JACOBS
UNIVERSITY

Modeling, Estimating, and Visualizing Spatial and Temporal Uncertainty for Image-Guided Therapy

by

Bojan Kocev

a Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy
in Computer Science

Approved Dissertation Committee

Prof. Dr. Horst Karl Hahn and Prof. Dr. Lars Linsen
Jacobs University Bremen

Prof. Dr. Ron Kikinis
Harvard Medical School

Prof. Dr. William M. Wells
Harvard Medical School

Jun.-Prof. Dr. Christian Hansen
University of Magdeburg

Prof. Dr. Andreas Birk
Jacobs University Bremen

Date of Defense: June 03 2020

Computer Science & Electrical Engineering

Statutory Declaration

Family Name, Given/First Name	Kocev, Bojan
Matriculationnumber	10000470
What kind of thesis are you submitting: Bachelor-, Master- or PhD-Thesis	PhD-Thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

The Thesis has been written independently and has not been submitted at any other university for the conferral of a PhD degree; neither has the thesis been previously published in full.

German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde in der vorliegenden Form weder einer anderen Prüfungsbehörde vorgelegt noch wurde das Gesamtdokument bisher veröffentlicht.

January 31 2020, 

Date, Signature

"We are all shaped by the tools we use, in particular: the formalisms we use shape our thinking habits, for better or for worse, and that means that we have to be very careful in the choice of what we learn and teach, for unlearning is not really possible."

E. Dijkstra.

JACOBS UNIVERSITY BREMEN

Abstract

Computer Science
School of Engineering and Science

PhD Thesis

by Bojan Kocev

We address different aspects of the problem of uncertainty in image-guided therapies. This is of great interest, because such uncertainty could have a substantial impact on public health. Significant uncertainties can occur in image segmentation and estimating deformations of anatomy. To model the uncertainty in the image segmentation aspect, we propose a novel stationary Gaussian process (GP)-based generative segmentation model. This segmentation model allows us to draw many possible image segmentations, which can be used for estimating and visualizing different aspects of the uncertainty in image-guided therapies. To enable drawing of many image segmentation samples efficiently, we propose a fast method for sampling from stationary GPs. To model the uncertainty in the aspect of estimating deformations of anatomy, we propose a novel spatiotemporal GP model for uncertainty-aware soft-tissue motion estimation using GP regression. The spatiotemporal GP formalism enables the estimation of anatomy displacements at any location, and for any time interval from measured motions that are sparse in space and time. The use of GP regression enables the quantification of uncertainty in the soft-tissue motion estimation result, which allows the amount of uncertainty in some aspects, e.g., registered planning medical images, of image-guided therapies or procedures governing the decisions of medical specialists to be conveyed. To convey the amount of uncertainty in the anatomy motion estimates, we propose novel motion uncertainty visualization methods. To showcase the use of the devised methods, we deploy them in the context of radiotherapy and image-guided soft-tissue intervention navigation. We expect that incorporating estimates of spatial and temporal uncertainty into the processing pipelines of image-guided therapy will eventually enable improved treatments, and thus, improved outcomes.

Contents

Chapter 1

Introduction

A therapy or medical procedure may cause unwanted adverse effects. Ideally, one would like to have means to predict the possibility of such effects. One reason for the aforementioned effects could be errors in the planning that takes place beforehand or due to changes in the planning factors after the actual planning has taken place. Therefore, it would be useful to provide information about these errors. For many medical cases, it is common to use medical imaging to create a visual representation of the interior of the whole human body or individual organs and use this data for planning. In this regard, many planning procedures require to partition the medical image into different meaningful segments (e.g., organs, tumors, etc.), so that the plan can be performed accordingly. To quantify the errors in the planning, it would therefore be useful to quantify the errors in the image segmentation result itself. In some cases, the planning image coupled with the plan is also used during the actual treatment or procedure for guidance. In the majority of those cases, the image coupled with the plan needs to be adapted to reflect the possible changes (e.g., deformation) in the anatomy. To make these adaptations one needs to estimate the motion of the human anatomy. To quantify the errors in the adapted image coupled with the plan, one needs therefore to quantify the errors in the human anatomy motion estimate. The overall above-discussed errors need to be conveyed to the medical specialists in an effective manner, so that they can make better-informed decisions about the quality of the plans.

Regarding the errors in the image segmentation of a region of interest (ROI), in Chapter 4, we model the uncertainty therein using marginal label probabilities (MLPs) and a suitable stationary covariance structure. The MLPs can be the output of a probabilistic image segmentation algorithm or based on (a) (multi-)rater segmentation(s) of the ROI. For the latter case, we devise different methods for estimating these probabilities. We engineer the covariance structure ourselves, but its underlying parameters can be trained

on the input segmentation(s). We then set up a generative segmentation model with the estimated MLPs and covariance structure. We model the segmentation boundary as a zero level set of a stochastic level set function. The stochastic level set function is modeled as a Gaussian process (GP). In this regard, we devise a new method for setting the mean function of the GP based on the estimated MLPs. To generate an image segmentation sample (ISS), we then draw a GP sample and threshold it. To be able to draw a large enough number of ISSs, we devise a novel method for efficient sampling from stationary GPs.

We deploy the above-described devised generative segmentation model in the context of radiotherapy. In this realm, it is common to delineate tumors and "organs at risk" (OARs) in an image and optimize the radiotherapy dose plan accordingly. In case a tumor is close to an OAR, the radiation field is normally optimized to have a high spatial gradient between the two structures. However, if there is uncertainty in the segmentation of the two structures, then there is some risk that the OAR will receive high dose and/or the tumor will receive low dose. This could lead to unwanted adverse effects such as destroying healthy tissue or tumor recurrence. We use the proposed generative segmentation model to draw many possible tumor and/or OAR segmentation samples and based on them estimate a posterior distribution on relevant dose metrics (e.g., generalized equivalent uniform dose (gEUD), tumor control probability (TCP), normal tissue control probability (NTCP), etc.). We believe that the estimated posterior distributions on relevant dose metrics could help medical specialists in predicting the possibility of radiotherapy adverse effects.

Regarding the quantification of the errors in the human anatomy motion estimate, in Chapter 2, we propose to estimate the motion in an uncertainty-aware fashion. For more accurate motion estimation, we fuse motion measurements (i.e., motion signal samples) with motion prior (e.g., motion dynamics, shape, etc.). We formulate the fusion problem as the problem of uncertainty-aware interpolation of motion signal samples that are randomly, non-uniformly scattered in the spatiotemporal domain. In this regard, we propose a novel motion field representation which allows us to formally define asynchronous motion measurements. To perform the interpolation in an optimal unbiased manner, we make use of GP regression. The motion prior is embedded in the GP prior mean and covariance functions. In this regard, we propose using a non-zero mean function and a covariance function with anisotropic distance measure. The GP prior is then conditioned on the motion measurements during the regression, which results in an estimate of the (full) posterior distribution over motion signal functions. In summary, the spatiotemporal GP formalism enables the estimation of anatomy displacements at any location, and for any time interval from measured motions that are sparse in space and time.

To convey the information about the estimated errors in the anatomy motion estimate, in Chapter 3, we propose to visualize the uncertainty therein. In this regard, we use the estimate of the (full) posterior distribution over motion signal functions based on the above-described method for summarizing and conveying the uncertainty. In the general setting, we devise methods for uncertainty visualization in time-varying, due to motion (including deformation), scalar fields. We visualize the maximum a posteriori (MAP) estimate of the moved scalar field. Furthermore, we propose to perform a spatially-varying (Gaussian) blur (according to the uncertain motion estimate) to an image (of the same size as the scalar field) containing a grid to convey the overall motion uncertainty. In this regard, the amount of blur visually encodes the uncertainty. We then combine the visual encoding of the scalar field motion uncertainty with the MAP estimate of the moved scalar field using "alpha blending and thresholding" or "masking and thresholding" and show 2D cross sections to the end user. Furthermore, we suggest to also perform a spatially-varying (Gaussian) blur (according to the uncertain motion estimate) to an input probabilistic segmentation of a ROI to estimate where the ROI is after the motion. We then propose to show visually mapped 2D cross sections of the resulting MLPs or MLP isocontours overlayed over the corresponding 2D cross section of the MAP estimate of the moved scalar field.

We deploy the above-described uncertainty-aware anatomy motion estimation and visualization methods in the context of image-guided soft-tissue (e.g., biopsy) intervention navigation. A planning (e.g., magnetic resonance imaging (MRI)) image is acquired and the planning performed prior to the intervention. For example, we segment the relevant structures (e.g., the tumor that needs to be biopsied). The same planning image coupled with the plan is also used during the actual intervention for guidance. As the soft tissue deforms, we need to adapt the planning image coupled with the plan to reflect the possible changes (e.g., deformation) in the anatomy. We compute the MAP estimate of the adapted planning image based on the GP posterior distribution over motion fields and visualize the uncertainty therein as described above. Furthermore, we also visualize in an uncertainty-aware fashion where a ROI (e.g., the tumor that needs to be biopsied) is after deformation using the above-described methods. We believe that our proposed uncertainty-aware motion estimation and visualization approaches would be particularly useful for (e.g., biopsy) intervention specialists when performing navigation in deforming planning images and may aid in avoiding unwanted adverse effects (e.g., wrong decisions made because of inaccurate biopsy specimens).

It is our expectation that incorporating estimates of spatial and temporal uncertainty into the processing pipelines of image-guided therapy will eventually enable improved treatments, and thus, improved outcomes.

Chapter 2

Uncertainty-aware Asynchronous Scattered Motion Interpolation Using Gaussian Process Regression[1]

Bojan Kocev[1,2,3], Horst Karl Hahn[2,3], Lars Linsen[4], William M. Wells[5], Ron Kikinis[1,2,5]

[1]University of Bremen, Bremen, Germany
[2]Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany
[3]Jacobs University, Bremen, Germany
[4]Westfälische Wilhelms-Universität Münster, Germany
[5]Harvard Medical School and Brigham and Women's Hospital, Boston, MA 02115, USA

Abstract

We address the problem of interpolating randomly non-uniformly spatiotemporally scattered uncertain motion measurements, which arises in the context of soft tissue motion estimation. Soft tissue motion estimation is of great interest in the field of image-guided soft-tissue intervention and surgery navigation, because it enables the registration of pre-interventional/pre-operative navigation information on deformable soft-tissue organs.

[1]This material originally appeared in Kocev et al. (2019).

To formally define the measurements as spatiotemporally scattered motion signal samples, we propose a novel motion field representation. To perform the interpolation of the motion measurements in an uncertainty-aware optimal unbiased fashion, we devise a novel Gaussian process (GP) regression model with a non-constant-mean prior and an anisotropic covariance function and show through an extensive evaluation that it outperforms the state-of-the-art GP models that have been deployed previously for similar tasks. The employment of GP regression enables the quantification of uncertainty in the interpolation result, which would allow the amount of uncertainty present in the registered navigation information governing the decisions of the surgeon or intervention specialist to be conveyed.

2.1 Introduction

Registration of pre-interventional/pre-operative navigation information on deformable soft-tissue organs requires estimating the motion that soft-tissue organs undergo during an intervention or surgery (Baumhauer et al. (2008)). The motion is to be estimated in an uncertainty-aware fashion, so as to allow conveying the amount of uncertainty present in the registered navigation information influencing the surgeon's or intervention specialist's decisions (Risholm et al. (2010)).

For accurate motion estimation, one needs to fuse motion measurements (i.e., signal samples) with motion prior (e.g., motion dynamics, shape, etc.) information. According to the proposed general *data centric* taxonomy by Khaleghi et al. (2013), the motion signal samples may be imperfect, correlated, inconsistent, and/or in disparate forms/modalities. The imperfection aspect is manifested through uncertainty (Irani and Anandan (2000); Kanazawa and Kanatani (2001); Leedan and Meer (2000); Zhou et al. (2004, 2005)), imprecision, and/or granularity. The inconsistency problem is in terms of conflict, outlier (spurious data), and/or disorder (out-of-sequence data). Furthermore, the imprecision is expressed in the form of vagueness, ambiguity, and/or incompleteness.

In soft tissue navigation, the motion measurements information source is, in general, composed variously of real-time surface (e.g., electromagnetic (EM) (Kocev et al. (2014); Zhang et al. (2006)), optical (OP) (Meinzer et al. (2008)), point cloud (Cash et al. (2005); Kocev et al. (2013)), etc.) and volumetric (e.g., ultrasound-ultrasound (US-US) correlation (Kocev et al. (2014); Wang et al. (2013)), etc.) data. In this regard, we believe that using such various multimodal tracking systems could allow for better sampling of the motion signal function and therefore could improve the motion estimation accuracy. In return, this could improve the soft tissue navigation accuracy and the overall outcome of the surgery or intervention. We assume that the tracked soft-tissue organ is represented

in a discrete (point-based) fashion (Lim et al. (2015)), while its motion state at any discrete time point is assumed to be directly defined by the positions of all points that constitute its discrete representation. In this context, the motion measurements impose three main challenges: (1) they capture the real-time motion of the tracked soft-tissue organ at, in general, randomly non-uniformly scattered points that do not necessarily correspond either in number or physically to the state points (Kocev et al. (2014)), in this case the information source is incomplete; (2) the temporal resolution of different measurement mechanisms may vary, i.e., the measurements may arrive at multiple rates, in which case the source contains out-of-sequence data; (3) the measurements are contaminated with noise, i.e., they are uncertain. For an abstract visualization of the nature of motion measurements and the motion estimation problem itself, please see Figure 2.1.

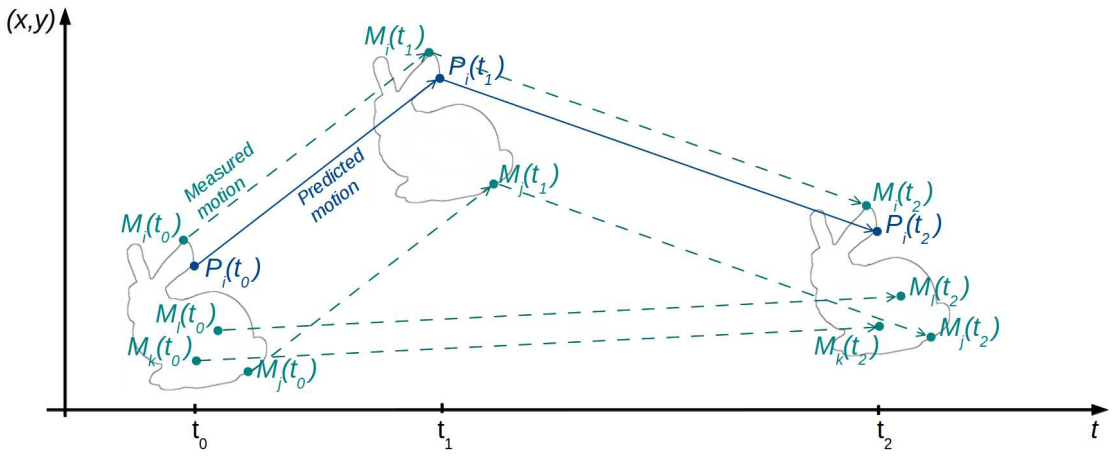

FIGURE 2.1: Predicting motion from asynchronous observations. M_i, M_j, M_k, and M_l are example measurement points, which are updated asynchronously at different points in time t_0, t_1, and t_2, while P_i is an example state point, which shall be predicted using uncertainty-aware asynchronous scattered motion interpolation.

To address the challenges imposed by the motion measurements information source in the context of soft tissue navigation, we propose an algorithm for uncertainty-aware interpolation of motion signal samples that are randomly, non-uniformly scattered in the spatiotemporal domain. To formally define out-of-sequence, i.e., asynchronous, motion signal samples, we propose a novel motion field representation (see Section 2.3). Our proposed algorithm employs Gaussian process (GP) regression (Rasmussen (2006)) (see Section 2.4) to perform the interpolation in an optimal unbiased fashion. The GP embeds the motion prior, which is then conditioned on the motion signal samples when performing the regression. The conditioning leads to an estimate of the (full) posterior distribution over motion signal functions, which could be directly used

for summarizing and conveying the uncertainty (Risholm et al. (2010)) in any registered pre-interventional/pre-operative soft-tissue navigation information. In this regard, we believe that providing information about the uncertainty in the registered pre-interventional/pre-operative soft-tissue navigation information to the surgeon or intervention specialist could improve their decisions and the overall outcome of the surgery or intervention. The work of Lüthi et al. (2011) is similar to ours on this point. They use a zero-mean GP to model the motion prior and perform regression on motion signal samples that are randomly non-uniformly scattered in the spatial domain. Their motion signal samples are, in contrast to ours, in-sequence, i.e., synchronous. Hence, they do not need to use any temporal information. In contrast to their work, we propose using a non-zero mean function (see Subsection 2.5.1) and enable the interpolation in the spatiotemporal domain. We demonstrate that non-zero-mean GPs outperform zero-mean ones independent of whether the data are randomly non-uniformly scattered in the spatial or in the spatiotemporal domain. To enable the interpolation in the spatiotemporal domain, we propose using a squared exponential covariance function with anisotropic distance measure (ADM) (Rasmussen and Williams (2006)) (see Subsection 2.5.2). The spatiotemporal GP formalism facilitates the prediction of tissue displacements at any location, and for any time interval, from observed deformations that are sparse in space and time. We evaluated our algorithm by interpolating simulated randomly non-uniformly spatiotemporally scattered uncertain motion signal samples and comparing the interpolation results with the respective simulated ground truth (see Section 2.6). As in some applications the randomly non-uniformly spatiotemporally scattered uncertain motion signal samples may be drawn over some restricted region of space, e.g., on the surface of the soft-tissue organ, we also evaluated our proposed method on such simulated measurements.

2.2 Related Work

Our proposed GP regression model can be seen as a deformation model (Sotiras et al. (2013)) because it is intended to be used for interpolation of soft, i.e., deformable, tissue motion signal samples. Sotiras et al. (2013) classified all deformation models according to what drives the geometric transformations that are computed by these models. According to their top-level classification, the geometric transformation could be: 1) inspired by physical models (Wassermann et al. (2014)), 2) inspired by interpolation/approximation theory (Ledesma-Carbayo et al. (2005); Perperidis et al. (2005); Rohr et al. (2001); Vandemeulebroucke et al. (2011); Wörz and Rohr (2008)), and/or 3) knowledge-based, i.e., embedding prior information regarding the sought deformation, (Glocker et al. (2009);

Kocev et al. (2014); Lüthi et al. (2011)). For a more detailed classification of all deformation models, we refer the reader to the work of Sotiras et al. (2013). In the following, we discuss in detail the above classified related work and contrast it against our own work.

Rohr et al. (2001) employed approximating thin-plate splines (TPS) to account for both isotropic and anisotropic landmark localization errors, where the corresponding TPS result from a minimizing functional with respect to the sought transformation. They weight the quadratic/squared Euclidean distance between the corresponding landmarks according to the landmarks localization uncertainty and, in this way, control the influence of the landmarks on the registration result. In follow-up work, Wörz and Rohr (2008) improved the accuracy of the approximating TPS (Rohr et al. (2001)) by incorporating Gaussian elastic body splines (GEBS) resulting in a new approximation approach with an improved underlying physical deformation model. They used an energy functional related to the Navier equation under Gaussian forces, while the landmarks are still individually weighted in a similar fashion according to their localization uncertainties. The use of Gaussian forces (physically more plausible as they do not diverge, and decrease with distance) provided them with a free parameter controlling the locality of the transformation. One problem in the two above approaches is that they treat the landmark and regularization terms independently, which could be improved because the landmarks can constrain the regularization term itself as well (Lüthi et al. (2011)). We incorporate the landmark information (in general, the motion signal samples) during the training phase as part of the regularization as a priori knowledge on the admissible deformations. Moreover, in contrast to our approach, they do not handle asynchronous data and do not estimate the uncertainty in the resulting transformation.

Several groups have investigated the problem of spatiotemporal image registration (Bersvendsen et al. (2016); Ledesma-Carbayo et al. (2005); Perperidis et al. (2005); Shi et al. (2013); Vandemeulebroucke et al. (2011)). Perperidis et al. (2005) addressed the spatiotemporal iconic, i.e., intensity-based, registration of 3D image sequences by devising 4D affine and free-form deformation (FFD) (based on a 4D B-Spline model) models that are separated into spatial and temporal components. Ledesma-Carbayo et al. (2005) investigated the spatiotemporal iconic registration of 2D image sequences by using a semi-local spatiotemporal parametric linear model, based on 3D B-splines, that is also separable in time and space. Both Perperidis et al. (2005) and Ledesma-Carbayo et al. (2005) place the space- and time-axis basis functions on a uniform rectangular spatial grid and regularly spaced time intervals, respectively. Ledesma-Carbayo et al. (2005) discuss in depth the choice of time- and space-axis scale parameters, governing the knot spacing, which, in our case, would be handled by the characteristic length-scales of our covariance function (see Subsection 2.5.2). Shi et al. (2013) revisited the classic FFD approach and

devised a novel sparse representation for FFD using the principles of compressed sensing. They reconstruct the deformation from a pair of images (or image sequences) and apply a sparsity constraint to the parametric space. They extended the sparsity constraint to the temporal domain and proposed a temporal sparse free-form deformation (TSFFD) model which enabled capturing of fine local details, e.g., motion discontinuities in the spatiotemporal domain. They addressed the trade-off between robustness and accuracy for FFD-based registration by deploying sparsity constraints as an additional regularization term. Bersvendsen et al. (2016) addressed the temporal alignment by optimizing the alignment of the normalized cross correlation (NCC)-over-time curves of the sequences (Perperidis et al. (2005) also employed NCC, in the optimization of their temporal component), within their proposed fully automatic method for spatiotemporal (spatially rigid) registration between two partially overlapping 3D image sequences. However, Bersvendsen et al. (2016), Ledesma-Carbayo et al. (2005), Perperidis et al. (2005), and Shi et al. (2013) do not estimate the transformation in an uncertainty-aware fashion. Moreover, in contrast to our approach, they do not address the problem of interpolating randomly non-uniformly scattered spatiotemporal motion signal samples, one that arises in the context of spatiotemporal geometric, i.e., landmark-based, image registration.

Our proposed deformation model falls into the group of knowledge-based statistically constrained geometric transformations. Similar to Glocker et al. (2009), we determine a probability density function modeling the prior distribution of the sought deformation field. In contrast to our approach, Glocker et al. (2009) use Gaussian mixture models (GMMs) to represent their probability density functions and employ a Markov random field (MRF)-based formulation of the registration problem when computing the maximum a posteriori (MAP) estimate of the deformation field on a regular grid of control points (as in FFDs). For any other point in the domain, they employ a B-spline basis functions-based interpolation between the estimated control point displacements. Regarding the MRF graph cost function, they set the edge penalty costs based on the negative log likelihood, which in return, is defined by their prior probability density functions. As a result, they compute a single optimal deformation estimate, i.e., the MAP estimate, while we estimate the full a posteriori deformation field distribution. Therefore, in addition, we estimate the uncertainty in the deformation estimate at any point in the domain. We would, in general, be able to incorporate their clustering idea by training different instances of our proposed GP model for each identified cluster.

