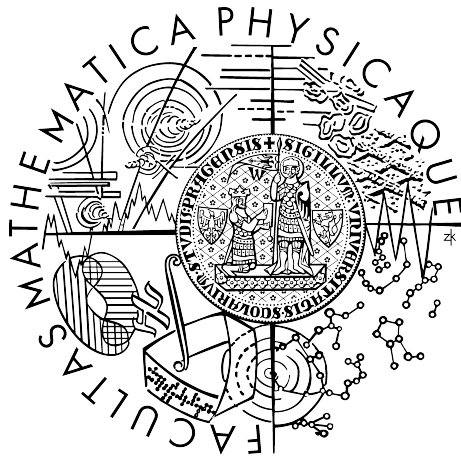Charles University in Prague

Faculty of Mathematics and Physics

MASTER THESIS

Marcel Krčah

Segmentation of Bones In 3D CT Images

Department of Software and Computer Science Education

Supervisor:	Prof. Jiří Matas
Study Programme:	Informatics
Specialization:	Software Systems

Prague 2011

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In .... date .... Signature:

Název práce: Segmentace kostí v 3D CT snímcích
Autor: Marcel Krčah
Katedra (ústav): Kabinet software a výuky informatiky
Vedoucí diplomové práce: Prof. Jiří Matas, Katedra Kybernetiky, Fakulta Elektrotechnická, České Vysoké Učení Technické, Praha

Abstrakt: V oblasti zpracování biomedicínskych dat se za poslední desetiletí výrazně zvýšil zájem o algoritmy, které dokážou přesně, automaticky a bez jakékoliv předcházející informaci o tvaru, vysegmentovat hledanou kost. Tato práce popisuje plně automatickou metodu pro segmentaci stehenní kosti z 3D snímků výpočetní tomografie (CT). Algoritmus kombinuje metodu graph-cut se speciálním filtrem, který zvýrazňuje okraje kosti pomocí analýzy zakřivení lokální iso-oblasti. Navrhované řešení bylo otestováno na souboru 197 CT snímků a porovnáno s dalšími třemi plně automatickými segmentačními metodami. Výsledky experimentů ukazují, že navrhovaná metoda dosahuje nejlepší výsledky ze všech čtyř testovaných metod a je schopna vysegmentovat stehenní kost v 81% případech bez jakékoliv interakce s uživatelem či předcházející informaci o tvaru.
Klíčova slova: segmentace, výpočetní tomografie, graph-cut, femur

Title: Segmentation of Bones in 3D CT Images
Author: Marcel Krčah
Department: Department of Software and Computer Science Education
Supervisor: Prof. Jiří Matas, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Prague

Abstract: Accurate and automatic segmentation techniques that do not require any explicit prior model have been of high interest in the medical community. We propose a fully-automatic method for segmenting the femur from 3D Computed Tomography scans, based on the graph-cut segmentation framework and the bone boundary enhancement filter analyzing second-order local structures. The presented algorithm is evaluated in large-scale experiments, conducted on 197 CT volumes, and compared to other three automatic bone segmentation methods. Out of the four tested approaches, the proposed algorithm achieved most accurate results and segmented the femur correctly in 81% of the cases.
Keywords: segmentation, computed tomography, graph-cut, femur

Contents

Chapter 1

Introduction

Due to its advancements and a wide spread in the medical community, the Computed Tomography (CT) modality has become common data source for many segmentation studies. However, as pointed out in Wang et al. (2006), despite several years of active research, bone segmentation from CT datasets remains in several aspects an open problem. Due to segmentation challenges coupled with the CT modality, established approaches are of limited use, including thresholding, region growing, watershed transformation, snakes and level-sets.

Recently, the graph-cut framework, introduced by Boykov and Jolly (2001), has shown to provide an elegant and efficient approach for segmentation. It has been applied to the problem of bone segmentation in many studies with promising results, e.g. in Fürnstahl et al. (2008).

This work presents a fully-automatic graph-cut-based method for segmenting the femur bone from 3D-CT volumes without employment of any shape prior knowledge. In particular, the study proposes new cost terms for the graph-cut framework which utilize a novel bone-boundary enhancement filter with prior knowledge of bone radiodensity properties. The applicability of the proposed method is subsequently evaluated in large-scale experiments, conducted on a set of 197 CT volumes.

The motivation of the proposal is to provide a tool for bootstrapping the construction of a statistical model of shape and appearance. These models integrate variability of bones within a population and have been subject to an increasing interest within the past decade in a wide variety of medical applications (Heimann and Meinzer, 2009). Besides anatomical studies, such models are particularly interesting in the domain of orthopedics, especially for intra-operative guidance, re-constructive surgery or implant design. However, learning such models requires that a large number of examples are available and pre-processed. Segmentation, in particular, remains a serious bottleneck in this context since manual processing is common practice. Automatic segmentation algorithms are therefore of high interest since they provide an automated tool for acquisition of initial samples.

1.1 Thesis structure

Chapter 2 introduces the problem of bone segmentation from CT modality and discusses manifold established segmentation methods which have been used for bone segmentation. Chapter 3 provides detailed description of the proposed method, with the focus on the graph-cut framework and proposal of graph-cut cost functions designed specifically for the purpose of femur segmentation. Chapter 4 expands on evaluation of the method in large-scale experiments and comparison with other existing approaches. Chapter 5 concludes the thesis with summary of results and discussion over future work.

Chapter 2

Segmentation of bones in CT

This chapter describes in more detail the problem of bone segmentation from computed tomography (CT) scans. Section 2.1 provides deeper explanation of the problem and discusses its manifold challenging aspects. Section 2.2 then expands on current approaches targeting the bone segmentation.

2.1 Problem statement

This section details the problem of bone segmentation from CT volumes. Section 2.1.1 describes the CT modality in general and briefs the process of acquiring a 3D-CT scan. The problem of bone segmentation from such a scan is defined in Section 2.1.2. The latter section also expands on typical challenges coupled with bone segmentation.

2.1.1 Computed tomography scans

In spiral CT, a 3D radiodensity image is created by reconstructing a stack of 2D images from successively acquired X-ray data (Lipson, 2006). A patient is first placed and fixed on a table, which is subsequently rolled into the CT gantry. A set of 2D scans is then acquired by using an X-ray source, which rotates continuously in helical fashion, and a detector array, determining current material attenuation. To acquire larger anatomical regions, the table mechanically moves and the scanning procedure is repeated for another set of 2D scans.

Raw output of a CT scanner is a 3D radiographic density map. Each box-like volumetric element, also called a *voxel*, represents average attenuation or density of the object at the corresponding area. In the last step, original radiographic density measurements are linearly transformed into a more descriptive scale – Hounsfield units (HU). By definition, the value of 0 HU corresponds to radiodensity of distilled water and -1000 HU to the radiodensity of air. HU values of other substances commonly present in human body are based on this definition and are listed in Table 2.1.

To view such a CT scan on a monitor, Hounsfield units are mapped to a gray scale. The transformation function is called a *transfer function* and is usually fully customizable in medical software. A typical example of a 3D-CT scan, containing human forearm, is depicted in Figure 2.1.

Substance	HU value
Air	-1000
Fat	-120 to -50
Water	0
Muscle	+40
Cortical bone	+400 to +2000

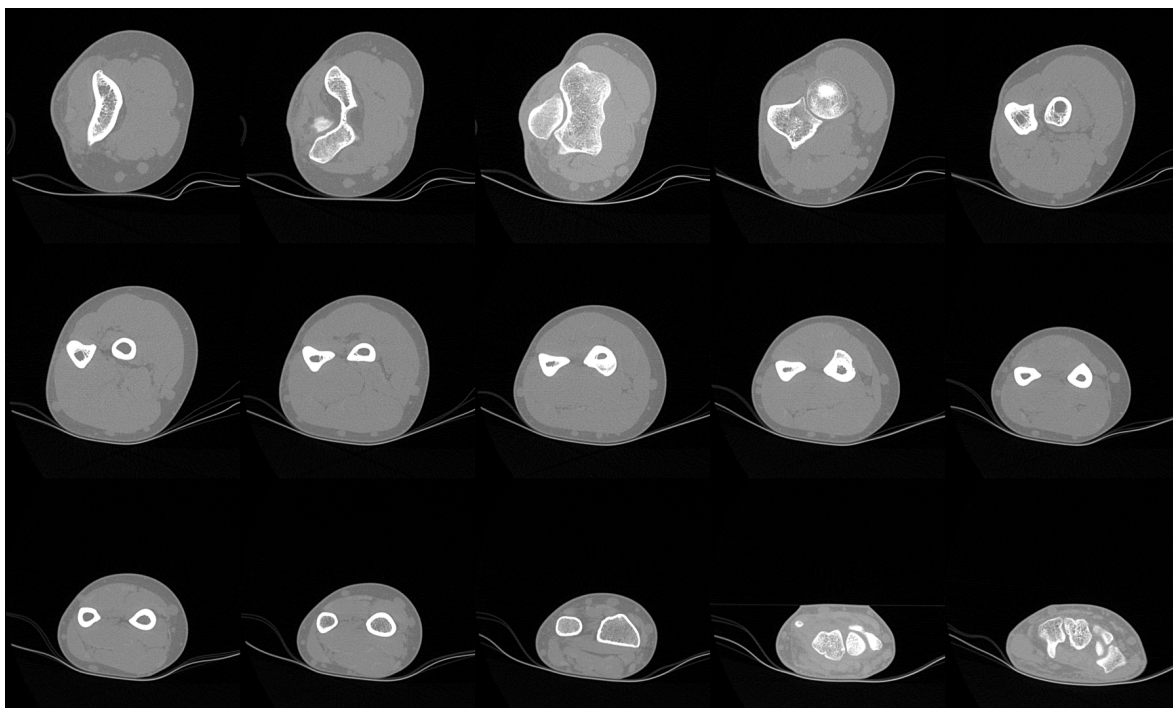Table 2.1: HU values for common substances

Figure 2.1: Selection of 15 axial slices (out of 660) from a CT volume of human forearm. The temporal axial ordering is top-to-bottom left-to-right. Voxel intensities correspond to HU values. Black regions represent air with HU value below -800. Soft-tissues, water and fat (HU from -200 to 400) is depicted by gray. Strong cortical bone layers with HU above 400 is colored white. Resolution of this volume is 320x320x660.

2.1.2 Femur segmentation

In general, a binary segmentation of an image can be formulated as follows: Assume that $\mathcal{I} : \Omega \to \mathbb{R}$ is an image on a (grid) domain Ω. Segmenting an object means to find a labelling function

$$A : \Omega \to \{obj, bkg\}$$

which classifies all pixels belonging to the object of interest as obj and the remaining pixels as bkg. Having the pixels classified, the image can be considered as a union of two disjoint sections $S_{obj}, S_{bkg} \subset \mathcal{I}$, where S_{obj} and S_{bkg} represent the object and background parts of the image, respectively.

In this study, the input image for segmentation is a product of a CT scanner, i.e. a

3D volume with voxels expressed in the Hounsfield scale. The goal is to automatically segment the femur bone without employment of any prior knowledge of the femur shape. Example of a desired result is illustrated in Figure 2.2.

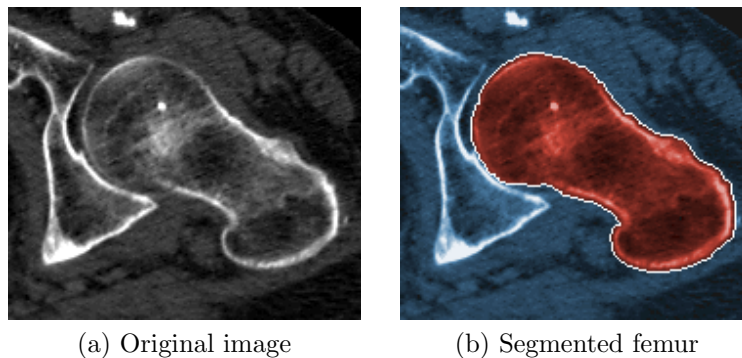|	
(a) Original image	(b) Segmented femur

Figure 2.2: Femur segmentation illustrated on an axial slice. The object of interest (femur) is depicted in red. Background (pelvis, fat and soft-tissues) is depicted blue.

Typical appearance of the femur in the CT modality is illustrated in Figure 2.3. The middle part of a long bone, also called a *shaft* or a *diaphyses*, consists of rigid, strongly attenuating and dense *cortical* tissue, which exhibits high HU values in CT scans. Due to smooth bone boundaries and high contrast between the cortical layer and encasing low-intensity soft tissue, segmentation of the shaft is a relatively simple tasks. Global thresholding (subsection 2.2.1) followed by basic morphological operations (Haralick et al., 1987) can deliver satisfactory results.

