Eurographics Workshop on Visual Computing for Biology and Medicine (2012)
T. Ropinski, A. Ynnerman, C. Botha, and J. B. T. M. Roerdink (Editors)

Impact of Physical Noise Modeling on Image Segmentation in Echocardiography

D. Tenbrinck[1], A. Sawatzky[2], X. Jiang[1], M. Burger[2], W. Haffner[1], P. Willems[1], M. Paul[3] and J. Stypmann[3]

[1]Institute of Computer Science, University of Münster, Germany
[2]Institute for Computational and Applied Mathematics, University of Münster, Germany
[3]Department for Cardiology and Angiology, University Hospital Münster, Germany

Abstract

Segmentation is an essential task in ultrasound image analysis. Recently, the trend in literature is towards incorporation of high-level information, e.g., shape priors, since many low-level segmentation techniques suffer from the characteristics of medical ultrasound images, i.e., speckle noise, scattering artifacts, and shadowing effects. However, the majority of these works implicitly assume an additive Gaussian noise model in ultrasound images, although a strong deviation from this assumption is well known, and the impact of correct physical noise modeling is not examined sufficiently until now. In this paper we investigate the influence of three different noise models from literature using a variational region-based segmentation framework, which allows for the incorporation of both low-level and high-level information. We demonstrate that correct physical noise modeling is of high importance for the computation of accurate segmentation results. The numerical results are validated on real patient datasets from echocardiographic examinations and compared to manual segmentations from echocardiographic experts.

Categories and Subject Descriptors (according to ACM CCS): I.4.6 [Image Processing and Computer Vision]: Segmentation —Region growing, partitioning I.4.7 [Image Processing and Computer Vision]: Feature Measurement —Moments

1. Introduction

Segmentation in ultrasound (US) imaging is used to separate regions-of-interest from background signals and to calculate borderlines and/or isosurfaces of structures within the given data. In the field of echocardiography segmentation is used to assess medical parameters of the cardiovascular system. In particular, physicians calculate medical parameters like left ventricular (LV) volume, ejection fraction or strain of LV by segmenting datasets from US examinations of a patient's myocardium [MvR*10]. Furthermore, segmentation is a fundamental technique to extract volumes-of-interest from 3D medical imaging data. It is naturally applied for visualization and operation planning by physicians [NNM11].

Automatic segmentation of US data is a hard task due to low contrast, shadowing effects, and speckle noise. In order to tackle these problems a huge variety of approaches has been proposed until today [NB06]. Recently, several authors proposed to explicitly model multiplicative noise characteristics in US images in order to improve segmentation results [LJBF*10, STJB11, TTB06] (cf. also references in [NB06]). Although this procedure is effective in the case of image noise, it is not sufficient for regions with structural artifacts, i.e., shadowing effects and low contrast regions in US data. This special situation occurs daily in clinical routine, e.g., when US waves get reflected by ribs during echocardiographic examinations of the human heart. Thus, finding a segmentation algorithm which can automatically segment the LV of the myocardium is of great interest to cardiologists. In order to tackle this difficult problem the incorporation of high-level information, such as prior knowledge about the shape to be segmented, is proved to be useful. Different works show that the use of shape priors for segmentation of echocardiographic data leads to improved and generally more robust segmentation results in the presence of image noise and structural artifacts [COS06, LJBF*06].

The contribution of this work is to investigate the impact of selected noise models from literature on the process of US image segmentation using shape priors. In contrast to related works (e.g., [COS06, LJBF*06]) we quantify the influence of appropriate noise modeling for high-level segmentation of

DOI: 10.2312/VCBM/VCBM12/033-040

ultrasound images and determine the best model for a global convex segmentation method. The paper is structured as follows: in Section 2 we describe segmentation tools needed to investigate the impact of correct noise modeling on high-level segmentation. In particular, we discuss different shape descriptors for segmentation from literature in Section 2.1 and present a shape prior energy based on the description of a given shape by Legendre moments in Section 2.2. We then propose a variational region-based segmentation framework in Section 2.3, which allows to easily switch between different noise models, which are discussed in Section 2.4, in the process of segmentation. In Section 2.5 we shortly present a possibility to incorporate the shape prior as regularization term to the described segmentation framework and highlight modifications in the numerical realization of the segmentation algorithm. Implementation details, suitable parameter settings, and the computational complexity of the proposed method are discussed in Section 3. To validate the performance of different physical noise models we present different segmentation results for real patient data from echocardiographic examinations in Section 4 and compare these to contours manually delineated by expert echocardiographers. In Section 5 this work is concluded by discussion.

2. Methods

In the following we shortly specify the tools needed to propose a unified variational segmentation framework incorporating low-level (noise models) and high-level (shape priors) information.