To the best of our knowledge, the work of Lüthi et al. (2011) is the most similar to ours. They model the deformations as a zero-mean vector-valued GP and regard the landmarks as additional information on the admissible deformations. In the general setting,

they reduce the vector-valued case to the scalar case by constructing a matrix-valued covariance function as the product of a scalar-valued covariance function and a symmetric positive definite $N \times N$ matrix (N is the number of output dimensions), encoding the a priori knowledge about the correlation between the output dimensions. Our assumption with respect to the independence between the motion signal output dimensions is equivalent to using an $N \times N$ diagonal matrix in this case. They make the same assumption in all of their test examples. Furthermore, in contrast to our approach, they work with input landmarks that define synchronous motion signal samples that are randomly non-uniformly scattered in the spatial domain. In contrast to their approach, we propose to use a non-zero mean GP and allow the motion signal samples to be randomly non-uniformly scattered in the spatiotemporal domain. In a follow-up work, Gerig et al. (2014) devised a new method for spatially-varying (allowing for different regularization properties in different regions) iconic registration using GP priors. As a result, they came up with a non-stationary GP, that allowed to model different amount of smoothness in different regions. In this way, they were able to differentiate between tissue types or to make the regularization stronger in regions with noisy data. In theory, their proposed GP can have any mean function (Lüthi et al. (2013)). However, in practice they use a zero-mean function in all of their test examples, except for their statistical shape model for which they estimate the mean function based on example shapes. To present their approach in the medical setting, they demonstrated a solution for the challenging problem of atlas-based skull registration of cone-beam CT images. Similar to Gerig et al. (2014), Zhao et al. (2017) presented an alternative method for spatially-varying registration which is also anisotropic. Zhao et al. (2017) presented physically realistic deformations for the task of surface registration, by modeling the surface as an orthotropic elastic thin shell. They devised a statistical framework (Physical-Energy-Based MRF model) that can be deployed for estimating spatially varying anisotropic shell elasticity parameters with the input being only a set of known surface deformations. In parallel to estimating the elasticity parameters, they estimate the registration as well. They applied their approach in the context of 3D endoscopy reconstruction, which requires to generate a 3D reconstruction surface from multiple endoscopic movie frames. They managed to register all single-frame 3D reconstruction surfaces into a single surface. We view the novelties presented by Gerig et al. (2014) and Zhao et al. (2017) as complementary to the novelties presented in this paper.

Wassermann et al. (2014) employed a stochastic differential equation (SDE) for modeling the deformations as the evolution of a time-varying velocity field. The SDE is defined based on a deterministic ordinary differential equation (ODE) and a covariance function that is calculated as the matrix Green's function of the linear differential operator that regularizes the deterministic velocity field. The linear differential operator restricts the

deformations to the space of diffeomorphisms. Hence, their framework is suitable for modeling large deformation diffeomorphisms. In principle, they place a zero-mean GP prior with the above covariance function on the stochastic velocity field. They focus on operators that regularize in space but not in time. In contrast to their approach, we regularize in the spatiotemporal domain and employ a non-zero-mean GP prior.

In other medical imaging contexts, GPs have been used to quantify the uncertainty in image segmentation of deformable objects by defining a probability distribution of image segmentation boundaries as a GP and measuring the effect of using various plausible segmentation samples therefrom (Lê et al. (2015)).

GPs have generally been employed in several areas of visual computing for interpolation of uncertain data (Schlegel et al. (2012); Stytz and Parrott (1993); Wachinger et al. (2014)). In all of these cases, the data are given on a uniform spatial grid, which renders the zero-mean GP prior assumption by Schlegel et al. (2012) and Wachinger et al. (2014) as not very critical. In other words, if one assumes a constant-mean latent process function and uniform data contaminated with Gaussian white noise, one could model a zero-mean process by subtracting the empirical mean from the input data and adding it back after processing, if necessary (Schlegel et al. (2012)). However, if the data are given on a structured grid (e.g., if one applies the method of Wachinger et al. (2014) in the context of non-rigid image registration), or on an unstructured grid, or are randomly non-uniformly scattered, we believe that the empirical mean of the data is then less representative of the latent process mean even if the latent process mean is constant throughout the domain mainly because of the presence of noise. If, in reality, there is a large deviation from the assumption of a constant-mean latent process function, then one cannot, by default, model, as described above, a zero-mean GP prior. To address these aspects, we propose non-constant-mean GP models that can handle data that are randomly non-uniformly scattered in the spatiotemporal domain. This imposes the challenges of identifying and training a suitable non-constant mean function as well as an appropriate covariance function that can handle the aspects of drift (Stytz and Parrott (1993)) and spatiotemporal distribution. The randomness and non-uniformity in the distribution of the given data in the spatiotemporal domain are also addressed during the training phase.

For a general overview of traditional/classical and generalized interpolation techniques, we refer to the work of Thévenaz et al. (2000a,b).

2.3 Motion Field Representation

We define the motion signal function as

$$u : \mathbb{R}^{N+2} \to \mathbb{R}^N, \tag{2.1}$$

where the first N dimensions in the input domain are used to specify the location of a given N-dimensional spatial point, and the last two are used to define the time interval in which the motion of this spatial point took place. Each $(N+2)$-dimensional spatiotemporal point is mapped to an N-dimensional displacement vector specifying the motion that the given spatial point underwent within the given time interval. The displacement vector can be seen as the position of the given spatial point at the end of the time interval t_{end}, where the position is defined in a local coordinate system with the origin at the position of the spatial point at the beginning of the time interval t_{begin}. In contrast to the classical Lagrangian specification of a flow field, our representation of the motion field allows for specifying the motion between any time interval (t_{begin}, t_{end}). This is especially required for specifying asynchronous motion signal samples.

2.4 Gaussian Process Regression

A GP is defined as a collection of random variables, any finite number of which have (consistent) joint Gaussian distributions (Rasmussen (2006)). A scalar-valued GP is uniquely defined through its mean $m(x) : \Omega \to \mathbb{R}$ and covariance $k(x, x') : \Omega \times \Omega \to \mathbb{R}$ functions, where Ω is an index set. GPs can be used to define distributions over functions. We use the following notation:

$$f \sim \mathcal{GP}(m, k) \tag{2.2}$$

to denote that the function f is distributed as a GP with mean function m and covariance function k. Moreover, GP models can be used to formulate a Bayesian framework for regression. In that context, a GP is used as a prior for Bayesian inference. In order to make predictions for unseen test cases x_*, one needs to compute the posterior by conditioning the prior on a given training data set of n observations

$$\mathcal{D} = \{(x_i, y_i) | i = 1, ..., n\}. \tag{2.3}$$

where x_i are the training data locations and y_i are the observations (usually noisy, see Subsection 2.5.3) of the function values $f(x_i)$. It is generally assumed that the noise is additive independent and identically distributed zero-mean Gaussian, i.e.,

$$y(x) = f(x) + \epsilon,\ \epsilon \sim \mathcal{N}(0, \sigma_n^2),\ y \sim \mathcal{GP}(m, k + \sigma_n^2 \delta_{ii'}) \tag{2.4}$$

where $\delta_{ii'}$ is the Kronecker's delta function, i.e., $\delta_{ii'} = 1$ if and only if $i = i'$. After conditioning the prior process on the observations, one obtains the posterior process as follows:

$$
\begin{aligned}
f|\mathcal{D} \sim \mathcal{GP}(m_{\mathcal{D}}, k_{\mathcal{D}}) \\
m_{\mathcal{D}}(x) = m(x) + \Sigma(X, x)^T(\Sigma + \sigma_n^2\mathcal{I})^{-1}(y - \mu) \\
k_{\mathcal{D}}(x, x') = k(x, x') - \Sigma(X, x)^T(\Sigma + \sigma_n^2\mathcal{I})^{-1}\Sigma(X, x') \quad (2.5)
\end{aligned}
$$

where $\Sigma(X, x)$ is a vector of covariances between every training case and x (analogous for $\Sigma(X, x')$), μ is the vector of the function mean values at the training data locations, i.e., $\mu_i = m(x_i), i = 1, ..., n$, Σ is the covariance between the training data, and $\mathcal{I}$ is the identity matrix.

In the context of motion estimation, we would generally need to use a vector-valued Gaussian process (Lüthi et al. (2011)). The displacement $u(x)$, at any input location x, would need to be modeled as an N-dimensional random vector. The mean and covariance functions would then need to have the following form:

$$
\begin{aligned}
m(x) &: \Omega \to \mathbb{R}^N \\
k(x, x') &: \Omega \times \Omega \to \mathbb{R}^{N \times N}.
\end{aligned} \quad (2.6)
$$

However, we assume that the motion signal output dimensions are independent (see Section 2.2 for a discussion on how our work is related to the work of Lüthi et al. (2011) on this point). Therefore, we are able to employ a separate scalar-valued GP for each output dimension (Chan (2013)). For a more general treatment of vector-valued GPs, see Hein and Bousquet (2004) and Micchelli and Pontil (2005).

2.5 Model Selection

The Gaussian process model we employ is a hierarchical non-parametric (i.e., it needs access to all training data in the process of making predictions) model with two levels. At the first level are the free (hyper-) parameters θ of the underlying modeling functions (mean function (see Subsection 2.5.1), covariance function (see Subsection 2.5.2), and likelihood function (see Subsection 2.5.3)). The (hyper-) parameters control the distribution of the target values. At the second/top level, we have a (discrete) set of possible model structures, $\mathcal{H}_i$ out of which we should choose. On a side note, one could consider a zero level with the noise-free latent function values f at the training inputs as the parameters of the Gaussian process model. We select the model structure $\mathcal{H}_i$ at the second/top level by specifying the function families to which we believe the mean

(see Subsection 2.5.1), covariance (see Subsection 2.5.2), and likelihood (see Subsection 2.5.3) functions belong. As we normally have only vague information about the (hyper-) parameters, we deploy a mechanism for learning them from the training data (see Subsection 2.5.4).

2.5.1 Mean Function

Generally, the mean function of a scalar-valued GP $f(x)$ is defined as follows

$$m(x) = \mathbb{E}[f(x)], \tag{2.7}$$

where $\mathbb{E}$ is the expectation. In order to keep notations simple, the mean function of the prior GP is often set to zero (Rasmussen and Williams (2006)). The zero-mean assumption for the prior GP does not imply that the posterior GP would be zero-mean, i.e., from this point of view, one could see this assumption as not being critical. However, several problems may arise with respect to the interpretability of the model, the expressiveness of prior information, etc. (Rasmussen and Williams (2006)). Schlegel et al. (2012) model a zero-mean process on their data by subtracting the empirical mean from the input data and adding it back after processing, if necessary. However, Kuss (2006) pointed out that, in general, the mean of the data is not necessarily the mean of the process. Furthermore, we argue that the use of the empirical mean in this way is even more critical when dealing with non-uniformly scattered training data in contrast to when the data are uniformly scattered or even on a Cartesian grid. Therefore, we propose to model the mean function of the prior GP explicitly in order to be able to specify an appropriate non-zero mean function. One could use a fixed (deterministic) mean function or alternatively a few fixed basis functions with a set of coefficients/parameters which would need to be inferred from the training data. Then, one normally optimizes over the hyperparameters of the covariance function (see Subsection 2.5.2) jointly with the parameters of the mean function when fitting the model (see Subsection 2.5.4) on the training data. We build upon the work of O'Hagan and Kingman (1978) and propose to couple their linear mean function $m_1(x) = \beta^T x$ with a constant mean function $m_2(x) = c$, when modeling the mean function of the GP that embeds the motion prior. The two functions are simply added, i.e.,

$$m(x) = m_1(x) + m_2(x) = \beta^T x + c \tag{2.8}$$

which results in a composite mean function (Rasmussen and Nickisch (2010)). Hence, the prior mean is realized as the sum of a linear and constant function. The parameters β and c are inferred from the data (see Subsection 2.5.4).

2.5.2 Covariance Function

In general, the covariance function of a scalar-valued GP $f(x)$ is defined as follows

$$k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]. \tag{2.9}$$

As such, the choice of a covariance function induces the properties, e.g., stationarity, isotropy, smoothness, periodicity, etc., of the functions that are likely under the GP prior (Barber (2012); Rasmussen and Williams (2006)). If the covariance function is a function only of $x - x'$ then it is stationary, while if it is a function only of $|x - x'|$ then it is isotropic (Rasmussen and Williams (2006)). Suitable properties for the covariance function generally need to be learned from the training data. Some of these properties are encoded through the hyperparameters (e.g., characteristic length-scale, variance, etc.) of the chosen covariance function. In the context of doing regression on spatiotemporal motion signal samples, we need to use covariance functions whose input domain Ω is a subset of $\mathbb{R}^D$. In this regard, we propose using the Squared Exponential (SE) covariance function with distance measure with a different characteristic length-scale (hyper-) parameter per input dimension

$$k(x, x') = \sigma_f^2 \exp(-\frac{1}{2}(x - x')^T P^{-1}(x - x')) \tag{2.10}$$

where P is a diagonal matrix with length-scale (hyper-) parameters $l_1^2, l_2^2, ..., l_D^2$ as entries on the diagonal, and σ_f^2 is the signal variance (Rasmussen and Williams (2006)). As this covariance function uses different length-scales l_i on different dimensions, it is an anisotropic covariance function. In simple words, P encodes how far, along individual axes in input space, the input locations need to be so that the function values at those locations become uncorrelated. In our case, this is particularly suited for anisotropically adjusting the distance measure along the spatial and temporal axes in our input space. In other words, we can handle the anisotropy between space and time. These (hyper-) parameters will be learned from the data (see Subsection 2.5.4). In general, the functions in the SE covariance function family are infinitely differentiable, i.e., a GP process with a covariance function from this family has mean square derivatives of all orders, i.e., it is very smooth. Therefore, the use of this covariance function could also be interpreted as a mechanism for regularization that penalizes non-smooth solutions (Lüthi et al. (2011)). We therefore believe that the SE covariance function is particularly suitable for modeling soft-tissue motion prior within our target application because soft-tissue organs undergo smooth motion.

2.5.3 Likelihood Function

We employ a Gaussian likelihood function

$$p(y|f,\sigma_n) = \frac{1}{\sigma_n\sqrt{2\pi}}\exp(-\frac{(y-f)^2}{2\sigma_n^2}) \tag{2.11}$$

for regression, where y is the actual observation/measurement of the true latent value (of a component of the displacement field) f (see Section 2.4), and σ_n^2 is the noise variance (hyper-) parameter. In other words, this function defines how the noisy observations/measurements are assumed to diverge from the noise-free function f. The incorporation of a Gaussian likelihood function together with the Gaussian process prior allows for a posterior Gaussian process over functions and keeps things analytically tractable (Rasmussen and Williams (2006)). Our Gaussian noise model assumes homoscedasticity, i.e., the noise variance σ_n^2 is assumed to be constant throughout the domain. For modeling heteroscedastic Gaussian noise, we refer to the work of Goldberg et al. (1997) who uses the noise variance as a function of x.

2.5.4 (Hyper-) parameters Training

The (hyper-) parameters are optimized by maximizing the probability of the model, given the training data. The probability of the model, given the training data, is computed based on the marginal likelihood (ML), also known as model evidence,

$$\begin{aligned}\log\ p(y|X,\theta,\mathcal{H}_i)\\ = -\frac{1}{2}(y-\mu)^T\Sigma_y^{-1}(y-\mu) - \frac{1}{2}log|\Sigma_y| - \frac{n}{2}log2\pi,\end{aligned} \tag{2.12}$$

where $\Sigma_y = \Sigma+\sigma_n^2\mathcal{I}$. The ML is a probability distribution over the observations y, conditioned on the input locations x, the (hyper-) parameters θ (i.e., $\theta = (\beta, c, l_1, l_2, ..., l_D, \sigma_f, \sigma_n)$, see Subsections 2.5.1, 2.5.2, and 2.5.3), and the chosen model structure $\mathcal{H}_i$. The use of log ML is appropriate because it automatically incorporates a trade-off between data-fit (first term of Eq. 2.12, i.e., $-\frac{1}{2}(y-\mu)^T\Sigma_y^{-1}(y-\mu)$) and model complexity (second term of Eq. 2.12, i.e., $\frac{1}{2}log|\Sigma_y|$), i.e., it does not require an external parameter for controlling this trade-off (Rasmussen and Williams (2006)). The third term of Eq. 2.12, i.e., $\frac{n}{2}log2\pi$, is a normalization constant. Generally, the mechanism of using the data to estimate the prior parameters is known as empirical Bayes (Gelman et al. (2014)). This approximation of the complete hierarchical Bayesian analysis eliminates the need to put a probability model over all (hyper-) parameters. In principle, instead of putting prior distributions over the (hyper-) parameters and marginalizing them out, the (hyper-) parameters are set with the values that maximize the ML. The use of ML avoids over-fitting, which is associated with the classical maximum likelihood approach, by marginalizing out the

model parameters and allows comparison of different models on all training data, i.e., it eliminates the need to use cross-validation (Bishop (2006)). However, due to possible sensitivity of the ML on the prior, it is still recommended to use an independent test dataset for the comparison of different models in practical applications (Bishop (2006)). To demonstrate the effect of the characteristic length-scale parameter on the probability of an example model given example training data, we try to fit only the length-scale parameter l_i while the remaining hyperparameters are set in accordance with the process from which the example training data are drawn. The training data are drawn from a zero-mean GP with an SE covariance function with ADM defined over a 2D index set, with (hyper-) parameters $(l_1, l_2, \sigma_f, \sigma_n) = (0.25, 0.125, 1, 0.1)$. In Figures 2.2 and 2.3, one can observe the fitting of the characteristic length-scales l_1 and l_2 respectively on the training data. The plots show the corresponding log ML and the decomposition into its constituents as a function of the respective characteristic length-scale hyperparameter. In both plots, the negative complexity term increases (i.e., the model loses complexity) as the respective length-scale hyperparameter l_i increases. Furthermore, in both plots, the data-fit term decreases monotonically as the length-scale hyperparameter l_i increases, because the model loses flexibility. The marginal likelihood in Figure 2.2 reaches its peak value for $l_1 = 0.25$, while in Figure 2.3 the peak of the marginal likelihood is reached for $l_2 = 0.125$. This agrees with the respective characteristic length-scales of the GP from which the training data are drawn.

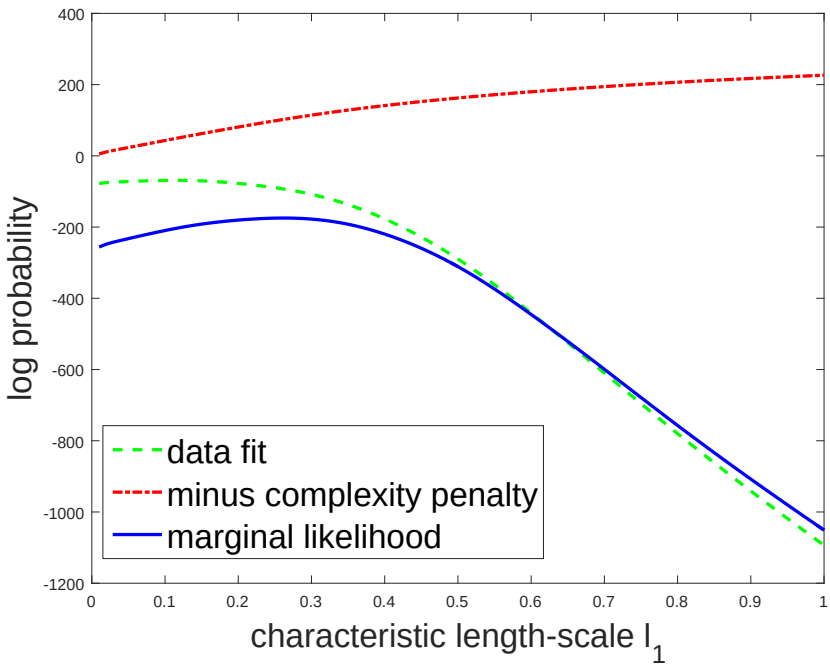

FIGURE 2.2: Log marginal likelihood decomposition into its constituents: data-fit and complexity penalty, as a function of the characteristic length-scale l_1.

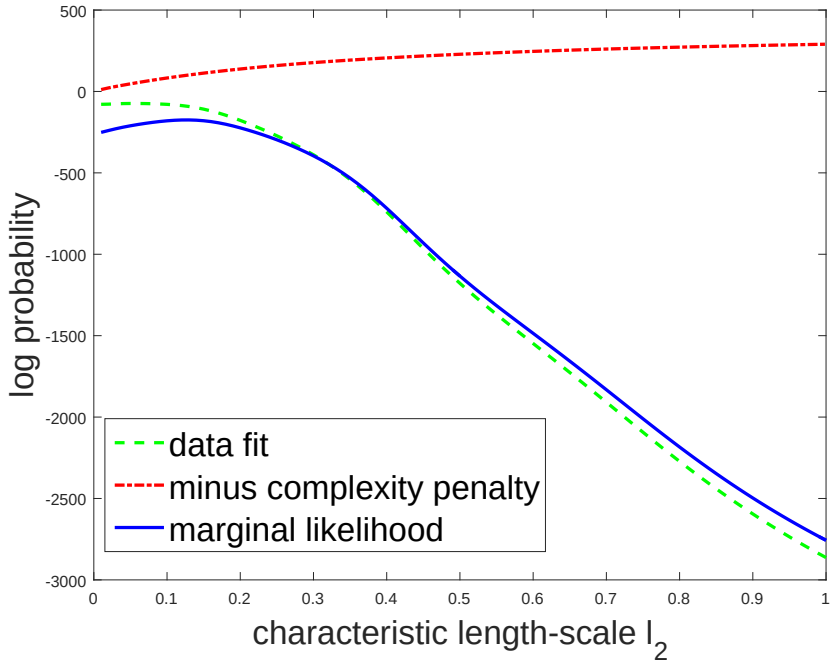

FIGURE 2.3: Log marginal likelihood decomposition into its constituents: data-fit and complexity penalty, as a function of the characteristic length-scale l_2.

In the general setting, we use conjugate gradients (Polak and Ribiere (1969)) to optimize the (hyper-) parameters jointly, while maximizing the probability of the model given the training data. To circumvent bad local extrema, we perform random restarts and select the (hyper-) parameters configuration that gives the maximum probability of the model given the training data.

2.6 Evaluation

In this section, we aim to demonstrate that, when doing regression on randomly non-uniformly scattered motion signal samples, employing a GP prior with the proposed non-zero-mean function yields better results than using a GP prior with a zero-mean function. Furthermore, we will show that, in the case that the observations of the function are randomly non-uniformly scattered in the spatiotemporal domain, our proposed covariance function with ADM is more suitable for defining the GP prior of a given motion signal function than the one of the same family with an isotropic distance measure (IDM). Therefore, we compare the obtained results when doing regression using each of the GP priors configurations in Table 2.1.

Each of the GP model configurations in Table 2.1 is used to perform regression on simulated soft-tissue motion data. We simulated ground-truth soft-tissue motion data (see Subsection 2.6.1) out of which a relatively small portion (contaminated with noise)

ID	Mean Func.	Cov. Func.	Lik. Func.
1	zero	SE w/ IDM	proposed
2	proposed	SE w/ IDM	proposed
3	proposed	proposed	proposed

TABLE 2.1: GP priors configurations.

was used as training data and the rest (noise-free) as ground-truth test data. The (hyper-) parameters of GP models having each of the above configurations are first trained (see Subsection 2.6.2) by maximizing the likelihood, as defined in Eq. 2.12, and then the prior is conditioned on the training dataset according to Eq. 2.5, when making predictions for unseen test cases. The predictions are then evaluated in terms of accuracy against the respective ground truth (see Subsection 2.6.3). We train three different scalar-valued GP models with the same configuration for each motion signal output dimension (assumed to be independent, see Section 2.4). Hence, predictions are made for each output dimension separately.

2.6.1 Data Simulation

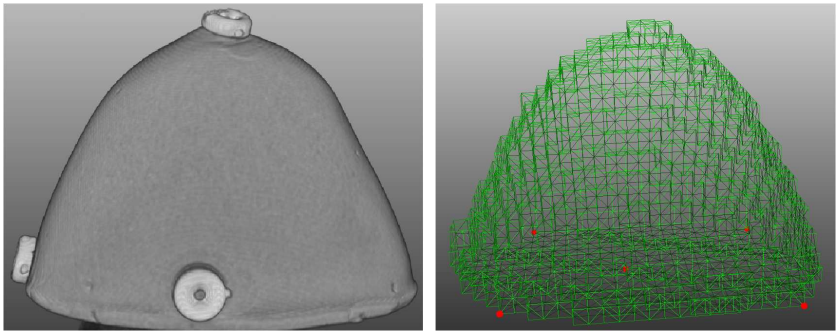

FIGURE 2.4: Left: Breast phantom with 4 markers (one on the back side). Right: FE model composed of tetrahedral elements which are extracted from the MRI scan data of the breast phantom. The points in red are fixed, i.e., the FE nonlinear motion prediction model considers these vertices as not moving. (Image courtesy of Kocev et al. (2014).)

