \kappa_{YYXX}^2 + \\ & \left. + \kappa_{XYYY}^2 + \kappa_{YXYY}^2 + \kappa_{YYXY}^2 + \kappa_{YYYX}^2 + \kappa_{YYYY}^2] + \dots\right)\end{aligned} \tag{A.20}$$

Here, the dimension is notated with either $X$ or $Y$ rather than 1 or 2 and for simplification superscript denoting distribution is dropped. From here, we can express $MI$ as:

$$
\begin{aligned}
& MI(X,Y) = J(X,Y) - J(X) - J(Y) \approx \frac{1}{2}\Big(\kappa_X^2 + \kappa_Y^2 + \\
& + \frac{1}{2!}[(\kappa_{XX} - \delta_{XX})^2 + (\kappa_{XY} - \delta_{XY})^2 + (\kappa_{YX} - \delta_{YX})^2 + (\kappa_{YY} - \delta_{YY})^2] + \\
& + \frac{1}{3!}[\kappa_{XXX}^2 + \kappa_{XXY}^2 + \kappa_{XYX}^2 + \kappa_{YXX}^2 + \kappa_{XYY}^2 + \kappa_{YXY}^2 + \kappa_{YYX}^2 + \kappa_{YYY}^2] + \\
& \frac{1}{4!}[\kappa_{XXXX}^2 + \kappa_{XXXY}^2 + \kappa_{XXYX}^2 + \kappa_{XYXX}^2 + \kappa_{YXXX}^2 + \\
& + \kappa_{XXYY}^2 + \kappa_{XYXY}^2 + \kappa_{YXXY}^2 + \kappa_{XYYX}^2 + \kappa_{YXYX}^2 + \kappa_{YYXX}^2 + \\
& + \kappa_{XYYY}^2 + \kappa_{YXYY}^2 + \kappa_{YYXY}^2 + \kappa_{YYYX}^2 + \kappa_{YYYY}^2] + \dots \\
& -(\kappa_X^2 + \frac{1}{2!}(\kappa_{XX} - \delta_{XX})^2 + \frac{1}{3!}\kappa_{XXX}^2 + \frac{1}{4!}\kappa_{XXXX}^2 + \dots) \\
& -(\kappa_Y^2 + \frac{1}{2!}(\kappa_{YY} - \delta_{YY})^2 + \frac{1}{3!}\kappa_{YYY}^2 + \frac{1}{4!}\kappa_{YYYY}^2 + \dots)\Big) \qquad \text{(A.21)}
\end{aligned}
$$

where $\delta_{ij}$ is Kronecker delta function, i.e. $\delta_{ij} = 0 \forall i \neq j$.

Now, the question is can we simplify this further, by canceling the factors of the negentropies $J(X)$ and $J(Y)$ with autocumulant factors from the negentropy $J(X,Y)$. The actual question is whether the following identity holds:

$$\kappa_{ij\dots k}^{p_X} = \kappa_{ij\dots k}^{p_{XY}} \, \forall i = j = \dots = k \qquad \text{(A.22)}$$

Since cumulants can be expressed in terms of moments (see Equation A.28), the Identity A.22 holds, due to the fact that:

$$\int\int p_{XY} X dx dy = \int p_X X dx \qquad \text{(A.23)}$$

Therefore, we cancel the factors in Equation A.21 to get the expression:

$$
\begin{aligned}
& MI(X,Y) = J(X,Y) - J(X) - J(Y) \approx \frac{1}{2}\Big( \\
& + \frac{1}{2!}[(\kappa_{XY} - \delta_{XY})^2 + (\kappa_{YX} - \delta_{YX})^2] + \\
& + \frac{1}{3!}[\kappa_{XXY}^2 + \kappa_{XYX}^2 + \kappa_{YXX}^2 + \kappa_{XYY}^2 + \kappa_{YXY}^2 + \kappa_{YYX}^2] + \\
& + \frac{1}{4!}[\kappa_{XXXY}^2 + \kappa_{XXYX}^2 + \kappa_{XYXX}^2 + \kappa_{YXXX}^2 + \kappa_{XXYY}^2 + \kappa_{XYXY}^2 + \kappa_{YXXY}^2 + \\
& + \kappa_{XYYX}^2 + \kappa_{YXYX}^2 + \kappa_{YYXX}^2 + \kappa_{XYYY}^2 + \kappa_{YXYY}^2 + \kappa_{YYXY}^2 + \kappa_{YYYX}^2] + \dots\Big) \qquad \text{(A.24)}
\end{aligned}
$$