However, the bone structure is significantly different in the *epiphyses* – part of a bone near joints, where more flexibility has to be provided due to pressure and
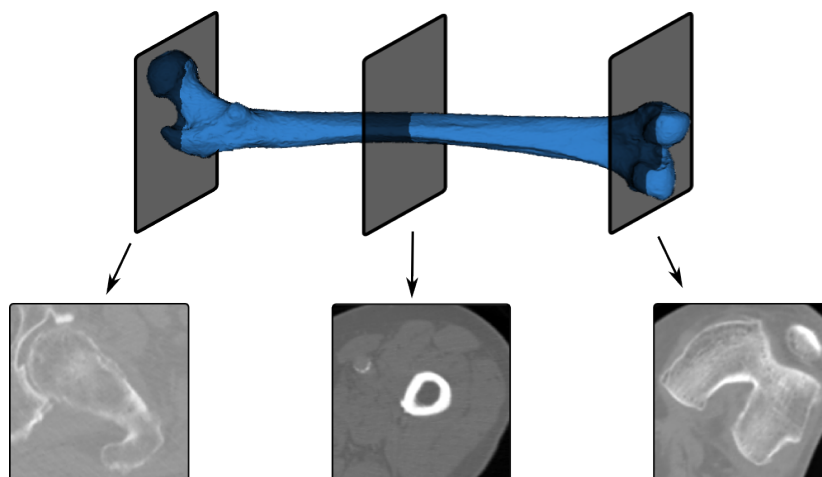
Figure 2.3: Typical appearance of the femur in a CT scan. The middle part of the femur consists of strongly attenuating and rigid cortical tissue. The corticalis diminishes towards the joint areas where it is replaced by low-attenuating vascular and hard-to-segment cancellous tissue.
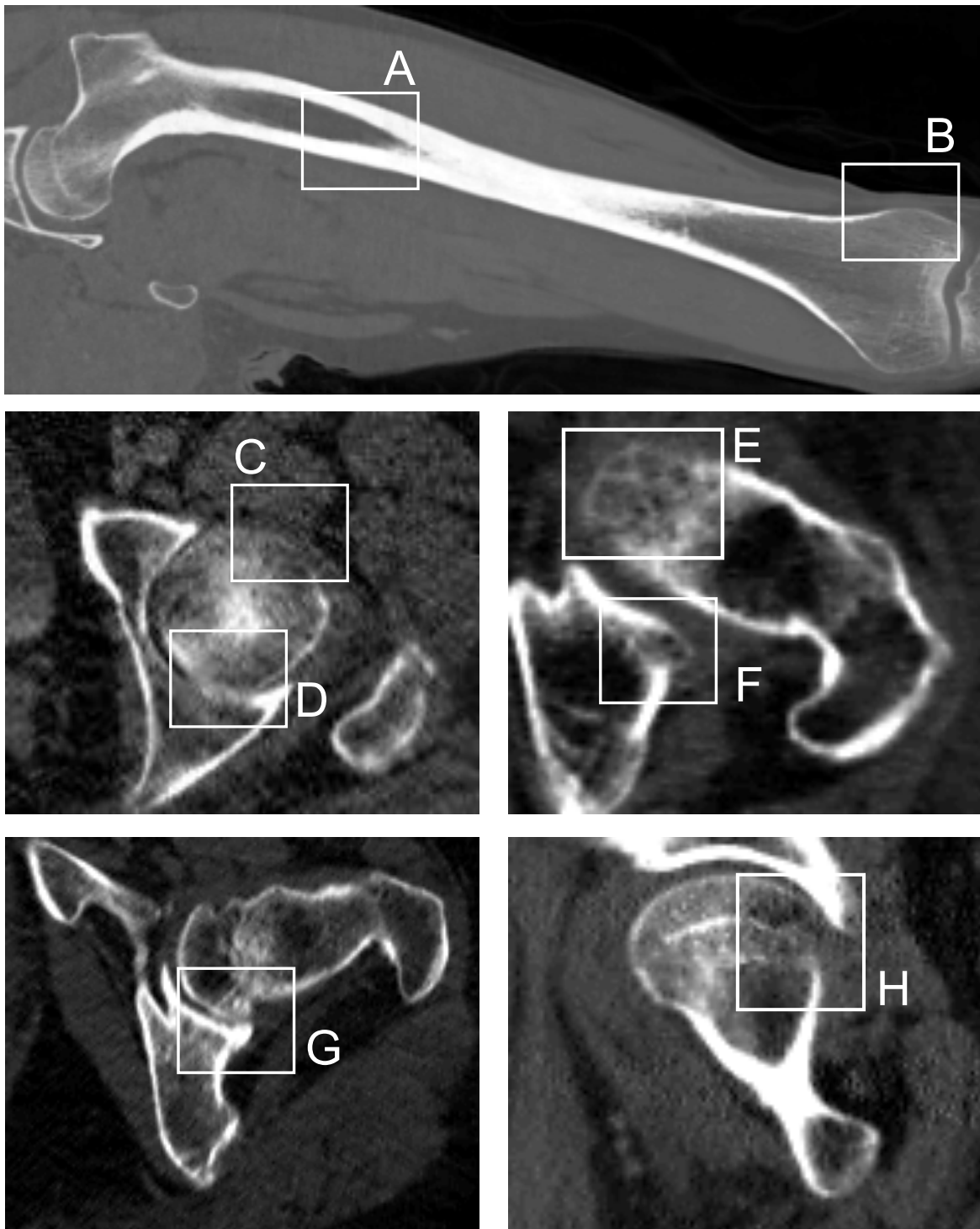
Figure 2.4: Challenges coupled with segmentation of the femur from CT volumes: different intensity distributions of the cortical and cancellous tissues (A,B), weak and diffused bone boundaries (C, E), missing bone boundaries (F,H), narrow inter-bone spacing (D,G) and noise (C).

body movement. The bone interior consists almost exclusively of low-attenuating and vascular *cancellous* tissue. As Sebastian et al. (2003) and Fürnstahl (2009) listed, there are manifold challenges regarding the segmentation of bone epiphyses from CT volumes (see Figure 2.4):

Low-contrast and weak bone boundaries Thickness of the cortical shell decreases rapidly towards the epiphyses area which results in low-contrast bone boundaries. In particular, this is the case for elderly patients with chronically decreased bone density. What is more, the bone boundary often appears to be heavily diffused or is not visible at all, due to the presence of a cartilage or a vessel nourishing the bone interior.

Varying density of cancellous tissues Cancellous bone tissue is vascular and spongy which causes the bone to appear textured and inhomogeneous. The density properties of bone tissues cannot be characterized uniformly. From the radiodensity perspective, the epiphyses interior partially resembles low-attenuating fat and soft-tissues; partially high-attenuating corticalis. Therefore, basic segmentation techniques like thresholding or region growing (section 2.2) inevitably yield insufficient results (Sebastian et al., 2003).

Narrow inter-bone spacing Articular space between adjacent bones is extremely narrow. Due to partial volumetric effect, inherently present in CT modality, the inter-bone regions are diffused and brighter and the bones appear as being in direct contact.

Low quality of CT scans Moreover, the overall quality of CT scans is generally rather low. Volumes suffer from noise and low-resolution. Voxels are usually anisotropic with dimensions ranging from 0.5mm to 2mm. The resolution of a CT scan is often insufficient for the image processing purposes. Conversely, memory requirements coupled with software processing are very high with respect to the current computer capacities. Consider a CT volume of a resolution 512x512x1024. If four bytes per voxel are used for the Hounsfield unit, the volume size reaches 1GiB.

2.2 Current approaches

This section discusses manifold established techniques which have been applied to the problem of bone segmentation in 2D/3D CT domain. The summary begins with the description of the most common and easy-to-implement methods: Section 2.2.1 focuses on thresholding and Section 2.2.2 briefs the region growing. In Sections 2.2.3 and 2.2.4 , the Watershed transformation and approaches employing first-order local structure analysis are explained, respectively. A family of segmentation methods based on deformable models are described in the remaining sections. Section 2.2.5 and Section 2.2.6 cover snakes and geometric active contours, respectively. The Section 2.2.7 expands on segmentation methods based on statistical models of shape and appearance.

2.2.1 Thresholding

Thresholding is one of the most common and simplest segmentation methods. In its most simplistic form, (global) thresholding classifies pixels according to the statistical distribution of their intensity values – histogram. Pixels with intensity values above a manually chosen HU value are interpreted as a part of an object, otherwise pixels are classified as the background.

The global thresholding performs well if the intensities of an object and a background do not vary significantly throughout the image. However, this is not the case for CT bone segmentation due to similarities in radiological density of bones and surrounding soft tissues. For example, the global thresholding cannot properly handle segmentation of vascular cancellous bones (see Figure 2.5a) nor segmentation of joint regions with diffused and weak bone contours (Sebastian et al., 2003; Wang et al., 2006). The method also struggles in narrow inter-bone regions due to the volumetric effect (see Figure 2.5b).

Additionally, to achieve satisfactory results the threshold value has to be chosen carefully; the operation which tends to be tedious and time-consuming in case of manual interactive large-scale segmentations (Sebastian et al., 2003). Also, incorporation of various constraints (as a continuous bone boundary) is also hardly feasible (Fürnstahl, 2009).

Several of the aforementioned limitations are suppressed in local thresholding – a method featuring local adaptability. Each pixel is classified according to a local threshold, calculated from local image characteristics. However, Sebastian et al. (2003) and Wang et al. (2006) pointed out that the local thresholding fails in narrow inter-bone regions since methods employed to determine the threshold typically use small local windows centered around a pixel.

Being one of the oldest segmentation techniques, the thresholding methods experienced an enormous scientific interest in the past. Many researchers incorporated prior knowledge and statistical methods (e.g. clustering, entropy, etc.) to increase robustness, resulting in numerous thresholding variations. Excellent reviews are given in Sahoo et al. (1988) and Sezgin and Sankur (2004). In the latter review, more than 40 thresholding methods are categorized, compared and unified under the same notation.

Despite its limitations, thresholding is often built into many segmentation pipelines as a preprocessing step or as a partial processing stage.

2.2.2 Region growing

Region growing methods partition an image into coherent areas, each of which satisfies a given homogeneity criterion. The most common region growing algorithms are based on *seeded region growing*, introduced by Adams and Bischof (1994).

Initially, several object seeds are chosen manually by a user. The seeds then start to grow into regions by annexing adjacent pixels. The annexing criterion is defined by a similarity measure, usually based on pixel intensities, gradient, local neighborhood characteristics, statistical tests, etc. When the growing stops, statistically similar adjacent regions are typically merged together, resulting in *region merging* algorithms.

Seeded region growing tends to be sensitive to noise, which results in undesirable

(a) Global thresholding: The lack of local adaptability produces holes in the segmented object since the tissues with similar intensities cannot be differentiated.

(b) Global thresholding: Due to the volumetric effect, thresholding merges adjacent bones with narrow inter-bone spacing.

(c) Seeded region growing: Leakages occur through weak and missing bone boundaries.

(d) Region competition: Segmentation contains leakages to soft tissues and other bones.

(e) Watershed transform: The transformation results in an oversegmented image.

(f) Local Structure Analysis: Adaptive Filtering controlled by local structure tensors (Westin et al., 1998).
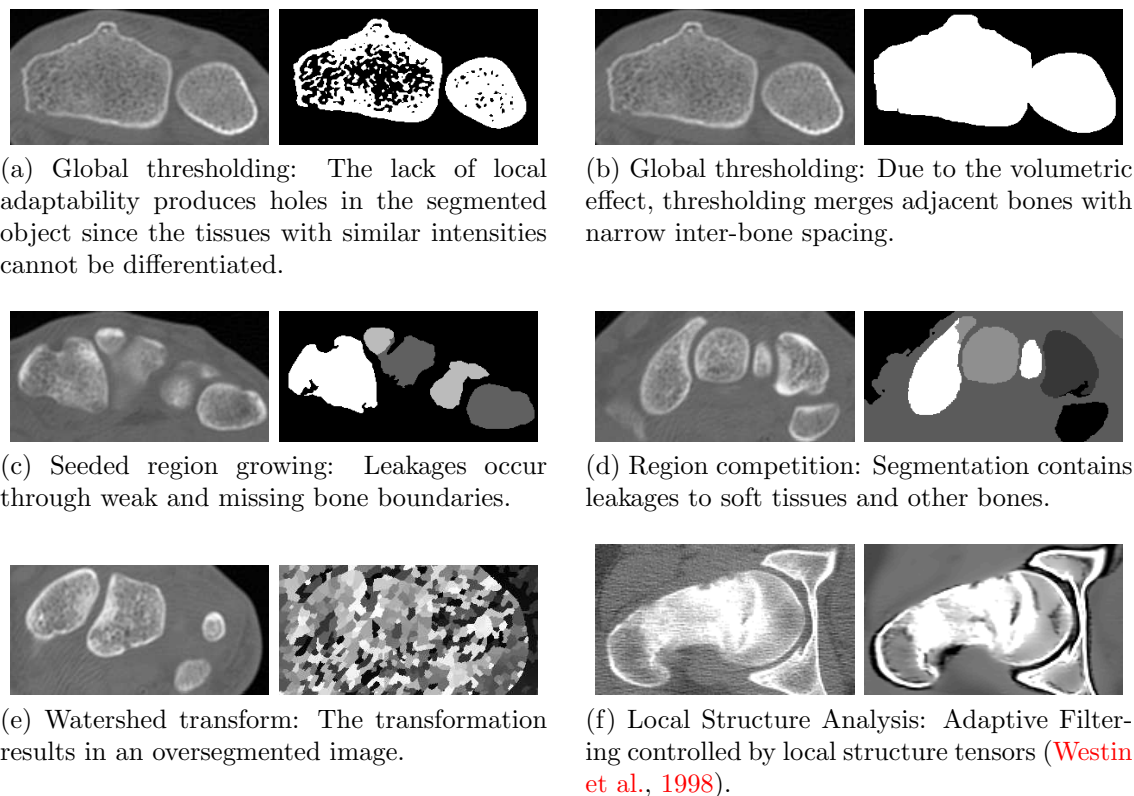
Figure 2.5: Common outputs and problems of the presented low-level segmentation methods. Each example contains an original unprocessed image (left) and the corresponding output (right). Image sources: Sebastian et al. (2003), Westin et al. (1998).

holes, and to initial seeds (Sebastian et al., 2003). Due to partial volumetric effect, regions with narrow inter-bone spacing tend to be merged together (Wang et al., 2006). Since no geometrical constraints are incorporated, leakages through gaps in bone surfaces occur as well (see Figure 2.5c).

Many authors extended the idea of region growing: Modayur et al. (1997) proposed an adaptive similarity measure to alleviate limitations caused by noise and spatially varying tissue intensities. Zhu and Yuille (1996) presented a *region competition* (see Figure 2.5d) – an algorithm which incorporates geometrical properties of active contours (Sections 2.2.5 and 2.2.6) into region growing. The method also incorporates local competition between regions allowing recovery-from-errors behavior.

2.2.3 Watershed transform

The Watershed transform (WT) is a morphological gradient-based technique pioneered by Vincent and Soille (1991). A gray image is interpreted as a height-map, where pixel values correspond to elevation. First, several water sources are initialized either as points with minimum gradient or as user-defined markers. Water emerging from these sources then gradually floods the map and creates a catchment basin for each source. Watersheds delineating filled basins represent segmented regions.

Due to presence of noise and high inhomogeneities in CT domain, general WT produces oversegmented images (Sebastian et al., 2003), as illustrated in Figure 2.5e. Moreover, being a gradient-based method, the watershed transform struggles in regions with low-contrast bone boundaries.

Hahn (2005) introduces an interactive WT for bone segmentation and gives an excellent overview of WT methods. Neubauer et al. (2005) combines manually initialized WT with thresholding to segment radius, ulna and carpus. Straka et al. (2003) proposes WT in combination with a probabilistic anatomic atlas.

2.2.4   Local structure analysis

To remove noise and increase contrast at bone boundaries, behavior of local iso-intensity regions for all locations of an image can be scrutinized. To analyze a local structure, the eigenvalue decomposition of the structure tensor $\mathcal{T}$ may be employed, where

$$\mathcal{T} = \begin{pmatrix} \mathcal{I}_x^2 & \mathcal{I}_x\mathcal{I}_y & \mathcal{I}_x\mathcal{I}_z \\ \mathcal{I}_x\mathcal{I}_y & \mathcal{I}_y^2 & \mathcal{I}_y\mathcal{I}_z \\ \mathcal{I}_x\mathcal{I}_z & \mathcal{I}_y\mathcal{I}_z & \mathcal{I}_z^2 \end{pmatrix}$$

and $\mathcal{I}$ denotes a 3D image. Westin et al. (1997) incorporated the eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$ of tensor $\mathcal{T}$ into a planarity measure

$$c_{plane} = \frac{\lambda_3 - \alpha\lambda_2}{\lambda_3}.$$

For each image voxel, the measure approaches 1 for planar structures (e.g. bones) and 0 otherwise. The measure is subsequently used in adaptive thresholding classifying each pixel according to the local threshold $t(x) = t_0 - \alpha c_{plane}(x)$, where t_0 is a manually chosen global threshold and α is a constant weight factor.

The algorithm was employed to segment thin-walled bone regions from 3D-CT volumes. Since the tensor $\mathcal{T}$ is based on the first-order variations only, the planarity measure is strong on the bone boundaries but weak inside the cortical bone, where the gradient approaches zero.

Extensions followed in (Westin et al., 1998) and (Westin et al., 2000). Westin et al. (1998) presented a tensor controlled adaptive filters (see Figure 2.5f). In Westin et al. (2000), they incorporated the adaptive components derived from local structure descriptors into the Geodesic Active Contours (see Section 2.2.6).

Westin's approach may be extended to Hessian matrices, which represent curvature of a local region via second-order derivatives. subsection 3.1.2 describes this technique in more detail as it is a building block of the proposed femur segmentation method.