The ground-truth motion data are simulated over time at the points that constitute the discrete representation of the given deformable object. For this purpose, we employed a finite element (FE) model (Georgii and Westermann (2008, 2005)) for physics-based simulation of motion data. For this evaluation, we created an FE model of the CIRS triple modality breast biopsy training phantom with 1,962 vertices (see Figure 2.4.) and used it to simulate a physically plausible non-linear motion that a soft-tissue organ is likely to undergo during a biopsy intervention. The motion that a soft tissue organ is likely to undergo during a biopsy intervention is basically the motion that takes place

upon pressing the target organ with a biopsy needle. We have defined a biopsy insertion point (the point in yellow in Figure 2.5) on the surface of the breast phantom and simulated an external point force acting on this point towards the breast phantom centroid. The largest simulated deformation is about 16.19 mm. It is important to note that it is

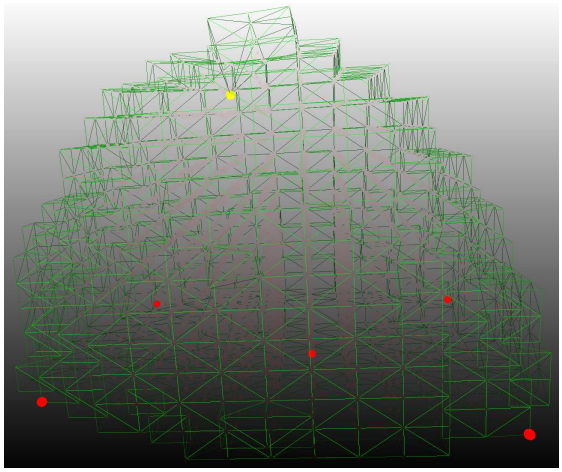

FIGURE 2.5: Deformed breast phantom FE model with biopsy insertion point visualized in yellow. The points in red are fixed, while the point in yellow depicts the location where the external point force is applied.

not necessary for the material properties of the model to match those of the CIRS phantom exactly, i.e., no possible deviations influence the accuracy of our evaluation on the specific simulated motion data, as long as the defined material properties describe some realistic (soft-tissue) material. In principle, one could use virtually any deformable FE model for simulating such non-linear motion data provided that the FE model simulates the defined (realistic) dynamics accurately. Formally, the simulated ground-truth test data are in the form of samples from the output of the motion signal function as defined in Eq. 2.1 with $N = 3$. We simulated a ground-truth test dataset

$$\mathcal{D}_{gt} = \{((x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-1)\Delta t, k\Delta t), f_i) \\ |i = 1, ..., n_*; k = 1, ..., s\} \quad (2.13)$$

with cardinality sn_*. Here, n_* corresponds to the number of vertices of the employed FE model, Δt corresponds to the time step size used by the time integration scheme, s is the total number of discrete time points, and $x_i^{(l)}$ is the l-th component of the ground-truth test location x_i. It is assumed that the state points and the FE vertices have 1-to-1 correspondence (Kocev et al. (2014)). For evaluating our algorithm on in-sequence (i.e., synchronous) uncertain motion signal samples, a training dataset

$$\mathcal{D}_{sync} = \{((x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-1)\Delta t, k\Delta t), y_i) \\ |i = 1, ..., n\} \quad (2.14)$$

is created, for a discrete time point k, $1 \leq k \leq s$, by selecting n ($n \ll n_*$) randomly non-uniformly scattered samples out of the sn_* simulated ground-truth test samples and then contaminating them with heteroscedastic Gaussian noise. In this regard, $y_i = f_i + \epsilon_i$ with $\epsilon_i \sim \mathcal{N}(0, \Sigma_i)$ being the Gaussian distributed noise value at x_i. For accuracy analysis of our algorithm on out-of-sequence (i.e., asynchronous) samples, another training data set

$$\begin{aligned}\mathcal{D}_{async} \\ &= \{((x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-1)\Delta t, k\Delta t), y_i) | i = 1, ..., n_1\} \\ &\cup \{((x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-2)\Delta t, k\Delta t), y_i) | i = 1, ..., n_2\} \\ &\cup \{((x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-3)\Delta t, k\Delta t), y_i) | i = 1, ..., n_3\} \end{aligned} \quad (2.15)$$

is created, for a discrete time point k, $3 \leq k \leq s$, by taking the union of three different subsets (with cardinalities n_1, n_2, and n_3 respectively) of the simulated ground-truth test data (i.e., in total $n = n_1 + n_2 + n_3 \ll n_*$ randomly non-uniformly scattered samples) and then contaminating the samples with heteroscedastic Gaussian noise. Note that each subset contains samples with different time intervals, i.e., the resulting training dataset, $\mathcal{D}_{async}$, contains asynchronous motion signal samples.

For the purpose of this evaluation, we set $n_* = 1,962$, $s = 6$, $\Delta t = 33\ ms$, $n = 26$, $n_1 = 9$, $n_2 = 9$, and $n_3 = 8$. We chose this setting because it is a realistic one for the medical context that drives our developments.

2.6.2 Training Results

Let us denote each trained GP model with $\mathcal{M}_{\mathcal{D}}^{(c,d)}$, where $\mathcal{D}$ is the dataset on which the model is trained, $c \in \{1, 2, 3\}$ corresponds to the ID of the employed GP prior configuration (see Table 2.1), and $d \in \{1, 2, 3\}$ is the index of the motion signal output dimension for which the model is trained. We optimized the (hyper-) parameters of each model by maximizing its evidence (estimated using Eq. 2.12, given the respective training dataset) as explained in Subsection 2.5.4. The training has been performed on example synchronous (see Eq. 2.14) and asynchronous (see Eq. 2.15) training datasets. The values of the optimized (hyper-) parameters of the proposed GP models $\mathcal{M}_{\mathcal{D}_{async}}^{(3,1)}$, $\mathcal{M}_{\mathcal{D}_{async}}^{(3,2)}$, and $\mathcal{M}_{\mathcal{D}_{async}}^{(3,3)}$ are presented in Tables 2.2 and 2.3. The resulting probability of each model $\mathcal{M}_{\mathcal{D}}^{(c,d)}$, given the respective training dataset, is presented in Table 2.4. Please note that we do not train GP models having the third configuration on synchronous datasets as in that case the temporal information is irrelevant. In the case of training on synchronous data, i.e., $\mathcal{D} = \mathcal{D}_{sync}$, the resulting model probabilities show that models having the second configuration, i.e., $\mathcal{M}_{\mathcal{D}_{sync}}^{(2,d)}$, are able to explain the example training data better than models having the first configuration, i.e., $\mathcal{M}_{\mathcal{D}_{sync}}^{(1,d)}$. In

	β	c
$d = 1$	[0.000105, 0.001851, -0.001966, 0.003670, -0.008100]	0.999858
$d = 2$	[-0.001354, 0.002312, 0.007315, -0.013983, 0.007166]	0.999935
$d = 3$	[-0.001398, -0.007534, -0.000155, 0.000535, -0.004999]	0.999874

TABLE 2.2: (Hyper-) parameters of the mean function of the proposed models $\mathcal{M}^{(3,d)}_{\mathcal{D}_{async}}$.

	l_1	l_2	l_3	l_4	l_5	σ_f	σ_n
$d = 1$	30.017140	30.021074	30.008921	600.005046	600.000000	0.005032	0.227006
$d = 2$	23.156668	23.155694	23.119362	300.062788	300.000000	0.015149	0.238498
$d = 3$	15.033298	15.028769	15.030861	600.007680	600.000000	0.010139	0.175856

TABLE 2.3: (Hyper-) parameters of the covariance and likelihood functions of the proposed models $\mathcal{M}^{(3,d)}_{\mathcal{D}_{async}}$.

regard to training on asynchronous data, i.e., $\mathcal{D} = \mathcal{D}_{async}$, models having the third configuration, i.e., $\mathcal{M}^{(3,d)}_{\mathcal{D}_{async}}$, accommodate the training data better than models having any of the other two prior configurations. Therefore, our proposed modeling of the mean (see Subsection 2.5.1) and covariance (see Subsection 2.5.2) functions allows for GP models that can be better trained to explain training data given in the form of randomly non-uniformly scattered motion signal samples.

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	-27.9233	-31.7146	-8.4969	-1.1611	-9.6756
$d = 2$	-9.3417	-13.1894	14.8096	13.9991	-8.6495
$d = 3$	-10.9099	-12.9253	-2.6988	-3.7344	-9.1918

TABLE 2.4: Negative log marginal likelihoods of models $\mathcal{M}^{(c,d)}_{\mathcal{D}}$, with $\mathcal{D}$ being the training dataset (either synchronous or asynchronous), c the ID of the GP prior configuration (see Table 2.1), and d the index of the motion signal output dimension.

2.6.3 Accuracy Analysis

Using Eq. 2.5, each trained GP model $\mathcal{M}^{(c,d)}_{\mathcal{D}}$ is conditioned on the training dataset $\mathcal{D}$ in order to make predictions, at the original n_* test locations $(x_i^{(1)}, x_i^{(2)}, x_i^{(3)}, (k-1)\Delta t, k\Delta t)$, for the d-th motion signal output dimension using the c-th GP prior configuration. The resulting predictions for the motion signal function mean values at the n_* test locations are represented as an n_*-dimensional vector $\bar{f}_*$, which is evaluated in terms of accuracy

against the respective n_*-dimensional vector of ground-truth function mean values $\bar{f}_{*_{gt}}$. The difference between the vectors $\bar{f}_*$ and $\bar{f}_{*_{gt}}$ is measured using:

1. Euclidean distance;
2. cosine distance;
3. (Pearson) correlation distance;
4. root-mean-square error (RMSE);
5. mean absolute error (MAE), and
6. Wilcoxon two-sided rank sum test (Wilcoxon (1945)), with the null hypothesis that data in $\bar{f}_*$ and in $\bar{f}_{*_{gt}}$ are samples from continuous distributions with equal medians.

By combining the predictions (obtained using the $c-th$ GP prior configuration and training dataset $\mathcal{D}$) $\bar{f}_*$ for each output dimension into an $n_* \times 3$ matrix, we obtain the estimate of the full motion signal function mean values at the given n_* test locations. Each such $n_* \times 3$ matrix of (3D) displacement predictions together with the respective n_* test locations define an $n_* \times 3$ matrix $\bar{S}_*$ with its rows being the (3D) position mean values of all displacement vectors' endpoints, which is then evaluated in terms of accuracy against the respective ground truth $\bar{S}_{*_{gt}}$ using:

1. $L_{2,1}$ norm of $(\bar{S}_* - \bar{S}_{*_{gt}})^T$, and
2. $L_{2,2}$, i.e., Frobenius, norm of $(\bar{S}_* - \bar{S}_{*_{gt}})$.

In the following, we present accuracy analysis results on both synchronous and asynchronous training datasets.

2.6.3.1 Accuracy Analysis Results

Synchronous Data In performing regression from space and time to motion signal function values using a synchronous training dataset, $\mathcal{D} = \mathcal{D}_{sync}$, all predictions $\bar{f}_*$ by models $\mathcal{M}^{(2,d)}_{\mathcal{D}_{sync}}$ are closer, in terms of Euclidean distance, cosine distance, and RMSE, to the ground truth $\bar{f}_{*_{gt}}$ than those by models $\mathcal{M}^{(1,d)}_{\mathcal{D}_{sync}}$ (see Tables 2.5, 2.6, and 2.8). In terms of (Pearson) correlation distance and MAE, the predictions for the second output dimension, $d = 2$, are closer to the ground truth when using the first GP prior configuration, $c = 1$. For the other two output dimensions, the second GP prior configuration,

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	3.48379	3.28559	3.77276	6.75917	2.83698
$d = 2$	15.61828	15.52604	17.40310	14.72825	5.36545
$d = 3$	11.27818	11.04141	4.66890	8.25508	5.92519

TABLE 2.5: Euclidean distances between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	0.19607	0.17226	0.10324	0.15673	0.89327
$d = 2$	0.12244	0.12214	0.26239	0.13359	0.32024
$d = 3$	0.61894	0.57702	0.03873	0.06904	0.00839

TABLE 2.6: Cosine distances between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

$c = 2$, allows for better predictions (see Tables 2.7 and 2.9). $\mathcal{M}^{(1,2)}_{\mathcal{D}_{sync}}$ is better, in terms of MAE but not in terms of RMSE, than $\mathcal{M}^{(2,2)}_{\mathcal{D}_{sync}}$ mainly because in that case the first configuration gives higher-value outliers in the component-wise absolute differences to which RMSE gives higher weights (Chai and Draxler (2014)). However, RMSE may be a more appropriate metric for deciding which algorithm is better suited for safety-critical applications (Knight (2002)), e.g., navigated surgery (Mezger et al. (2013)) where large errors are to be avoided. An additional argument supporting that RMSE is more appropriate to represent model performance than MAE in this case is that the distribution of the error $\bar{f}_* - \bar{f}_{*_{gt}}$ yielded by both configurations is Gaussian (Chai and Draxler (2014)). To confirm that the error is Gaussian-distributed, we used a t-test with the null hypothesis being that the data in $\bar{f}_* - \bar{f}_{*_{gt}}$ come from a normal distribution with unknown variance and mean value equal to the error sample set mean. Moreover, we visually analyzed the histogram of $\bar{f}_* - \bar{f}_{*_{gt}}$ yielded by both configurations (see Figures 2.6 and 2.7). As the means of the error sample sets yielded by both configurations are not exactly zero, the error sample sets are slightly biased (Chai and Draxler (2014)). Therefore, we provide, in addition, the respective standard error (SE) information (Chai and Draxler (2014)) (see Table 2.10). Note that when the error distribution and sample set are unbiased, the SE is equivalent to the RMSE (Chai and Draxler (2014)). While the results from the Wilcoxon two-sided rank sum tests show that there is no strong evidence supporting that there is a significant (at the 5% significance level) difference between the predictions yielded by $\mathcal{M}^{(2,d)}_{\mathcal{D}_{sync}}$ and the respective ground truth, however,

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	0.63771	0.51480	1.05600	1.30947	0.53873
$d = 2$	0.58169	0.60213	0.97160	0.15555	0.04888
$d = 3$	0.90247	0.81330	0.07474	0.15025	0.01750

TABLE 2.7: (Pearson) correlation distances between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	0.07865	0.07418	0.08517	0.15260	0.06405
$d = 2$	0.35260	0.35052	0.39290	0.33251	0.12113
$d = 3$	0.25462	0.24927	0.10541	0.18637	0.13377

TABLE 2.8: Root-mean-square errors between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	0.03910	0.02879	0.08428	0.14592	0.05122
$d = 2$	0.07216	0.08021	0.38632	0.32558	0.10092
$d = 3$	0.05766	0.05222	0.08918	0.15518	0.10830

TABLE 2.9: Mean absolute errors between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

there is strong evidence supporting that there is significant difference (except for $d = 1$) between the predictions made using GP prior configuration 1 (i.e., $c = 1$) and the respective ground truth. Hence, there is stronger evidence that, in general, the second GP prior configuration allows for better estimation of the true motion function median values (see Table 2.11). Furthermore, both the $L_{2,1}$ norm of $(\bar{S}_* - \bar{S}_{*_{gt}})^T$ and $L_{2,2}$ (i.e., Frobenius) norm of $(\bar{S}_* - \bar{S}_{*_{gt}})$ are smaller when using the second GP prior configuration (see Tables 2.12 and 2.13), i.e., the overall motion estimation error is larger when using the first GP prior configuration.

Asynchronous Data When mapping from spatiotemporal locations to motion signal function values using an asynchronous training dataset (i.e., $\mathcal{D} = \mathcal{D}_{async}$), all predictions $\bar{f}_*$ by models $\mathcal{M}_{\mathcal{D}_{async}}^{(3,1-2)}$ are closer, in terms of Euclidean distance, (Pearson) correlation

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	0.07863	0.07407	0.012308	0.044653	0.055790
$d = 2$	0.35084	0.34961	0.071571	0.067533	0.099143
$d = 3$	0.25241	0.24850	0.057916	0.131253	0.087739

TABLE 2.10: Standard error between $\bar{f}_*$ and $\bar{f}_{*_{gt}}$.

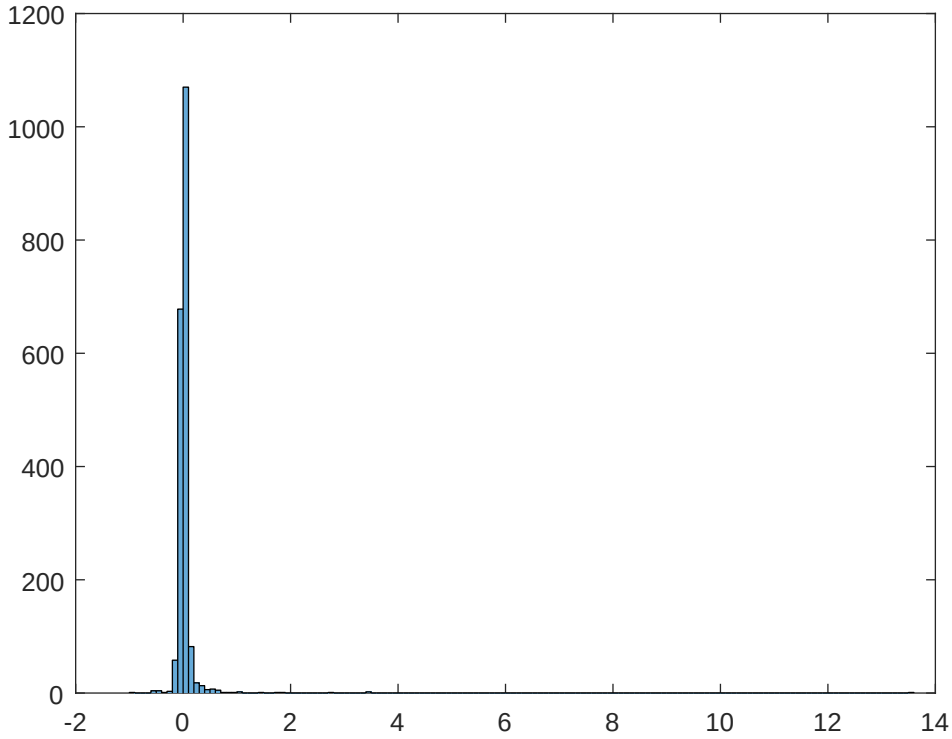

FIGURE 2.6: Histogram of $\bar{f}_* - \bar{f}_{*_{gt}}$ yielded by $\mathcal{M}^{(1,2)}_{\mathcal{D}_{sync}}$.

distance, RMSE, and MAE, to the ground truth $\bar{f}_{*_{gt}}$ than those by models $\mathcal{M}^{(1,1-2)}_{\mathcal{D}_{async}}$ and $\mathcal{M}^{(2,1-2)}_{\mathcal{D}_{async}}$ (see Tables 2.5, 2.7, 2.8, and 2.9). Furthermore, the predictions by model $\mathcal{M}^{(3,3)}_{\mathcal{D}_{async}}$ are better, in terms of cosine and (Pearson) correlation distances, than those by models $\mathcal{M}^{(1,3)}_{\mathcal{D}_{async}}$ and $\mathcal{M}^{(2,3)}_{\mathcal{D}_{async}}$ (see Tables 2.6 and 2.7). In all cases where the third GP prior configuration fails to produce better predictions than the other two configurations, the first GP prior configuration is the one that outperforms (except in terms of cosine distance for the second output dimension) the other configurations (see Tables 2.5, 2.6, 2.8, and 2.9). The results from the Wilcoxon two-sided rank sum tests show that there is strong evidence supporting that there is a significant (at the 5%

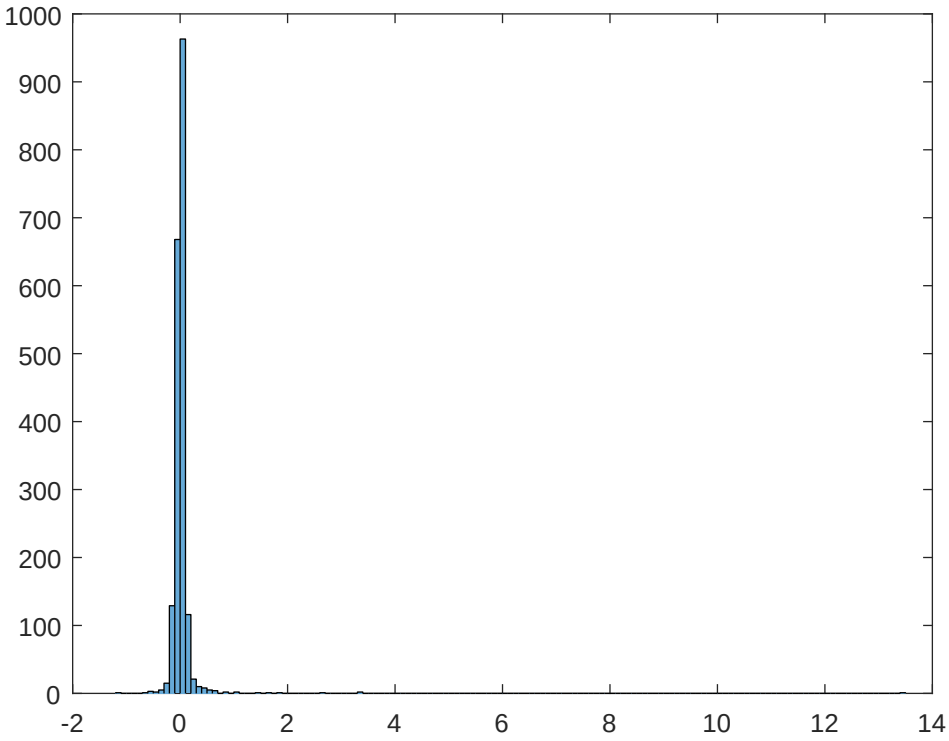

FIGURE 2.7: Histogram of $\bar{f}_* - \bar{f}_{*_{gt}}$ yielded by $\mathcal{M}^{(2,2)}_{\mathcal{D}_{sync}}$.

	$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
	$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
$d = 1$	(0.110516, 0)	(0.360029, 0)	(0, 1)	(0, 1)	(0, 1)
$d = 2$	(0.000556, 1)	(0.077498, 0)	(0, 1)	(0, 1)	(0, 1)
$d = 3$	(0, 1)	(0.326926, 0)	(0, 1)	(0, 1)	(0, 1)

TABLE 2.11: Wilcoxon two-sided rank sum test null hypothesis (that data in $\bar{f}_*$ and in $\bar{f}_{*_{gt}}$ are samples from continuous distributions with equal medians) results ((p, h): "h=1" indicates a rejection of the null hypothesis, while "h=0" indicates a failure to reject the null hypothesis, at the 5% significance level, based on the estimated p-value p).

significance level) difference between the predictions yielded by any $\mathcal{M}^{(c,d)}_{\mathcal{D}_{async}}$ and the respective ground truth (see Table 2.11). However, the Wilcoxon two-sided rank sum test does not provide strong evidence that any of the three different GP prior configuration allows for better estimation of the true motion function median values when compared to the others. On the other hand, the $L_{2,1}$ norm of $(\bar{S}_* - \bar{S}_{*_{gt}})^T$ and $L_{2,2}$ (i.e., Frobenius) norm of $(\bar{S}_* - \bar{S}_{*_{gt}})$ are clearly in favor of using the third GP prior configuration. They are smallest when using the third GP prior configuration (see Tables 2.12 and 2.13),

$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
228.0773	219.4586	803.9132	783.8536	343.4022

TABLE 2.12: $L_{2,1}$ norm of $(\bar{S}_* - \bar{S}_{*_{gt}})^T$.