Each cumulant can be expressed in terms of centralized moments (which can be derived from their generating functions, see e.g. [105] p.62 for univariate case, or [142] p.60 for multivariate case):

$$
\begin{aligned}
\kappa_i^2 &= \mu_i'^2 = 0 && \text{(A.25)}\\
\kappa_{ij}^2 &= (\mu'_{ij} - \mu'_i\mu'_j)^2 \\
&= \mu_{ij}'^2 && \text{(A.26)}\\
\kappa_{ijk}^2 &= (\mu'_{ijk} - \mu'_{ij}\mu'_k[3] + 2\mu'_i\mu'_j\mu'_k)^2 \\
&= \mu_{ijk}'^2 && \text{(A.27)}\\
\kappa_{ijkl}^2 &= (\mu'_{ijkl} - \mu'_{ijk}\mu'_l[4] - \mu'_{ij}\mu'_{kl}[3] + 2\mu'_i\mu'_j\mu'_{kl}[6] - 6\mu'_i\mu'_j\mu'_k\mu'_l)^2 \\
&= (\mu'_{ijkl} - \mu'_{ij}\mu'_{kl}[3])^2 \\
&= (\mu'_{ijkl} - \mu'_{ij}\mu'_{kl} - \mu'_{ik}\mu'_{jl} - \mu'_{il}\mu'_{jk})^2 && \text{(A.28)}
\end{aligned}
$$

Here, bracket notation (see [142]) is used to denote all indices permutation (e.g. $\mu'_{ij}\mu'_{kl}[3] = \mu'_{ij}\mu'_{kl} + \mu'_{ik}\mu'_{jl} + \mu'_{il}\mu'_{jk}$).

$\mu_{ijk}$ is invariant to index perimutation, due to the product associativity property, i.e. $\mu'_{XXY} = E[X'X'Y'] = E[X'Y'Y'] = E[Y'X'X'] = \mu'_{YXX}$ (here prime denotes centralized variable $X$ or $Y$). This leads to following simplifications:

$$
\begin{aligned}
\kappa_{XY} &= \mu'_{XY} = \kappa_{XY} && \text{(A.29)}\\
\kappa_{XXY} &= \mu'_{XXY} = \kappa_{XYX} = \kappa_{YXX} && \text{(A.30)}\\
\kappa_{YYX} &= \mu'_{YYX} = \kappa_{YXY} = \kappa_{XYY} && \text{(A.31)}\\
\kappa_{XXXY} &= \mu'_{XXXY} - 3\mu'_{XX}\mu'_{XY} = \kappa_{XXYX} = \kappa_{XYXX} = \kappa_{YXXX} && \text{(A.32)}\\
\kappa_{XXYY} &= \mu'_{YYXX} - \mu'_{XX}\mu'_{YY} - 2\mu'_{XY}\mu'_{XY} = \\
&\quad = \kappa_{YYXX} = \kappa_{YXYX} = \kappa_{XYXY} = \kappa_{YXXY} = \kappa_{XYYX} && \text{(A.33)}\\
\kappa_{YYYX} &= \mu'_{YYYX} - 3\mu'_{XY}\mu'_{YY} = \kappa_{YYXY} = \kappa_{YXYY} = \kappa_{XYYY} && \text{(A.34)}
\end{aligned}
$$

Combining Equations A.29 to A.34 and A.24 and truncating series after forth order cumulant, gives following $MI$ approximation:

$$
\begin{aligned}
&\widehat{MI}(X,Y) \approx \frac{1}{2}\left(\frac{1}{2!}2\mu_{XY}'^2 + \frac{1}{3!}[3\mu_{XXY}'^2 + 3\mu_{YYX}'^2] + \right.\\
&+\frac{1}{4!}[(\mu'_{XXXY} - 3\mu'_{XX}\mu'_{XY})^2 + (\mu'_{YYXX} - \mu'_{XX}\mu'_{YY} - 2\mu_{XY}'^2)^2 \\
&+(\mu'_{YYYX} - 3\mu'_{XY}\mu'_{YY})^2]) && \text{(A.35)}\\
&= \frac{1}{2}\left(\mu_{XY}'^2 + \frac{1}{2}[\mu_{XXY}'^2 + \mu_{YYX}'^2] + \right.\\
&+\frac{1}{6}(\mu_{XXXY}'^2 - 6\mu'_{XXXY}\mu'_{XX}\mu'_{XY} + 9\mu_{XX}'^2\mu_{XY}'^2) + \\
&+\frac{1}{4}(\mu_{YYXX}'^2 + \mu_{XX}'^2\mu_{YY}'^2 + 4\mu_{XY}'^4 - 2\mu'_{YYXX}\mu'_{XX}\mu'_{YY} - \\
&-4\mu'_{YYXX}\mu_{XY}'^2 + 4\mu'_{XX}\mu'_{YY}\mu_{XY}'^2) + \\
&\left.+\frac{1}{6}(\mu_{YYYX}'^2 - 6\mu'_{YYYX}\mu'_{YY}\mu'_{YX} + 9\mu_{YY}'^2\mu_{YX}'^2)\right) && \text{(A.36)}
\end{aligned}
$$