First, we will introduce some preliminary definitions to describe the segmentation model accurately. Let $\Omega \subset \mathbb{R}^2$ be the image domain and $f : \Omega \to \mathbb{R}$ the given image to be segmented. The task of image segmentation is to find a partition of Ω in pairwise disjoint regions $\Omega_i, i = 1, \dots, m$, i.e.,

$$\bigcup_{i=1}^{m} \Omega_i = \Omega \,, \qquad \Omega_i \bigcap_{i \neq j} \Omega_j = \emptyset \,. \tag{1}$$

In the case of echocardiographic images we will restrict the problem to separation of the inner region of LV denoted by Ω_{in} from the surrounding tissue and background denoted by Ω_{out}, i.e., $m = 2$. This bimodal model is sufficient, since the endocardial border of the myocardium is traced in most medical assessments in order to diagnose and quantify possible cardiovascular diseases. In order to characterize the resulting shape of region Ω_{in} we introduce an indicator function χ, which is defined on Ω by

$$\chi(x) \;=\; \begin{cases} 1 \,, & \text{if } x \in \Omega_{in} \,, \\ 0 \,, & \text{else} \,. \end{cases} \tag{2}$$

2.1. Encoding high-level information with descriptors

There are many different ways to encode high level information of objects within images, e.g., shape, with the help of descriptors (cf. [HM09, ZL04]). In general, one can divide the proposed solutions in literature into region-based and contour-based shape descriptors. Within these two classes there are different paradigms to describe objects by corresponding feature vectors. On the one hand, contour-based methods try to describe the shape of an object by its boundary information. While structural methods try to break the contour into sub-parts and analyze them with respect to certain criteria, global approaches calculate a feature vector of the integral boundary directly. On the other hand, region-based techniques take all the pixels within a shape region into account to obtain the shape representation and hence are more robust to noise compared to contour-based approaches.

One popular approach for region-based shape descriptors is based on moments. Let Ω_{in} be the inside region of a shape which we identify by its characteristic function χ. Then the geometric moments of the shape are defined as,

$$m_{p,q} \;=\; \int_{\Omega} \chi(x,y)\, x^p y^q \, dxdy \;=\; \int_{\Omega_{in}} x^p y^q \, dxdy \,, \tag{3}$$

$(p,q) \in \mathbb{N}^2$ and $N = p + q$ is the order of the moments. As known from mathematics any arbitrary shape can be reconstructed from its infinite set of moments $m_{p,q}$ and thus encoding the characteristic function of a given shape by moments leads to region-based shape descriptors. In practice the order N of the moments used is a finite number, large enough to encode the given shape without losing important details. Note that reconstructing a shape from a finite number of non orthogonal, geometric moments involves inverting an ill-conditioned Gram matrix [Tal87] as opposed to orthogonal moments.

Next to geometric moments a variety of alternative features exist in literature. By exchanging the basis functions in (3) different moments can be computed for a given shape. Using radial, orthogonal polynomials one obtains *Zernike moments* [SU12]. The appealing feature of Zernike polynomials is the separable nature of their radial and angular components. However, their computation is rather complicated and can lead to numerical errors if not performed correctly. Another possibility of orthogonal polynomials are *Legendre polynomials*. Legendre moments are less complicated to implement compared to Zernike moments and are optimal with respect to the minimal mean square error between the original object and its reconstruction from moments [FCH09]. For this reason, we use Legendre moments for incorporation of shape descriptors into our segmentation framework as described in Section 2.5. We give a detailed introduction to Legendre moments in Section 2.2 below.

2.2. Legendre moments as shape prior

From the different approaches for shape description listed in Section 2.1 we use central-normalized Legendre moments as proposed in [FCH09]. By this we derive a multi-reference shape model from manual segmentations done by

two echocardiographic experts. This shape descriptor allows a compact, parametric representation of shapes and is capable of dealing with arbitrary shape topologies. Furthermore, the according shape prior is independent of the underlying segmentation framework, as opposed to methods based on signed distance functions (e.g., [COS06]).

In order to use Legendre moments as descriptor we use an orthogonal basis of Legendre polynomials P_n of order n defined on the interval $[-1,1]$ given by,

$$P_n(x) \;=\; \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n\,. \tag{4}$$

Then the Legendre moments $L_{pq}(\chi)$ of order $N=p+q$ for a shape χ are computed by

$$L_{pq}(\chi) \;=\; C_{pq}\int_{\Omega_{in}} P_p(x)P_q(y)\,dxdy\,, \tag{5}$$

where $C_{pq}=(2p+1)(2q+1)/4$ is a normalization factor. However, we refrain to use the presentation in (5) and follow the idea in [FCH09] in order to achieve scale and translation invariance. Hence, we introduce intrinsically invariant shape descriptors $\lambda_{pq}(\chi)$,

$$\lambda_{pq}(\chi) \;=\; C_{pq}\sum_{u=0}^{p}\sum_{v=0}^{q} a_{pu}a_{qv}\eta_{uv}(\chi)\,, \tag{6}$$

where a_{ij} are the Legendre coefficients. Here, the $\eta_{u,v}$ represent the normalized central moments, which can be deduced by aligning the shape at its centroid $(\bar{x},\bar{y})$ and normalizing its area, i.e.,

$$\eta_{uv}(\chi) \;=\; \int_{\Omega}\frac{(x-\bar{x})^u(y-\bar{y})^v}{(\alpha|\Omega_{in}|)^{(u+v+2)/2}}\,\chi(x,y)\,dxdy\,. \tag{7}$$