$\mathcal{D} = \mathcal{D}_{sync}$		$\mathcal{D} = \mathcal{D}_{async}$		
$c = 1$	$c = 2$	$c = 1$	$c = 2$	$c = 3$
19.5771	19.3330	18.4092	18.1866	8.4820

TABLE 2.13: $L_{2,2}$ (i.e., Frobenius) norm of $(\bar{S}_* - \bar{S}_{*_{gt}})$.

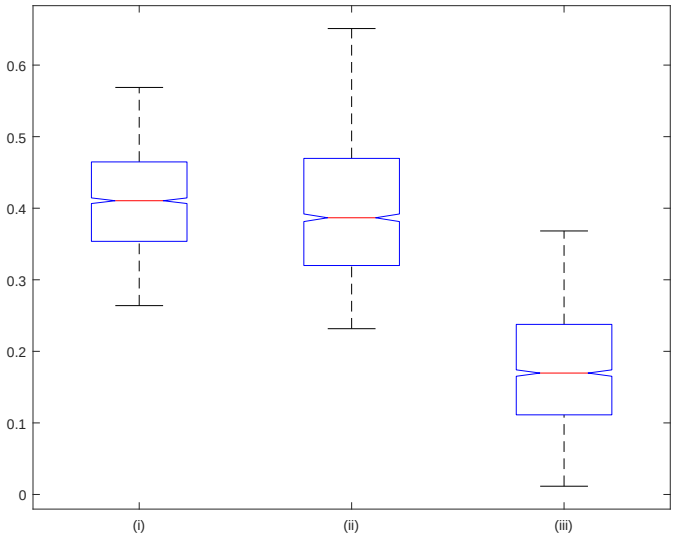

FIGURE 2.8: Boxplots of the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by different GP prior configurations trained on $\mathcal{D}_{async}$: (i) using models $\mathcal{M}^{(1,d)}_{\mathcal{D}_{async}}$, (ii) using models $\mathcal{M}^{(2,d)}_{\mathcal{D}_{async}}$, (iii) using models $\mathcal{M}^{(3,d)}_{\mathcal{D}_{async}}$.

i.e., the overall motion estimation error is smallest when using our proposed GP prior configuration. In this regard, note that the overall motion estimation error is decreased by more than 50% in case of using our proposed mean and covariance functions (i.e., the third GP prior configuration) when doing regression on asynchronous training data. This major improvement is achieved mainly by using our proposed covariance function. For a visual depiction of the overall median errors yielded by different GP prior configurations trained on the same asynchronous dataset, we boxplot the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by each GP prior configuration. In this regard, the boxplot notches in Figure 2.8 offer evidence of a statistically significant difference (at the 5% significance level) between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP

$\mathcal{D} = \mathcal{D}_{sync}$	$\mathcal{D} = \mathcal{D}_{async}$		
$(c = 1, c = 2)$	$(c = 1, c = 2)$	$(c = 1, c = 3)$	$(c = 2, c = 3)$
(0.006665, 1)	(0, 1)	(0, 1)	(0, 1)

TABLE 2.14: Wilcoxon two-sided rank sum test null hypothesis (that the values of the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by one GP prior configuration trained on a given dataset and those yielded by another configuration trained on the same dataset are samples from continuous distributions with equal medians) results ((p, h): "h=1" indicates a rejection of the null hypothesis, while "h=0" indicates a failure to reject the null hypothesis, at the 5% significance level, based on the estimated p-value p).

prior configurations. To formally confirm that there are statistically significant differences between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP prior configurations, we employ a Wilcoxon two-sided rank sum test (Wilcoxon (1945)). The null hypothesis is that the values of the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by one GP prior configuration trained on a given dataset and those yielded by another configuration trained on the same dataset are samples from continuous distributions with equal medians. Provided that there are statistically significant differences between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP prior configurations (see Table 2.14) and the fact that the median L_2 norm of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the third GP prior configuration is the smallest, we conclude that there is strong evidence that our proposed modeling significantly decreases the overall motion estimation error when performing regression using an asynchronous training dataset.

In some applications, the measurements may be restricted to some region, e.g., on the surface of the soft-tissue organ. To test whether our proposed method also improves the accuracy even under such constraints, we performed regression on an additional training dataset $\mathcal{D}_{asyncSurf}$ composed of randomly non-uniformly spatiotemporally scattered uncertain surface motion measurements. In Figure 2.9, similar to Figure 2.8, one can observe the overall median errors yielded by the different GP prior configurations trained on $\mathcal{D}_{asyncSurf}$. In this regard, the boxplot notches in Figure 2.9 also offer evidence of a statistically significant difference (at the 5% significance level) between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP prior configurations. To formally confirm that there are statistically significant differences between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP prior configurations trained on $\mathcal{D}_{asyncSurf}$, we again employ a Wilcoxon two-sided rank sum test (Wilcoxon (1945)). The null hypothesis is again that the values of the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by one GP prior configuration trained on $\mathcal{D}_{asyncSurf}$ and those yielded by another configuration trained on the same dataset are samples from continuous distributions with equal medians. The results from the Wilcoxon two-sided rank sum test are the same

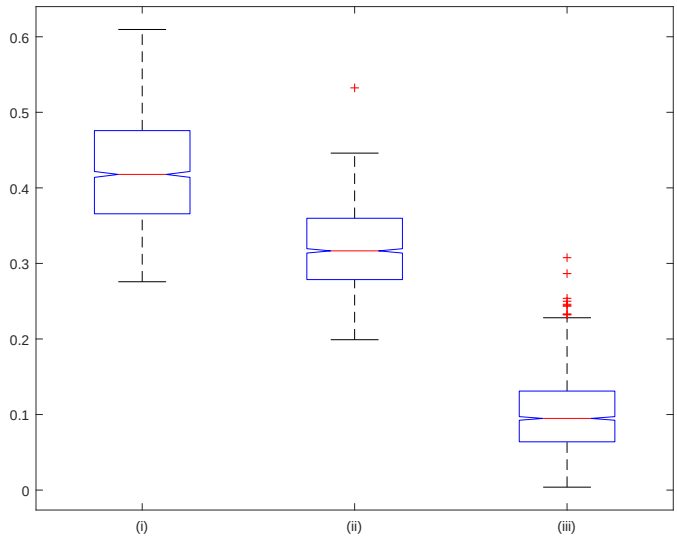

FIGURE 2.9: Boxplots of the L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by different GP prior configurations trained on $\mathcal{D}_{asyncSurf}$: (i) using models $\mathcal{M}^{(1,d)}_{\mathcal{D}_{asyncSurf}}$, (ii) using models $\mathcal{M}^{(2,d)}_{\mathcal{D}_{asyncSurf}}$, (iii) using models $\mathcal{M}^{(3,d)}_{\mathcal{D}_{asyncSurf}}$.

as those reported in Table 2.14 for the case $\mathcal{D} = \mathcal{D}_{async}$, i.e., it is formally confirmed that there are statistically significant differences between the median L_2 norms of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the three GP prior configurations. Provided this and the fact that the median L_2 norm of the rows of $(\bar{S}_* - \bar{S}_{*_{gt}})$ yielded by the third GP prior configuration is the smallest, we conclude that there is strong evidence that our proposed modeling significantly decreases the overall motion estimation error also when performing regression on randomly non-uniformly spatiotemporally scattered uncertain surface motion measurements.

2.7 Conclusion and Future Work

We presented a novel algorithm for uncertainty-aware interpolation of randomly non-uniformly spatiotemporally scattered motion signal samples. By employing GP regression, we were able to perform the interpolation in an optimal unbiased fashion. The use of a composite (as the sum of a constant and linear function) prior mean function enabled the learning of global and local drifts, present in the latent process mean function, from randomly non-uniformly spatiotemporally scattered samples. By means of a squared exponential covariance function with ADM, we were able to model the nearness or similarity between pairs of random motion function values over a spatiotemporal domain. Through estimating the full a posteriori motion field distribution, we were able to quantify the uncertainty in the resulting MAP estimate of the soft tissue motion at any location in the spatiotemporal domain.

The evaluation of our devised interpolation algorithm on simulated randomly non-uniformly spatiotemporally scattered uncertain motion signal samples revealed that our proposed GP model is able to, at the same time, learn more and yield statistically significantly better predictions than the state-of-the-art GP models that employ a zero-mean function and do not make use of ADM. We identified strong evidence supporting our contention that our proposed modeling significantly decreases the overall motion estimation error when performing regression both in the case of using synchronous and asynchronous motion signal samples as training data.

In future work, we plan to identify appropriate formalisms and, if needed, approximation approaches to optimize the conditioning (see Eq. 2.5) of the proposed GP model on a given training dataset, especially required when the training dataset is large. We would like to apply our proposed model for registering real pre-interventional/pre-operative navigation data on deformable soft-tissue organs during a real intervention or surgery. Furthermore, we intend to deploy our proposed GP model for modeling organ(s) deformation in the context of radiotherapy.

Acknowledgments

This work was supported by the Fraunhofer Internal Programs under Grant No. MAVO 823 287. It was also supported by the DFG Creative Unit "Intra-Operative Information: What Surgeons Need, When They Need It", and the NIH grants P41 EB015902, U24 CA180918, and P41 EB015898.

Chapter 3

Motion Uncertainty Visualization Based on Gaussian Process Regression

Bojan Kocev[1,3], Horst Karl Hahn[1,3], Ron Kikinis[2,3,5], William M. Wells[2], Lars Linsen[4]

[1]Jacobs University, Bremen, Germany
[2]Harvard Medical School and Brigham and Women's Hospital, Boston MA 02115, USA
[3]Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany
[4]Westfälische Wilhelms-Universität Münster, Germany
[5]University of Bremen, Bremen, Germany

Abstract

We address the problem of propagating the uncertainty from a Gaussian process (GP)-based soft tissue motion estimation result to a deformed/registered pre-operative/pre-interventional medical image and visualizing it in an uncertainty-aware fashion thereafter, which arises in the context of image-guided soft-tissue intervention and surgery navigation. Visualizing the uncertainty in the deformed image is of great interest, because it conveys the uncertainty in the registered pre-interventional/pre-operative navigation information on deformable soft-tissue organs governing the decisions of the surgeon or intervention specialist. To visualize the deformed medical image itself, we show its maximum a posteriori (MAP) estimate. To convey the overall uncertainty in the MAP deformed image estimate, we perform a spatially-varying blur (according to the

uncertain motion estimate) to an image containing a grid and combine it with the MAP deformed image itself and show 2D cross sections of the final image to the end user. To visualize where a point/region of interest (POI/ROI) (e.g., a tumor) is after deformation, we estimate marginal (i.e., per voxel) label (e.g., for being a tumor) probabilities (MLPs) based on an input probabilistic segmentation of the ROI in the non-deformed pre-interventional/pre-operative image and the input GP posterior distribution over deformation fields and show 2D cross sections of visually mapped MLPs or MLP isocontours. The proposed visualization of the amount and various forms/types of uncertainty in the registered pre-interventional/pre-operative navigation information should enable surgeons and intervention specialists to make better informed safety-critical decisions and therefore could potentially improve the overall outcome of the surgery or intervention.

3.1 Introduction

Motion describes how objects change their spatial position (or location) over time. In the case of deformable objects they may also change their shape over time. Capturing the state of an object, i.e., its position and shape, at some points in time allows for a reconstruction of the motion. Dynamic imaging is, for example, one widely used approach to measure the objects' state by a series of 2D or 3D images taken at different points in time. Motion reconstruction of an object of interest from such a time-varying field may, however, suffer from measurement inaccuracies such as imaging artifacts or insufficient sampling such as low spatial and/or temporal resolutions. These issues make the motion predictions uncertain. In this paper, we capture the uncertainties in motion predictions using a Gaussian process regression approach.

Conveying uncertainties in motion predictions during visualization can be of utmost importance. A prominent example is the visualization of position and shape of organs or tumors during a medical intervention. Possibly relevant uncertainty visualization goals for motion predictions address the following questions: Where did a single sample (of the initial state) move to (after a given period of time) and what are the probabilities for different locations? How did a given region of interest change its position and shape over time and what are the uncertainties? How did the entire captured field change over time and where in the field are uncertainties high/low? We present an approach where the listed motions and the respective uncertainties are estimated and visualized.

Our individual contributions comprise:

1. Computation and visualization of a probability map encoding the probability where a point of interest (sample) moved to.

2. Computation of marginal label probabilities (MLP) of regions of interest (ROIs) after deformation and their visualization using isocontours.

3. Computation of maximum a posteriori (MAP) estimates of deforming (time-varying) 2D/3D scalar fields and their visualization using grid-based approaches.

3.2 Related Work

Bonneau et al. (2014) pointed out to the work of Deitrick (2007) in which it is identified that visualizing uncertainty changes the decisions of the end users. In the medical domain for example, showing visual depiction of the uncertainty in the outcome of medical interventions to patients is suggested to possibly lead to better satisfaction in the shared decision process and a lower likelihood of regret about a decision (Politi et al. (2007); Bonneau et al. (2014)).

Bonneau et al. (2014) divide data uncertainty sources into three broad classes: uncertainty observed in sampled data, uncertainty measures generated by models or simulations, and uncertainty introduced by the data processing or visualization processes. Our work focuses on motion interpolation data uncertainty visualization, while the uncertainty in motion measurement data and their uncertainty-aware interpolation are addressed in a recently published work (Kocev et al. (2019)). In this regard, the output from the interpolation model (Kocev et al. (2019)) includes information about estimated error in the form of predicted distributions for values (Bonneau et al. (2014)). In principle, we propagate the uncertainty from the motion interpolation result to the uncertainty in the deformed pre-interventional/pre-operative image and visualize it adequately.

Next, we present related work about visual encoding/representation of uncertainty. The uncertainty can be visually encoded in various ways. A simple and well-known method is to map the uncertainty to color and overlay this color on top of the underlying visualization (Botchen et al. (2005)). Bonneau et al. (2014) point out that it is sensible to visualize uncertainty in a way consistent with our cognitive models of which perceptual elements contain variability or uncertainty. Example uncertainty visualization approaches that follow this principle represent the uncertainty as blur, flicker, reduced saturation, sketched outlines, or transparency (Bonneau et al. (2014)). Botchen et al. (2005) presented texture-based techniques to visualize uncertainty in time-dependent 2D flow fields. They show flow direction by streaklines and convey uncertainty by blurring these streaklines. As a survey over uncertainty visualization approaches that treat

uncertainty like an unknown or fuzzy quantity, Bonneau et al. (2014) pointed out the paper by Pang et al. (1997). We also follow this principle by applying a spatially-varying (Gaussian) blur to a 3D grid image of the same size as the input pre-interventional/pre-operative image, where the amount of blur visually encodes the uncertainty in the deformed image (see Section 3.6). Bonneau et al. (2014) stated that blurring or fuzzing a visualization accurately indicates the lowered confidence in the data, however they emphasized that this does not lead to more informed decision making. It is important to emphasize at this point that we are not blurring the actual data but an artificial grid, i.e., we do not obfuscate directly the underlying data. Chlan and Rheingans (2005) presented a distribution glyph that addresses the concept of how multivariate aggregated data is distributed. On the use of glyphs for visualizing uncertainty in vector fields, we point to the work of Wittenbrink et al. (1996). Allendes Osorio and Brodlie (2008) extended the concept of contouring to uncertainty by developing techniques for the visualization of uncertain contours. Grigoryan and Rheingans (2004) presented a method for visualization of surfaces with uncertainties using points as display primitives and applied their method in the context of visualizing tumor formations with uncertainty of tumor boundaries. Lodha et al. (1996) presented techniques for visualizing geometric uncertainty of surface interpolants. Torsney-Weir et al. (2011) use uncertainty-aware visualization methods to explore the response surface (computed using a Gaussian process model) of an automatic image segmentation algorithm over the space of parameters that impact the quality of the segmentation algorithm. Their end user is expected to navigate through the parameter space with the purpose to identify regions where the response value (goodness of segmentation) is high.

The method presented in Section 3.6 can be used to visualize uncertain scalar fields in an uncertainty-aware fashion. Therefore, next we present related work about visualization of uncertain scalar fields. Zehner et al. (2010) presented a visualization method for rendering scalar fields with a probability density function at each data point. They render isosurfaces and use a color scheme to encode which parts of the surface are more uncertain than others. In addition, they also augment the isosurfaces with additional geometry to visualize the envelope that indicates the volume in which the isosurface lies with a certain confidence. Djurcilov et al. (2001, 2002) presented methods for visualizing scalar volumetric data with uncertainty by incorporating the uncertainty information directly into the volume rendering equation and by post-processing (by inserting speckles/holes, adding noise, or adding texture) information of volume rendered images and then also enhanced these methods by showing the depth cues for the uncertainty. Lundström et al. (2007) proposed animation methods to convey uncertainty in direct volume rendering of medical datasets. They use a probabilistic transfer function model and animate the rendering by sampling the probability domain over time and in this way vary

the appearance for uncertain regions. I.e., certain image regions are static, while uncertain ones change over time. Sakhaee and Entezari (2016) devised a statistical framework for uncertainty quantification and uncertainty propagation through the main stages of the visualization pipeline. They presented a probabilistic transfer function classification model with which they incorporate probability density functions (modeling the accumulated uncertainties from interpolation and reduction stages) into the (direct) volume rendering integral. In principle, they estimate expected color and opacity values and convey uncertainties through fuzziness and transparency. The input data to the direct volume rendering algorithm is a field of random continuous variables. The probability density functions of the random variables are modeled in a non-parametric fashion. In our case, we could provide our random deformed 3D image as input to their direct volume rendering algorithm and perform the visualization of the deformed image using their algorithm. In this regard, we expect to get a blurred deformed image as a result where the amount of blur will convey the amount of uncertainty. In contrast to their work, we approximate the integral in discrete space and do not directly obfuscate the input image data but blur a grid image instead and combine it with the MAP estimate of the deformed image.

The method presented in Section 3.7 can be used for uncertainty-aware visualization of probabilistic segmentation results. Therefore, next we present related work about visualization of probabilistic segmentation results. Kniss et al. (2005) directly visualize combined "fuzzy" classification results from multiple segmentations. In other words, their visualization system provides access to the quantitative information computed during "fuzzy" segmentation. In this regard, they presented different methods for color mapping multi-class probabilities. Saad et al. (2010b,a) presented an interactive analysis and visualization tool for probabilistic medical image segmentation results. They devised different visualization tools, incorporating multidimensional transfer functions, for analyzing multivariate probabilistic field data with direct volume rendering. They guide the data exploration by shape and appearance knowledge learned from expert-segmented training images. In this regard, their multidimensional transfer functions make use of the output acquired from the deployed probabilistic segmentation algorithm and the probabilistic prior knowledge learned from previous expert segmentations. In the practical setting, their tool allows research clinicians to explore the uncertainty in the segmentation (w.r.t the population).

The method presented in Section 3.8 can be used for uncertainty-aware computation and visualization of the position of a moving particle. Therefore, we point to the work of Lodha et al. (2002) in which they compute and visualize the uncertainty associated with the positions of moving particles. They compute the probability distribution describing the position of a particle moving in 2D or 3D space, based on the individual

probability distributions that characterize the initial position, speed and direction of the particle. Regarding the visualization of the resulting positional probability distribution, they presented three visualization approaches (spherical glyphs or galaxy, transparency, and pseudo-color). Their work is, in nature, closely related to our work about uncertainty visualization of point motion (see Section 3.8), however in our case the resulting positional probability distribution is estimated based on a set of uncertain motion measurements that are interpolated using a GP regression method. Furthermore, the initial position of the point in our case is described using only MLPs, i.e., we do not define any covariance structure as when specifying probability distributions. With respect to the visualization, their transparency visualization approach bears some similarity to our visualization technique that maps the estimated POI MLPs to grayscale (see Section 3.8). Similar to their work, we also allow the user to scale the visualization by any factor, so that areas of low but still significant probability are visible.

The criteria for evaluation of uncertainty visualization methods have been also addressed in previous work. Bonneau et al. (2014) pointed out to the work of Harrower (2003), in which a number of evaluations of methods for representing uncertainty in map-based visualizations are surveyed, and in this regard stated that two principles may be derived from those evaluations. The first one is the superiority of displays that integrate value and certainty information over those that show each in a separate display. The second one is the preference for static displays over those that toggle between certainty and value. In our work, we integrate value and certainty information and use a static display with a parameter that controls how much of the blurred grid image is shown overall which is different from toggling between certainty and value. Furthermore, Sanyal et al. (2009) presented a user study that evaluates the perception of uncertainty amongst four of the most commonly used methods for visualizing uncertainty in 1D and 2D data.

3.3 Uncertainty-aware Asynchronous Scattered Motion Interpolation

In a recent work (Kocev et al. (2019)), an uncertainty-aware method for interpolating randomly non-uniformly spatiotemporally scattered uncertain motion measurements has been presented. They presented a novel Gaussian process (GP) regression model with a non-constant-mean prior and an anisotropic covariance function. In this regard, the use of GP regression makes it possible to estimate the uncertainty in the interpolation result. In principle, they fuse motion measurements with motion prior in an uncertainty-aware fashion. In terms of modeling the uncertainty in the measurements, a Gaussian likelihood function is employed. The motion prior is embedded in a GP prior and conditioned on

the motion measurements during the regression. The output of the regression is a (full) posterior distribution over motion signal functions, which is normally discretized at test locations (e.g., the voxels of a pre-operative/pre-interventional image that needs to be deformed). They view their GP regression model as a deformation model (Sotiras et al. (2013)), because they developed it for interpolation of soft (i.e., deformable) tissue motion measurements.

The key equation of a scalar-valued GP regression is:

$$
\begin{aligned}
f|\mathcal{D} \sim \mathcal{GP}(m_{\mathcal{D}}, k_{\mathcal{D}}) \\
m_{\mathcal{D}}(x) = m(x) + \Sigma(X,x)^T(\Sigma + \sigma_n^2\mathcal{I})^{-1}(y-\mu) \\
k_{\mathcal{D}}(x,x') = k(x,x') - \Sigma(X,x)^T(\Sigma + \sigma_n^2\mathcal{I})^{-1}\Sigma(X,x') \quad (3.1)
\end{aligned}
$$

where f is the function prior that is distributed as a GP with mean function $m(x) : \Omega \to \mathbb{R}$ (Ω is an index set with dimension D) and covariance function $k(x,x') : \Omega \times \Omega \to \mathbb{R}$, i.e., $f \sim \mathcal{GP}(m,k)$, $\mathcal{D} = \{(x_i, y_i)|i = 1, ..., n\}$ (x_i are the training data locations and y_i are the observations of the function values $f(x_i)$) is a training data set of n (usually noisy) observations, σ_n^2 is the noise variance, X is $D \times n$ matrix of the training inputs (i.e., the design matrix (Rasmussen and Williams (2006))), $\Sigma(X,x)$ is a vector of covariances between every training case and x (analogous for $\Sigma(X,x')$), μ is the vector of the function mean values at the training data locations, i.e., $\mu_i = m(x_i), i = 1, ..., n$, Σ is the covariance between the training data, and $\mathcal{I}$ is the identity matrix. In the recently published work (Kocev et al. (2019)), the authors assume that the motion signal output dimensions are independent and use a separate scalar-valued GP for each output dimension.

The above mentioned mean function, covariance function, and likelihood function all have (hyper-) parameters that control the distribution of the target values. These (hyper-) parameters are optimized by maximizing the probability of the model given the input training data.

We use the above described recently published method (Kocev et al. (2019)) to compute a GP posterior distribution over deformation fields based on which we register pre-interventional/pre-operative navigation information (e.g., an MRI) on a given deformable soft-tissue organ. We summarize and visually convey the uncertainty in the registered pre-interventional/pre-operative soft-tissue navigation information based on the information in the computed GP posterior distribution over deformation fields.