This $MI$ approximation is to some extent different then the one given in [22]. The main difference is that $MI$ approximation given in Equation A.36 does not contain autocumulants. This has an intuitive explanation. Only joint cumulants measure mutual dependence between two variables, while autocumulants are simply an offset (or even noise) added to that measure.

$AJM$ proposed in our paper is defined as:

$$AJM(X,Y) = |\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{1}{n!}\frac{1}{m!}E[(X')^n(Y')^m]| = \tag{A.37}$$

$$= |\mu'_{XY} + \frac{1}{2!}\mu'_{XXY} + \frac{1}{2!}\mu'_{XYY} + \frac{1}{3!}\mu'_{XXXY} + \frac{1}{2!2!}\mu'_{XXYY} + \frac{1}{3!}\mu'_{XYYY} + \ldots| \tag{A.38}$$

We could also propose a modified $AJM$ in the following way:

$$AJM^*(X,Y) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{1}{n!}\frac{1}{m!}|E[(X')^n(Y')^m]| = \tag{A.39}$$

$$\begin{aligned} &= |\mu'_{XY}| + |\frac{1}{2!}\mu'_{XXY}| + |\frac{1}{2!}\mu'_{XYY}| + |\frac{1}{3!}\mu'_{XXXY}| + \\ &+ |\frac{1}{2!2!}\mu'_{XXYY}| + |\frac{1}{3!}\mu'_{XYYY}| + \ldots \end{aligned} \tag{A.40}$$

Due to the triangle property of the Euclidean norm, the following identity holds:

$$AJM \leq AJM^* \tag{A.41}$$

A modification of the $\widehat{MI}$ can be given in the form:

$$\begin{aligned} \widehat{MI}^*(X,Y) \approx \frac{1}{2}\Big(&\mu'^2_{XY} + \frac{1}{2}\mu'^2_{XXY} + \frac{1}{2}\mu'^2_{YYX} + \\ &+ \frac{1}{6}(\mu'_{XXXY})^2 + \frac{1}{4}(\mu'_{YYXX})^2 + \frac{1}{6}(\mu'_{YYYX})^2\Big) \end{aligned} \tag{A.42}$$

Notice that $\widehat{MI}^* \geq \widehat{MI}$ iff the following conditions are met:

$$6\mu'_{XXXY}\mu'_{XX}\mu'_{XY} \geq 9\mu'^2_{XX}\mu'^2_{XY} \tag{A.43}$$

$$6\mu'_{YYYX}\mu'_{YY}\mu'_{YX} \geq 9\mu'^2_{YY}\mu'^2_{YX} \tag{A.44}$$

$$2\mu'_{YYXX}(\mu'_{XX}\mu'_{YY} + 2\mu'^2_{XY}) \geq (\mu'_{XX}\mu'_{YY} + 2\mu'^2_{XY})^2 \tag{A.45}$$

We can cancel some terms in condition A.43 (and similarly in A.44). Since $\mu_{XX} > 0$ we can cancel this term without change in the inequality. Regarding $\mu_{XY}$ there are two cases:

(a) $X, Y > 0 \Rightarrow \mu_{XY} > 0$ (and $\mu_{XXXY} > 0$)

(b) $X, Y < 0 \Rightarrow \mu_{XY} < 0$ (and $\mu_{XXXY} < 0$)

Leading to two inequities:

(a) $6\mu'_{XXXY} > 9\mu'_{XX}\mu'_{XY}$

(b) $6\mu'_{XXXY} < 9\mu'_{XX}\mu'_{XY}$

Since in the first case both $\mu_{XXXY}$ and $\mu_{XY}$ are positive numbers, and in the second case both of them are negative numbers, this reduces to only one condition:

$$|6\mu'_{XXXY}| > |9\mu'_{XX}\mu'_{XY}| \tag{A.46}$$

Similarly we get:

$$|6\mu'_{YYYX}| > |9\mu'_{YY}\mu'_{YX}| \tag{A.47}$$

Since $\mu'_{XX}\mu'_{YY} + 2\mu'^2_{XY} > 0$, the condition A.45 reduces down to:

$$2\mu'_{YYXX} > (\mu'_{XX}\mu'_{YY} + 2\mu'^2_{XY}) \tag{A.48}$$

Inequalities A.46, A.47, and A.48 are satisfied if $X$ and $Y$ are non-Gaussian random variables. If this is the case $\widehat{MI}^*$ will be the upper bound of the $\widehat{MI}$ and will have the same quantitative behavior as $AJM^*$ (one utilizes squared function, while other utilizes absolute value), which can also be interpreted as the upper bound of the proposed $AJM$.

This throws some lights on the behavior of the $AJM$ since to some extent $AJM$ could be interpreted as a rough approximation of the $MI$. The overall performance of $AJM$ situated between $CC$ and various instances of $MI$ algorithm is expected according to our previous discussion.

Interestingly, when the results of $\widehat{MI}$ are compared to the results of other ISMs, we can see that $\widehat{MI}$ behaves similar to $MI_{256}$ for translation and scaling, but significantly worse for rotation. Or in other words $AJM$ shows better results than $\widehat{MI}$ (see Table A.1).

Table A.1: The average absolute error for translation, rotation and scaling.

| | $\overline{\epsilon_T}$ | $\overline{\epsilon_R}$ | $\overline{\epsilon_S}$ |
|---|---|---|---|
| $CC$ | 51.27 | 72.17 | 0.49 |
| $AJM$ | 19.57 | 31.04 | 0.24 |
| $MI_8$ | 15.48 | 23.70 | 0.15 |
| $MI_{256}$ | 33.01 | 37.73 | 0.45 |
| $\widehat{MI}$ | 46.99 | 89.23 | 0.46 |

## A.2 AJM Properties

In this section we show some examples which give insights into some AJM properties. Some properties are well known to be relevant for any similarity measure, and were already presented in the literature. For example, Studholme et al. ([205], Fig. 6.) designed a test to show that a similarity measure is overlap invariant. Additionally, a similarity measure is more reliable if it is invariant to image interpolation. This was pointed out as one of the disadvantages of MI, by Roche et al. [171], while Pluim et al. [162] investigated optimal interpolation techniques to cope with this problem. Finally, approximation of PDF through histogram may cause that the performance of the MI-based image SM be deteriorated, depending on the selected number of histogram bins or the amount of noise. Therefore, in the rest of the Section we will investigate whether AJM is:

- Overlap invariant
- Sensitive to number of bins
- Sensitive to interpolation

and compare the results to the results of the CC and MI.

**The Studholme test** [205] has been used to evaluate the overall invariance property. In the test, an image from the medical image set is registered onto itself rather than onto a model (as done by Studholme et al.). The results of the Studholme test for one of the images from the medical data set is given in Figure A.1. From there it can be noticed that, at some point, AJM is sensitive towards the change of the field of view (FOV).

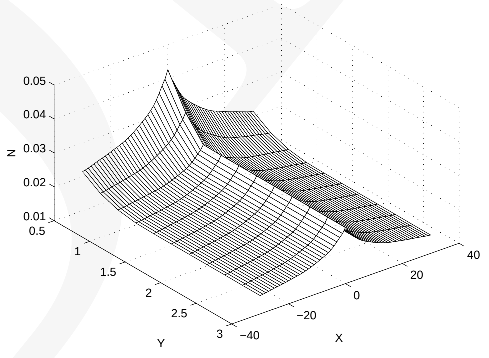


Figure A.1: Studholme test for *AJM*. Notice the declination of maximum with the increase of FOV (as defined in [205]). x-axis shows rotation from -30 to +30 degrees, and y-axis shows the change of FOV from 0.5 to 3.

**Interpolation artifacts** may greatly affect the performance of MI, as already shown in [162] and [171], but this is not the case with CC. In order to demonstrate this a following experiment was constructed. The test image shown in Figure A.2 (left) is rotated around center and interpolated with three different techniques (nearest neighbor, bilinear, and bicubic), followed by registration onto its original version.