Hence, using formula (6) we are able to encode a shape χ into a feature vector $\lambda=\{\lambda_{pq}\in\mathbb{R}\,|\,p+q\leq N\}$ whose dimension is $d=(N+1)(N+2)/2$. A given training set of reference shapes χ_k^{ref}, $k=1,\ldots,n$, is transformed into their respective Legendre moments $\lambda(\chi_k^{ref})$. In order to use only the most discriminative shape features to measure the similarity of objects one can perform a principal component analysis and use only the first $m\leq d$ principal components. Note that choosing d too small inevitably leads to a loss of details and thus to undiscriminable shape representations. Following the idea of [FCH09] we use a mixture-of-Gaussians to measure the similarity of a shape to a trained set of reference shapes as this is reasonable from a statistical point of view. Then, the multi-reference shape prior energy can be defined for n reference shapes χ_k^{ref} as

$$R_{sh}(\chi) = -log\left(\sum_{k=1}^{n} exp\left(-\frac{||\lambda(\chi)-\lambda(\chi_k^{ref})||_2^2}{2\sigma^2}\right)\right)\,, \tag{8}$$

where σ^2 is the fixed variance of the Gaussian functions.

To achieve rotational invariance we have to perform an additional step. Before building the shape model in the training phase as well as during the segmentation process, all shapes are rotated according to angles obtained by principal component analysis. This enhances the robustness of our segmentation algorithm proposed in Section 2.5 and enables us to segment ultrasound images obtained from different examination protocols for which the orientation of the LV varies.

2.3. Low-level information using variational region-based segmentation framework

In order to investigate the influence of different noise models on results of image segmentation we use a variational region-based segmentation framework recently proposed in [STJB11, STJB12]. This framework allows a flexible incorporation of different noise models occurring in medical imaging and a-priori knowledge about the subregions to be segmented using statistical (Bayesian) modeling. In contrast to comparable segmentation frameworks (e.g., [COS06]) this method allows for the modeling of fore- and background signal separately. Furthermore, it uses recent algorithms from the field of global convex segmentation to perform minimization of the corresponding energy functional and hence overcomes several drawbacks of methods based on level sets and signed distance functions. The corresponding energy functional is given by,

$$\begin{aligned} E(u_b,u_t,\chi) \;=\; & \\ \int_{\Omega} & \chi D_b(f,u_b) \;+\; (1-\chi)D_t(f,u_t)\,dx \\ & +\; \alpha_b R_b(u_b) \;+\; \alpha_t R_t(u_t) \;+\; \beta|\chi|_{BV(\Omega)}\,. \end{aligned} \tag{9}$$

In this context χ denotes an indicator function of the segmented region which can be used to represent the shape of the segmented object (cf. Section 2.2). The data fidelity terms D_b and D_t are negative log-likelihood functions, which are chosen according to the assumed noise model for the given data f. The regularization terms R_b and R_t are used to incorporate a-priori knowledge about the expected unbiased signals. Finally, the total variation $|\chi|_{BV(\Omega)}$ of χ (i.e., the perimeter of Ω_{in} in Ω) allows to regulate the level-of-details in the segmentation results and hence the smoothness of the contour. Since we want to concentrate on the influence of noise models on high-level segmentation results we restrict the framework (9) to a Chan-Vese-like [CV01] energy functional $E(c_{in},c_{out},\chi)$ (discarding the regularization terms R_b and R_t), which assumes constant approximations c_{in} and c_{out} in subregions Ω_{in} and Ω_{out}, respectively,

$$\begin{aligned} E(c_{in},c_{out},\chi) \;=\; & \\ \int_{\Omega} & \chi D_{in}(f,c_{in}) + (1-\chi)D_{out}(f,c_{out})\,dx + \beta|\chi|_{BV(\Omega)}\,. \end{aligned} \tag{10}$$

Although the assumption of a constant approximation is not valid for Ω_{out} (due to the inhomogeneity of regions surrounding LV) we will restrict on this case for the sake of simplicity. Computation of more realistic approximations u_b and u_t in (9) would increase the computational effort drastically and thus would complicate our evaluation of different

noise models for shape prior segmentation. For the details of a more realistic approximation of intensities in Ω_{out} we refer to [STJB11, STJB12] in order to incorporate a-priori knowledge about the smoothness of approximations.

2.4. Noise models for medical ultrasound data

Before we combine the shape prior energy introduced in (8) with the energy functional in (10) we will discuss three popular noise models used in literature for segmentation of medical ultrasound images. We compare and evaluate the performance of these three noise models in Section 4.

2.4.1. Additive Gaussian noise

The simplest assumption about the image noise is the presence of an *additive Gaussian noise* model of the form,

$$f = u + \eta \,, \tag{11}$$

in which η is a Gaussian distributed random variable with mean 0 and variance σ^2. This noise model applies for most data investigated in computer vision literature and is often used implicitly in standard segmentation methods, e.g., the Mumford-Shah formulation [MS89]. For this case the energy functional (10) corresponds exactly to the popular Chan-Vese model [CV01]. The negative log-likelihood functions in (10), i.e., the data fidelity terms $D_i(f,c_i), i \in \{in, out\}$, are given by (cf. [STJB11, STJB12])

$$D_i(f,c_i) \;=\; -\log p_i(f(x)\,|\,u_i(x)) \;=\; \frac{1}{2\sigma^2}\,(u_i(x) - f(x))^2 \,. \tag{12}$$

It is well known that in presence of additive Gaussian noise the optimal constants for an approximation of Ω_{in} and Ω_{out} can be computed as mean intensities of the respective regions,

$$c_i \;=\; \frac{1}{|\Omega_i|}\int_{\Omega_i} f\,dx \,, \quad i \in \{in, out\} \,, \tag{13}$$

where $|\Omega_i|$ denotes the amount of pixels in Ω_i. Note that the model (11) corresponds to the noise model in Equation (14) below for the special case of $\gamma = 0$.