3.4 Expected Deformed Image

Let $\mathbb{A}$ be the input 3D image that needs to be deformed and $\mathbb{B}$ be the deformed image. $\mathbb{A}$ can contain any information (e.g., MRI intensities, marginal label probabilities, any 3D texture information, etc.) about the organ that moves. The expected deformed image, i.e., $\bar{\mathbb{B}} = \mathbb{E}[\mathbb{B}]$, is then estimated as follows:

$$\bar{\mathbb{B}}_{i,j,k} \approx \sum_{u,v,w} \mathbb{P}^{i,j,k}_{u+s,v+s,w+s} \mathbb{A}_{i-u,j-v,k-w} \tag{3.2}$$

$$= \sum_{u,v,w} \mathbb{P}^{i,j,k}_{u,v,w} \mathbb{A}_{i-u+s,j-v+s,k-w+s} \tag{3.3}$$

$$= \sum_{u,v,w} \mathbb{P}^{i,j,k}_{u,v,w} \mathbb{A}_{i+s-u,j+s-v,k+s-w} \ , \tag{3.4}$$

where $(2s+1) \times (2s+1) \times (2s+1)$ is the size of the kernel $\mathbb{P}^{i,j,k}$ that defines the local neighborhood of the voxel (i,j,k), $\mathbb{P}^{i,j,k}_{u,v,w}$ holds the probability that voxel $(i+s-u, j+s-v, k+s-w)$ moves to voxel (i,j,k). The transition from Eq. (3.2) to Eq. (3.3) is made by applying the substitutions $u \leftarrow u-s$, $v \leftarrow v-s$, and $w \leftarrow w-s$, which is possible because u, v, and w in the sum on the right-hand side of Eq. (3.2) range over all legal subscripts for $\mathbb{P}^{i,j,k}_{u+s,v+s,w+s}$ and $\mathbb{A}_{i-u,j-v,k-w}$. Note that the probability $\mathbb{P}^{i,j,k}_{u,v,w}$ is computed based on the posterior displacement mean and covariance at the voxel $(i+s-u, j+s-v, k+s-w)$ that are obtained as output from the above-described recently published GP-based method (Kocev et al. (2019)).

3.5 Maximum a Posteriori Deformed Image Estimate

The estimated mean of the GP posterior distribution over deformation fields is also its mode, which is also called the *maximum a posteriori* (MAP) estimate (Rasmussen and Williams (2006)) of the deformation field. In our case, the estimated GP posterior mean is discretized at the voxels of the input image $\mathbb{A}$ and transformed into voxel space. The discretized transformed mean is stored in a 3D array $\bar{\mathbb{U}}$, of the same size as $\mathbb{A}$, that holds the estimate of the posterior mean displacement of every voxel in voxel space. Then, the MAP estimate of the deformed image is computed as follows:

$$\hat{\mathbb{B}}_{(i,j,k)+\bar{\mathbb{U}}(i,j,k)} = \mathbb{A}_{i,j,k} \ . \tag{3.5}$$

As we use forward mapping, the MAP estimate $\hat{\mathbb{B}}$ may contain holes in case no voxel moves to a given location. We consider this as a feature of our visualization as it allows us to observe the phenomenon of no voxel moving to a given location.

3.6 MAP Deformed Image Global Uncertainty Visualization

In this section, we present a method for visual encoding of the global uncertainty present in the MAP deformed image estimate $\hat{\mathbb{B}}$. We visually encode the uncertainty using the algorithm presented in Section 3.4. In this regard, we use an input 3D image $\mathbb{A}$, of the same size as $\hat{\mathbb{B}}$, that contains a 3D grid (see Figure 3.1 for an example of such a 3D grid 2D cross section) and compute $\bar{\mathbb{B}}$ using Eq. (3.4). In principle, the algorithm presented

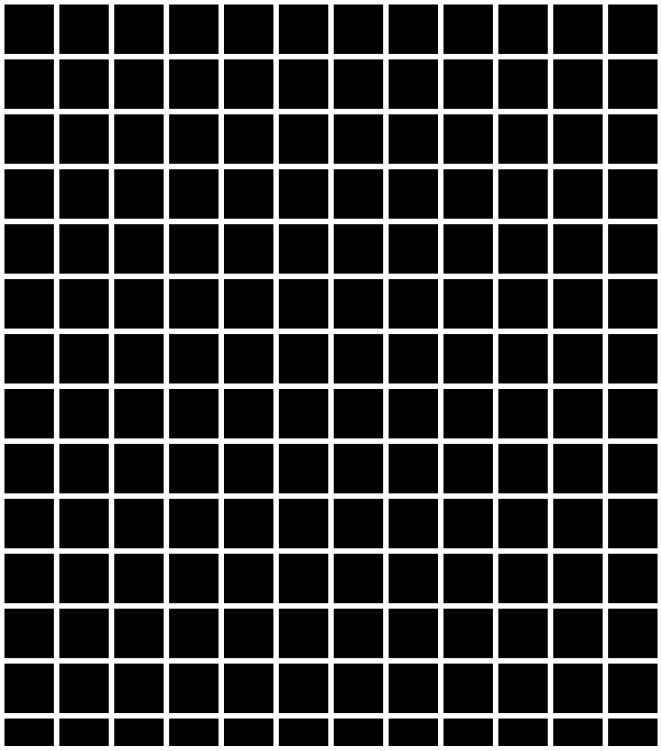

FIGURE 3.1: Example 3D grid 2D cross section.

in Section 3.4 applies a spatially-varying (Gaussian) blur (Chakrabarti et al. (2010)) to the input image $\mathbb{A}$ to compute the output image $\bar{\mathbb{B}}$. In this regard, Eq. (3.4) is analogous to a 3D array convolution with $\mathbb{P}_{u,v,w}^{i,j,k}$ being the spatial location-dependent (Gaussian) convolution kernel. The amount of blur in the output image $\bar{\mathbb{B}}$ visually encodes the uncertainty in the MAP deformed image estimate $\hat{\mathbb{B}}$. We then propose two methods for combining $\bar{\mathbb{B}}$ and $\hat{\mathbb{B}}$ into a final uncertainty-aware visualization output image $\mathbb{C}$. The first method uses alpha blending and thresholding:

$$\mathbb{C}_{i,j,k} = \alpha \hat{\mathbb{B}}_{i,j,k} + (1-\alpha)\mathcal{T}(\bar{\mathbb{B}}_{i,j,k}, \tau) \ , \tag{3.6}$$

where α is the opacity which ranges from 0.0 (fully transparent) to 1.0 (fully opaque), and $\mathcal{T}(z,t)$ is defined as follows:

$$\mathcal{T}(z,t) = \begin{cases} z & \text{if } z \geq t \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

The second method uses masking and thresholding:

$$\mathbb{C}_{i,j.k} = \mathcal{M}(\hat{\mathbb{B}}_{i,j,k}, \mathcal{T}(\bar{\mathbb{B}}_{i,j,k}, \tau)) \ , \tag{3.8}$$

where $\mathcal{M}(z,z')$ is defined as follows:

$$\mathcal{M}(z,z') = \begin{cases} z' & \text{if } z' > 0 \\ z & \text{otherwise.} \end{cases} \tag{3.9}$$

Note that it is intended to show 2D cross sections of the final uncertainty-aware visualization output image $\mathbb{C}$ to the end user.

3.7 Deformed Region of Interest Uncertainty Visualization

In this section, we present a method for uncertainty-aware visualization of where a region of interest (ROI) (e.g., a tumor, a whole organ, etc.) is after deformation. In general, the input to this method are a probabilistic 3D binary segmentation of the ROI in the pre-interventional/pre-operative 3D image (e.g., an MRI) and the estimated GP posterior distribution over deformation fields. The output from a probabilistic 3D binary segmentation algorithm is normally a 3D array, here denoted with Θ, which contains the segmented ROI marginal label probabilities (MLPs) (Ciresan et al. (2012); Ronneberger et al. (2015)) $\theta(x)$ for each voxel with position x, i.e., $\theta(x) = p(\zeta(x) = 1)$ where $\zeta(x) : \Omega \rightarrow \{0,1\}$ is a function which assigns binary labels to all voxels. In case of a deterministic input binary segmentation, we treat it as a special case of a probabilistic segmentation where $\theta(x) = 1.0$ for each voxel with position x that has been labeled one and $\theta(x) = 0.0$ otherwise. By modeling the input segmentation in a probabilistic fashion, we already allow for uncertainty in the answer to the question where the ROI is in the pre-interventional/pre-operative 3D image. In principle, we could visualize in an uncertainty-aware fashion where the ROI is in the pre-interventional/pre-operative 3D image by mapping the input probabilities $\theta(x)$ to grayscale values and showing 2D cross sections of the resulting grayscale 3D image. The more bright a pixel is the higher the probability that it belongs to the ROI and vice versa. To be able to perform such an uncertainty-aware visualization of where a ROI is after deformation, we would need to

estimate $\theta(x)$ for every voxel with position x in the deformed image. It turns out that we can estimate the expected ROI MLP for each voxel in the deformed image using the method presented in Section 3.4. In this regard, we set $\mathbb{A} = \Theta$ as the input 3D image of the method presented in Section 3.4 and estimate the expected deformed image $\bar{\mathbb{B}}$ using Eq. (3.4) which then holds the approximate expected ROI MLPs for each voxel in the deformed image. The uncertainty-aware visualization of where the ROI is after deformation is then obtained by converting $\bar{\mathbb{B}}$ to a grayscale image. Note again that it is intended to show 2D cross sections of the resulting grayscale image to the end user. As an alternative to this visualization, we allow for selecting a 2D cross section of the MAP deformed image estimate $\hat{\mathbb{B}}$ and overlay it with ROI MLP isocontours extracted from the corresponding slice in $\bar{\mathbb{B}}$.

3.8 Uncertainty Visualization of Point Motion

In this section, we present a method for uncertainty-aware visualization of where a point of interest (POI) (e.g., a tumor centroid) is after deformation. The input to this method is a POI in the pre-interventional/pre-operative 3D image (e.g., an MRI) and the estimated GP posterior distribution over deformation fields. It turns out that we can also define the input POI in a probabilistic fashion in the same way as the input ROI is defined in Section 3.7. This means we can perform the uncertainty-aware visualization of where the POI is after deformation using the method presented in Section 3.7. In this regard, we only need to set Θ accordingly. In case the input POI is specified in a deterministic fashion, we set the entry of Θ that corresponds to the POI equal to 1.0 and all other entries equal to 0.0. The resulting approximate expected POI MLPs for each voxel in the deformed image are then mapped to grayscale and 2D cross sections of the resulting grayscale image are shown to the end user as explained in Section 3.7. As the expected use case scenario for this method is one in which the end user is allowed to select a POI on a 2D cross section of the pre-interventional/pre-operative 3D image (e.g., an MRI), we propose to show the selected 2D cross section of the non-deformed pre-interventional/pre-operative 3D image (e.g., an MRI) overlayed with the corresponding 2D cross section of the selected POI MLPs (first mapped to grayscale and then to a distinctive color) and the corresponding 2D cross section of the approximate expected moved POI MLPs (first mapped to grayscale and then to a distinctive color) as a final visualization to the end user.

3.9 Results and Discussion

In this section, we present and discuss the results that have been generated using the above-proposed visualization methods.

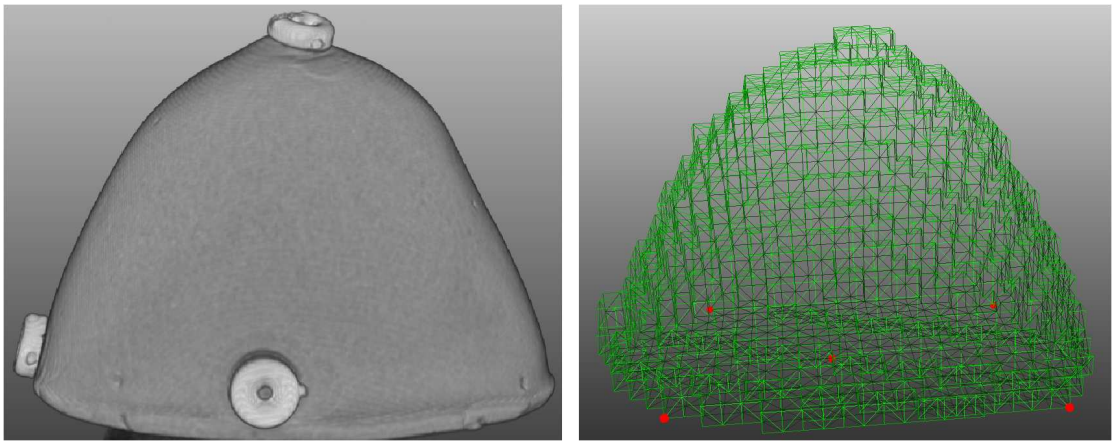

FIGURE 3.2: Left: Breast phantom with 4 markers (one on the back side). Right: FE model composed of tetrahedral elements which are extracted from the MRI scan data of the breast phantom. The points in red are fixed, i.e., the FE nonlinear motion prediction model considers these vertices as not moving. (Image courtesy of Kocev et al. (2014).)

The input to our visualization algorithm is a GP posterior distribution over motion signal functions (Kocev et al. (2019)), which we discretize it at the voxels of a pre-operative/pre-interventional MRI (see Figure 3.2 left) that needs to be deformed and the uncertainty therein visualized. The authors of the recently published work (Kocev et al. (2019)) compute the GP posterior distribution over motion signal functions by conditioning the GP prior on simulated randomly non-uniformly spatiotemporally scattered uncertain soft-tissue motion measurements during the regression. The measurements are simulated using a finite element (FE) model (Georgii and Westermann (2008, 2005)), which is created of the CIRS triple modality breast biopsy training phantom (see Figure 3.2). The FE model is used to simulate motion measurements that capture physically plausible non-linear motion that a soft-tissue organ is likely to undergo during a biopsy intervention.

Given the pre-operative/pre-interventional MRI (see Figure 3.2 left), we first compute the MAP estimate of the deformed MRI image using Eq. (3.5). In this regard, Figure 3.3 shows a 2D cross section of the MAP estimate of the deformed breast phantom 3D MRI based on a moderately uncertain deformation field estimate. On the other hand, Figure 3.4 shows the same 2D cross section of the MAP estimate of the deformed breast phantom 3D MRI based on a very uncertain (along the second output dimension) deformation field estimate. A possible clue about the amount of uncertainty in Figures 3.3 and 3.4 could be the number of holes within the breast. In this regard, it is extremely

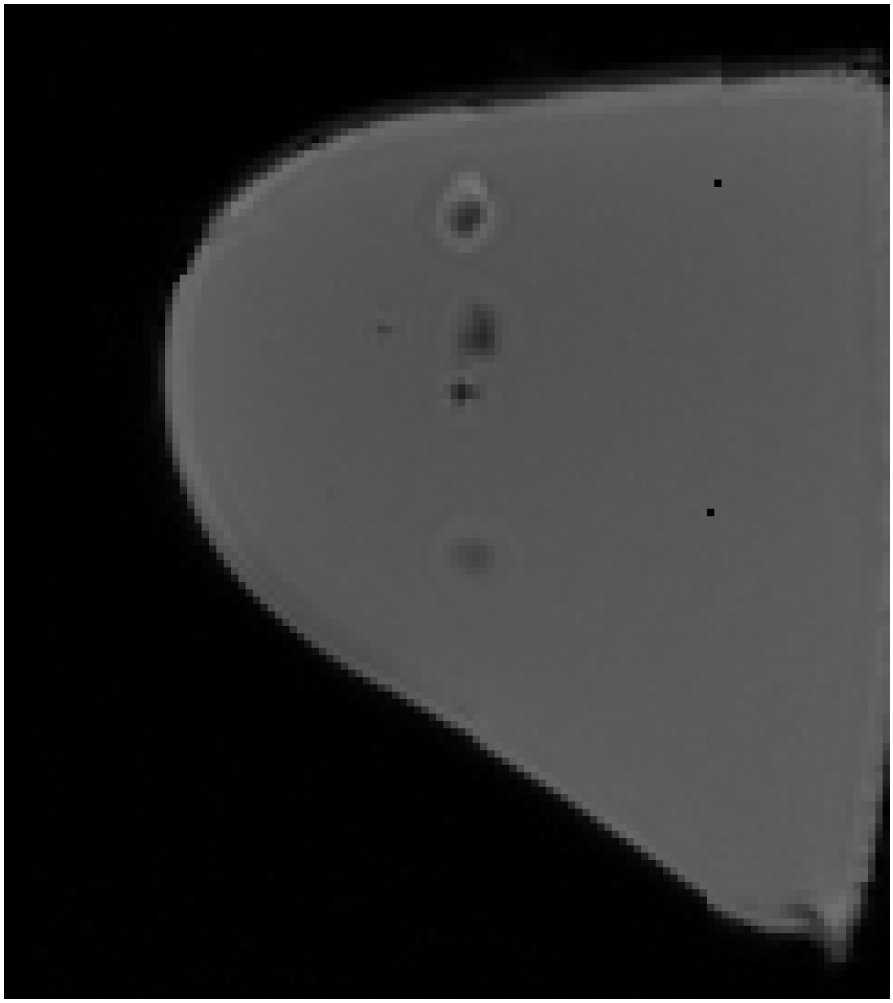

FIGURE 3.3: MAP deformed breast phantom 3D MRI cross section estimate based on a moderately uncertain deformation field estimate.

unlikely that such holes would appear inside the breast in reality. Therefore, as Figure 3.4 has more holes, one may conclude that it is very likely to be more uncertain than Figure 3.3. However, one cannot safely conclude that an image is certain if it does not contain such holes. Hence, in general we cannot judge or convey the amount of uncertainty in the images based on the number of such holes.

To visualize the amount of uncertainty in Figures 3.3 and 3.4, we deploy the method presented in section 3.6. The cross section of the blurred (based on the same GP posterior over deformation fields that was used to compute Figure 3.3) 3D grid image that visually encodes the uncertainty in Figure 3.3 is shown in Figure 3.5. Furthermore, the cross section of the blurred (based on the same GP posterior over deformation fields that was used to compute Figure 3.4) 3D grid image that visually encodes the uncertainty in Figure 3.4 is shown in Figure 3.6. As the amount of blur in Figure 3.6 is higher than the amount of blur in Figure 3.5, we can conclude that Figure 3.4 is more uncertain than Figure 3.3. Note that in Figure 3.6 the amount of blur along the horizontal axis is higher

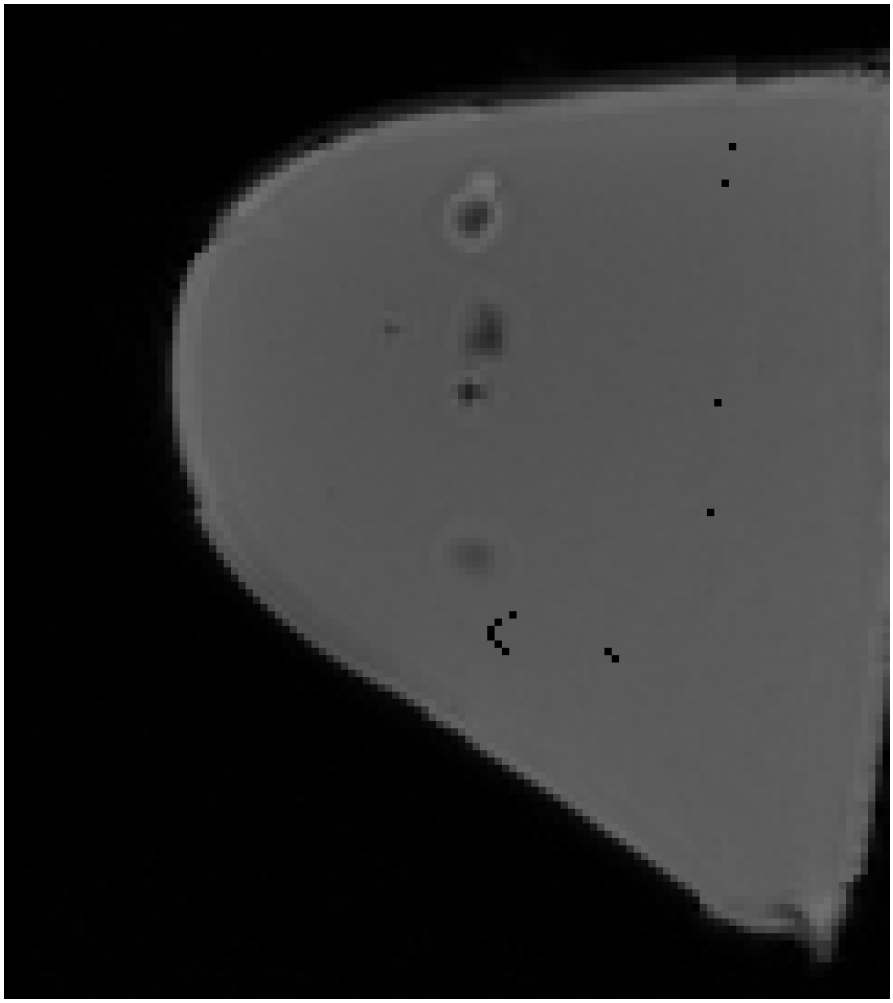

FIGURE 3.4: MAP deformed breast phantom 3D MRI cross section estimate based on a very uncertain (along the second output dimension) deformation field estimate.

than the amount of blur along the vertical axis in pixel space. This is in agreement with the fact that the deformation field estimate that has been used to compute Figures 3.4 and 3.6 has higher uncertainty along the second output dimension, because the second output dimension axis in world space corresponds approximately to the horizontal axis in pixel space.

The rest of the visualization results focus on the case with higher uncertainty. In this regard, we first combine Figures 3.4 and 3.6 into a final uncertainty-aware visualization output using Eq. (3.6) (see Figures 3.7, 3.8, and 3.9). We also combine Figures 3.4 and 3.6 into a final uncertainty-aware visualization output using Eq. (3.8) (see Figures 3.10, 3.11, and 3.12). In Figures 3.7, 3.8, 3.9, 3.10, 3.11, and 3.12, one can observe that the parameter τ controls how much of the blurred grid image is shown overall.