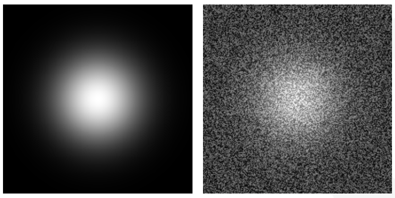
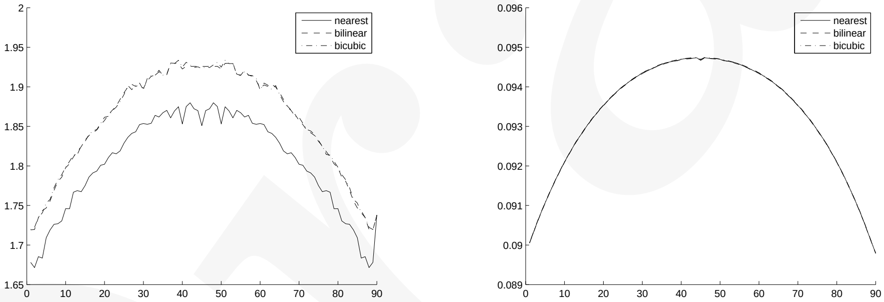


Figure A.2: The left image is used to produce results depicted in Figures A.3, A.4, while both images were used to produce the result in Figure A.5.

The results shown in Figure A.3 demonstrate that $MI_8$ (index is explained later) is affected by the type of interpolation used, while AJM is not. So in this case we can say that AJM inherited property of the CC, rather than MI.

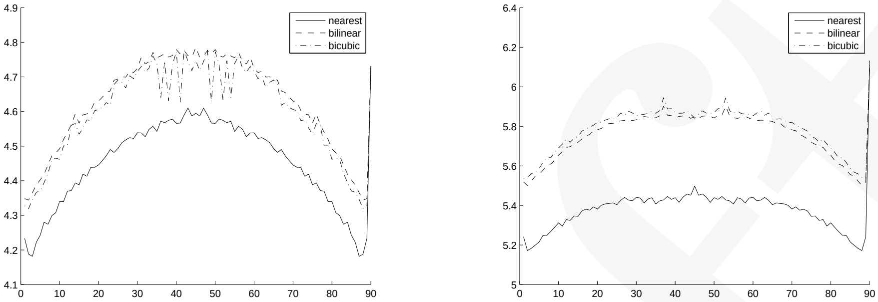


(a) $MI_8$ graphs almost overlap for bilinear and bicubic interpolation, and significant difference is visible when nearest neighbor interpolation is used.

(b) AJM graphs show the same performance for all types of interpolation.

Figure A.3: The graph shows interpolation artifact for $MI_8$ (a) and AJM (b). y-axis represents MI and AJM, while x-axis represents angle $\alpha$ in range [1, 90] degrees. Notice how MI is affected by the type of interpolation used, while AJM is not.

**Sensitivity to the number of bins** (and bin width) is expected, whenever a histogram is used for PDF approximation. To demonstrate this sensitivity of the MI we repeated the previous experiment with different numbers of bin used for MI calculation. The number of bins is used as index to differentiate one implementation from another, therefore we have: $MI_8$, $MI_{100}$, and $MI_{256}$. How the number of bins affects the MI is demonstrated in Figure A.4. The results are again given for a rotation of the image A.2 (left) by angle $\alpha$ in range 1 to 90 degrees.



(a) All three interpolation techniques gave different $MI_{100}$ graph. Bicubic interpolation gives rather noisy MI graph.

(b) $MI_{256}$ also behaves differently for different interpolation selected. Bicubic and bilinear interpolation are relatively close to one another.

Figure A.4: Performance of MI with respect to variable number of histogram bins and interpolation used. (a) $MI_{100}$ - 100 bins, (b) $MI_{256}$ - 256 bins. Compare results with Figure A.3a, which is given for 8 bins. It can be observed that MI behaves differently, both qualitative and quantitative, for different number of bins.

The results from this experiments demonstrate how MI is affected by the number of bins selection. In the same time they fortify the conclusion from the previous experiment, since the type of interpolation used leads to different MI graph. Both of this conclusions can be observed as an advantage of AJM over MI.

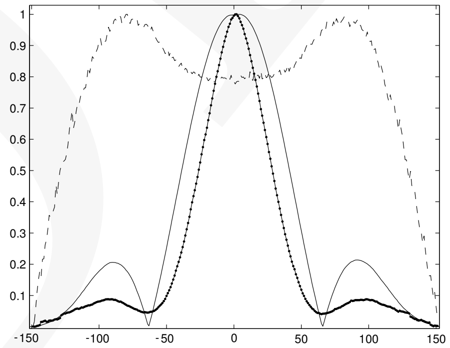


Figure A.5: Performance of AJM (solid line), $MI_8$ (dot-dashed line), and $MI_{256}$ (dashed line) for translation from -150 to 150 by 1 pixel.