2.4.2. Multiplicative speckle noise

In contrast to this additive signal-independent noise in Section 2.4.1 speckle noise is of *multiplicative* nature, i.e., the noise variance directly depends on the underlying signal intensity. This type of noise is characteristic for diagnostic ultrasound imaging and can be described mathematically by a statistical model for multiplicative noise [LMA89, RLO03],

$$f = u + u^{\frac{\gamma}{2}}\eta \,, \tag{14}$$

In this context u is the unbiased image intensity, η is Gaussian distributed random noise with mean 0 and variance σ^2, and f is the observed image. The parameters γ determines the degree of signal dependency and hence the characteristics of the multiplicative noise. Speckle noise leads to heavy distortions in the image, especially in regions with high intensity values. Typical values for γ can be found in literature, e.g., in an experimentally derived model for multiplicative speckle noise [LMA89] the authors investigate the case $\gamma = 1$, whereas in [RLO03] the authors choose $\gamma = 2$ to model the noise in medical US imaging. In general, the parameters γ and σ depend on the imaging system, the application settings, and the examined tissue. For our investigations we focus on the case $\gamma = 1$ as proposed by [LMA89] for log-compressed ultrasound images which we call *Loupas noise model* in the following, i.e.,

$$f = u + \sqrt{u}\eta \,. \tag{15}$$

In this case the negative log-likelihood functions in (10), i.e., the data fidelity terms $D_i(f,c_i), i \in \{in, out\}$, are given by (cf. [STJB11, STJB12])

$$\begin{aligned} D_i(f,c_i) \;&=\; -\log p_i(f(x)\,|\,u_i(x)) \\ &=\; \frac{1}{2\sigma^2}\,\frac{(u_i(x) - f(x))^2}{u_i(x)} \;+\; \frac{1}{2}\log u_i(x) \,, \end{aligned} \tag{16}$$

for which the optimal constants for Ω_{in} and Ω_{out} can be computed by

$$c_i = \frac{1}{2}\left(\sqrt{\sigma^4 + \frac{4\int_{\Omega_i} f^2\,dx}{|\Omega_i|}} - \sigma^2\right) dx \,, i \in \{in, out\} \,. \tag{17}$$

2.4.3. Rayleigh noise

Finally, we investigate a noise model based on *Rayleigh distribution*, which has been also used in [LJBF*06] in combination with shape priors described in Section 2.2. Here the data are modeled by

$$f = u\eta \,, \tag{18}$$

where η is a Rayleigh distributed random variable with the probability density function,

$$p(\eta) \;=\; \frac{\eta}{\sigma^2}\,e^{-\frac{\eta^2}{2\sigma^2}} \,, \tag{19}$$

and is assumed to be appropriate for US B-scans in the presence of a large number of randomly located scatters [WSSL83]. In this case the negative log-likelihood functions in (10), i.e., the data fidelity terms $D_i(f,c_i), i \in \{in, out\}$, are given by (cf. [STJB11, STJB12])

$$\begin{aligned} D_i(f,c_i) \;&=\; -\log p_i(f|c_i) \;=\; -\log\left(p\left(\frac{f}{c_i}\right)\frac{1}{c_i}\right) \\ &=\; -\log\left(\frac{f}{\sigma^2 c_i^2}\right) \;+\; \frac{f^2}{2\sigma^2 c_i^2} \end{aligned}$$

and the optimal constants for Ω_{in} and Ω_{out} can be computed by

$$c_i \;=\; \sqrt{\frac{1}{|\Omega_i|}\int_{\Omega_i}\frac{f^2}{2\sigma^2}\,dx} \,, \quad i \in \{in, out\} \,. \tag{20}$$

Note that the formulas for the optimal constants in (17)

and (20) contain a free parameter σ which determines the corresponding probability distribution of η. Since σ is unknown in real data, we choose this parameter according to the noise inherent in a given image and try to optimize its value with respect to the segmentation performance.

2.5. Incorporation of high-level information into segmentation framework

In order to incorporate both low-level and high-level information in a unified framework, we combine the variational segmentation framework in (10) with the shape prior energy in (8),

$$E(c_{in}, c_{out}, \chi, \chi_{sh}) = \int_\Omega \chi D_{in}(f, c_{in}) + (1-\chi) D_{out}(f, c_{out})\, dx + \beta |\chi|_{BV(\Omega)} + \gamma R_{sh}(\chi_{sh}) + \frac{\delta}{2} ||\chi - \chi_{sh}||^2 . \quad (21)$$

In addition we add a l^2-similarity distance between the image region-driven segmentation χ and the shape space-driven segmentation χ_{sh}, since we want to ensure that $\chi = \chi_{sh}$ in case of convergence. For minimization of (21) we follow the numerical realization proposed in [STJB11, STJB12], where an alternating minimization scheme is used as illustrated in Algorithm 1.