2.2.5   Parametric deformable models (Snakes)

Parametric deformable models (PDMs) segment an object via evolution of an explicitly defined contour. The most popular implementation of this approach was proposed by Kass et al. (1988) in their seminal publication on *snakes*:

Assume that $\Omega \subset \mathbb{R}^2$ denotes a continuous image plane. The snake contour is then defined as an explicit (parametric) curve $C : [0, 1] \rightarrow \Omega$. The curve is associated with an energy functional $E(C)$, defined as

$$E(C) = -\int |\nabla\mathcal{I}(C)|^2 ds + \nu_1 \int |C_s|^2 ds + \nu_2 \int |C_{ss}|^2 ds, \qquad (2.1)$$

where C_s, C_{ss} denote first-order and second-order derivatives of the curve C, respectively, values ν_1, ν_2 denote nonnegative constants and $\nabla\mathcal{I}$ denotes an image gradient.

In the beginning, the snake is initialized manually or as a result of a preprocessing step. Using a variational framework, the curve is subsequently evolved by minimizing the functional $E(C)$. More technically, the finite differences scheme is employed on the Euler-Lagrange equation derived from (2.1) to compute the local minimum.

The functional E is composed of two components: *external* and *internal* energy. The first term of Equation 2.1 represents external energy as it attracts the evolving curve towards large image gradient locations rather than towards local homogeneous regions. The second and third terms comprise internal energy; they measure the length and rigidity of the snake and therefore provide regularization of the evolving contour.

According to Cremers et al. (2007), the original snake technique is highly efficient and easily incorporates prior knowledge. However, the algorithm is sensitive to initial conditions and often gets stuck in a local undesired minimum. Moreover, due to the explicit contour representation, snakes lack the feature of topological adaptability (Cremers et al., 2007; Maa et al., 2009). Therefore, the initial contours must share the same topology as the object of interest.

There are many publications alleviating the aforementioned hindrances. To name a few, Cohen (1991) added a *balloon* force to the external force to prevent contours from shrinking or stopping on local minimum. Xu and Prince (1998) lowered dependence on initialization by replacing the gradient field with the gradient vector field. User interaction was increased by McInerney and Terzopoulos (1996). A detailed review coupled by an excellent comparative study is given in He et al. (2008).

2.2.6   Geometric deformable models

This section focuses on geometric active contours – methods designed to inherently handle topological changes during the contour evolution and thus alleviate issues coupled with snakes (Section 2.2.5). Geometric active contours are based on the level-set framework, pioneered by Osher and Sethian (1988), who applied level-sets to model ocean waves and burning flames. The level-set framework was first employed to the problem of image segmentation by Malladi et al. (1995).

As opposed to snakes, where the contour is moved explicitly, Osher and Sethian proposed to implicitly embed a contour into a higher-dimensional surface and track the contour as the zero level-set of this evolving embedding surface.

Contour evolution with level-sets

Assume that $\Omega \subset \mathbb{R}^2$ denotes a continuous image plane and a contour $\mathcal{C} \subset \Omega$ of arbitrary topology is manually initialized. The evolution operates with a time-dependent
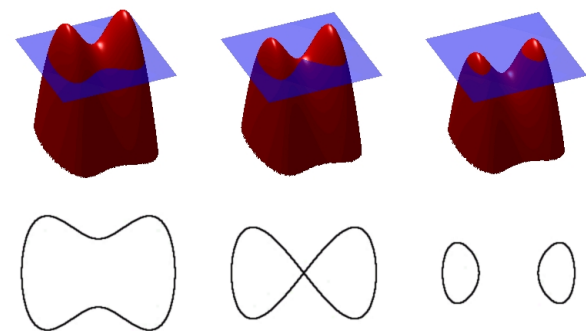
Figure 2.6: A sketch illustrating the idea behind the level-set technique. A lower-dimensional curve (bottom row) is embedded into a higher dimensional surface (top row) as a zero-level set. Level-set segmentation methods naturally feature topological adaptability during the contour evolution.

embedding surface $\phi : \Omega \rightarrow \mathbb{R}$, defined as

$$\phi(\mathbf{x}, t) = \pm d(\mathbf{x}),$$

where $\mathbf{x}$ is position in the image, t is time and $d(\mathbf{x})$ is a signed distance between $\mathbf{x}$ and the contour $\mathcal{C}$. The sign in front of the distance is positive for $\mathbf{x}$ outside the contour and negative for interior regions. In this manner, the contour $\mathcal{C}$ may be represented as a zero level-set of the surface ϕ, i.e.

$$\mathcal{C} = \{\mathbf{x} \in \Omega | \phi(\mathbf{x}) = 0\},$$

as depicted in Figure 2.6. Assumption that the contour motion is perpendicular to the boundary leads to the evolution model

$$\frac{\partial \mathcal{C}}{\partial t} = F \cdot \mathbf{n}, \tag{2.2}$$

where $\mathbf{n}$ denotes a normal vector and F denotes a speed function. Cremers et al. (2007) pointed out that the tangential component in the evolution model (2.2) is usually not included as the latter does not affect the contour but only the parametrization.

Since $\phi(\mathcal{C}(t), t) = 0$ all the time, the total time derivative of ϕ for points lying on the contour must vanish:

$$\begin{aligned}
0 &= \phi_t + \phi_x x_t + \phi_y y_t \\
&= \phi_t + \nabla \phi \cdot (x_t, y_t) \\
&= \phi_t + \nabla \phi \cdot \frac{\partial \mathcal{C}}{\partial t} \\
&= \phi_t + \nabla \phi \cdot F \mathbf{n} \\
&= \phi_t + F |\nabla \phi|,
\end{aligned}$$

using the chain rule and the definition of the normal vector $\mathbf{n} = \frac{\nabla \phi}{|\nabla \phi|}$. Hence, the resulting evolution equation for the embedding surface ϕ is

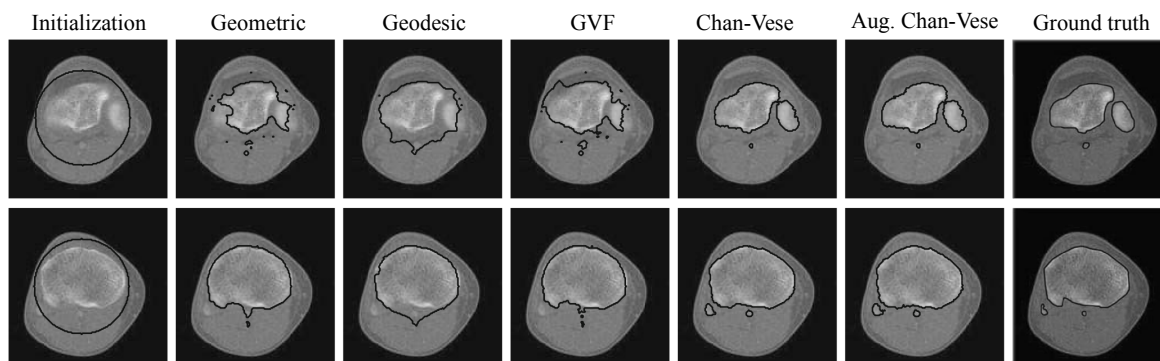$$\frac{\partial \phi}{\partial t} = -|\nabla \phi| F \tag{2.3}$$

Figure 2.7: Performance of various level-set methods applied to bone segmentation as surveyed by Truc et al. (2009).

with a given initial condition $\phi(\mathbf{x}, t = 0)$. This can be solved using finite differences on a discrete spatial grid of the Ω domain. In order to numerically implement the evolution equation (2.3), the speed function F, initially defined only for contour locations, needs to be extended to the whole image domain. For technical details see Malladi et al. (1995).

(Original) Geometric Active Contours

In the original geometric active contours (Caselles et al., 1993; Malladi et al., 1995) the speed function F is based on a function $g(\mathbf{x})$ which vanishes to zero on edges and approaches one in homogeneous regions. For example,

$$g(\mathbf{x}) = \exp\left(-\alpha|\nabla(\mathcal{G} * \mathcal{I})(\mathbf{x})|^2\right)$$

is such a function, where $\mathcal{G} * \mathcal{I}$ denotes convolution of an input image $\mathcal{I}$ with a Gaussian kernel $\mathcal{G}$ and α denotes a constant parameter. The corresponding contour evolution is given by

$$\frac{\partial \phi}{\partial t} = (\kappa + V_0)g|\nabla\phi|, \tag{2.4}$$

where κ is an Euclidean curvature and V_0 is a real constant. The intuitive idea behind geometric active contours (ACs) is the following: As the contour approaches an object boundary, the increased gradient magnitude decreases the speed of the propagating curve.

While the geometric ACs usually perform well in high contrast images, the algorithm struggles in regions with weak edges and gaps (Suri et al., 2002; Truc et al., 2009; Maa et al., 2009). In regions with diffused bone boundaries and holes, the gradient magnitude is too small to stop the evolving curve and causes leakages in resulting segmentations. The method also lacks a pulling-back feature (Suri et al., 2002), meaning that once an object boundary is crossed the propagating front cannot return.

Geodesic Active Contours   To overcome the leakage problem of the geometric ACs, Caselles et al. (1997) proposed *Geodesic ACs*, a method inspired by classical parametric

energy minimization techniques (snakes), briefed in Section 2.2.5. The contour flow of Geodesic ACs can be described as

$$\frac{\partial \phi}{\partial t} = (\kappa + V_0)g|\nabla \phi| + \nabla g \cdot \nabla \phi.$$

In comparison with Malladi's equation (2.4), an extra stopping term $\nabla g \cdot \nabla \phi$ is added to attract the propagating front towards weak edges.

Siddiqi et al. (1998) pointed out that Geodesic ACs still yield leakages for complex structures. Therefore, they incorporated another component to the flow – *area forces*. Suri et al. (2002) concluded in their review that albeit the latter approach performed better than the previous two models, the method is still not robust enough to handle the convolutedness of medical structures.

Other methods and Reviews A number of researchers extended the aforementioned ideas. For example, Paragios et al. (2004) incorporated gradient vector flow (Xu and Prince, 1998) into Geometric ACs. An alternative method of ACs based on Mumford-Shah functional (Mumford and Shah, 1989) was proposed by Chan and Vese (2001).

An excellent reviews and comparative studies of many level-set methods were presented by Suri et al. (2002), He et al. (2008) and Truc et al. (2009). A summary of statistical approaches to level-set segmentation is deeply detailed by Cremers et al. (2007). Truc et al. (2009), Sebastian et al. (2003) and He et al. (2008) employed the geometric active contours to the problem of bone segmentation. Performance of various AC methods applied to bone segmentation from CT volumes is illustrated in Figure 2.7.

2.2.7   Statistical shape models

Statistical shape models (SSMs) form a family of established methods for robust segmentation utilizing prior shape knowledge. This section scrutinizes 3D SSMs employed for purposes of medical image processing. A typical work-flow of a SSM is detailed and a brief description of the two most established methods is given: Active Shape Models (ASMs) and Active Appearance Models (AAMs). Refer to Heimann and Meinzer (2009) for an excellent review of SSMs for 3D medical image segmentation. Cootes et al. (2004) presents an extensive technical report of ASMs and AAMs.

A typical outline of a SSM segmentation is depicted in Figure 2.8. Initially, a collection of training samples is obtained by (manual) segmentation of numerous medical volumetric images. A statistical shape model, which extracts the mean shape and the most significant modes of variations, is subsequently constructed from this training data set. The training set is also utilized to encode object appearance into appearance statistical model. To extract an object of interest from a previously unseen image, the search algorithm employs both the shape and appearance model. The following text discusses each aforementioned aspect of the segmentation work-flow.

Shape representation

The majority of shape models in the medical image domain is based on *Point Distribution Models* (PDMs), the term coined by Cootes et al. (1992). A shape is represented
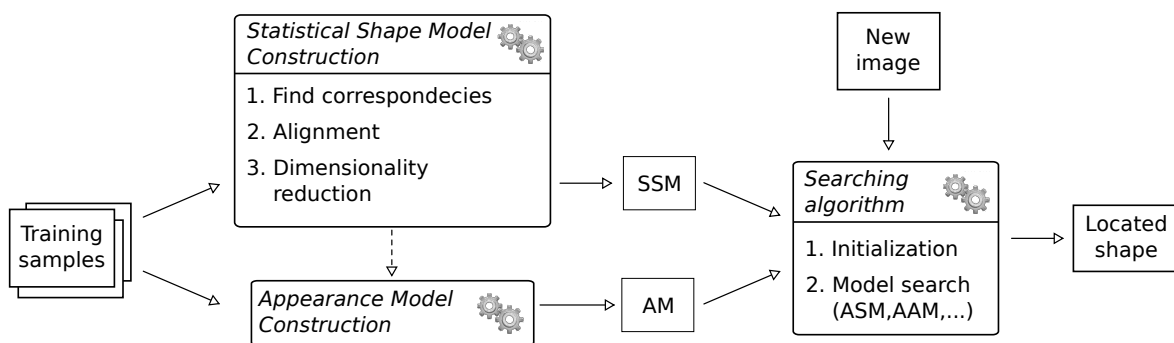
Figure 2.8: Outline of a segmentation algorithm based on Statistical Shape Models (SSMs). First, the statistical model reflecting the mean shape and the most significant modes of variation is extracted from the training dataset. Subsequently, the appearance model encoding shape boundary features or inner object region features is constructed. These two models are then employed in the search phase to locate a shape in a previously unseen image.

as a set of n points distributed over the surface. The points are concatenated into a vector $\mathbf{x}$ describing the shape:

$$\mathbf{x} = (x_1, y_1, z_1, \ldots, x_n, y_n, z_n)^T.$$

In literature, these points are often referred to as *landmarks*. A PDM coupled with connectivity information between the points is called a *mesh*.

Construction of a statistical shape model

A SSM extracts statistical properties like a mean shape and significant modes of variations from a collection of training samples. To build a SSM, landmarks on all training samples must be initially located at corresponding positions. This task is called a *correspondence problem* and forms the most challenging part of 3D model reconstruction. Basically, all algorithms for finding correspondences perform a registration between the involved shapes (Heimann and Meinzer, 2009). For a comparison of an exhaustive number of publications regarding the correspondence problem for SSM, refer to Heimann and Meinzer (2009).

Once the correspondences between the shapes are found, the volumetric samples are aligned in a common coordinate frame. *Generalized Proscrustes alignment* (Gower, 1975; Goodall, 1991) and *tangent space scaling* (Dryden and Mardia, 1998) are the most popular methods for solving this problem.