Despite the visualization of the global uncertainty present in Figure 3.4, we also visualize in an uncertainty-aware fashion where a ROI is after deformation using the

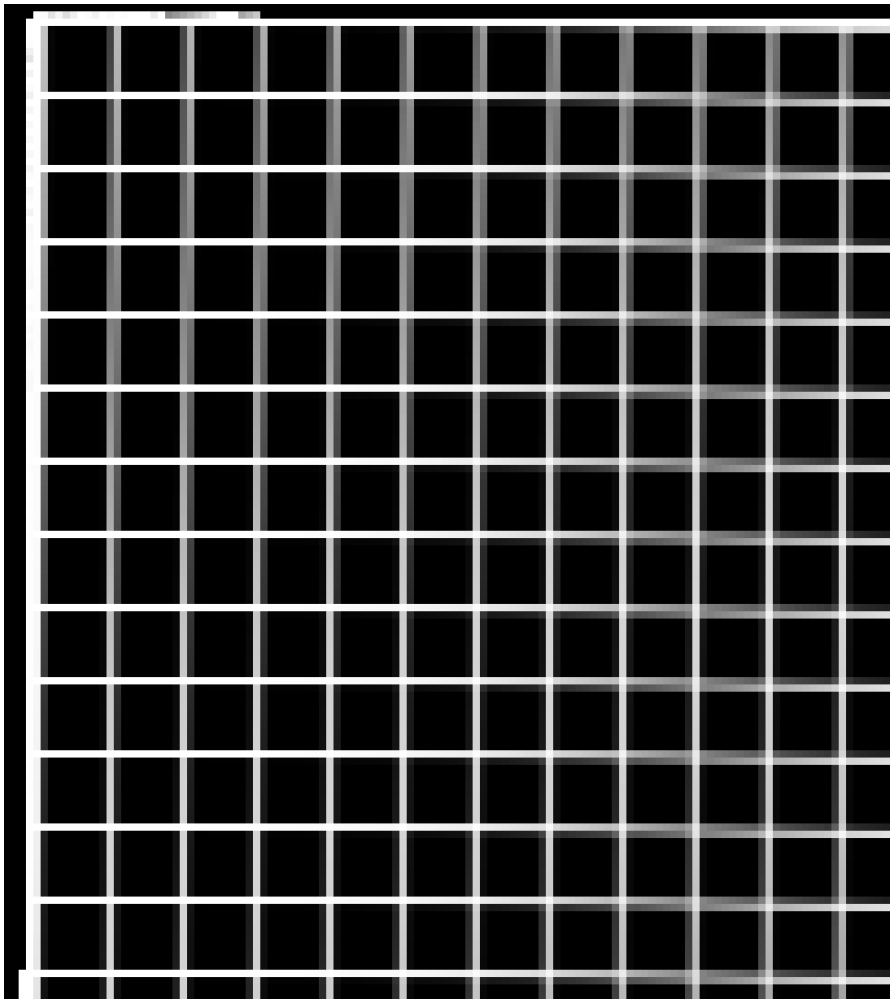

FIGURE 3.5: Blurred 3D grid image cross section encoding the uncertainty in Figure 3.3.

methods presented in Section 3.7. An example of segmented (in the non-deformed pre-operative/pre-interventional MRI) ROI MLPs 2D cross section mapped to grayscale is shown in Figure 3.13. Furthermore, the corresponding 2D cross section of the non-deformed pre-operative/pre-interventional MRI overlayed with the example ROI MLPs 2D cross section isocontours is shown in Figure 3.14. As can be observed in Figures 3.13 and 3.14, there is no uncertainty in the segmentation of the example ROI in the non-deformed pre-operative/pre-interventional MRI. This is the case because in this example we have a very accurate deterministic segmentation that we treat as a special case of a probabilistic segmentation as explained in Section 3.7. In simple words, this means that there is no uncertainty in the answer to question where the example ROI is in the non-deformed pre-operative/pre-interventional MRI. However, as the input deformation field estimate that was used to compute Figure 3.4 is uncertain, we expect that there will be uncertainty in the answer to the question where the example ROI is after deformation. In this regard, we can assume that the uncertainty that will be conveyed

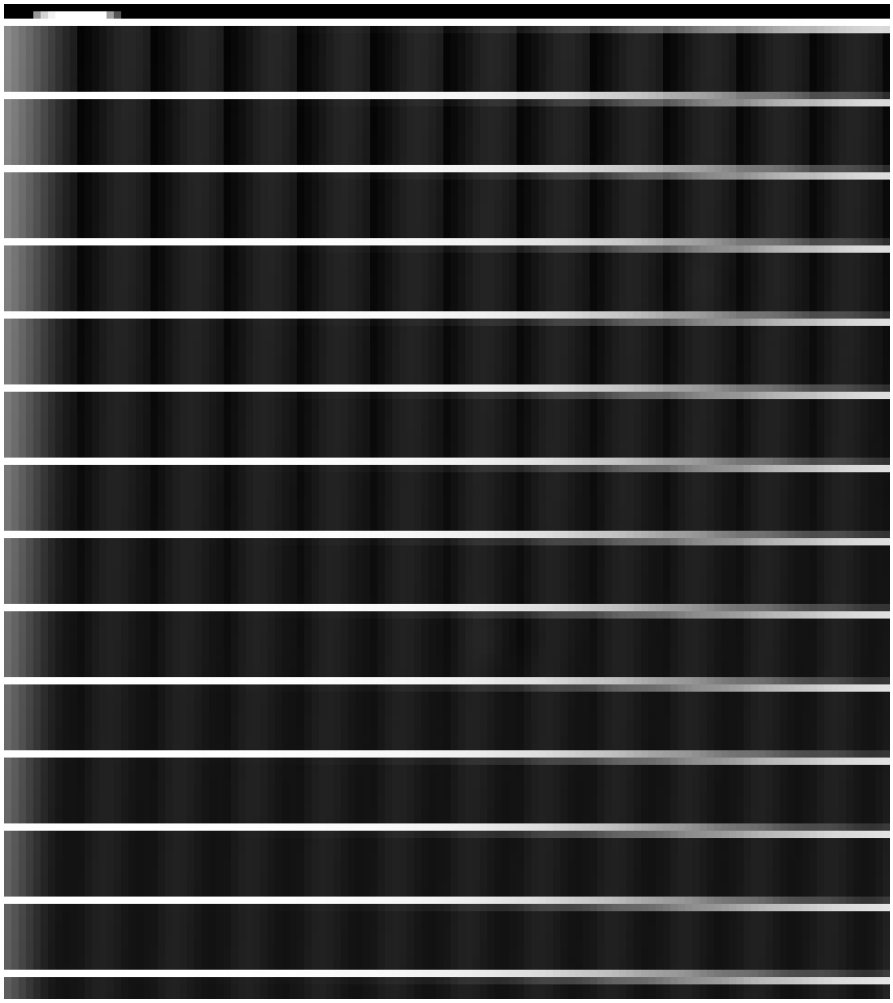

FIGURE 3.6: Blurred 3D grid image cross section encoding the uncertainty in Figure 3.4.

in the visualization, of where the example ROI is after deformation, will originate from the uncertainty in the input deformation field estimate. We proceed with setting Θ equal to the non-deformed example ROI MLPs and based on the input GP posterior distribution over deformation fields that was used to compute Figure 3.4, we perform an uncertainty-aware visualization of where the example ROI is after deformation as explained in Section 3.7. In this regard, the 2D cross section (corresponding to Figure 3.4) of the approximate expected deformed example ROI MLPs mapped to grayscale is shown in Figure 3.15. Furthermore, the isocontours extracted from the approximate expected deformed example ROI MLPs 2D cross section that is visualized in Figure 3.15 overlayed over the MAP deformed breast phantom 3D MRI cross section estimate from Figure 3.4 is shown in Figure 3.16. As can be observed in Figures 3.15 and 3.16, the uncertainty-aware visualization conveys that there is uncertainty in the answer to the question where the example ROI is after deformation. Furthermore, it is evident that there is higher uncertainty along the horizontal axis than along the vertical axis in pixel

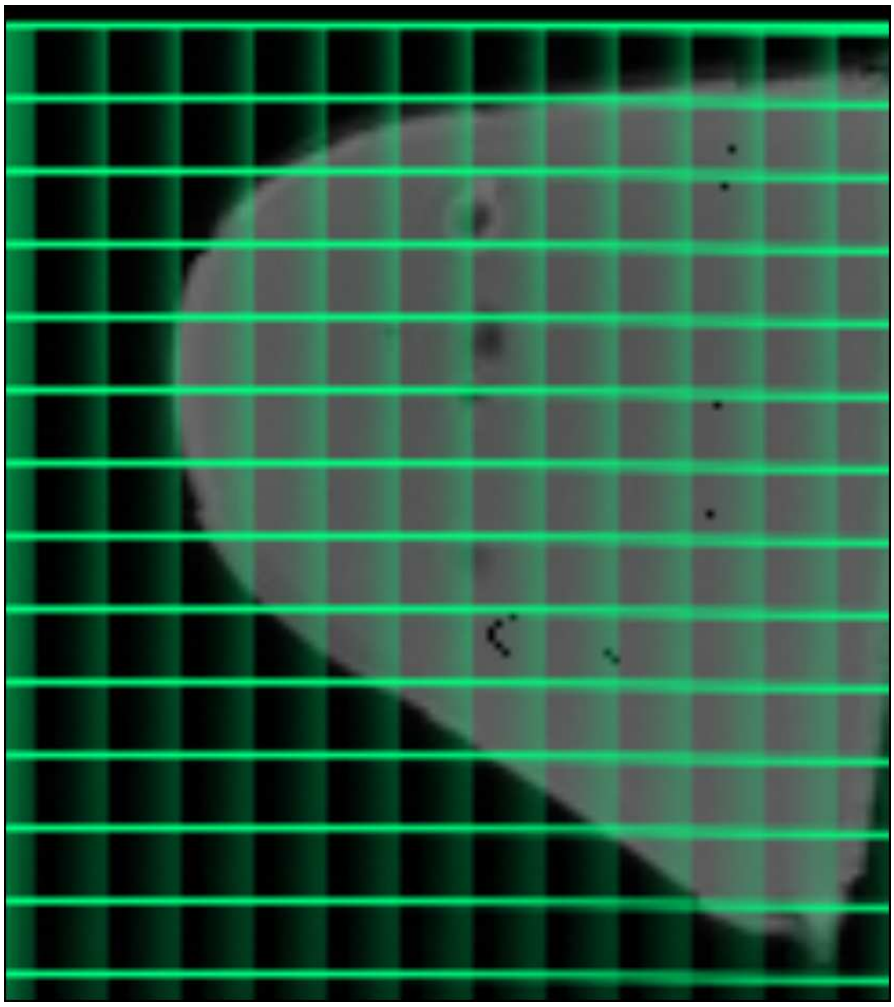

FIGURE 3.7: Generated by combining Figures 3.4 and 3.6 using alpha blending and thresholding ($\tau = 1$) according to Eq. (3.6).

space. This is in agreement with the visualization in Figure 3.6. We believe that the type of uncertainty-aware visualization presented in Figure 3.16 would be particularly useful for biopsy intervention specialists when performing ROI (e.g., a tumor) biopsy navigation in deforming pre-interventional images.

Despite the uncertainty-aware visualization of where a ROI is after deformation, we also visualize in an uncertainty-aware fashion where a POI is after deformation using the method presented in Section 3.8. An example of selected (in the non-deformed pre-operative/pre-interventional MRI) POI MLPs 2D cross section (first mapped to grayscale and then to green color) overlayed over the corresponding 2D cross section of the non-deformed pre-operative/pre-interventional MRI is shown in Figure 3.17. As can be observed in Figure 3.17, there is no uncertainty in the position of the example POI in the non-deformed pre-operative/pre-interventional MRI. This is the case because in this example we have a very accurate deterministic specification of the position of the example POI that we treat as a special case of a probabilistic specification as explained

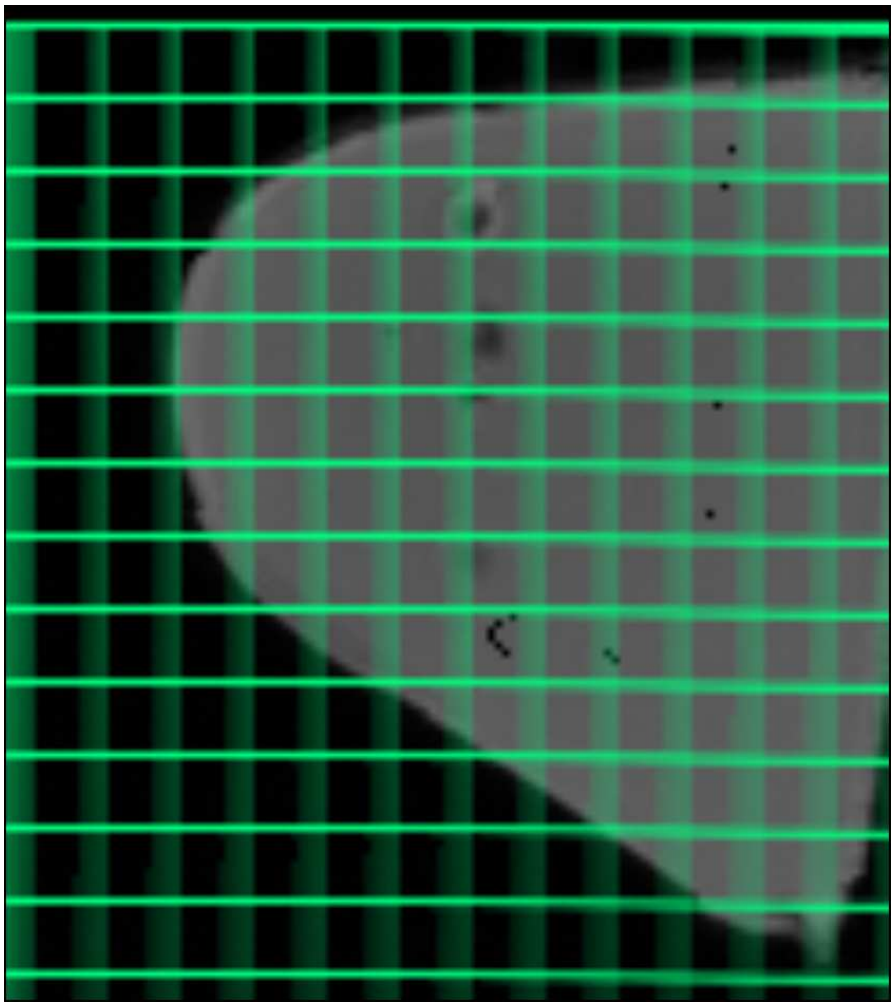

FIGURE 3.8: Generated by combining Figures 3.4 and 3.6 using alpha blending and thresholding ($\tau = 25$) according to Eq. (3.6).

in Section 3.8. However, the corresponding 2D cross section of the approximate expected moved (according to the input GP posterior distribution over deformation fields that was used to compute Figure 3.4) example POI MLPs (first mapped to grayscale and then to magenta color) overlayed over the non-deformed pre-operative/pre-interventional MRI 2D cross section in Figure 3.17 clearly shows that there is uncertainty in the position of the example POI after deformation. Furthermore, it is evident that there is higher uncertainty along the horizontal axis than along the vertical axis in pixel space. This is in agreement with the visualization in Figures 3.6, 3.15, and 3.16.

3.10 Conclusion and Future Work

We presented a novel method for soft tissue motion uncertainty propagation and uncertainty-aware visualization of deformed/registered pre-operative/pre-interventional medical images. By performing a spatially-varying blur to an image containing a grid, we were able

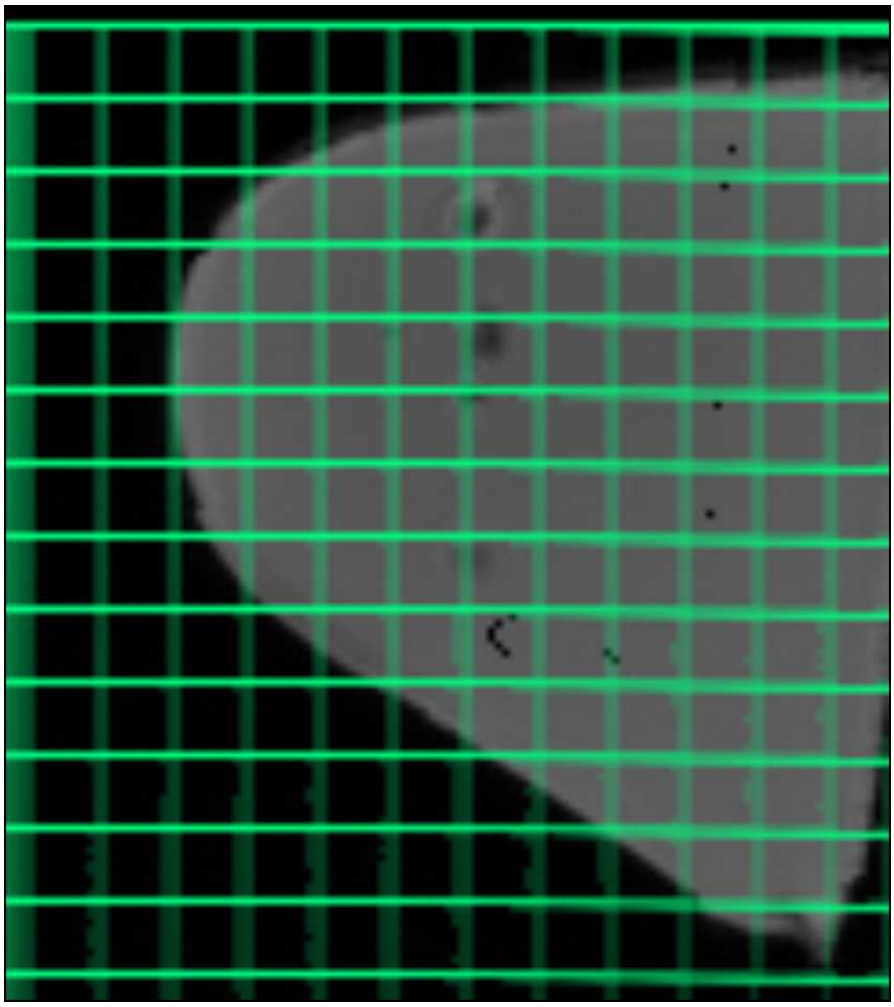

FIGURE 3.9: Generated by combining Figures 3.4 and 3.6 using alpha blending and thresholding ($\tau = 50$) according to Eq. (3.6).

to convey the overall uncertainty in the displayed MAP deformed image estimate in way consistent with our cognitive models of which perceptual elements contain variability or uncertainty. By estimating POI/ROI (e.g., tumor) MLPs and showing visually mapped 2D cross sections thereof or alternatively visualizing MLP isocontours, we were able to visualize where a POI/ROI (e.g., a tumor) is after deformation.

In future work, we would like to apply our proposed method for uncertainty-aware visualization of registered pre-operative/pre-interventional patient images during a real intervention or surgery.

Acknowledgments

This work was supported in part by a grant from DFG ()...

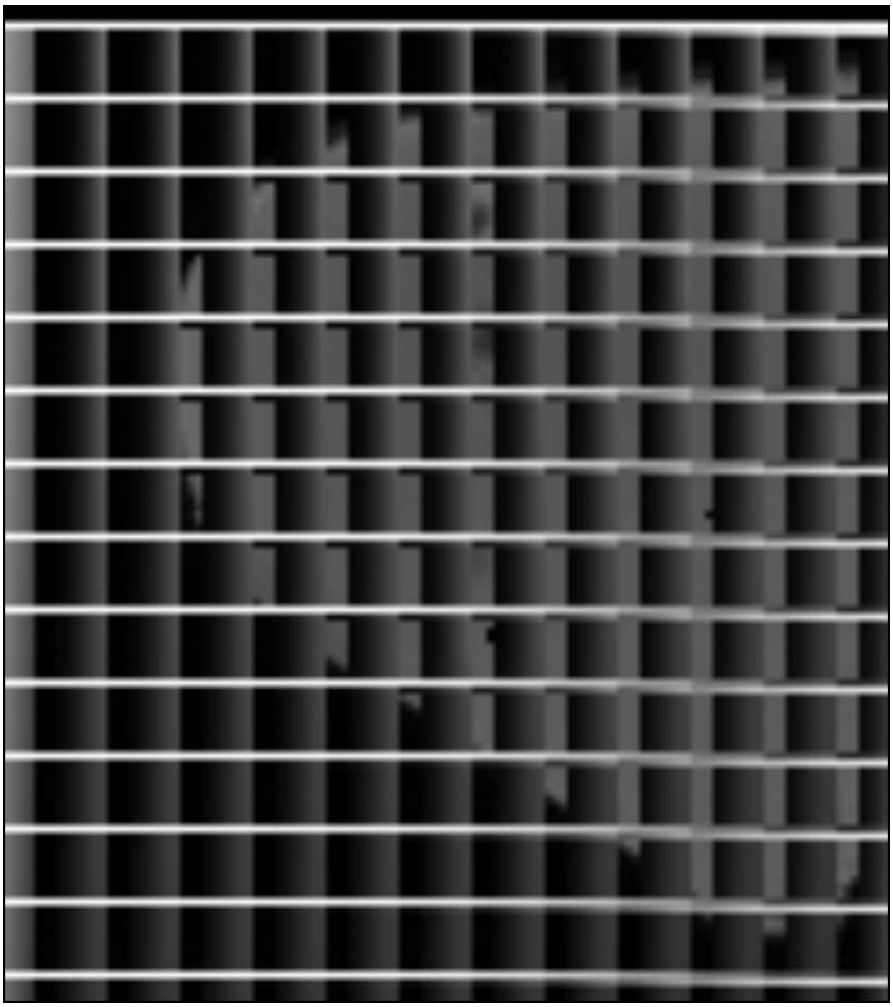

FIGURE 3.10: Generated by combining Figures 3.4 and 3.6 using masking and thresholding ($\tau = 1$) according to Eq. (3.8).

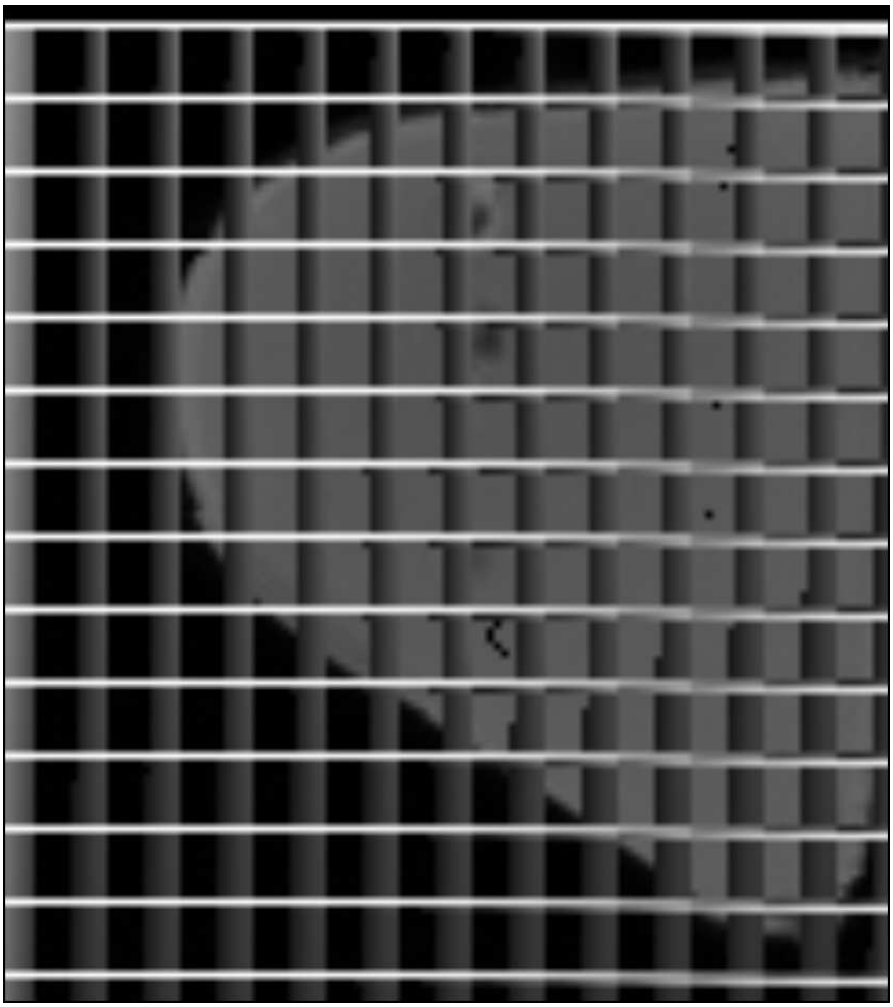

FIGURE 3.11: Generated by combining Figures 3.4 and 3.6 using masking and thresholding ($\tau = 25$) according to Eq. (3.8).