**Alignment** between two images may be heavily affected by this type of problem, as we will demonstrate in the following example. A similarity between images from the Figure A.2 is calculated for a shift of one pixel in the interval

[-150, 150]. The similarity was defined as $AJM$, $MI_8$, or $MI_{256}$. All three measures are scaled to codomain [0,1] and plotted on the same graph (see Figure A.5) for easier comparison.

**As an additional motivation** for the use of the AJM as an image SM the following examples are given. First synthetic example is motivated by the example from [171]. For the two images from the Figure A.6 we try to determine the correct alignment between them by scaling one of the images and calculating the similarity. Notice that both of the images differ only by the amount and type of noise added to them, therefore, the correct alignment should be for scaling factor of one. To ease the comparison between similarity measures the graphs for all three similarity measures are scaled to the interval $[0,1]$. Notice how the MI and CC image SM are unable to find the location of the correct image alignment, while AJM is. This example is also a good illustration of the problems reported in the earlier paragraph.

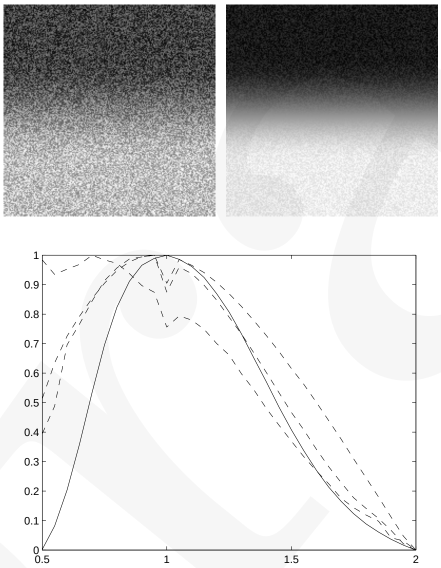


Figure A.6: Two synthetic images are aligned using scaling only on interval [0.5, 2] with step 0.05. The graphs below show the $CC$, $MI_8$, and $MI_{256}$ as dashed lines and $AJM$ as solid line. All image SMs are scaled to codomain [0,1] for easier comparison. The x-axes show the scaling factor. The correct alignment of images unity scaling.

To demonstrate that AJM is capable to find the correct alignment for a complex relationship between images, another simulation is done. The similarity between images in Figure A.7 is calculated for a shift of one pixel in the interval [-150, 150]. As the images differ only by different color mapping the correct alignment should be for zero translation. Since this example may be observed as a simulation of multimodality registration problem it is clear that MI performs well in this case. For this reason, and for clarity of presentation, MI graph is

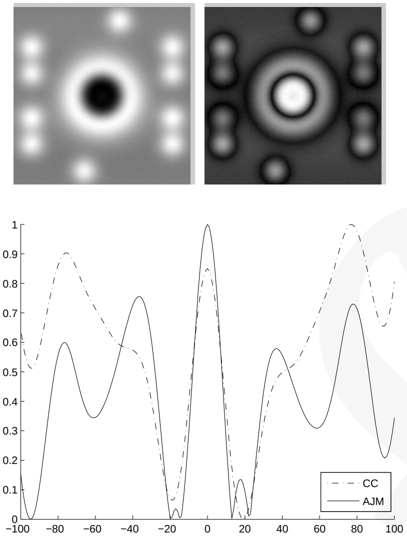


Figure A.7: Two synthetic images are registered using translation only. For clarity, only CC and AJM graphs are shown. The x-axes show the amount of translation in pixels. The correct alignment of images is for zero translation.

omitted in Figure A.7.