The PCA analysis, e.g. Jolliffe (2002), is subsequently employed to extract the mean shape and the modes of shape variation to globally describe the training data in lower dimension. The mean shape $\overline{\mathbf{x}}$ is approximated by simply averaging over the sample set

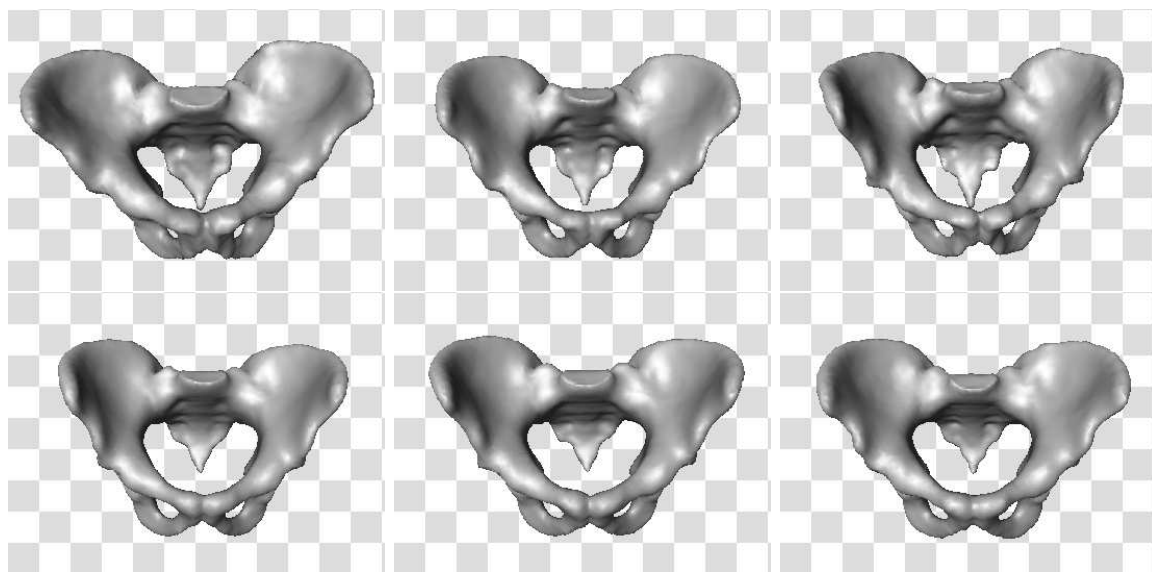$$\overline{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x_i},$$

Figure 2.9: Principal modes of variation for an SSM of the pelvic mode as presented by Lamecker et al. (2004): In the left column, the eigenmode with the largest variance λ_1 is varied between $\pm 3\sqrt{\lambda_1}$, the medium and right column show the variation of the second and third largest eigenmode, respectively.

where N denotes the number of samples and $\mathbf{x_i}$ encodes the aligned PDM of i-th training sample. The eigenvalue decomposition of a covariance matrix S, defined as

$$S = \frac{1}{N-1} \sum_{i=1}^{N} (\mathbf{x_i} - \overline{\mathbf{x}})(\mathbf{x_i} - \overline{\mathbf{x}})^T,$$

delivers the c largest principal modes of variation $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_c}$. Each valid shape can be then approximated by a linear combination of the first c modes as

$$\mathbf{x} \approx \overline{\mathbf{x}} + \sum_{i=1}^{c} \alpha_i \mathbf{v_i}.$$

In this manner, a shape is represented by only a small number of shape parameters $\alpha_1, \ldots, \alpha_c$ instead of a large PDM vector $\mathbf{x}$. An illustration of a constructed SSM using PCA is given in Figure 2.9.

Construction of the appearance model

To apply the SSM for object segmentation, the appearance model (AM) for the object of interest must be built as well. In principle, there are two well-established families of AM: *boundary-based* and *region-based*.

To construct a boundary-based AM, the profiles perpendicular to the object surface are scrutinized for corresponding landmarks. Various features may be employed to describe the profiles, e.g. (normalized) intensities or derivatives. Machine learning techniques like PCA or kNN-classifier are subsequently applied to globally describe the profile for each landmark.

In the region-based AM, the entire inner object region is analyzed. One established approach is the Active Appearance model, explained in the next paragraph. All training samples are first transformed into a common shape, e.g. into the mean shape $\overline{\mathbf{x}}$ of the SSM. Then, a feature vector encoding properties of an inner object region is computed for each transformed image. The appearance model is then obtained by applying a PCA on the feature vectors.

Searching algorithm

The segmentation is performed by locating the constructed SSM in a new image. Since most popular methods use local search algorithms, initial estimate of the model pose in necessary. In the most simple case, a user can manually estimate rough alignment and rotation of the shape. Alternatively, histogram analysis (Soler et al., 2001), affine registration with an atlas (Fripp et al., 2005), generalized Hough-transform (Ecabert et al., 2008), genetic algorithms (Hill and Taylor, 1992; Stegmann et al., 2001; Pitiot et al., 2002) or partical filtering (De Bruijne and Nielsen, 2005; Heimann et al., 2005) may be employed for initialization.

One of the most popular searching algorithms is *Active Shape Model Search* (ASM), proposed by Cootes et al. (1992, 1995). Location of the object in a new image can be approximated by the model $\mathbf{m}$ defined as

$$\mathbf{m} = T(\overline{\mathbf{x}} + \sum_{i=1}^{c} b_i \mathbf{v_i}),$$

where $(\overline{\mathbf{x}}, \mathbf{v}_1, \dots, \mathbf{v}_c)$ denotes the SSM as described earlier, $b_1, \dots, b_c$ are the shape parameters and T is a similarity transform. The object is found by iteratively adjusting the transformation T and parameters $b_1, \dots, b_c$. In each update step, the fit of a local boundary-based AM at different positions perpendicular to the surface is evaluated for each landmark.

In *Active Appearance Models* (AAMs), pioneered by Cootes et al. (1998), the region-based appearance models are employed. Shape and appearance variations are combined into one vector which is iteratively adjusted according to a certain criterion. AAMs belongs to a generative models, which can synthesize realistic images of the modeled data (Heimann and Meinzer, 2009).

Since there are numerous publications regarding various search techniques using SSMs, refer to Heimann and Meinzer (2009) for an exhaustive survey of the current approaches.

Application

SSMs have been widely used in 3D medical image analysis, including the domain of bone segmentation. Active Shape Models have been recently employed to segment pelvic bones (Lamecker et al., 2004), femur bones (Josephson et al.) and knee bones (Fripp et al., 2006). A hybrid shape model for the reconstruction of femur bones was presented by Tang and Ellis (2005). Active Appearance Models are rarely employed in medical image context due to its excessive memory requirements for the 3D texture model.

In general, SSMs are suitable for shapes which are relatively stable over populations. This makes SSMs ideal for segmentation of bones or deep brain and cardiac structures. Conversely, shapes with highly varying structures like liver or vessels are much harder to analyze with SSMs. Structures with mainly random shapes, e.g. tumors, are therefore unsuitable for statistical modeling.

Chapter 3

Algorithm proposal

The proposed algorithm is based on the graph-cut framework and a novel bone boundary enhancement filter. First, the graph-cut framework and the basic idea behind the enhancement filter are described in Section 3.1. Section 3.2 details the volume preprocessing used to increase low contrast at bone boundaries. Section 3.3 explains how the graph-cut framework is initialized to segment all bone tissues from input CT volumes. Section 3.4 expands on the postprocessing step which handles bones appearing to be in directed contact in CT volumes.

3.1 Prerequisites

This section covers the prerequisites required to understand the proposed method. Section 3.1.1 describes the graph-cut segmentation framework and Section 3.1.2 introduces the sheetness score as a measure of similarity of local iso-intensity regions to a planarity structure based on Hessian analysis.

3.1.1 Graph-Cut framework

The family of graph-cut methods are relatively recent. Pioneered by Greig et al. (1989), who applied the algorithm to binary image restoration, many authors followed his work and used graph-cuts in many areas of computer vision, including image restoration, multi-view reconstruction, texture synthesis, etc. Graph-cut methods were applied to the problem of image segmentation by Boykov and Jolly (2001). Boykov and Funka-Lea (2006) give deeper explanation of the presented algorithm and gives an excellent overview of current graph-cut methods. This section introduces basic ideas of the graph-cut and describes main features offered by the framework: incorporation of hard constraints and directed boundary terms.

Basic Overview

The graph-cut framework is an energy-minimization segmentation method based on the combinatorial graph theory. For the graph-cut, an image (of arbitrary dimension) is represented as a set of pixels $\mathcal{P}$ within a neighborhood system $\mathcal{N} \subseteq \binom{\mathcal{P}}{2}$.

In general, the method requires to define two components of a cost function: the *per-pixel* term and the *boundary-term*. The per-pixel term

$$R_p : \mathcal{P} \times \{obj, bkg\} \to R_0^+$$

specifies for each label $L \in \{obj, bkg\}$ and each pixel $p \in \mathcal{P}$ an individual penalty for assigning the label L to the pixel p. Normally, $R_p(\cdot)$ encodes how the intensity of the pixel p fits into a known intensity model of the object and the background (Boykov and Jolly, 2001). For example, the per-pixel term can be defined as:

$$
\begin{aligned}
R_p(obj) &= -\ln P(I_p|obj), \\
R_p(bkg) &= -\ln P(I_p|bkg),
\end{aligned}
\tag{3.1}
$$

where $P(I_p|obj)$ denotes probability that the pixel p of intensity I_p belongs to an object. The two distributions can be calculated either from prior segmentation results or employing prior information about the object and the background. Typically, the per-pixel penalty $R_p(obj)$ ($R_p(bkg)$) should be low for pixels which *fit* with an object (background) model and high otherwise.

The *boundary* term $B(\{p,q\}) : \mathcal{N} \to R_0^+$ defines a penalty for classifying adjacent pixels p and q with different labels. The boundary term penalizes discontinuities between adjacent pixels and thus encourages spatial coherence within object and within background. Typically, the penalty is high for "similar" pixels and low otherwise. Various similarity measures may be employed to define the boundary term; based e.g. on local intensity gradient, Laplacian zero-crossing, gradient direction or others (Mortensen and Barrett, 1998). As Boykov and Funka-Lea (2006) noted, it is often sufficient to set the boundary penalties from a simple function like

$$B(\{p,q\}) \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \frac{1}{dist(p,q)}. \tag{3.2}$$

The function penalizes a lot for discontinuities between pixels of similar intensities, i.e. when $|I_p - I_q| < \sigma$. Conversely, the function encourages boundaries if adjacent pixels are very different, i.e. $|I_p - I_q| > \sigma$.

With the two terms defined, the graph-cut computes a binary labelling A as a global minimum of an energy function $E(A)$ defined as:

$$E(A) = \sum_{p \in \Omega} R_p(A_p) + \lambda \sum_{\{p,q\} \in \mathcal{N}} \delta(A_p, A_q) B(\{p,q\}), \tag{3.3}$$

where $A_p \in \{obj, bkg\}$ is a label assigned to a pixel p in the segmentation A, the function δ is the Kronecker delta and λ is a parameter defining relative importance of the per-pixel and the boundary term. The purpose of $\delta(A_p, A_q)$ is to restrict discontinuity penalties for cases where pixels p, q are assigned different labels, since it is not meaningful to penalize for discontinuity if discontinuity is not present.

Hard Constraints

Typically, regional properties of the object of interest in real examples are not distinct enough and thus the object cannot be segmented from the background without further
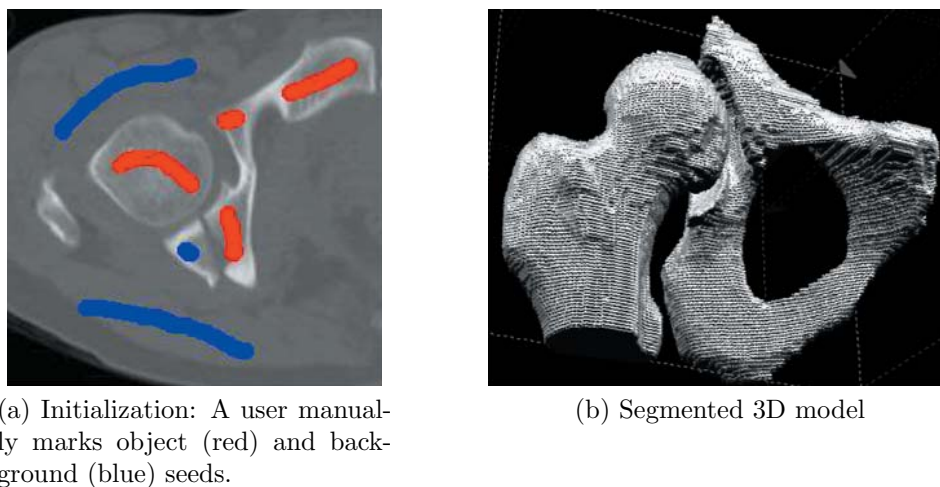
(a) Initialization: A user manually marks object (red) and background (blue) seeds.

(b) Segmented 3D model

Figure 3.1: The graph-cut algorithm applied for segmentation of bones from a 3D CT volume Boykov and Funka-Lea (2006). Note that only one slice (a) was used in the initialization.

constraining the search space of possible solutions before computing the optimal one (Boykov and Funka-Lea, 2006). The graph-cut framework provides flexible support for incorporating *prior* topological information to segmentations.

Assume that subsets $\mathcal{O}, \mathcal{B} \subset \mathcal{P}$ are *a priori* known to be part of the object and the background, respectively. The sets $\mathcal{O}, \mathcal{B}$ can be either determined manually as user-defined seeds or computed automatically in highly specialized applications such as organ segmentation from medical volumes. Incorporating this knowledge into a resulting segmentation means to restrict the search space for all possible segmentations with the following hard constraints:

$$\begin{aligned} \forall p \in \mathcal{O} : A_p = obj, \\ \forall p \in \mathcal{B} : A_p = bkg. \end{aligned} \tag{3.4}$$

Since $\mathcal{O} \cap \mathcal{B} = \emptyset$, the definition is correct. These hard constraints can be integrated into the per-pixel term as:

$$R_p(obj) = \begin{cases} 0 & \text{for } p \in \mathcal{O}, \\ K & \text{for } p \in \mathcal{B}, \end{cases} \qquad R_p(bkg) = \begin{cases} K & \text{for } p \in \mathcal{O}, \\ 0 & \text{for } p \in \mathcal{B}, \end{cases}$$

where K represents an infinitely large value. For the purpose of graph-cut segmentation, the value of K should be at least

$$K \geq 1 + \max_{p \in \mathcal{P}} \sum_{q:\{p,q\} \in \mathcal{N}} B(\{p, q\})$$

as suggested by Boykov and Funka-Lea (2006). The power of hard constraints is demonstrated in Figure 3.1, where the graph-cut framework was applied to segment the femur and the pelvis from a 3D-CT volume.
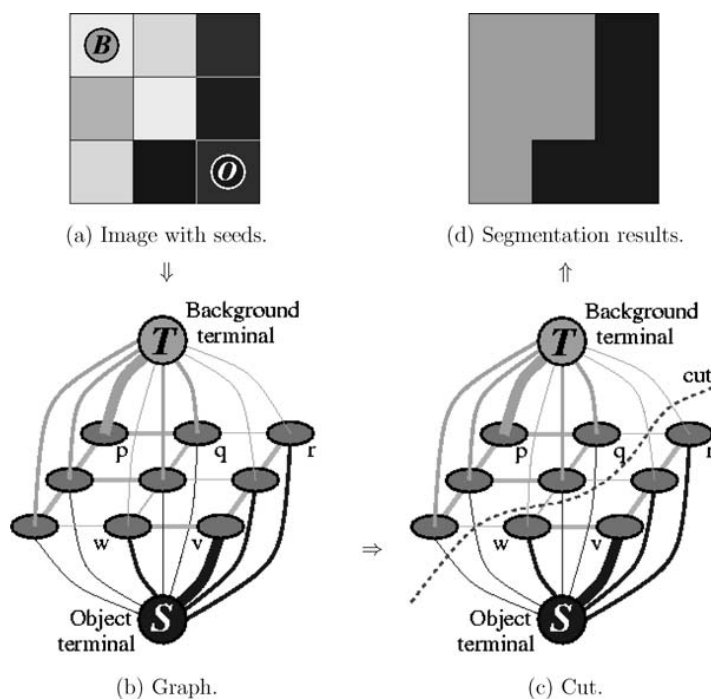
Figure 3.2: Typical work-flow of a graph-cut algorithm applied to a simple 3x3 image (Boykov and Funka-Lea, 2006). Detailed description is given in the text.