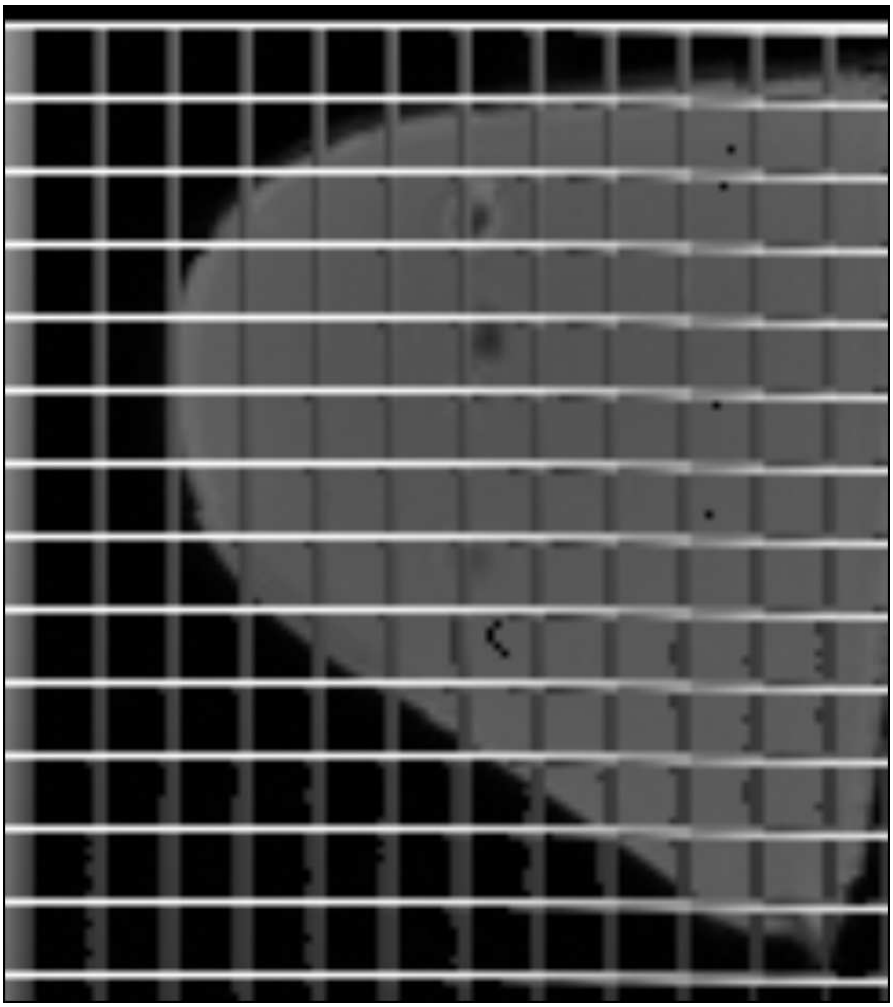

FIGURE 3.12: Generated by combining Figures 3.4 and 3.6 using masking and thresholding ($\tau = 50$) according to Eq. (3.8).

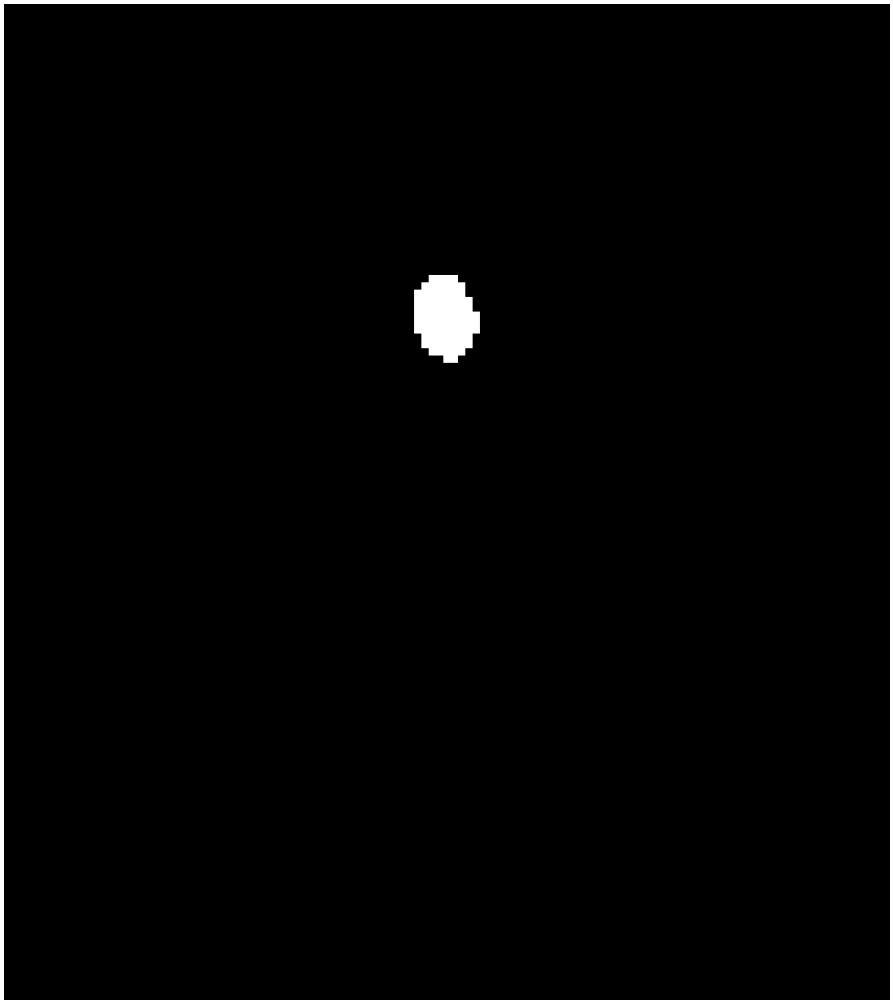

FIGURE 3.13: A segmented (in the non-deformed pre-operative/pre-interventional MRI image) ROI MLPs 2D cross section mapped to grayscale.

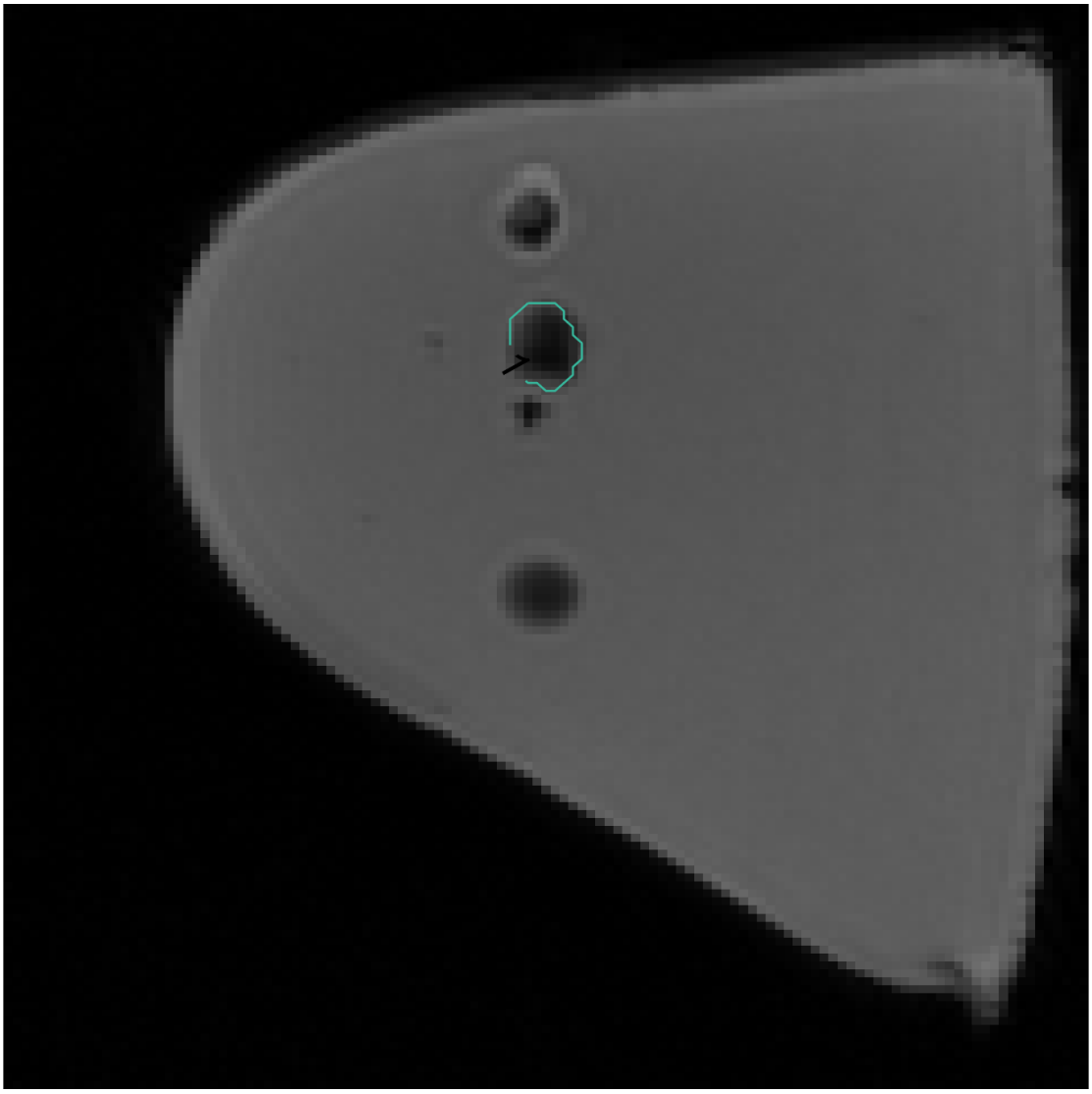

FIGURE 3.14: Isocontour at level/height 1.0 extracted from the segmented ROI MLPs 2D cross section that is visualized in Figure 3.13 overlayed over the corresponding 2D cross section of the non-deformed pre-operative/pre-interventional MRI.

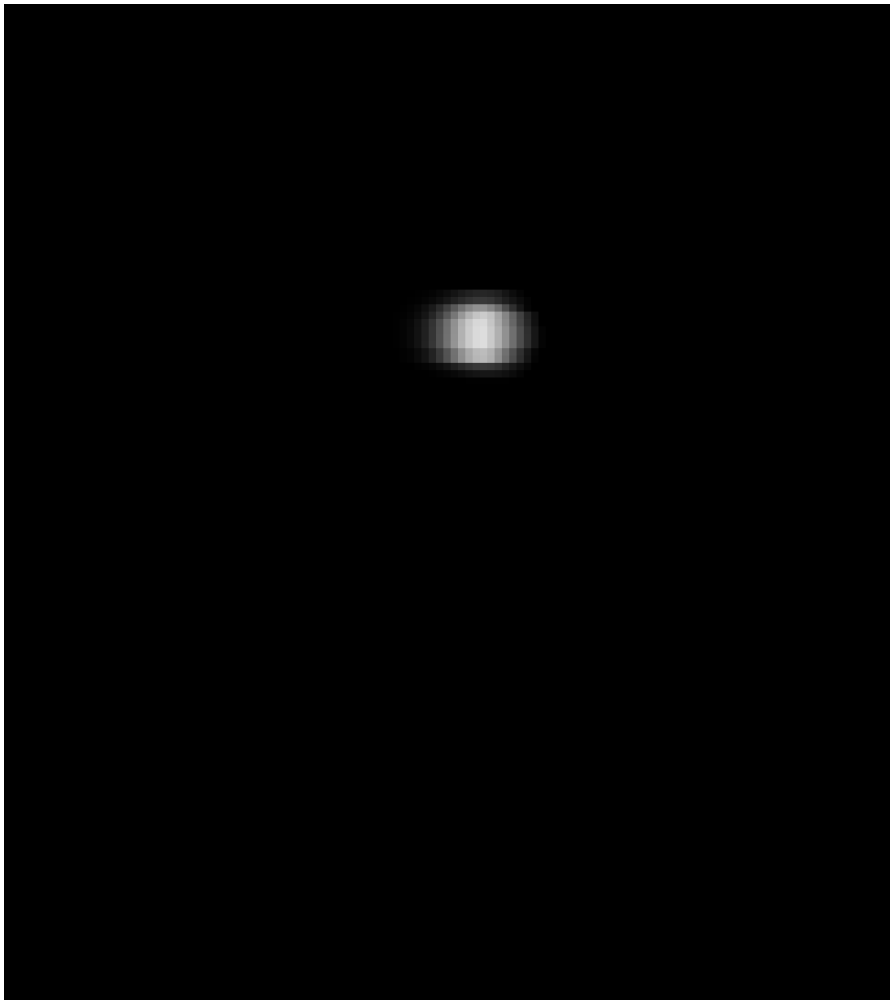

FIGURE 3.15: Approximate expected deformed example ROI MLPs 2D cross section (corresponding to Figure 3.4) mapped to grayscale.

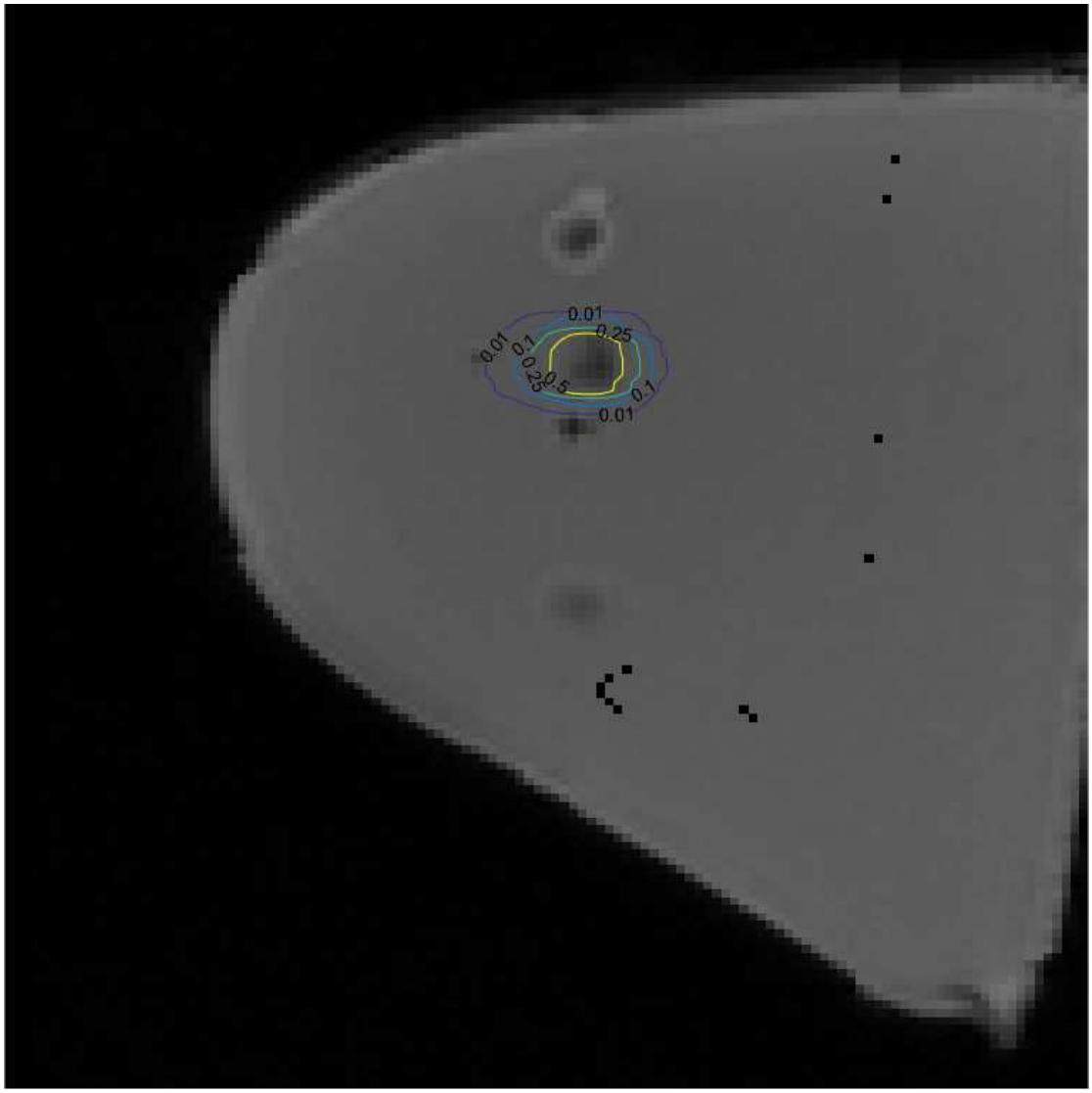

FIGURE 3.16: Isocontours at levels/heights $0.01, 0.1, 0.25$, and 0.5 extracted from the approximate expected deformed example ROI MLPs 2D cross section that is visualized in Figure 3.15 overlayed over the MAP deformed MRI cross section estimate from Figure 3.4.

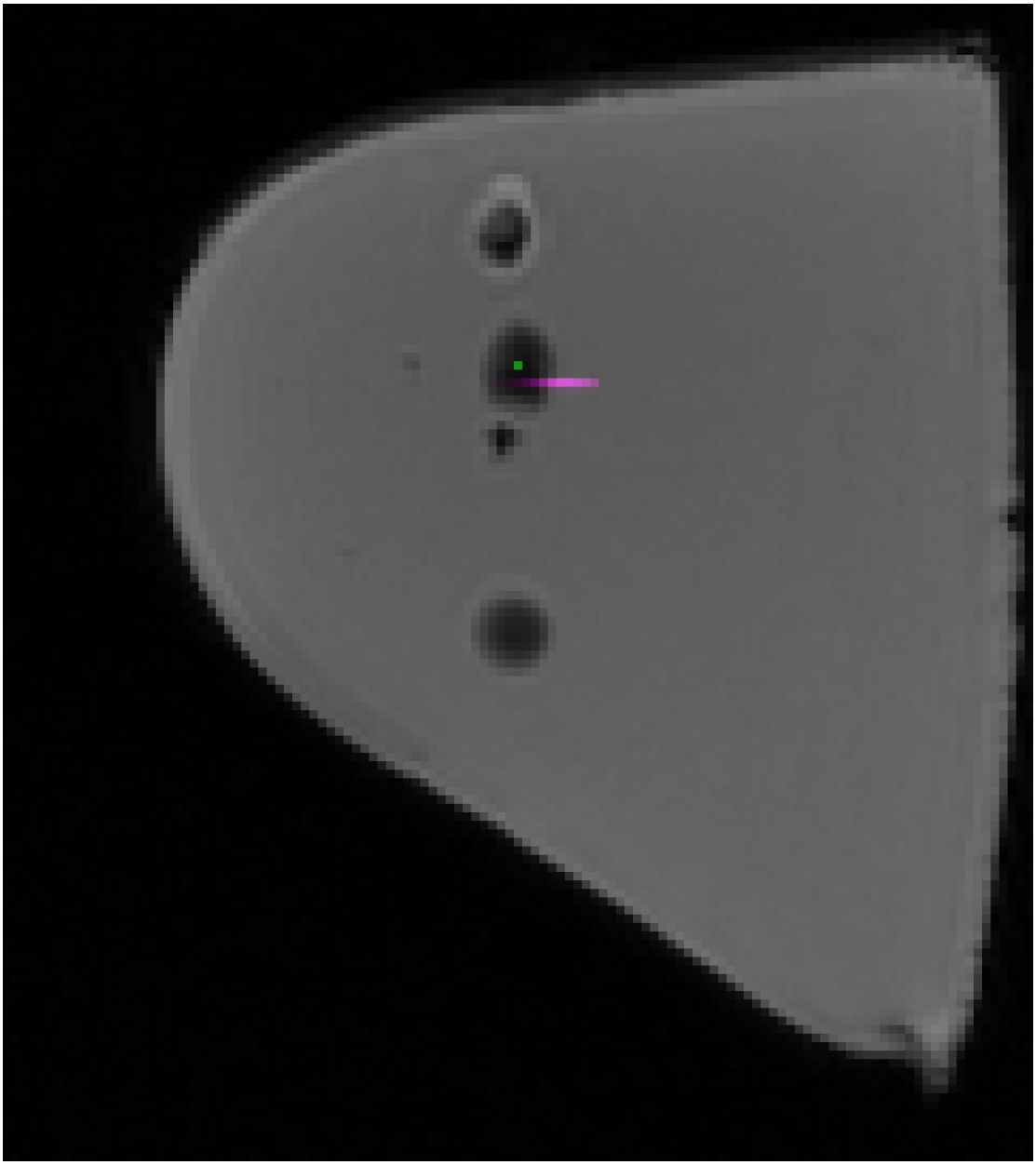

FIGURE 3.17: Non-deformed pre-operative/pre-interventional MRI 2D cross section overlayed with the example POI MLPs corresponding 2D cross section (first mapped to grayscale and then to green color) and the approximate expected moved example POI MLPs corresponding 2D cross section (first mapped to grayscale and then to magenta color).

Chapter 4

Gaussian Process-based Generative Segmentation Model for Uncertainty Quantification in Radiotherapy

Bojan Kocev[1,2,3], Horst Karl Hahn[2,3], Lars Linsen[4], Robert Cormack[5], Ron Kikinis[1,2,5], William M. Wells[5]

[1]University of Bremen, Bremen, Germany
[2]Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany
[3]Jacobs University, Bremen, Germany
[4]Westfälische Wilhelms-Universität Münster, Germany
[5]Harvard Medical School and Brigham and Women's Hospital, Boston, MA 02115, USA

Abstract

We address the problem of automatic efficient generation of plausible image segmentation samples that are similar to a single or multiple rater segmentations or in agreement with the output class label probabilities of a probabilistic segmentation, which arises in the field of uncertainty-aware image-guided radiotherapy. This is of great interest because it allows to quantify the uncertainty in various radiotherapy dose metrics, which can be used for evaluating the quality of the computed radiotherapy dose plan in terms of possible adverse effects on the patient. To model the segmentation boundary in an

uncertainty-aware fashion, we use a zero level set of a stochastic level set function which itself is modeled as a stationary Gaussian process (GP). In this regard, we devise a novel GP mean function which is computed based on input marginal label probabilities (MLPs). For the case of non-probabilistic (e.g., rater) input segmentation(s), we devise novel methods for estimating the MLPs based on a single or multiple (e.g., rater) input segmentations. To efficiently draw realizations from a stationary GP with a desired covariance structure on a 2D/3D image grid, we use process convolution to blur a 2D/3D white noise image with a 2D/3D image smoothing kernel in Fourier space. To showcase the devised methods, we apply them on multi-rater segmentation datasets for which a corresponding radiotherapy dose plan is available and estimate the uncertainty in computed dose metrics (e.g., generalized equivalent uniform dose (gEUD)) based on a large number of automatically generated plausible segmentation samples.