Algorithm 1 An alternating minimization scheme for the numerical realization of (21).

repeat

computeOptimalConstants(χ^k) $\hat{=}$

$$(c_{in}^{k+1}, c_{out}^{k+1}) = \arg\min_{c_i} E(c_{in}, c_{out}, \chi^k, \chi_{sh}^k)$$

segmentImage($c_b^{k+1}, c_t^{k+1}, \chi_{sh}^k$) $\hat{=}$

$$\chi^{k+1} = \arg\min_{\chi} E(c_{in}^{k+1}, c_{out}^{k+1}, \chi, \chi_{sh}^k)$$

updateShape(χ^{k+1}) $\hat{=}$

$$\chi_{sh}^{k+1} = \arg\min_{\chi_{sh}} E(c_{in}^{k+1}, c_{out}^{k+1}, \chi^{k+1}, \chi_{sh})$$

until Convergence

The optimal constants c_{in}^{k+1} and c_{out}^{k+1} are computed for each assumed noise model individually as described in Section 2.4. The shape prior update χ_{sh}^{k+1} based on the current segmentation χ^{k+1} is performed as proposed in [FCH09] by a steepest gradient step in the shape space of the form

$$\chi_{sh}^{k+1} = \chi^{k+1} - \frac{\gamma}{\delta} \frac{\partial R_{sh}}{\partial \chi_{sh}}(\chi^{k+1}) . \quad (22)$$

Note that we approximate the gradient descent vector using the given segmentation χ^{k+1} in order to avoid the inversion of the gradient of R_{sh}.

The image region-driven segmentation step in Algorithm 1 requires only a minor modification compared to [STJB11, STJB12]. In the case of the energy functional (21), the segmentation step can be rewritten to

$$\chi^{k+1} \in \arg\min_{\chi \in BV(\Omega;\{0,1\})} \left\{ \int_\Omega \chi\, g\, dx + \beta |\chi|_{BV(\Omega)} \right\} \quad (23)$$

with

$$g = D_{in}(f, c_{in}^{k+1}) - D_{out}(f, c_{out}^{k+1}) - \delta \left(\chi_{sh}^k - \frac{1}{2} \right)$$

using the identity $(\chi_{sh}^k)^2 = \chi_{sh}^k$ for characteristic functions. Then the problem (23) can be solved analogously to [STJB11, STJB12] using the well known Rudin-Osher-Fatemi (ROF) model in image processing. The advantage of this approach is the strict convexity of the ROF model and thus the avoidance of local minima, in contrast to, e.g., level set methods.

3. Implementation

We implemented the alternating minimization scheme for our segmentation algorithm in the numerical computing environment MathWorks MATLAB (R2010a) on a 2×2.2GHz Intel Core Duo processor with 2GB memory and a Microsoft Windows 7 (64bit) operating system.

We optimized the selection of regularization parameters β, γ, and δ in Equation (21) with respect to the segmentation performance as described in Section 4 below. For *additive Gaussian noise* (cf. Section 2.4.1) we used $\beta \in [0.02, 1.5], \gamma \in [0.01, 0.05]$, and $\delta \in [10^{-4}, 0.9]$. In the case of *multiplicative speckle noise* (cf. Section 2.4.2) we chose $\beta \in [0.015, 0.02], \gamma \in [0.01, 0.05]$, and $\delta \in [0.8, 0.9]$. For *Rayleigh noise* (cf. Section 2.4.3) the best parameters were $\beta \in [0.1, 0.5], \gamma \in [10^{-4}, 10^{-3}]$, and $\delta \in [0.1, 0.2]$. During our experiments we observed that a noise variance parameter of $\sigma = 0.19$ is the best choice in the case of multiplicative speckle noise, while $\sigma = 0.27$ showed best results for Rayleigh noise. We use Legendre polynomials of order $N = 40$ to encode shapes since this was the minimal number without losing important details during reconstruction.

For the image-based segmentation step of Algorithm 1 we observed that $850 - 1400$ inner iterations are enough to reach a stationary state for Rayleigh and multiplicative speckle noise, i.e., no more changes between two consecutive inner iteration steps. For additive Gaussian noise $1200 - 2400$ inner iterations were needed. For the outer iterations we observe between $25 - 35$ iteration steps until convergence of the segmentation process.

3.1. Computational complexity

In order to understand the computational complexity of our proposed segmentation framework in Equation (21) and the overall time needed to compute a segmentation we give a detailed discussion of the substeps of Algorithm 1 with respect to their runtime in the following.