Optimal Solution via Minimal Graph Cuts

This section briefs algorithmic details behind the graph-cut segmentation method. For more details, discuss Boykov and Funka-Lea (2006).

Based on a set of pixels $\mathcal{P}$ within a neighboring system $\mathcal{N}$ and the components $R_p(\cdot)$ and $B(\{p, q\})$, a graph $G = (V, E)$ with weighted edges is constructed as follows:

$$V = \mathcal{P} \cup \{S, T\},$$
$$E = \mathcal{N} \bigcup_{p \in \mathcal{P}} \{\{p, S\}, \{p, T\}\},$$

where S and T are additional nodes not present in $\mathcal{P}$. The edge costs are defined according to Table 3.1.

Edge	Weight (cost)	For
$\{p, q\}$	$\lambda \cdot B(p, q)$	$\{p, q\} \in \mathcal{N}$
$\{p, S\}$	$R_p(bkg)$	$p \in \mathcal{N}$
$\{p, T\}$	$R_p(obj)$	$p \in \mathcal{N}$

Table 3.1: Edge weights of the graph used in the graph-cut segmentation.

Apart from nodes representing image pixels, the graph contains two additional nodes: the node S representing an *object* source terminal and the node T representing a *background* sink terminal. The neighborhood system and the boundary term is

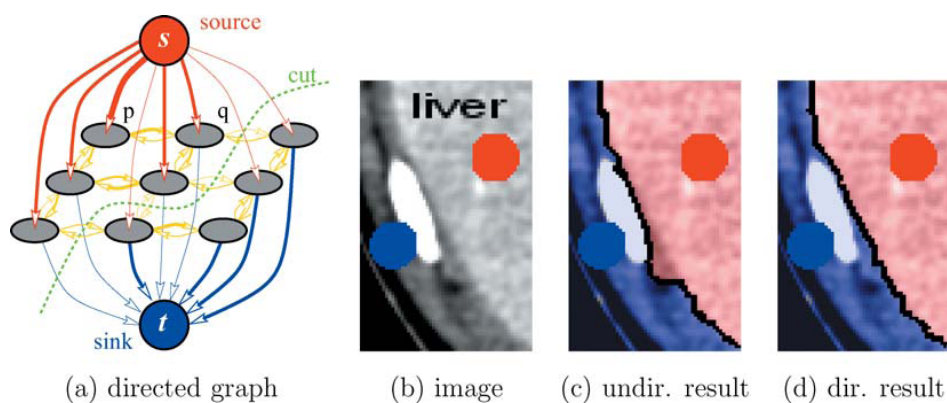(a) directed graph (b) image (c) undir. result (d) dir. result

Figure 3.3: Graph-cut segmentation with directed boundary term. User seeds are depicted as red (object) and blue (background) circles. Compare the results on an undirected graph (c) with the results on a directed graph (d) (Boykov and Funka-Lea, 2006).

captured in the graph by a set of edges $\mathcal{N}$, also called *n-links*. The per-pixel term is encoded in the graph via *t-links*, which connect each pixel to the sink and the source.

The segmentation boundary optimal with respect to Equation 3.3 can be determined by finding the minimum cost cut on the graph G (see Boykov and Funka-Lea (2006) for the technical proof). The cut divides the graph into two components – one connected to the source and one to the sink. The two components then represent the object and the background in the result segmentation, respectively (see Figure 3.2).

There are several well-established algorithms, which can be employed to compute the minimal cut efficiently in low-order polynomial time, e.g. Ford and Fulkerson (1962) or Goldberg and Tarjan (1988). Recently, an algorithm specialized for minimization of the energy function 3.3 has been introduced by Boykov and Kolmogorov (2004). As authors claim, their algorithm performs several times faster than the aforementioned methods for many cases of input images.

Using Directed Edges

Consider the medical image in Figure 3.3(b): a medium-bright object of interest (liver) is separated from a very bright bone by a thin layer of dark muscle tissue. Assume that the goal is to segment the liver. A user manually places object and background seeds as depicted in Figure 3.3(b) and defines components of the cost function according to Equation 3.1 and Equation 3.2. The graph-cut results in the segmentation depicted in Figure 3.3(c).

Visibly, the liver boundary is not precise, since the segmentation sticks to the bright bone instead of following the true liver boundary. Since the contrast between the bone and surrounding muscle tissue is very high, it is *cheaper* (the boundary term is lower) to lead the boundary along the bone rather than along the liver. To improve precision of the liver boundary, a user must add more seeds in the initialization step. Alternatively, the feature of directed edges offered by the graph-cut framework can be exploited to automatically correct this error:

So far, n-links of the graph G have been considered undirected; the value of $B(\{p, q\})$

represented a penalty of assigning pixels p, q two different labels. However, the graph can connect each pair of adjacent voxels p, q with two directed edges of weights $w_{(p,q)}$ and $w_{(q,p)}$, i.e.

$$E = \bigcup_{\{p,q\}\in\mathcal{N}} \{(p,q),(q,p)\} \bigcup_{p\in\mathcal{P}} \{\{p,S\},\{p,T\}\}.$$

If an optimal cut separates nodes p, q such that p is connected to the source and q is connected to the sink, then the cost of the minimum cut includes $w_{(p,q)}$ and ignores $w_{(q,p)}$. In the context of graph-cut segmentation, the directed boundary term $B(p,q)$ would represent a penalty of assigning p to the object and label q to the background. Formally, the *directed* version of the energy which is minimized by the graph-cut may be presented as:

$$E(A) = \sum_{p\in\Omega} R_p(A_p) + \lambda \sum_{(p,q)\in\mathcal{N}'} B(p,q)\cdot\delta_{A_p=obj,A_q=bkg},$$

where $\mathcal{N}'$ contains all ordered pairs (p,q) of neighboring pixels and the value of δ_x is one if the condition x is fulfilled and zero otherwise.

To achieve more precise liver boundaries, the directed boundary term may be exploited to specifically favor transitions from the object to the background in regions with decreasing intensity, e.g.:

$$B(p,q) \propto \begin{cases} 1 & \text{if } I_p \leq I_q, \\ \exp\left(-\dfrac{(I_p - I_q)^2}{2\sigma^2}\right) & \text{otherwise.} \end{cases}$$

This definition encourages transitions from the liver to background in areas where liver boundary goes from bright tissue to darker tissue. Therefore, the result segmentation (Figure 3.3(d)) does not stick to the bone, where the transition from liver to background goes from darker tissue to brighter tissue.

3.1.2 Sheetness measure

As detailed in Section 2.1.2, boundaries of cancellous bone in CT images are often weak and suffers from low-contrast. To enhance the contrast in these regions, behavior of a local iso-intensity region for each location of the input image may be analyzed. The filters based on first-order derivatives has been discussed in Section 3.1.2. This section describes analysis based on second-order derivatives.

Consider a 3D image $\mathcal{I} : \Omega \to \mathbb{R}$, a voxel of interest v and the Hessian matrix $\mathcal{H}_\sigma(v)$ defined as

$$\mathcal{H}_\sigma(v) = \begin{pmatrix} \mathcal{I}_{xx} & \mathcal{I}_{xy} & \mathcal{I}_{xz} \\ \mathcal{I}_{xy} & \mathcal{I}_{yy} & \mathcal{I}_{yz} \\ \mathcal{I}_{xz} & \mathcal{I}_{yz} & \mathcal{I}_{zz} \end{pmatrix},$$

where $\mathcal{I}_{ab}$ is a second-order derivative at v computed at scale σ along axes a and b. The eigenvectors and eigenvalues of the locally computed Hessian $\mathcal{H}_\sigma(v)$ correspond to the local principal directions and the principal region curvatures, respectively. Hence, according to the eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, local structures can be classified into
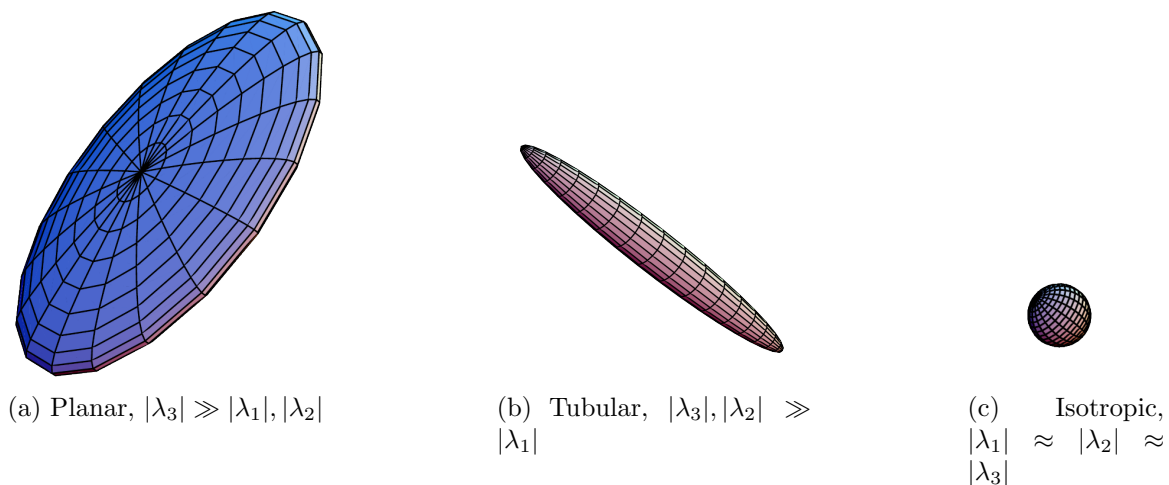
(a) Planar, $|\lambda_3| \gg |\lambda_1|, |\lambda_2|$ (b) Tubular, $|\lambda_3|, |\lambda_2| \gg |\lambda_1|$ (c) Isotropic, $|\lambda_1| \approx |\lambda_2| \approx |\lambda_3|$

Figure 3.4: Illustration of the local structure based on eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ of the Hessian matrix.

Eigenvalues	Local structure	Example
$|\lambda_1| \approx |\lambda_2| \approx 0, |\lambda_3| >> 0$	Sheet-like	bone,skin
$|\lambda_1| \approx 0, |\lambda_2| \approx |\lambda_3| >> 0$	Tube-like	vessels, nerves
$|\lambda_1| \approx |\lambda_2| \approx |\lambda_3| >> 0$	Blob-like	nodule
$|\lambda_1| \approx |\lambda_2| \approx |\lambda_3| \approx 0$	Noise-like	background, noise

Table 3.2: Local structure classification based on eigenvalue decomposition of the Hessian matrix.

sheet-like, tube-like and blob-like shape categories (Lorenz et al., 1997; Sato et al., 1998), as described in Table 3.2 and illustrated in Figure 3.4

Inspired by Westin et al. (1997) and Frangi et al. (1998), Descoteaux et al. (2005) incorporated the eigenvalues of the Hessian into a *sheetness score* $S(v)$, which for each voxel v quantifies similarity of a local iso-area to a planar structure. The score is defined as:

$$S_\sigma(v) = \exp\left\{-\frac{R_{sheet}^2}{\alpha^2}\right\} \exp\left\{-\frac{R_{tube}^2}{\beta^2}\right\} \left(1 - \exp\left\{-\frac{R_{noise}^2}{\gamma^2}\right\}\right), \qquad (3.5)$$

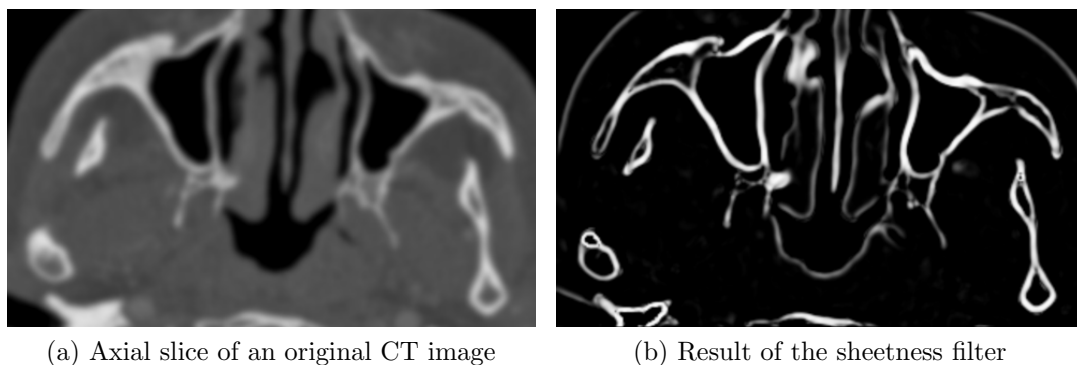where α, β, γ are constant parameters and the terms R_{sheet}, R_{tube} and R_{noise} are defined as

$$R_{sheet} = \frac{|\lambda_2|}{|\lambda_3|},$$

$$R_{tube} = \frac{|\lambda_1|}{|\lambda_2||\lambda_3|},$$

$$R_{noise} = \frac{|\lambda_1| + |\lambda_2| + |\lambda_3|}{T},$$

with T denoting the average trace of the Hessian at each image voxel. Each term in Equation 3.5 has a specific function.

(a) Axial slice of an original CT image   (b) Result of the sheetness filter

Figure 3.5: Parasinal bone structures enhanced with a multi-scale sheetness filter, proposed by Descoteaux et al. (2005).

- $\exp\left\{-\dfrac{R_{sheet}^2}{\alpha^2}\right\}$ is a sheet enhancement term. The maximum value occurs for sheet-like structures and the minimum for noise.

- $\exp\left\{-\dfrac{R_{tube}^2}{\beta^2}\right\}$ is a tube enhancement term. Similarly, the term reaches the maximum value for tube-like structures and diminishes for noise.

- $\left(1 - \exp\left\{-\dfrac{R_{noise}^2}{\gamma^2}\right\}\right)$ is a blob suppression term. The value is maximal for inhomogeneous regions and minimal for blob-like areas.

To increase the robustness of the sheetness score, Descoteaux et al. (2005) proposed to compute the Hessian matrix at different scales Σ and retaining, for each voxel, the maximum sheetness response:

$$S(v) = \max_{\sigma \in \Sigma} S_\sigma(v).$$

See Figure 3.5 for an example of a multi-scale sheetness filter applied to enhance parasinal bone structures.

Recently, similar versions of the sheetness filter have been i included into various segmentation pipelines targeting the CT modality. Harders et al. (2007) used the region growing after applying the sheetness filter. Fürnstahl et al. (2008) combined the sheetness score with intensity maps to initialize a slice-wise 2D graph-cut method. Fornaro et al. (2010) enhanced the filter to increase the contrast of true positives compared to false positives in the resultant image.