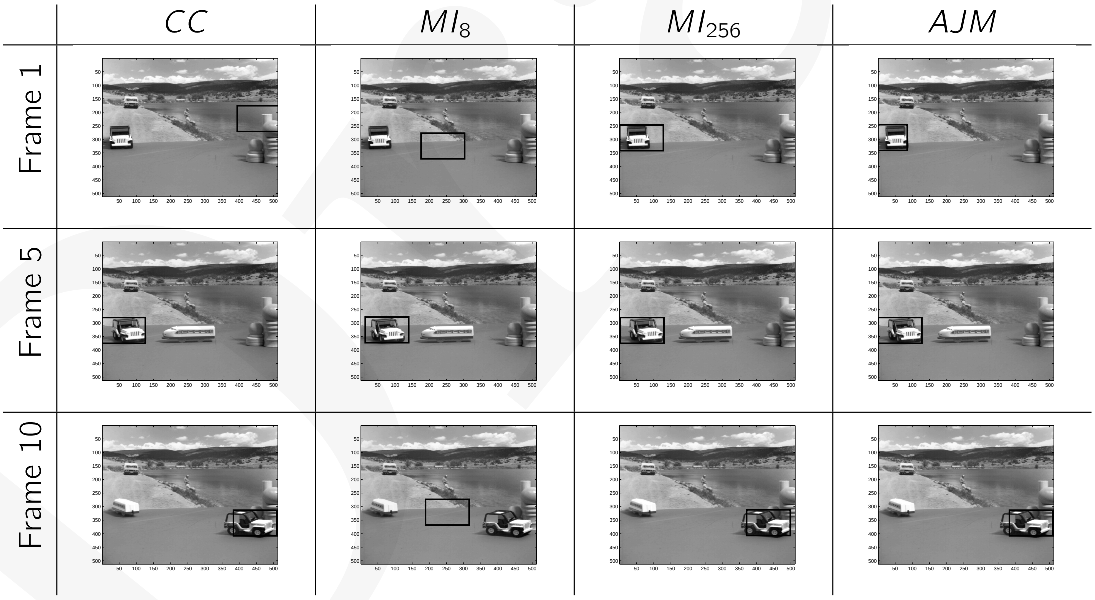


Figure A.8: Figure shows object detection using three different image similarity measures (CC, $MI_8$, $MI_{256}$ and AJM), for the first (first row), fifth (second row) and final image from the sequence (third row).

As a final example, a well-known and publicly available sequence[1] is used. The performance of the image similarity measures (CC, MI, AJM) for

---

[1] Toy Car sequence from `http://sipi.usc.edu/database`

the purpose of object tracking within a sequence of images is demonstrated in Figure A.8. From the sixth image in the sequence, the image of an object is taken and used as a model for object detection in the image sequence. The AJM performs well in all image frames. For the sake of brevity only the first, fifth and the final image from the sequence are depicted in Figure A.8. The results demonstrate that AJM finds the correct object in all ten images of the image sequence, while CC finds the object correctly in five out of ten images, and $MI_8$ finds the object in just four images. The reason for such poor performance of MI should be attributed to the large homogeneous region which misguides the search for optimal alignment due to low joint entropy. However, MI performance can be significantly improved (9 out of 10 correct detections) if 256 bins are used instead of 8. However, this may deteriorate the MI performance in some other cases, since it increases sensitivity to noise and computation time (both of which will also be demonstrated in further sections).

## A.3 AJM existence conditions

There are two conditions for existence of the proposed AJM similarity measure: i) the existence of the expectation, and ii) the convergence of the sum in Equation 7.3. We first demonstrate that the first condition holds.

In practical digital image processing applications, we deal with images whose intensities are finite values and have a finite region of support over the set $S \subset \mathbb{R}^2$. Therefore, for any digital image $y$, the following holds:

$$\int_{\mathbb{R}^2} y \cdot f_y dy = \int_S y \cdot f_y dy \tag{A.49}$$

If we denote the supremum of the set $S$ as $y_s$ then we may write: $\int_S y \cdot f_y dx \leq y_s \cdot \int_S f_y dy$ and since, by definition: $\int_S f_y dy = 1$, we see that all functions have a bounded integral given by Equation A.49, i.e. their expectation is defined.

Now, we show that the second condition (convergence of the sum) holds. Since the expectation in the Equation 7.3 exists for all $m$ and $n$, the following statement holds:

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{\omega_n} \frac{1}{\omega_m} \cdot E[(T - \mu_T)^n (O - \mu_O)^m] \leq E_{max} \cdot \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{\omega_n} \frac{1}{\omega_m} \tag{A.50}$$

with $E_{max} = max(E[(T - \mu_T)^n (O - \mu_O)^m], \forall m, n)$. Since we may easily select $\omega_n$ and $\omega_m$ so that the right side of the Equation A.50 converges, by a comparison test we conclude that the left side converges as well. In other words, when all the moments exist (expectation is defined) the convergence of the sum (and therefore the existence of AJM) depends only on the selection of $\omega_n$ and $\omega_m$. The selection used for the proposed AJM ($\omega_n = n!$ and $\omega_m = m!$) is just one of many possibilities satisfying the condition for convergence of the sum.

Since $E$ is a linear operator, the equation may be rearranged in the form given by Equation 7.5. With the proper selection of $\omega_n$ and $\omega_m$, the sums from the Equation 7.3 will converge. With the particular selection of $\omega_n$ and $\omega_m$ proposed, the sums will converge to the result given by the Equation 7.5.