Let us assume that we have k outer iterations of our segmentation process. In each of these iterations we have to compute the optimal constants for Ω_{in} and Ω_{out} as described in Section 2.4. Furthermore, we have to perform the image-based segmentation of Ω in Equation (23) with respect to the updated optimal constants c_{in} and c_{out}. The last step is the update of the shape χ_{sh} according to its similarity to the training set of shapes as described in Equation (22). The computation of the optimal constants can be performed in $\mathcal{O}(|\Omega|)$, since the intensity values of all pixels are used to perform these calculations. The image-based segmentation step is rather complex, since efficient solver schemes from numerical mathematics are used. For the sake of clarity we focus on the part with highest computational costs and refer to [STJB11, STJB12] for a detailed discussion of the algorithm. Let us assume we need p inner iteration steps, then the computational complexity of the segmentation step is in $\mathcal{O}(p \cdot |\Omega| \, log(|\Omega|))$ (there is a discrete cosine transformation in every inner iteration step). Finally, we discuss the shape update using a single steepest gradient step. Let N be the degree of the used Legendre polynomials and let us assume we use all principal components of the feature vectors, i.e., $m = N$. Furthermore, let $d = (N+1)(N+2)/2$ be the dimension of the vector of central normalized Legendre moments λ. To encode the current shape χ by Legendre-moments (cf. Section 2.2) we have a complexity of $\mathcal{O}(d \cdot |\Omega|)$. The steepest gradient step for the optimization of the shape prior in Equation (22) is performed in $\mathcal{O}(d)$. To reconstruct the updated shape from Legendre moments we need $\mathcal{O}(d \cdot |\Omega|)$ operations. Hence, the total computational complexity of the proposed segmentation algorithm is in $\mathcal{O}(k \cdot (p \, |\Omega| \, log(|\Omega|)))$.

For a 108×144 pixel image we measured the following runtime: the computation of the optimal constant approximations for fore- and background takes approximately 1ms and the shape update only 60ms. As described above the image-based segmentation has the highest computational complexity and needs 5.1s. The overall time for the segmentation process with 35 outer iterations takes 149.2s.

4. Results

In this section we investigate the influence of different noise models proposed in Section 2.4 on high-level segmentation of ultrasound data. In order to evaluate the segmentation results we obtained manual delineations of the endocardial contour for each dataset by two different echocardiographic experts, who are familiar to this task due to daily clinical routine. The performance of each segmentation process is measured using the Dice index, which compares two given segmentations A and B by

$$D(A,B) = \frac{2\,|A \cap B|}{|A| + |B|} \, . \tag{24}$$

The Dice index is normalized between 0 (no similarity) and 1 (exact match). For each image in the test dataset we list the inter-observer variability between the two physicians and compute the average similarity between the automatically segmented images and the two observers.

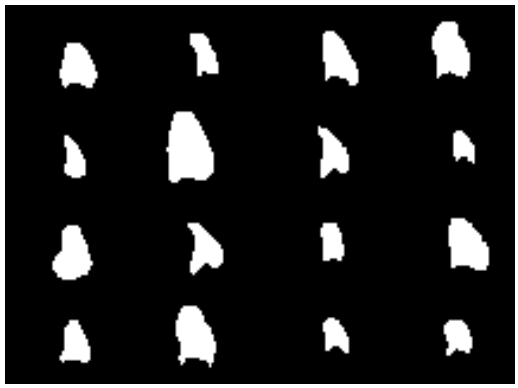


Figure 1: Part of the training data set used to build the shape prior energy (8). The masks show manually segmented shapes of LV of the human heart.

4.1. 2D B-mode echocardiographic images

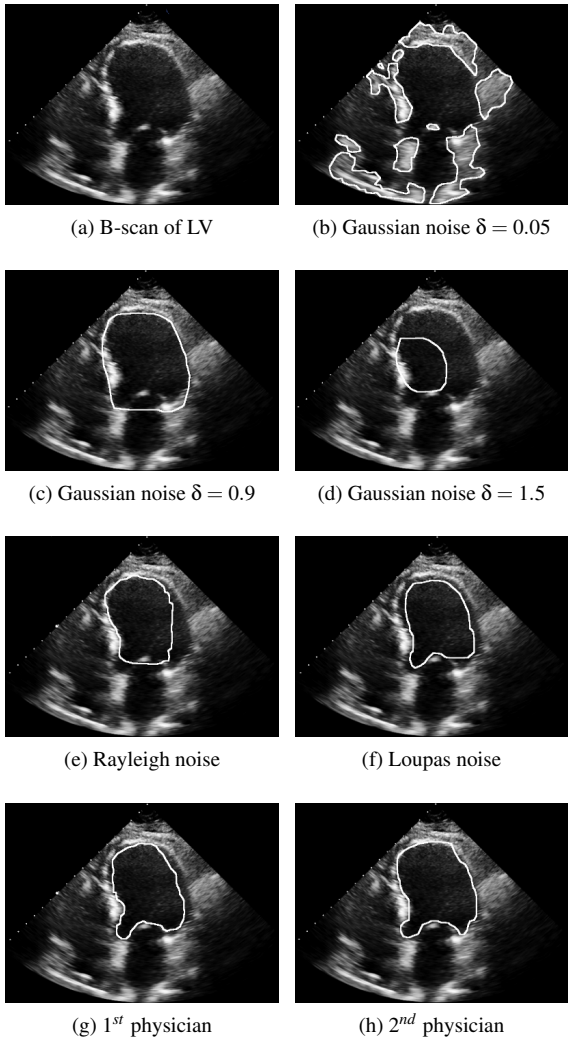
We use 30 datasets from echocardiography (Philips iE33) containing images of LV of the human heart from different acquisition angles, i.e., apical two-chamber and four-chamber views of the human myocardium, which were manually segmented by two experts. Since the shape of the left ventricle depends on the acquisition angle, we get a significant inter-shape diversity within the training data set as can be seen in Figure 1. Instead of specializing our algorithm for one certain US imaging protocol, we train our method for different echocardiographic examinations for the sake of flexibility. For quantification we choose 8 images from the dataset which cover all challenges for image processing we observed in the given data, e.g., speckle noise and shadowing effects. The parameters in the proposed segmentation framework (21) are optimized for each test image with respect to the Dice index. For training of the shape prior energy we use a leave-one-out strategy, i.e., $N_{sh} = 58$ manual delineations, and use the two excluded delineations of the experts for validation. This procedure is necessary since the training set needs to be large enough to provide robustness in the presence of large variance of the discussed effects.