3.2 Preprocessing - Enhancement of bone boundaries

In the preprocessing step, the contrast at bone boundaries is enhanced. First, the input image $\mathcal{I}$ (Figure 3.6a) is enhanced by the unsharp masking (Figure 3.6b), defined as

$$\mathcal{I}^U = \mathcal{I} + k(\mathcal{I} - \mathcal{I} \star G_s),$$
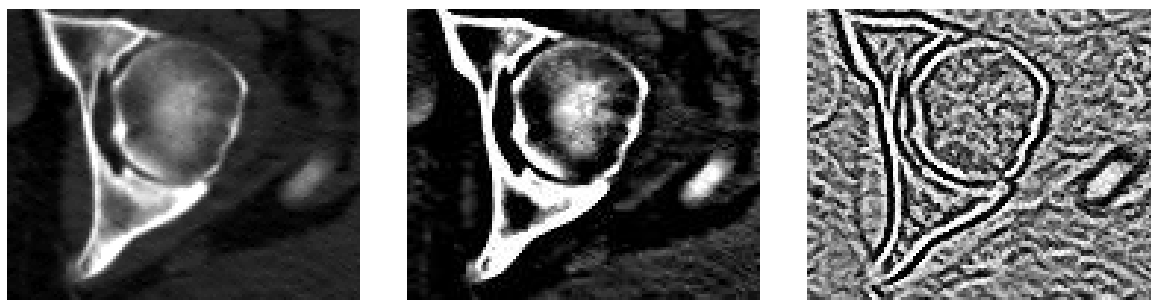
(a) Axial slice of an acetab-
ulofemoral joint

(b) Unsharp masking

(c) Sheetness filter

Figure 3.6: Preprocessing step: Two filters are successively applied to the original image (a) to enhance contrast at the bone boundaries: unsharp masking (b) and the sheetness measure (c).

where $\star$ denotes convolution, G_s a Gaussian kernel with variance s^2 and k a scaling constant.

Subsequently, the image is processed with a bone boundary enhancement filter. The filter is a modified version of the sheetness score, described in Section 3.1.2: The sign of the largest eigenvalue has been incorporated into the sheetness measure to further increase the contrast at bone boundaries:

$$S'_\sigma(v) = -\operatorname{sgn}(\lambda_3)S_\sigma(v). \tag{3.6}$$

Moreover, the multi-scale implementation has been altered to retain the response which corresponds to the highest absolute value:

$$S(v) = \operatorname*{argmax}_{S'_\sigma(v):\sigma \in \Sigma} |S'_\sigma(v)|.$$

The output of these filters (Figure 3.6c), denoted by S in the following, greatly enhances contrast at the bone boundaries. Note that the filter is especially efficient in narrow inter-bone regions, where it robustly separates two closely adjacent bones. In the following sections, the novel sheetness score is utilized for definition of the per-pixel and the boundary term used in the graph-cut segmentation.

3.3 Graph-cut bone segmentation

To segment all bone tissues from an input CT volume, the graph-cut framework utilizing the preprocessed volume (Section 3.2) is employed. The definition of the per-pixel term is given in Section 3.3.1. Section 3.3.2 covers definition of the boundary term. Since CT volumes are regular 3D raster grids, the graph-cut uses standard 6-neighborhood system to capture adjacency information.

3.3.1 Per-Pixel term

As detailed in Section 3.1.1, the per-pixel term $R_p(A_p)$ reflects a penalty for assigning the voxel p the label A_p. Therefore, the cost should be

(a) Background exclusion region $E_{\neg bkg}$
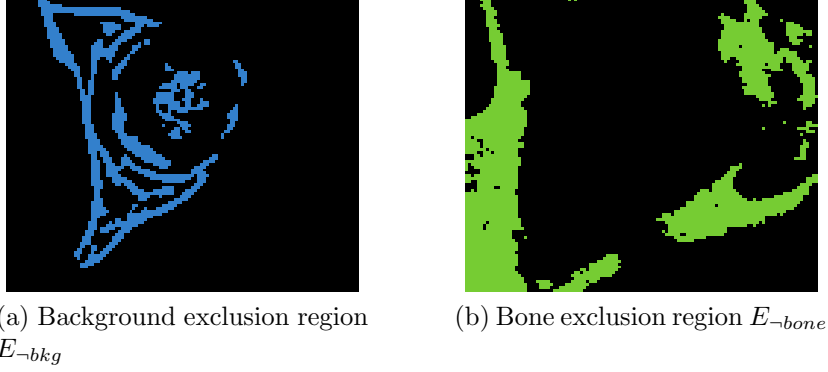

(b) Bone exclusion region $E_{\neg bone}$

Figure 3.7: Bone and background exclusion regions used for the graph-cut initialization.

- low if p belongs to a bone and $A_p = bone$,

- low if p belongs to background and $A_p = bkg$,

- otherwise, the cost should be high.

Weak bone boundaries, narrow inter-bone space and low intensities in the trabecular bone make image intensity alone a relatively poor feature for discriminating bone from background. Nevertheless, exclusion regions $E_{\neg bkg}$ and $E_{\neg bone}$ (Figure 3.7) can safely be estimated as:

$$E_{\neg bkg} = \{x \in \Omega | \mathcal{I}(x) \geq 400 \text{ HU} \wedge S(x) > 0\},$$
$$E_{\neg bone} = \mathrm{lcc}(\{x \in \Omega | \mathcal{I}(x) < -50 \text{ HU}\}),$$

where lcc denotes the largest connected component of the binary argument. The term $\mathcal{I}(x) \geq 400$ HU selects high-intensity voxels present mostly in cortical bone while $S(x) > 0$ discards soft-tissue voxels in narrow inter-bone areas. The term $\{x | \mathcal{I}(x) < -50 \text{ HU}\}$ retains mostly fat and air voxels of low intensity values, and the lcc operation discards low-intensity voxels encountered in the bone interior.

Based on the computed exclusion regions, the per-pixel term of the graph-cut cost function (Equation 3.3) is defined as

$$R_p(A_p) \propto \begin{cases} 1 & \text{if } A_p = bone \text{ and } p \in E_{\neg bone}, \\ 1 & \text{if } A_p = bkg \text{ and } p \in E_{\neg bkg}, \\ 0 & \text{otherwise.} \end{cases}$$

This initial classification can be considered as the initialization of the graph-cut segmentation. Since the value of the per-pixel term is 0 for all voxels outside $E_{\neg bkg} \cap E_{\neg bone}$, it is the boundary term that arbitrates on the classification of these voxels.

3.3.2   Boundary term

To encourage spatial coherency and to automatically achieve more precise bone boundaries, the proposed graph-cut segmentation exploits the directed version of the boundary term. As detailed in Section 3.1.1, the value of $B(p, q)$ represents a penalty for
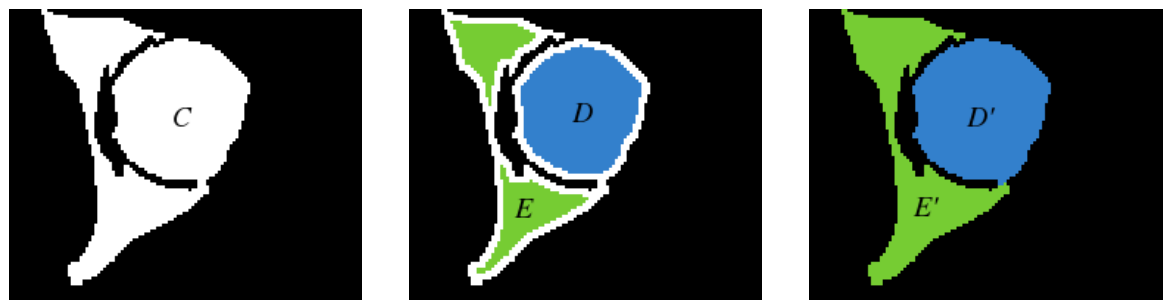
(a) Input: all bone tissues correctly segmented

(b) Individual bones identified using morphological erosion.

(c) Final results after applying a simplified graph-cut.

Figure 3.8: Postprocessing step: Separation of adjacent bones

assigning the voxel p to bone and the voxel q to the background. Employing the modified sheetness measure S, introduced in Section 3.2, the boundary term is defined as

$$B(p,q) \propto \begin{cases} \exp\left\{ -\dfrac{|S(p) - S(q)|}{\sigma_s} \right\} & \text{for } S(p) \geq S(q), \\ 1 & \text{otherwise,} \end{cases}$$

where σ_s is a constant scaling parameter. This definition encourages boundaries in regions with abrupt variations of the sheetness measure. Moreover, by exploiting directed edges of the underlying combinatorial graph, transition from bone to background is encouraged in regions with decreasing sheetness score.

3.4   Postprocessing - Separation of adjacent bones

Applying the graph-cut method described in section 3.3 results in binary images similar to Figure 3.8a. The method correctly discriminates voxels of bone tissues from background. However, in many cases, the method cannot properly separate two closely adjacent bones due to narrow low-contrast inter-bone regions. Albeit the sheetness filter greatly enhances contrast in the joint channels, adjacent bones often remains connected with several voxels in the output segmentation. This is a typical case for the acetabulofemoral joint at the upper femur extremity and for the knee-joint at the lower extremity, where the femur appears to be in direct contact with adjacent bones. The goal of the postprocessing step is to automatically determine if such a bottleneck exists and possibly identify individual bones within a segmented component.

To determine existence of a bottleneck, a morphological erosion (Haralick et al., 1987) with a spherical element of radius R is applied to the binary image (Fig. 3.8b). If the erosion divides a component $C \subset \Omega$ into two subcomponents $D, E \subset C$, a bottleneck has to be present in C. Otherwise, the component C correctly represents a bone or the error is too large to be detected.

After the bottleneck is detected, the next step is to find disjoint parts $D', E' \subset C$ such that $D' \cup E' = C, D \subset D', E \subset E'$ and the number of voxels on the boundary between both sets is minimal. To accurately find D', E', a simplified graph-cut is

employed:

Instead of Ω, only the part C of a volume is used as a domain for the simplified graph-cut, with a standard 6-neighborhood system. The hard-constraints $D \subset D', E \subset E'$ are imposed on the classification via the per-pixel term, as described in subsection 3.1.1:

$$\forall p \in C : \quad R_p(A_p) \propto \begin{cases} \infty & \text{if } A_p = \text{"D" and } p \in E, \\ \infty & \text{if } A_p = \text{"E" and } p \in D, \\ 0 & \text{otherwise.} \end{cases}$$

To encourage minimal boundary between D' and E', the symmetric boundary term

$$B(\{p, q\}) = 1$$

is used. This procedure can be employed to either segment all bones in the image by assigning them different labels or to extract only one of them (Fig. 3.8c).

Chapter 4

Experiments

To visually and quantitatively evaluate performance, the proposed method was tested and compared to other three fully-automatic bone segmentation approaches in a large-scale experiment conducted on a set of 197 femur samples. Section 4.1 describes the sample set in more detail, introduces the quantitative measures recorded for each sample, presents the three alternative approaches and lists concrete parameter settings for each method. Section 4.2 expands on the results and discusses impact of each specific aspect of the proposed method to the challenges coupled with the femur segmentation.

4.1 Experiment design

The method was evaluated on a dataset comprising 197 CT volumes. The scans were cropped around the femur, with voxel intensities expressed in Hounsfield units. The volume spacing of the scans ranged from 0.6 to 1.17mm in-plane and from 0.8 to 1.25mm inter-slice. The slice resolution ranged from 83x83 to 260x289 pixels and the number of slices varied from 314 to 658. For all datasets, manual segmentation by a medical expert was available and served as a reference for quantitative evaluations.

All volumes were processed automatically in a batch mode with a pilot application (Appendix A) which implements the proposed method. The output of the application is a multi-label volume where voxels of zero value correspond to the background and voxels of positive value represent bone tissue. Individual bones were numbered in ascending order according to their size from the largest bone to the smallest. It was assumed that the largest bone (of label one) corresponds to the femur.

For each sample, the following measures were recorded:

- True Positive Rate

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}},$$

- False Positive Rate

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}},$$

- Hausdorff distance (HD) between the estimated (X) and the reference (Y) surface:

$$d_{\mathrm{H}}(X, Y) = \max\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \}, \qquad (4.1)$$
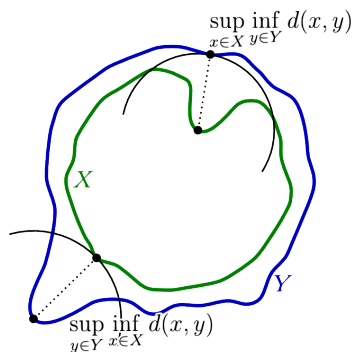
Figure 4.1: Components of the calculation of the Hausdorff distance (Equation 4.1) between the green surface X and the blue surface Y (wikipedia.org).

- the processing time and

- the memory consumption.

The terms TP and TN denote number of correctly classified bone and background voxels, respectively. Similarly, terms FP and FN denote number of misclassified bone and background voxels, respectively. Since the term TP+FN represents true size of the bone according to the ground truth, the TPR ratio measures size of correctly segmented bone with respect to the size of the whole bone. Conversely, the FPR rate measures the ratio of number of misclassified bone voxels to true background size. The Hausdorff distance (Figure 4.1) represents the maximum distance error between the segmented and the reference surface. Generally, a good segmentation algorithm achieves TPR close to one (bone interior segmented) and FPR and HD close to zero (no leakages to the background).

The proposed method was compared to other three state-of-the-art existing fully-automatic 3D-based bone segmentation approaches, none of which employs any prior shape knowledge:

Gradient-based Geometric Active Contour (GeomAC) The geometric active contour method is detailed in subsection 2.2.6. The method has been implemented with the following settings: the initial level-set initialized as the signed distance of the binary image $E_{\neg bkg}$, propagation parameter 5.0, curvature parameter 1.0 and max. number of iterations 400.

Zhang iterative adaptive thresholding (ZIAT) The method proposed by Zhang et al. (2010) first performs binary thresholding with a manually chosen global threshold and then gradually increases quality of the segmentation by local iterative steps. The method has been implemented with the window size of 15x15x3 pixels and the initial threshold value of 70 HU.

Intensity-based graph-cut method (IBGC) For IBGC, the graph-cut algorithm

was used with the following terms:

$$R_p(bone) = \begin{cases} 1 & \text{for } p \in E_{\neg bone}, \\ 0 & \text{otherwise}, \end{cases}$$

$$R_p(bkg) = \begin{cases} 1 & \text{for } \mathcal{I}(p) > 500 \text{ HU}, \\ 0 & \text{otherwise}, \end{cases}$$

$$B(\{p, q\}) = \exp\left\{\frac{|\mathcal{I}(p) - \mathcal{I}(q)|}{100}\right\}.$$

Based on the image gradient only, it is expected that the IBGC method will produce leakages due to weak and diffused bone boundaries and will not be able to separate bones which appear to be in direct contact.

Parameters of all methods were chosen experimentally and represent settings which, according to visual evaluation, produce the most satisfying segmentation results. The parameters for the proposed method were also manually selected (see Table 4.1) and kept constant for all datasets.

Parameter	Description	Value
k	Scaling constant of the unsharp masking	10
s^2	Variance of the Gaussian kernel in the unsharp masking	1
α	Scaling parameter of the sheetness measure	0.5
β	Scaling parameter of the sheetness measure	0.5
γ	Scaling parameter of the sheetness measure	0.25
Σ	Scales for the multi-scale sheetness measure	$\{0.75, 1.0\}$
λ	Relative importance of the per-pixel and the boundary term	5
σ_s	Scaling parameter of the boundary term	0.2
R	Radius of the spherical element used in the morphologic erosion	3

Table 4.1: Parameters of the pilot graph-cut application.

4.2   Results

Based on visual evaluation of 197 results, it can be concluded that the method provides satisfactory segmentation results in most of the cases (see Figure 4.2) and robustly handles segmentation challenges coupled with the CT imaging:

The problem of large intensity difference between the cortical layer and the bone interior in the shaft area is eliminated due the definition of the exclusion region $E_{\neg bone}$ employed in the per-pixel term, especially due to the operation of largest connected component. Via the definition of $E_{\neg bkg}$, the per-pixel term also alleviates the issue of high intensity variance within the vascular cancellous tissues. The segmentation in these regions is robustly handled also due the bias of the graph-cut towards smaller object boundaries. Moreover, the boundary term based on the sheetness score visibly handles the challenge of weak and diffused bone boundaries. Also, the directed boundary term increases precision of bone boundaries.