## A.4 Numerical computations consideration

Due to the exponential nature of the Equation 7.3 the range of the output value can be very wide. If we use 64-bit double precision floating point arithmetic we will avoid the overflow and underflow as long as AJM value is within the range $[2^{-1022}, 2^{1023}]$. This will be the case whenever the pixel values are approximately within the range [-354, 354], which will be true for all 8-bit images. However, when working with floating point arithmetic, the gap between two successive numbers ($\epsilon$) is not fixed. For example the machine epsilon ($\epsilon_m$) defined as the half of the distance between 1 and the next larger floating point number has the property: $1 \oplus \epsilon_0 = 1, \forall \epsilon_0 < \epsilon_m$ where symbol $\oplus$ denotes the machine adding. As the numbers increase, the gap between numbers gets larger and larger. During the expansion of the sum from the Equation 7.3 the machine may get the following result: $e^n \oplus 1 \ominus e^n = 0, \forall n > m$ where $m$ depends on the machine architecture (precision used) and $\ominus$ denotes the machine subtraction. These effects happen in all machine calculation concerning floating point numbers but they are more apparent as the range of the values (and also gaps between numbers) get larger, and due to the exponential property of the AJM measure, the measure will be more sensitive to this effects.

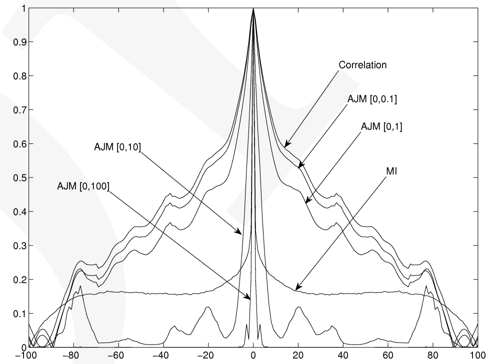


Figure A.9: Figure shows the behavior of AJM with respect to change of the input values range. The intervals on which the images pixel values are scaled are given in square brackets. For comparison correlation and $MI_{255}$ graphs are also shown.

The latter discussion is related to the input image pixel values, which ought to be roughly within the interval [-354, 354]. This can be easily achieved by mapping the image values to a subinterval of [-354, 354]. In our work we used images whose pixel values are scaled to the interval [0,1], however, we will briefly discuss how the range of the pixel values affect the behavior of the AJM. In the Figure A.9, an illustrative example of how the AJM behaves if the range of input values are increased, is given.

For the construction of the Figure A.9 an image is registered onto itself using translation. Registration is done on the same image, changing only the range of pixel values. For easier comparison all graphs are scaled to the interval [0,1] and correlation and $MI_{255}$ graphs are also plotted. Notice how the local maxima of the AJM get sharper and sharper as the range of input (and therefore output) values increase. For the smaller input values of the AJM gets really close to the correlation, which is due to the fact that for small values we may use the approximation $e^x \approx 1 + x$. As the range of the input values increases this approximation does not hold anymore since the higher order moments play a more important role.

# Acknowledgment

The work presented in this thesis was a part of research and development projects done at the University of Zagreb, in the first place: ”Intelligent methods for image processing and analysis” and ”An integrated, model based, approach for the quantification of cardiac function based on cardiac imaging”.

As a collaborative work many people were involved: Bart Bijnens from K.U. Lueven, Maja Čikeš, Davor Miličić and Ivo Čikeš from Medical School Zagreb, and Sven Lončarić from Faculty of Electircal Engineering and Computing. My thanks go to them.

Many thanks to Neven Elezović and Ivica Kopriva for their valuable advices and for their help with the literature.

Some of the results were calculated on the brain MR data using elastix. The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/. The elastix was initially developed at the Image Sciences Institute (ISI), University Medical Center Utrecht, The Netherlands, and can be found at http://elastix.isi.uu.nl/. My thanks go to them for providing the data and software.

Due to large computational complexity of the problem investigated in the thesis, many preliminary results and calculations were done on our cluster. My thanks go to my mentor professor Lončarić for providing the resources for constructiong the first and the biggest supercomputer at the University. Also, my thanks go to Kristina Wanous for her webpage www.debianclusters.org on which she provided many valuable ”hands on” information that helped me start up the cluster.

Many thanks to my family, especially my wife Jelena, for patience and given support. And to all other friends and supporters unmentioned but by no means forgotten I give my most heart-full thanks.

# List of Figures

# List of Tables