Figure 2 illustrates one representative result on a rather easy dataset. It was not possible to find a good parameter setting for the case of additive Gaussian noise. Figure 2b-2d demonstrate our typical observations for this noise model on all datasets as discussed in detail below. Using the Loupas or Rayleigh noise model leads to reasonable results as can be seen in Figure 2e- 2f. In order to illustrate the challenges occurring in our chosen test dataset, in Figure 3 the typical noise artifacts present at the lateral wall of LV are illustrated, which leads to significant problems in the process of segmentation. Additionally, one of the mitral valve leaflets

can be seen within the cavity of LV, which also produces erroneous segmentations. The Loupas noise model in Figure 3d is able to deal with these noise artifacts and produces the best segmentation results again compared to the other noise models in Figure 3b-3c.

The Dice indices of numerical results for our experiments are presented in Table 1. As expected, the segmentation with the additive Gaussian noise model using the optimal constants in (13) failed on all test images. We observed two different ways of possible behavior of our segmentation framework during optimization of parameters in this model. The results vary extremely between a segmentation of the US image without any similarity to the prior shapes on the one side and the convergence towards a mean shape of the training dataset without influence of image information on the other side. However, for an initialization of Ω_{in} inside LV and an extensive parameter search it is also possible to obtain results comparable to the case of Rayleigh modeling in some cases. But we disregard these results due to instability of the algorithm with respect to the choice of the initialization shape and segmentation parameters. However, we note that the mentioned observations in the case of Gaussian noise modeling are only valid in our case due to the globally convex segmentation approach proposed in Section 2.5 and might not be appropriate for methods based on level set methods, due to the existence of local minima.

The cases of Rayleigh noise and the speckle noise model proposed by Loupas et al. [LMA89] lead to significantly better results (see Table 1). In particular, they are robust with respect to the choice of shape initialization and segmentation parameters compared to the additive Gaussian noise case discussed above. We observe also that the model of Loupas has an average Dice index of 0.87 and thus is more suitable for high-level segmentation of US B-scan images than the Rayleigh modeling, which has an average segmentation performance of 0.78 on our 8 test datasets. This can also be seen visually as for example illustrated in Figure 3, which shows all three automatic segmentation results of dataset 6 in Table 1 compared with the manual delineations of the two physicians.



(a) B-scan of LV (b) Gaussian noise $\delta = 0.05$

(c) Gaussian noise $\delta = 0.9$ (d) Gaussian noise $\delta = 1.5$

(e) Rayleigh noise (f) Loupas noise

(g) 1^{st} physician (h) 2^{nd} physician

Figure 2: US B-scan of LV with manual delineations of physicians and segmentation results using noise models described in Section 2.4.

5. Discussion

In this work we investigate the impact of three different noise models from literature on high-level segmentation of echocardiographic data. The numerical results show that the incorporation of proper noise models for US have a significant influence on segmentation performance when using shape priors. Although the additive Gaussian noise model is not a good choice under these conditions, its impact on other segmentation frameworks, e.g., level set methods, has yet to be explored. Furthermore, it would be interesting to include temporal information using consecutive US frames to increase the robustness of segmentation results, since experts from echocardiography also heavily depend on this information when evaluating the examination data.

Acknowledgements

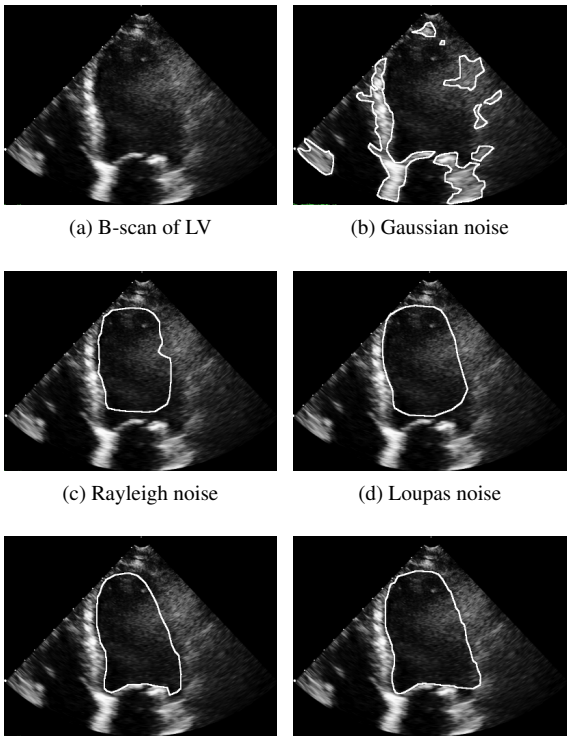
This study has been supported by the German Research Foundation DFG, SFB 656 MoBil (project C3), as well as DFG project BU 2327/1.