In general, the method segmented the femur correctly because the graph-cut is initialized with an accurate estimation of true bone tissues and true background tissues and the framework employs the sheetness score to precisely delineate the bone boundary. Moreover, the sheetness score in the definition of $E_{\neg bkg}$ minimizes the number of false positives in the narrow inter-bone regions. With this support, the postprocessing step is able to automatically identify individual bones and separate the femur accordingly.

However, the method did not produce satisfactory results in several cases (see Figure 4.3), where major parts of bone boundaries were missing or two adjacent bones seems to touch each other in larger areas. In particular, this is a case for elderly patients, where the inter-bone channels in the acetabulofemoral joint may shrunk. Due to the volumetric effect present in CT images, the acetabulum part of the pelvis and the femur appear to be in touch, which makes femur segmentation without any shape prior knowledge hardly feasible.

The comparative results (Figure 4.4 and Figure 4.5) show that all tested approaches, except ZIAT, are able to correctly discriminate the true femur voxels with satisfactory precision (TPR > 0.85). Due to the low contrast in narrow inter-bone channels and weak bone boundaries, leakage (GeomAC, IBGC, ZIAT) and contour discontinuity (ZIAT) always occurred in segmentations from these three methods. None of them could separate the femur accurately (HD > 5cm), nor segment the challenging weak bone contours (visual evaluation).

Our approach is clearly superior in detecting bone boundaries in narrow inter-bone regions and in suppressing leakage into adjacent bones. The proposed method correctly segmented and separated the femur from adjacent bones (TPR > 0.85, FPR < 0.001, HD < 8mm) in 81% of the cases, with average HD of 5.4mm within this set. The postprocessing procedure (Section 3.4) separated the femur from at least one neighboring bone in 57% of cases.

In comparison with ZIAT and GeomAC, the graph-cut-based approaches have also proven to be much faster, requiring only a few minutes per datasets. Conversely, the memory requirements of the graph-cut library are much higher.
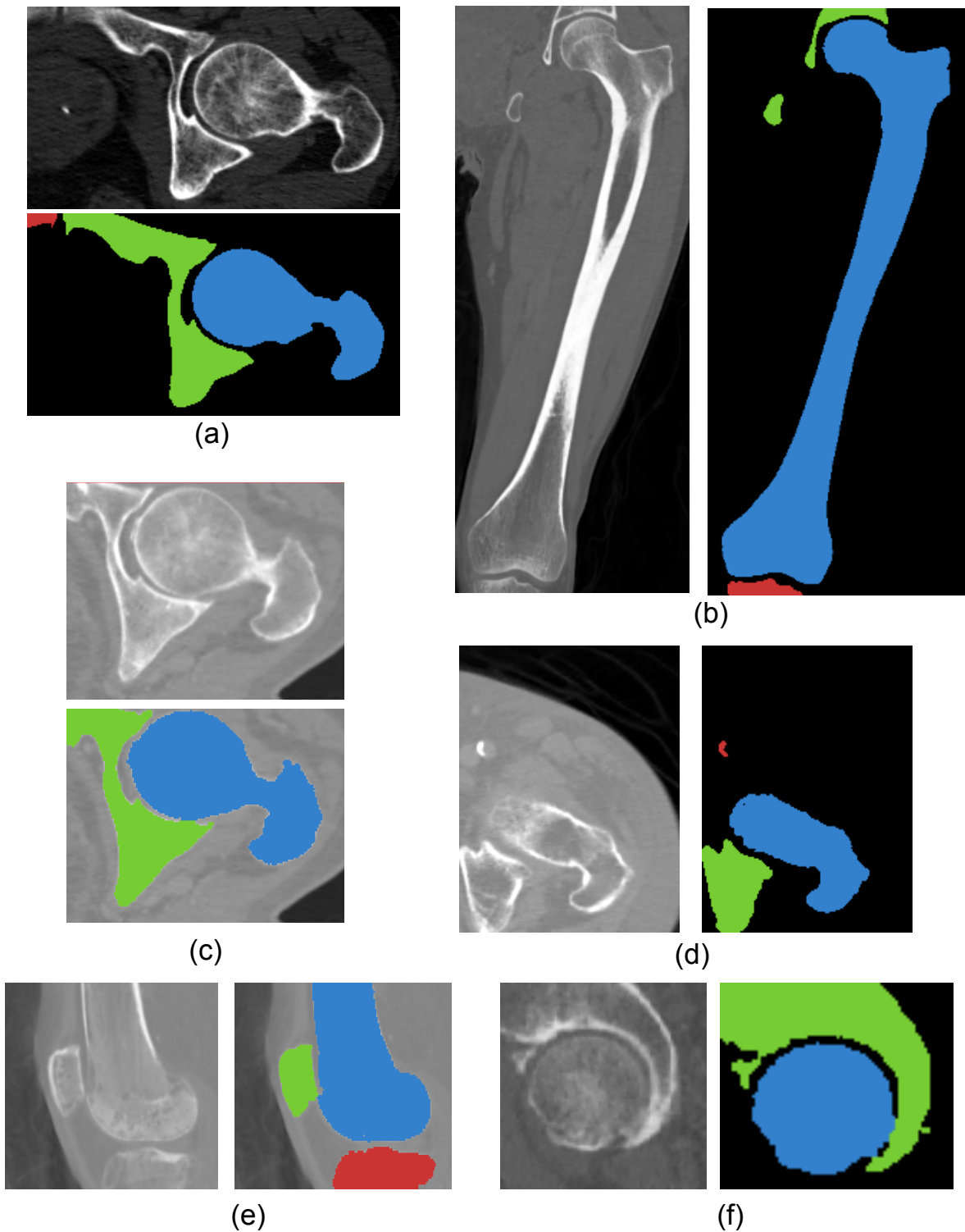
Figure 4.2: Satisfactory results produced with the proposed method. The method robustly handles high density variability of cancellous bone (a) as well as different intensity distributions between the shaft and joint regions (b). Due to the bone boundary enhancement filter, the bone surface is segmented well without any significant discontinuity (d,e). The postprocessing step is able to separate adjacent bones which appear to be in direct contact (c,e,f).
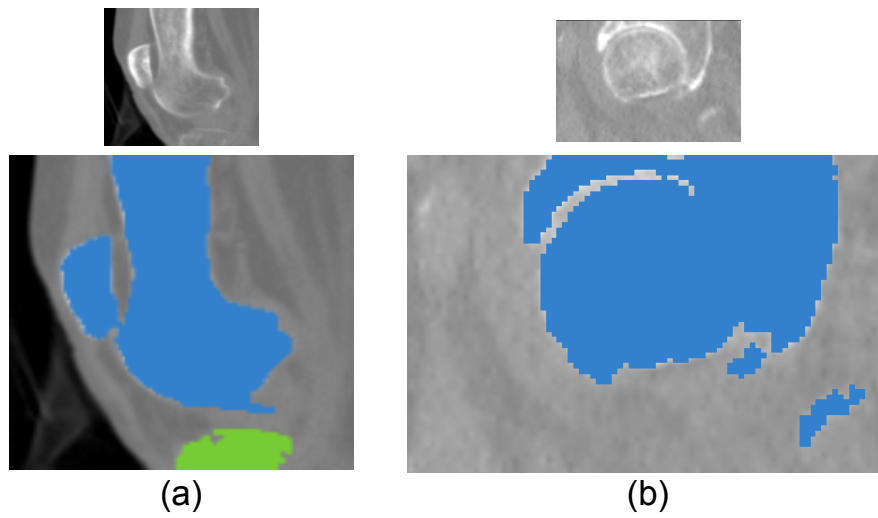
Figure 4.3: Unsatisfactory results produced by the proposed method.

(a) Original image

(b) Reference
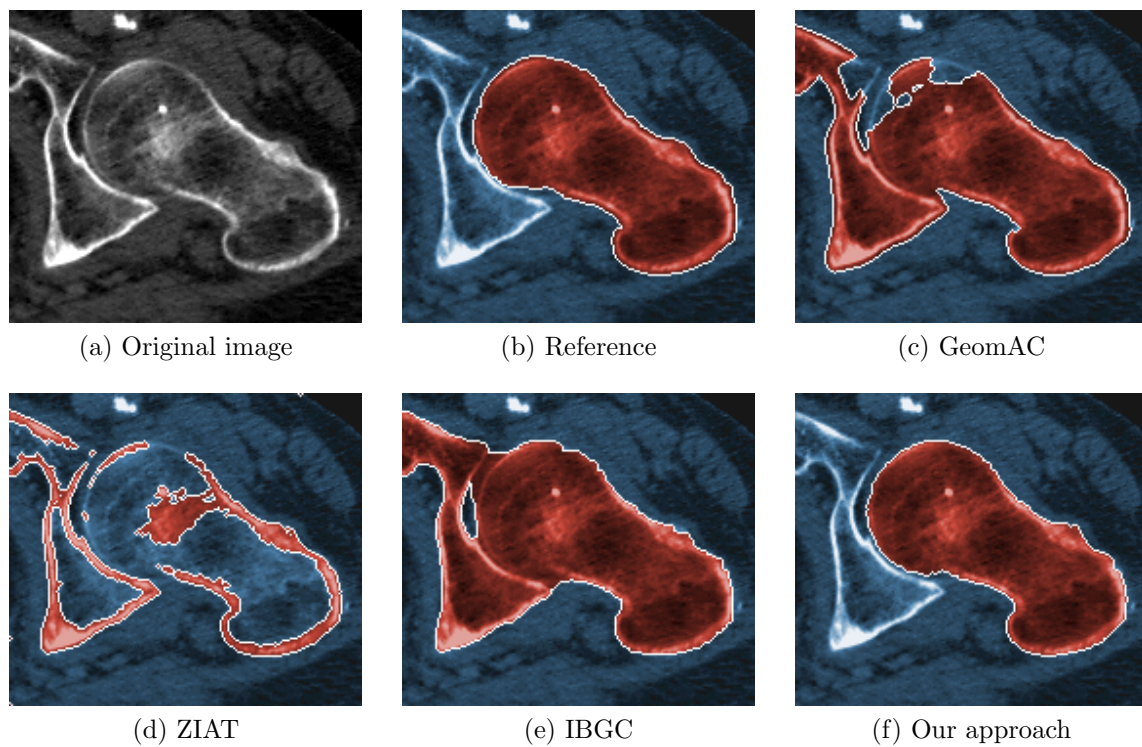
(c) GeomAC

(d) ZIAT

(e) IBGC

(f) Our approach

Figure 4.4: Segmentation results of four tested methods illustrated on axial slice of the acetabulofemoral joint.
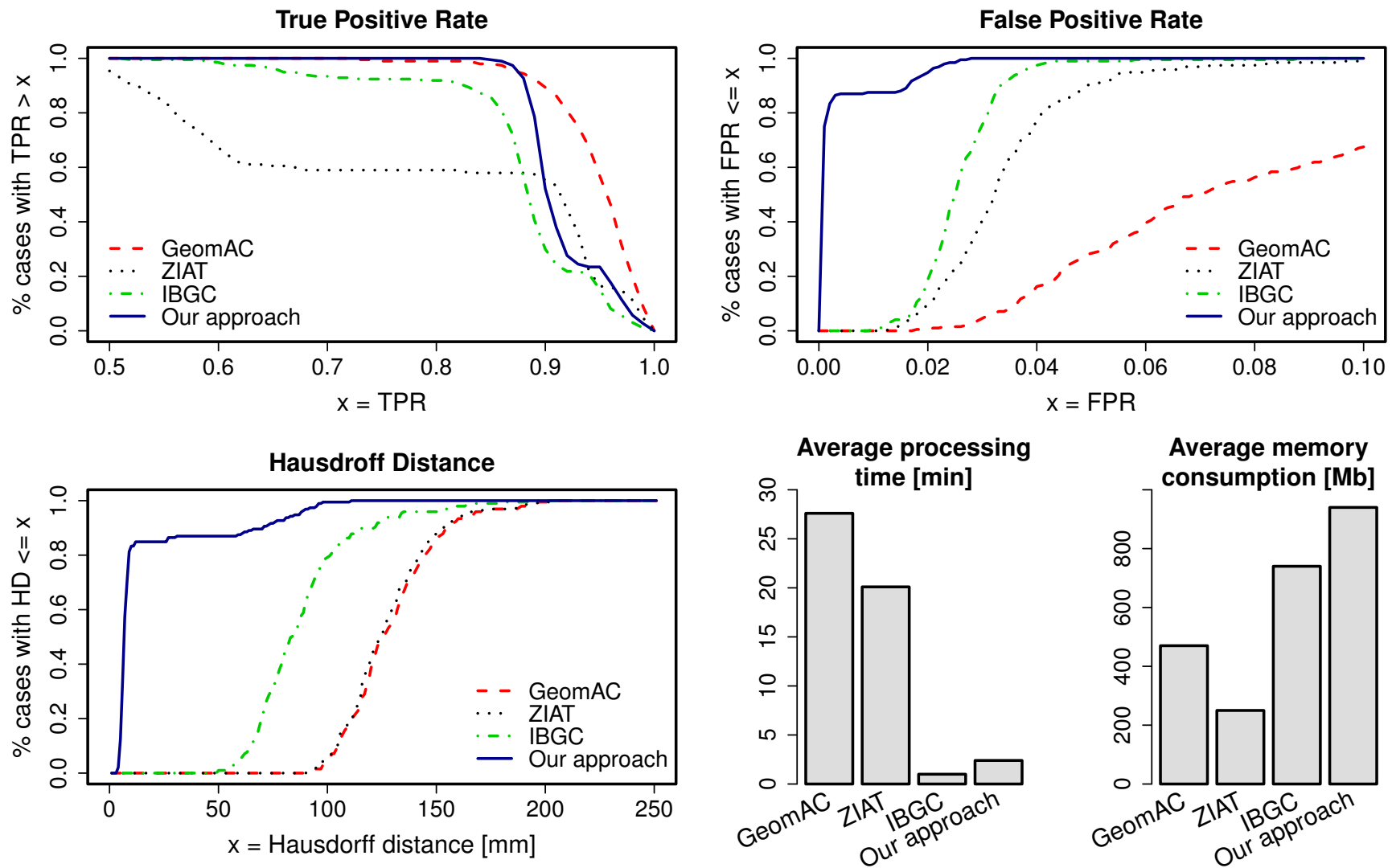
Figure 4.5: Quantitative comparison of the proposed method with three other fully-automatic approaches based on segmentation of 197 samples.

Chapter 5

Conclusion

Accurate and automatic segmentation techniques that do not require any explicit prior model have been of high interest in the biomedical imaging community. Despite several years of active research, bone segmentation from 3D Computed Tomography (CT) scans remains in several aspects an open problem, due to inherent challenges present in the CT modality.

The first contribution of this thesis is a summary of established methods for segmentation of bones from medical images. The review details the low-level approaches, incl. thresholding, region growing, watershed transform and local structure analyses, as well as high-level, more sophisticated, algorithms such as snakes, geometric active contours, statistical shape models and the graph-cut framework. For each method, the basic idea is explained and the applicability to bone segmentation is discussed. The review also contains references to further extensions and summaries, which focus on specific families of segmentation algorithms.

However, the main contribution is the proposal of a novel method for automatic femur segmentation from CT volumes. The work introduces a bone boundary enhancement filter based on the second-order analysis of local iso-intensity regions. Utilizing an improved version of the *sheetness score*, this filter greatly increases contrast at the bone boundary and effectively attacks the problem of narrow inter-bone spacing. Based on the modern graph-cut framework, the method has been optimized and designed to segment all bone tissues within the input CT volume without any shape prior knowledge. Also, the method features a postprocessing step automatically distinguishing adjacent bones which appear to be in direct contact in a CT scan.

The third contribution of the study is an extensive evaluation of the presented method. To compare performance, three other fully-automatic state-of-the-art methods were chosen and implemented. The four approaches were tested on a set of 197 femur CT volumes and compared both visually and quantitatively. Results show that out of the four methods, the proposed algorithm achieved most accurate results and segmented the femur correctly in 81% of the cases. Other methods would require manual postprocessing to achieve comparable results.

In summary, the goals of this study have been fulfilled: the thesis proposes a method which robustly segment the femur from a 3D-CT volumes without any shape prior knowledge and human interaction. Results of this study were also accepted and pre-
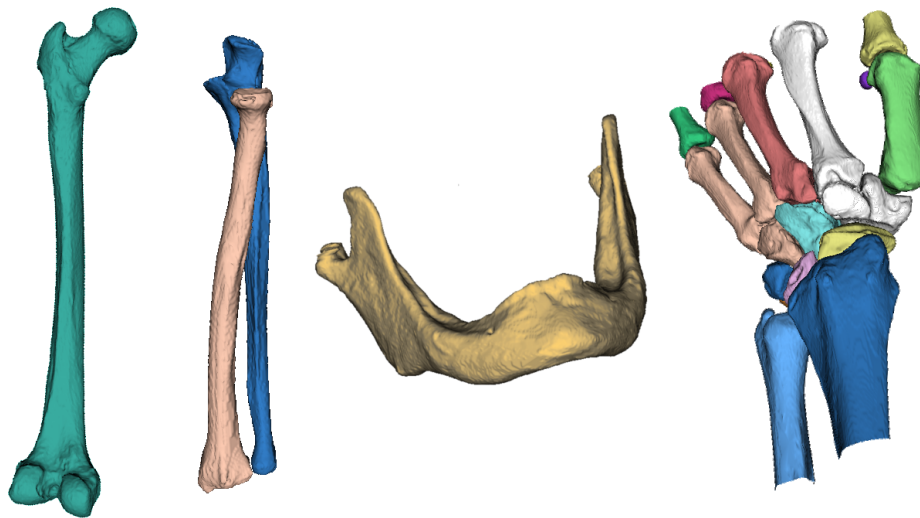
Figure 5.1: 3D models generated from an automatically segmented femur, radius and ulna, mandible and carpal bones (from left to right).

sented on The IEEE International Symposium on Biomedical Imaging, April 2011, held in Chicago, USA. The corresponding publication can be found in Appendix A.

By design, the proposed method does not require any explicit shape prior knowledge. Therefore, it is potentially suitable for segmenting other bones from CT images. Indeed, preliminary experiments on forearms, (toothless) mandibles and wrists showed promising results (Figure 5.1) without any parameter tuning.

As a consequence, the proposed method is suitable for bootstrapping the construction of bone statistical shape models. These models integrate the variability of shape and appearance within a population and have been of particular interest in the past decades. However, construction of such models requires a large set of pre-segmented samples. This remains a serious bottleneck due to the lack of accurate and fully automatic segmentation methods. Manual processing, which is often tedious, remains a common practice. Therefore, algorithms providing an automated mechanism for acquisition of initial samples are of high interest.

Future work aims at embedding the proposed automatic segmentation tool into a statistical model generation approach, by iteratively improving the model and refining the segmentations. In this manner, the method can potentially segment the bone of interest from more challenging cases.

Moreover, demand for specialized max-flow algorithms for massive large-scale N-D grids grows rapidly in the computer vision community. Several recent publications already target the problem of high-memory graph-cut requirements (e.g. Delong and Boykov (2008)). Replacement of the Boykov-Kolmogorov graph-cut library with a more specialized approach can substantially improve performance of the proposed segmentation method and is therefore also planned as future work.

Appendix A

Attached CD

The CD attached to this thesis contains:

thesis.pdf	Document containing electronic version of this thesis.
ISBI11.pdf	Publication regarding the proposed method, accepted to IEEE International Symposium of Biomedical Imaging: From Nano to Macro, April 2011, Chicago, USA
src/	Folder containing source files of the four tested segmentation methods.
sample-volumes/	Folder containing several sample CT volumes with segmented results.

All applications were developed in C++ using the ITK library (www.itk.org) for general image processing. Compile the programs using the CMake utility (www.cmake.org). The target platform is Linux x86. Due to legal restrictions, the full set of 197 CT femur samples could not be included in the CD. Sample CT volumes can be viewed using e.g. the free application 3D Slicer (www.slicer.org).

Bibliography

R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994. 12

Y. Boykov and G. Funka-Lea. Graph cuts and efficient nd image segmentation. *International Journal of Computer Vision*, 70(2):109–131, 2006. 23, 24, 25, 26, 27

Y. Boykov and M.P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in ND images. In *International Conference on Computer Vision*, volume 1, pages 105–112. Citeseer, 2001. 5, 23, 24

Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1124–1137, 2004. 27

V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66(1):1–31, 1993. 17

V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International journal of computer vision*, 22(1):61–79, 1997. 17

T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277, 2001. 18

L.D. Cohen. On active contour models and balloons. *CVGIP: Image understanding*, 53(2):211–218, 1991. 15

T.F. Cootes, C.J. Taylor, DH Cooper, and J. Graham. Training models of shape from sets of examples. In *Proc. British Machine Vision Conference*, volume 557. Citeseer, 1992. 18, 21

T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, et al. Active shape models-their training and application. *Computer vision and image understanding*, 61(1):38–59, 1995. 21

TF Cootes, GJ Edwards, and CJ Taylor. Active appearance models. *Computer Vision—ECCV'98*, page 484, 1998. 21

T.F. Cootes, C.J. Taylor, et al. Statistical models of appearance for computer vision. 2004. 18

D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International Journal of Computer Vision*, 72(2):195–215, 2007. 15, 16, 18

M. De Bruijne and M. Nielsen. Multi-object segmentation using shape particles. In *Information Processing in Medical Imaging*, pages 762–773. Springer, 2005. 21

A. Delong and Y. Boykov. A scalable graph-cut algorithm for nd grids. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008. 43

M. Descoteaux, M. Audette, K. Chinzei, and K. Siddiqi. Bone enhancement filtering: Application to sinus bone segmentation and simulation of pituitary surgery. *Lecture Notes in Computer Science*, 3749:9, 2005. 29, 30

I.L. Dryden and K.V. Mardia. *Statistical shape analysis*. Wiley New York, 1998. 19

O. Ecabert, J. Peters, H. Schramm, C. Lorenz, J. Von Berg, MJ Walker, M. Vembar, ME Olszewski, K. Subramanyan, G. Lavi, et al. Automatic model-based segmentation of the heart in CT images. *IEEE Transactions on Medical Imaging*, 27(9): 1189–1201, 2008. 21

LR Ford and DR Fulkerson. Flows in networks. 1962. 27

J. Fornaro, G. Székely, and M. Harders. Semi-automatic Segmentation of Fractured Pelvic Bones for Surgical Planning. *Biomedical Simulation*, 5958:82–89, 2010. 30

A.F. Frangi, W.J. Niessen, K.L. Vincken, and M.A. Viergever. Multiscale vessel enhancement filtering. *Lecture Notes in Computer Science*, pages 130–137, 1998. 29

J. Fripp, S. Crozier, S. Warfield, and S. Ourselin. Automatic Initialization of 3D Deformable Models for Cartilage Segmentation. In *Digital Image Computing: Techniques and Applications, 2005. DICTA'05. Proceedings 2005*, pages 74–74, 2005. 21

J. Fripp, SK Warfield, S. Crozier, and S. Ourselin. Automatic segmentation of the knee bones using 3D active shape models. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 1, 2006. 21

P. Fürnstahl, T. Fuchs, A. Schweizer, L. Nagy, G. Székely, and M. Harders. Automatic and robust forearm segmentation using graph cuts. *Biomedical Imaging: From Nano to Macro*, 2008. 5, 30

Philipp Fürnstahl. *Computer Assisted Planning for Orthopedic Surgery*. PhD thesis, ETH Zürich, Switzerland, 2009. 11, 12

A.V. Goldberg and R.E. Tarjan. A new approach to the maximum-flow problem. *Journal of the ACM (JACM)*, 35(4):921–940, 1988. 27

C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991. 19

J.C. Gower. Generalized procrustes analysis. *Psychometrika*, 40(1):33–51, 1975. 19

DM Greig, BT Porteous, and AH Seheult. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(2):271–279, 1989. 23

H.K. Hahn. *Morphological Volumetry: Theory, Concepts, and Application to Quantitative Medical Imaging*. PhD thesis, Staats-und Universitätsbibliothek Bremen, 2005. 14

R.M. Haralick, S.R. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (4): 532–550, 1987. 9, 33

M. Harders, A. Barlit, C. Gerber, J. Hodler, and G. Székely. An optimized surgical planning environment for complex proximal humerus fractures. In *MICCAI Workshop on Interaction in Medical Image Analysis and Visualization*, 2007. 30

L. He, Z. Peng, B. Everding, X. Wang, C.Y. Han, K.L. Weiss, and W.G. Wee. A comparative study of deformable contour methods on medical image segmentation. *Image and Vision Computing*, 26(2):141–163, 2008. 15, 18

T. Heimann, I. Wolf, T. Williams, and H.P. Meinzer. 3D active shape models using gradient descent optimization of description length. In *Information Processing in Medical Imaging*, pages 566–577. Springer, 2005. 21

Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis*, 13(4):543 – 563, 2009. 5, 18, 19, 21

A. Hill and CJ Taylor. Model-based image interpretation using genetic algorithms. *Image and Vision Computing*, 10(5):295–300, 1992. 21

IT Jolliffe. *Principal component analysis*. Springer verlag, 2002. 19

K. Josephson, A. Ericsson, and J. Karlsson. Segmentation of medical images using three-dimensional active shape models. *Image Analysis*, pages 719–728. 21

M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988. 14

H. Lamecker, M. Seebass, H.C. Hege, and P. Deuflhard. A 3D statistical shape model of the pelvic bone for segmentation. In *Proceedings of SPIE*, volume 5370, page 1341. Citeseer, 2004. 20, 21

S.A. Lipson. Mdct and 3d workstations. 2006. 7

C. Lorenz, IC Carlsen, T.M. Buzug, C. Fassnacht, and J. Weese. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. *Lecture Notes in Computer Science*, pages 233–242, 1997. 29

Z. Maa, J.M.R.S. Tavaresa, R.N. Jorgea, and T. Mascarenhasb. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Computer Methods in Biomechanics and Biomedical Engineering*, (1):1–1, 2009. 15, 17

R. Malladi, J.A. Sethian, B.C. Vemuri, et al. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995. 15, 17

T. McInerney and D. Terzopoulos. Deformable models in medical image analysis. In *Mathematical Methods in Biomedical Image Analysis, 1996., Proceedings of the Workshop on*, pages 171–180, 1996. 15

B. Modayur, J. Prothero, G. Ojemann, K. Maravilla, and J. Brinkley. Visualization-based mapping of language function in the brain. *Neuroimage*, 6(4):245–258, 1997. 13

E.N. Mortensen and W.A. Barrett. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60(5):349–384, 1998. 24

D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math*, 42(5):577–685, 1989. 18

A. Neubauer, K. Bühler, R. Wegenkittl, A. Rauchberger, and M. Rieger. Advanced virtual corrective osteotomy. In *International Congress Series*, volume 1281, pages 684–689. Elsevier, 2005. 14

S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988. 15

N. Paragios, O. Mellina-Gottardo, and V. Ramesh. Gradient vector flow fast geometric active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):402–407, 2004. 18

A. Pitiot, A.W. Toga, and P.M. Thompson. Adaptive elastic segmentation of brain MRI via shape-model-guided evolutionary programming. *IEEE transactions on medical imaging*, 21(8):910–923, 2002. 21

PK Sahoo, S. Soltani, AKC Wong, and YC Chen. A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2):233–260, 1988. 12

Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig, and R. Kikinis. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Medical Image Analysis*, 2(2): 143–168, 1998. 29

T.B. Sebastian, H. Tek, J.J. Crisco, and B.B. Kimia. Segmentation of carpal bones from CT images using skeletally coupled deformable models. *Medical Image Analysis*, 7(1):21–45, 2003. 11, 12, 13, 14, 18

M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–168, 2004. 12

K. Siddiqi, YB Lauziere, A. Tannenbaum, and SW Zucker. Area and length minimizing flows for shape segmentation. *IEEE Transactions on Image Processing*, 7(3):433–443, 1998. 18

L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, et al. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142, 2001. 21

M.B. Stegmann, R. Fisker, and BK Ersboll. Extending and applying active appearance models for automated, high precision segmentation in different image modalities. In *Proceedings of The Scandinavian Conference on Image Analysis*, pages 90–97. Citeseer, 2001. 21

M. Straka, A. La Cruz, A. Köchl, M. Šrámek, M.E. Gröller, and D. Fleischmann. 3D watershed transform combined with a probabilistic atlas for medical image segmentation. *Proc. MIT 2003*, pages 1–8, 2003. 14

J.S. Suri, K. Liu, S. Singh, S.N. Laxminarayan, X. Zeng, and L. Reden. Shape recovery algorithms using level sets in 2-D/3-D medical imagery: A state-of-the-art review. *IEEE Transactions on Information Technology in Biomedicine*, 6(1):8–28, 2002. 17, 18

T.S.Y. Tang and R.E. Ellis. 2d/3d deformable registration using a hybrid atlas. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pages 223–230, 2005. 21

P.T.H. Truc, T.S. Kim, S. Lee, and Y.K. Lee. A Study on the Feasibility of Active Contours on Automatic CT Bone Segmentation. *Journal of Digital Imaging*, 1:1–13, 2009. 17, 18

L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE transactions on pattern analysis and machine intelligence*, 13(6):583–598, 1991. 13

L.I. Wang, M. Greenspan, and R. Ellis. Validation of bone segmentation and improved 3-D registration using contour coherency in CT data. *IEEE transactions on medical imaging*, 25(3):324, 2006. 5, 12, 13

C.F. Westin, A. Bhalerao, H. Knutsson, and R. Kikinis. Using local 3D structure for segmentation of bone from computer tomography images. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 794–800. Institute of Electrical Engineers Inc (IEEE), 1997. 14, 29

C.F. Westin, S.K. Warfield, A. Bhalerao, L. Mui, J.A. Richolt, and R. Kikinis. Tensor controlled local structure enhancement of CT images for bone segmentation. *Lecture Notes in Computer Science*, pages 1205–1212, 1998. 13, 14

CF Westin, J. Richolt, V. Moharir, and R. Kikinis. Affine adaptive filtering of CT data. *Medical Image Analysis*, 4(2):161–172, 2000. 14

C. Xu and JL Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3):359–369, 1998. 15, 18

J. Zhang, C.H. Yan, C.K. Chui, and S.H. Ong. Fast segmentation of bone in ct images using 3d adaptive thresholding. *Computers in Biology and Medicine*, 40(2):231–236, 2010. 36

S.C. Zhu and A. Yuille. Region competition. *IEEE trans. on pattern analysis and machine intelligence*, 18(9):884–900, 1996. 13