References

[COS06] CREMERS D., OSHER S., SOATTA S.: Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision 69*, 3 (2006), 335–351. 1, 3

Dataset	1	2	3	4	5	6	7	8
Observer variability	0.9228	0.9354	0.9034	0.9310	0.9151	0.9246	0.9391	0.8435
Gaussian noise	0.3444	0.4470	0.3306	0.3595	0.3439	0.4754	0.2953	0.3689
Loupas noise	**0.8245**	0.7559	**0.9106**	**0.8891**	**0.9030**	**0.8862**	**0.8855**	**0.8942**
Rayleigh noise	0.8123	**0.7838**	0.7539	0.8017	0.7999	0.7693	0.7689	0.7368

Table 1: Dice index values for comparison with manual segmentation.



(a) B-scan of LV (b) Gaussian noise

(c) Rayleigh noise (d) Loupas noise

(e) 1^{st} physician (f) 2^{nd} physician

Figure 3: US B-scan of LV (dataset 6 in Table 1) with manual delineations of physicians and segmentation results using noise models described in Section 2.4.

[CV01] CHAN T. F., VESE L. A.: Active contours without edges. *IEEE Transactions on Image Processing 10*, 2 (2001), 266–277. 3, 4

[FCH09] FOULONNEAU A., CHARBONNIER P., HEITZ F.: Multi-reference shape priors for active contours. *International Journal of Computer Vision 81* (2009), 68–81. 2, 3, 5

[HM09] HEIMANN T., MEINZER H.: Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis 13*, 4 (2009), 543–563. 2

[LJBF*06] LECELLIER F., JEHAN-BESSON S., FADILI J., AUBERT G., REVENU M., SALOUX E.: Region-based active contour with noise and shape priors. *IEEE International Conference on Image Processing* (2006), 1649–1652. 1, 4

[LJBF*10] LECELLIER F., JEHAN-BESSON S., FADILI J., AUBERT G., REVENU M.: Region-based active contours with exponential family observations. *Journal Mathematical Imaging and Vision 36* (2010), 28–45. 1

[LMA89] LOUPAS T., MCDICKEN W. N., ALLAN P. L.: An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *IEEE Transactions on Circuits and Systems 36*, 1 (1989), 129–135. 4, 7

[MS89] MUMFORD D., SHAH J.: Optimal approximation by piecewise smooth functions and associated variational problems. *Communication on Pure and Applied Mathematics 42*, 5 (1989), 577–685. 4

[MvR*10] MA M., VAN STRALEN M., REIBER J., BOSCH J., LELIEVELDT B.: Model driven quantification of left ventricular function from sparse single-beat 3D echocardiography. *Medical Image Analysis 14* (2010), 582–593. 1

[NB06] NOBLE J., BOUKERROUI D.: Ultrasound image segmentation: A survey. *IEEE Transactions on Medical Imaging 25*, 8 (2006), 987–1010. 1

[NNM11] NYSTRÖM I., NYSJÖ J., MALMBERG F.: Visualization and haptics for interactive medical image analysis: Image segmentation in cranio-maxillofacial surgery planning. In *Visual Informatics: Sustaining Research and Innovations* (2011), no. 7066 in Lecture Notes in Computer Science, Springer-Verlag, pp. 1–12. 1

[RLO03] RUDIN L., LIONS P., OSHER S.: Multiplicative denoising and deblurring: Theory and algorithms. In *Geometric Level Set Methods in Imaging, Vision, and Graphics*, Osher S., Paragios N., (Eds.). Springer, 2003, pp. 103–122. 4

[STJB11] SAWATZKY A., TENBRINCK D., JIANG X., BURGER M.: *A Variational Framework for Region-Based Segmentation Incorporating Physical Noise Models*. Tech. Rep. CAM11-81, UCLA, December 2011. 1, 3, 4, 5, 6

[STJB12] SAWATZKY A., TENBRINCK D., JIANG X., BURGER M.: A variational framework for region-based segmentation incorporating physical noise models. *Journal of Mathematical Imaging and Vision* (2012), under revision. 3, 4, 5, 6

[SU12] SINGH C., UPNEJA R.: Fast and accurate method for high order zernike moments computation. *Applied Mathematics and Computation 218*, 15 (2012), 7759–7773. 2

[Tal87] TALENTI G.: Recovering a function from a finite number of moments. *Inverse Problems 3*, 3 (1987), 501–517. 2

[TTB06] TAO Z., TAGARE H. D., BEATY J. D.: Evaluation of four probability distribution models for speckle in clinic cardiac ultrasound images. *IEEE Transactions on Medical Imaging 25*, 11 (2006), 1483–1491. 1

[WSSL83] WAGNER R. F., SMITH S. W., SANDRIK J. M., LOPEZ H.: Statistics of speckle in ultrasound B-scans. *IEEE Transactions on Sonics and Ultrasonics 30*, 3 (1983), 156–163. 4

[ZL04] ZHANG D., LU G.: Review of shape representation and description techniques. *Pattern Recognition 37* (2004), 1–19. 2