Notation

$\mathbb{c}$	a 2D/3D array of the (world) positions (in mm) of the pixels/voxels of the 2D/3D image that is segmented into background and foreground. In general, $\mathbb{c} = \text{VEC}^{-1}(c)$	76, 86
$\mathbb{u}$	a matrix or 3D array with n elements whose values, e.g., in 2D, are set as follows $\mathbb{u}_{q,r} = u_{h(q,r)} = u_i$. In general, $\mathbb{u} = \text{VEC}^{-1}(u)$	78
$\mathbb{K}$	2D/3D array (currently of the same size as the 2D/3D image that is segmented into background and foreground) whose values, e.g., in 2D, are set as follows $\mathbb{K}_{u,v} = k\left(\mathbb{c}_{u,v}, \mathbb{c}_{m,n}\right)$, where (m, n) is the 2D image center in pixel coordinates	76
$\mathbb{S}$	analogous to $\mathbb{K}$, i.e., e.g., in 2D, $\mathbb{S}_{u,v} = s(\mathbb{c}_{u,v}, \mathbb{c}_{m,n})$. Note that we do not identify s, we only assume that it exists, and compute $\mathbb{S}$ based on $\mathbb{K}$. The relationship with S is: $S = \text{MAT}(\mathbb{S})$.	76
$\mathbb{v}$	a matrix or 3D array with n elements whose values, e.g., in 2D, are set as follows $\mathbb{v}_{q,r} = v_{h(q,r)} = v_i$. In general, $\mathbb{v} = \text{VEC}^{-1}(v)$	62, 76, 93
$\mathbf{M}$	linear transformation component of the affine transformation that transforms the pixel/voxel coordinates in mm	86
$\mathbf{t}$	translation component of the affine transformation that transforms the pixel/voxel coordinates in mm	86

c	1D array of the (world) positions (in mm) of the pixels/voxels of the 2D/3D image that is segmented into background and foreground. In general, $c = \mathrm{VEC}(\mathbb{c})$.	86
l	characteristic length-scale (hyper-) parameter used for all input dimensions	70
$*$	the convolution operator	72, 76
$\mathrm{cov}(w)$	covariance matrix of the random vector w	75
D	a real-valued random variable which takes as values the signed distance values of all pixels/voxels with respect to all rater segmentations.	72
D_b	a real-valued random variable which takes as values the signed distance values (w.r.t. the "true" segmentation) of all pixels/voxels on the boundary of all rater segmentations.	74, 81
u	discretized stochastic level set function sample	74
$\mathbb{E}[a]$	expected value of a	75
$\sqrt[\circ]{\mathbb{A}}$	element-wise square root of the 2D/3D array $\mathbb{A}$	76
$F(\mathbb{A})$	multidimensional Fourier transform of the 2D/3D array $\mathbb{A}$. (Implemented using a fast Fourier transform algorithm)	76
$F^{-1}(\mathbb{A})$	multidimensional discrete inverse Fourier transform of the 2D/3D array $\mathbb{A}$. (Implemented using a fast Fourier transform algorithm)	76
$\mathcal{GP}(m, k)$	a Gaussian Process (GP) with a mean function m and a covariance function k	69, 74
H	the Heaviside step function	69
h	a function which maps from matrix or 3D array indices to 1D array (i.e., vector) indices, e.g., in 2D, $[\mathrm{VEC}(\mathbb{v})]_{h(q,r)} = \mathbb{v}_{q,r}$	76

$\odot$	the Hadamard product operator	77
$\mathcal{I}$	the identity matrix (of size n)	75
Ω	input space and also the index set for the GP	69
K	$n \times n$ covariance (or Gram) matrix	74
$k(x, x')$	covariance (or kernel) function evaluated at x and x'	69, 76, 87, 89, 92
$\text{MAT}(\mathbb{A})$	an operator that converts (relatively small) 2D/3D arrays into (relatively large) 2D (covariannce/standard codeviation?) matrices. For 2D arrays, it is defined as follows: $[\text{MAT}(\mathbb{A})]_{i,j} = \mathbb{A}_{q-s+m,r-t+n}$, where $h(q,r) = i$ and $h(s,t) = j$. Note that $K = \text{MAT}(\mathbb{K})$ and $S = \text{MAT}(\mathbb{S})$.	63, 76, 88
$m(x)$	the mean function of a/the Gaussian process (GP)	69
$\theta(x)$	the probability that the index location x is assigned the label 1	69
Θ	a 3D array that holds the estimated marginal label probabilities $\theta(x)$ for each voxel with world position x in mm.	72
μ	mean vector (of length n) obtained by evaluating the mean function of the GP at the n index set locations	74
$f(\mathbf{y} \mid a, \Sigma)$	Multivariate normal probability density function with a mean vector a and a covariance matrix Σ.	70
n	the number of index set locations of the GP, which in our case corresponds to the number of pixels/voxels of the 2D/3D image that gets segmented into background/foreground	75
Φ	the cumulative distribution function (CDF) of the standard normal distribution	70
n_{rs_*}	the total number of rater segmentations during the test phase	71

P	a diagonal matrix whose entries along the diagonal are set as follows $P_{i,i} = l^2$	70
$R^{a,b}(\mathbb{A})$	2D array shift operator, which is defined as $\left[R^{a,b}(\mathbb{A})\right]_{u,v} = \mathbb{A}_{u+a,v+b}$. Note that $R^{a,b}$ is a linear operator.	64, 76, 91, 94
ρ_{i*}	the binary segmentation map by the i-th rater provided during the test phase.	71
S	"matrix square root" of K, which fulfills $SS = K$. Note that K is positive definite symmetric (covariance) matrix. Hence, S is also symmetric and therefore it fulfills $SS^T = K$	75
σ_f^2	variance of the (noise free) signal. In our case, the signal is the level set function $\phi(x)$	70
ϕ	stochastic level set function modeled as a GP	69
T	a Bernoulli random variable which takes as values the "true" binary labels of all pixels/voxels.	72
v	standard normal random (column) vector of n elements, i.e., $v \sim \mathcal{N}(\mathbf{0}, \mathcal{I})$	61, 74
$\text{VEC}(\mathbb{A})$	Reshapes/converts the 2D/3D array $\mathbb{A}$ into a 1D array (i.e., vector)	76, 91
$\text{VEC}^{-1}(a)$	Reshapes/converts the 1D array (i.e., vector) a into a 2D/3D array	79
$z(x)$	binary segmentation label function, which assigns a label to each index location x	69
$\mathbf{0}$	(column) vector of all 0's (of length n)	75

4.1 Introduction

One of the problems that radiation treatment planners face is the situation where a tumor is close to an "organ at risk" (OAR). In this case, the goal is to create a radiation field that has a high spatial gradient between the two structures. The concern with

respect to uncertainty is that if the shape and position of any region of interest (ROI) is uncertain, then there is some risk that part of the OAR will receive high dose that leads to complications while part of the tumor will receive low dose that leads to recurrence. We believe that estimating this uncertainty and making effective visualizations of this situation would improve the radiation treatment quality.

The uncertainty in shape and position of the ROI normally originates from uncertainties in the setup or in the output from the deployed image segmentation algorithm (Lê et al. (2015, 2016)) or the human rater. To quantify the impact of uncertainty in segmentation on radiotherapy dose delivered to the patient, one could use *Image Segmentation Sampling* (ISS) (Lê et al. (2015, 2016)). During this generative segmentation process, one draws multiple segmentation samples and estimates the effect on the delivered dose and accompanying dose metrics. The easy-to-apply sampling from generative models on segmentation boundaries is particularly attractive because the binary segmentation output is at the same time the required input to the systems which compute different dose metrics.

We automate the generation of plausible image segmentation samples that are similar to a single or multiple rater segmentations or in agreement with the output label probabilities of a probabilistic segmentation. We model the segmentation boundary as a zero level set of a stochastic level set function which itself is modeled as a stationary Gaussian process (GP) (Lê et al. (2015, 2016)). In other words, the segmentation boundary is modeled as an implicit surface/contour (Gerardo-Castro et al. (2015); Turk and O'brien (1999); Williams and Fitzgibbon (2007)). This means that image segmentation samples can be generated by thresholding realizations of the GP, which is reasonably fast in comparison to more complex Markov Chain Monte Carlo (MCMC) approaches.

The first novelty of our approach stems from a generalization of the current state-of-the-art GP mean function (Lê et al. (2015, 2016)), by setting it based on input marginal label probabilities (MLPs). In other words, we identify the relationship between the GP mean function value at an index location and the MLP (i.e., the probability that the index location is labeled with a given label) (see Subsection 4.3.2). This means that if the deployed image segmentation algorithm can output MLPs, we can directly use them to determine the mean function. In return, this broadens the applicability of the proposed approach. We also, as a second novelty, propose different methods for estimating MLPs based on a single or multiple (e.g., rater) input segmentations (see Section 4.4). This further broadens the applicability of the proposed method to cases when (a) non-probabilistic (e.g., rater) input segmentation(s) is/are provided. Regarding the choice of the GP covariance function, we follow up on the idea by Lê et al. (2015, 2016) and set it with a squared exponential covariance function with an isotropic distance

measure (Rasmussen and Williams (2006)). We use process convolution (Higdon (2002)) to convole/blur 2D/3D white noise images with a 2D/3D image smoothing kernel in Fourier space and in this way efficiently draw samples/realizations from a stationary GP on a 2D/3D image grid. The image smoothing kernel is computed based on the deployed GP covariance function. In regard to the presented method for efficient sampling (see Section 4.5), we view the detailed derivation of how to apply the process convolution idea on 2D/3D image grids as our third novelty. We apply the devised methods on multi-rater segmentation datasets for which a corresponding radiotherapy dose plan is available and find out that there is significant uncertainty in the computed dose metrics (e.g., generalized equivalent uniform dose(gEUD)) based on a large number of automatically generated plausible image segmentation samples (see Section 4.6). We view this finding as our fourth novelty, which may be considered of utmost value in clinical practice. In summary, our contributions to the state of the art comprise:

1. a novel mean function for the GP over level set functions whose zero level sets are the segmentation boundaries;
2. novel methods for estimating MLPs based on a single or multiple (e.g., rater) input segmentations;
3. a detailed derivation of how to apply the process convolution idea on 2D/3D image grids, and
4. application on multi-rater segmentation datasets coupled with radiotherapy dose plans and identifying that there is a significant uncertainty in the computed dose metrics.

4.2 Related Work

Warfield et al. (2004) compute a probabilistic estimate of the true segmentation based on a collection of segmentations generated by human raters or automated segmentation algorithms. In this regard, they estimate an optimal combination of the segmentations by weighting each segmentation accordingly and combining it with a prior model and spatial homogeneity constraints. In more detail, they evaluate the conditional probability density of the hidden true segmentation given the segmentations and estimates of the performance level of the segmentation generators. The probabilistic estimate of the true segmentation can be given as an input to our image segmentation sampling algorithm, which then can be used to draw many possible true segmentation samples.

Warfield et al. (2006, 2008) estimate the performance characteristics of segmentation generators and the true segmentation from segmentations with labels that may be ordered or continuous measures. They support, for example, surface, distance transform or level set representations of segmentations. In this regard, they generalize their previous method (Warfield et al. (2004)) to segmentations with ordered labels. We instead use a Gaussian process to model a level set representation of segmentations based on input binary segmentations and then draw many possible segmentations, similar to the input segmentations, for uncertainty quantification.

Pohl et al. (2007) pointed out to different explicit and implicit approaches to shape presentation. In this regard, they propose a novel implicit shape representation. They identified the relationship between signed distance maps and the logarithm of the odds ratio (LogOdds) of probabilistic atlases (analogous to marginal label probabilities in our context). Their novel implicit shape representation is therefore called LogOdds, which embeds signed distance maps in a vector space. In more detail, they pointed out that signed distance maps can always be interpreted as LogOdds maps. Furthermore, they stated that the reverse is in general not true. The LogOdds, also known as the logit function, is related to the probit function, which we use for setting the mean function of our GP model over level set functions whose zero-level sets are the segmentation boundaries. The logit and the probit functions are both quantile functions, i.e., they are both the inverse of the cumulative distribution function of a probability distribution.

Sabuncu et al. (2010) fuse multiple expert (training) segmentations to compute the final segmentation using a generative model. Their model assumes that the final segmentation is generated from one or more training segmentations, however from which training segmentations is unknown. In more detail, they define a latent random field which specifies the index of the training image that is used to generate the label of each voxel in the test image. However, they assume that the labels at each voxel are conditionally independent. Furthermore, they do not draw many possible final segmentations, but compute a single final segmentation estimate.

Fan et al. (2007) define complex probability distributions on the space of curves and generate samples therefrom using a Markov chain Monte Carlo (MCMC) algorithm. They use a GP sample to randomly perturb the previous curve sample when generating a candidate curve sample from the proposal distribution. To generate a zero-mean GP sample, they circularly convolve a white Gaussian noise sample with a smoothing (e.g., Gaussian) kernel. This is related to how we generate GP samples, however their GP is defined over a 1D domain and they do not discuss how to draw GP samples efficiently. Furthermore, they do not provide information about how to engineer a desired GP

covariance structure using the smoothing kernel. In addition, they perform segmentation sampling on 2D images.

Niethammer et al. (2017) presented an active mean fields (AMF) approach that allowed an efficient approximation of the posterior distribution on segmentation labels using a variational mean-field (VMF) approach. In this regard, they approximate the posterior distribution using a simpler variational distribution, while minimizing the Kullback-Leibler divergence between the two distributions. In more detail, their mean-field approximation of the posterior distribution is defined as a field of independent Bernoulli random variables. Their work, however, does not focus on (efficient) sampling from the (posterior) distribution on segmentation labels. Instead of representing the segmentation boundary as a zero level set, they use a dense logit ("log odds") representation of label probabilities. They estimate a level-set function in a probabilistic fashion, which they interpret it as an approximate posterior on the segmentation. In regard to our work, their estimated level-set function is analogous to the mean function of our GP model over level-set functions. Their estimated parameters of the Bernoulli posterior distribution on labels can be used for setting the mean function of our GP. They use a prior to regularize the length of the segmentation boundary, which in our case is controlled by the characteristic length-scale (hyper-) parameter of the covariance function of our GP and the GP mean function itself.

The inspiring work of Lê et al. (2015, 2016) is the closest to ours. They presented a method for an efficient generation of various plausible image segmentation samples of the same structure based on a single expert segmentation. They also model the segmentation boundary as a zero level set of a stochastic level set function. In this regard, their stochastic level set function is modeled as a GP with a mean function equal to the signed geodesic distance from the expert segmentation and with a squared exponential covariance function. We view their mean function as a particular assumption of marginal label probabilities based on image structure, i.e., our proposed GP mean function is more general because it can accommodate for any specified marginal label probabilities. Regarding efficient sampling from a GP, they point out to three different methods. The first two methods use properties of block circulant with circulant blocks matrices and depend on the discrete Fourier transform. In this regard, these two methods are closely related to our method for efficient sampling, however, we use process convolutions to generate the samples. Their third method uses the property that their chosen covariance function is separable, depends on the Kronecker product, and performs Cholesky decomposition in an efficient manner but still slower than the other two methods. This is their method of choice.

Regarding the presented method for efficient sampling from stationary GPs, we were inspired by the work of Higdon (2002). They pointed out that a GP can be constructed by convolving a continuous white noise process with a smoothing kernel. In this regard, they described the relationship, based on the convolution theorem for Fourier transforms, between the smoothing kernel and the covariance function of the GP. We use this relationship to compute a 2D/3D image smoothing kernel with which we convolve/blur a white noise 2D/3D image in order to draw samples/realizations of a zero-mean GP with a desired covariance structure on a 2D/3D image grid. In principle, we work out the derivations of how to apply the process convolution idea on 2D/3D image grids.

4.3 Gaussian Process-based Segmentation Contour Sampler

We model the segmentation contour as a zero level set of a stochastic level set function ϕ (Lê et al. (2016, 2015)). Hence, we use an implicit shape representation (Pohl et al. (2007)). The stochastic level set function is modeled as a GP with a novel mean function $m(x) : \Omega \to \mathbb{R}$ and a covariance function $k(x, x') : \Omega \times \Omega \to \mathbb{R}$, i.e., $\phi \sim \mathcal{GP}(m, k)$, where Ω is an index set. Note that the notation $\phi \sim \mathcal{GP}(m, k)$ means that ϕ is distributed as a GP with mean and covariance functions m, and k, respectively. We require $k(x, x) = 1$ (see Subsection 4.3.2). The novel mean function is defined based on pre-computed marginal label probabilities.

4.3.1 Marginal Label Probabilities

We define a label function $z(x) : \Omega \to \{0, 1\}$, which assigns binary labels to all index locations. The labels are computed by thresholding the level set function samples using the Heaviside step function H (with the convention $H(0) = 1$), i.e.,

$$z(x) \doteq H(\phi(x)). \tag{4.1}$$

We then define the marginal label probability $\theta(x) : \Omega \to [0, 1]$ as the probability that the index location x is assigned the label 1, i.e., $\theta(x) = p(z(x) = 1)$. We have different means for pre-computing the desired $\theta(x)$ (see Section 4.4).

4.3.2 Mean Function

The GP shall be set such that when drawing ϕ samples the following condition $p(z(x) = 1) = p(\phi(x) \geq 0) = \theta(x)$ is fulfilled. In this regard, the marginal label probability $\theta(x)$

can be interpreted as an excursion probability (Adler and Taylor (2009)). Note that this condition is on the marginal distribution on $\phi(x)$, which we model it as a univariate Gaussian with mean $m(x)$ and variance 1, i.e., $\phi(x) \sim \mathcal{N}(m(x), 1)$. Hence, this condition is on the mean function value $m(x)$. Formally, this condition can be expressed as the following integral equality

$$\int_0^\infty f\left(y \mid m(x), 1\right)\, dy \;=\; \theta(x),$$

where $f\left(\mathbf{y} \mid a, \Sigma\right)$ in this case is the normal probability density function. The above integral equality is therefore equivalent to

$$\int_0^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{(y - m(x))^2}{2})\, dy \;=\; \theta(x).$$

Let $y' = m(x) - y$. This yields $dy' = -dy$. Furthermore, the lower and upper limits of the integral become $m(x) - 0 = m(x)$ and $m(x) - \infty = -\infty$ respectively. Therefore, the above integral equality becomes

$$-\int_{m(x)}^{-\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y'^2}{2})\, dy' \;=\; \theta(x),$$

which is equivalent to

$$\int_{-\infty}^{m(x)} \frac{1}{\sqrt{2\pi}} \exp(-\frac{y'^2}{2})\, dy' \;=\; \theta(x).$$

The left-hand side of the above equality is the cumulative distribution function of the standard normal distribution evaluated at $m(x)$, i.e., we obtain

$$\Phi(m(x)) = \theta(x),$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. As a result, the mean function of the GP evaluated at x shall be set equal to the inverse of the standard normal CDF evaluated at $\theta(x)$, i.e.,

$$m(x) = \Phi^{-1}(\theta(x)). \tag{4.2}$$

The mean function $m(x)$ defined as such is known in the literature as the probit function. In summary, given specified marginal label probability $\theta(x)$, we can construct the mean function value $m(x)$ such that the excursion probability $p(\phi(x) \geq 0)$ is equal to $\theta(x)$.

4.3.3 Covariance Function

We set the covariance function of the GP with a squared exponential covariance function with isotropic distance measure, i.e.,

$$k\left(x, x'\right) = \sigma_f^2 \exp(-\frac{1}{2}(x - x')^T P^{-1}(x - x')) \tag{4.3}$$

where P is a diagonal matrix whose entries along the diagonal are set as $P_{i,i} = l^2$ (l is a characteristic length-scale (hyper-) parameter), and σ_f^2 is the signal variance (Rasmussen and Williams (2006)). To fulfill the requirement $k\left(x, x\right) = 1$, we set $\sigma_f = 1$.

4.4 Estimation of Marginal Label Probabilities

In this section, we present four different methods for estimating the marginal label probability $\theta(x)$. The first two methods do not require training, however they require multi-rater segmentations during the testing phase. The other two methods require some sort of training, however they could be used to estimate $\theta(x)$ based on a single test rater segmentation. In the following, each of these four methods is presented in detail.

4.4.1 α-smoothed Empirical Probabilities

Provided that there is large enough test multi-rater segmentations dataset, the simplest approach is to compute empirical probabilities as follows:

$$\theta(x) = \frac{\sum_{i=1}^{n_{rs_*}} \rho_{i_*}(x)}{n_{rs_*}}, \tag{4.4}$$

where n_{rs_*} is the total number of test rater segmentations and ρ_{i_*} is the i-th test rater segmentation. (Note that, in order to simplify notation, $\rho_{i_*}(x)$ is treated as an integer (either zero or one) instead of a Boolean in Eq. (4.4).) However, especially in case of low-probability events (i.e., in case of voxels having a low probability of being labeled one by the raters) and with small test datasets (i.e., for small n_{rs_*}), there is the possibility of a possible event not occurring. Therefore, the observed frequency $\sum_{i=1}^{n_{rs_*}} \rho_{i_*}(x)$ would be zero, which will imply a probability $\theta(x)$ equal to zero. This simplification is inadequate. We therefore artificially adjust the probability of rare (but not impossible) events, so that the probabilities of these rare events are not exactly zero and in this way circumvent the zero-frequency problems. To do that, we deploy additive/Laplace/Lidstone smoothing (Schütze et al. (2008)) and compute α-smoothed empirical probabilities as follows:

$$\theta(x) = \frac{\sum_{i=1}^{n_{rs_*}} \rho_{i_*}(x) \ + \alpha}{n_{rs_*} + 2\alpha}, \tag{4.5}$$

where $\alpha > 0$ is a smoothing parameter that is usually called a "pseudocount". The factor 2 in 2α in the denominator corresponds to the number of categories, i.e., a voxel can be either marked one or zero. The "pseudocount" amount α, not necessarily an integer, is added to the number of observed cases, so that the expected probability is not zero. Basically, for each category we have an additional count of α a priori. The α-smoothed empirical probability is usually called a posterior probability, given rater data and prior.

4.4.2 Blurred α-smoothed Empirical Probabilities

Let $\mathbb{A}$ be a 3D array which holds the marginal label probabilities $\theta(x)$, estimated using Eq. (4.5), for each voxel with world position x in mm. In this subsection, we argue that,

especially in case of small test datasets (i.e., for small n_{rs_*}), neighboring entries in $\mathbb{A}$ may not be segmented similarly enough despite the use of α-smoothing. One could argue that one could use a larger α to increase the smoothing effect, however in that case one would also increase the probability of zero-frequency events for all voxels/entries which might not be wanted. To alleviate this problem, we propose to complement the α-smoothing with a Gaussian blur on the 3D array $\mathbb{A}$. To achieve this, we convolve the 3D array $\mathbb{A}$ with a 3D Gaussian convolution kernel $\mathbb{G}$ which then results in another 3D array Θ with entries:

$$\Theta_{i,j,k} = [\mathbb{G} * \mathbb{A}]_{i,j,k} = \sum_{u,v,w} \mathbb{G}_{u,v,w} \mathbb{A}_{i-u,j-v,k-w} \ , \tag{4.6}$$

where $*$ is the convolution operator, and u, v, and w range over all legal subscripts for $\mathbb{G}_{u,v,w}$ and $\mathbb{A}_{i-u,j-v,k-w}$. (See also Eq. (4.C.1).)

4.4.3 Bayesian Probabilities

In this subsection, we present a method that aims at estimating realistic marginal label probabilities $\theta(x)$ based on a single test rater segmentation. The motivation for this is the assumption that it is more likely in practice to have a single rater segmentation during testing. The probability $\theta(x)$ is formulated as a conditional probability with the conditioning being performed on the event that a random variable takes the value $\mathcal{D}(\rho_*, x)$, where $\mathcal{D}(\rho_*, x)$ is the signed distance transform (SDT) of the single test rater binary segmentation image ρ_* evaluated at the voxel with world position x in mm. Formally, we have

$$\theta(x) = p(T = 1 \mid D = \mathcal{D}(\rho_*, x)), \tag{4.7}$$

where T is a Bernoulli random variable which takes as values the "true" binary labels of all voxels and D is a real-valued random variable which takes as values the signed distance values of all voxels with respect to all rater segmentations. To evaluate the right-hand side of Eq. (4.7), we use Bayes' theorem to obtain an expression that contains probabilities that we can realistically estimate:

$$p(T = 1 \mid D = \mathcal{D}(\rho_*, x)) = \frac{p(D = \mathcal{D}(\rho_*, x) \mid T = 1)p(T = 1)}{p(D = \mathcal{D}(\rho_*, x))} \ , \tag{4.8}$$

where the denominator in the expression of the right-hand side of Eq. (4.8) will be estimated by marginalizing out T:

$$\begin{aligned} p(D = \mathcal{D}(\rho_*, x)) = \ & p(D = \mathcal{D}(\rho_*, x) \mid T = 0)p(T = 0) \\ & + p(D = \mathcal{D}(\rho_*, x) \mid T = 1)p(T = 1) \end{aligned} \tag{4.9}$$

As can be observed in Eq. (4.8) and Eq. (4.9), in order to estimate $\theta(x)$, we need to estimate these four probabilities: $p(T = 0)$, $p(T = 1)$, $p(D = \mathcal{D}(\rho_*, x) \mid T = 0)$, and

$p(D = \mathcal{D}(\rho_*, x) \mid T = 1)$. To estimate the prior probability $p(T = 1)$, we require some prior knowledge about the ratio of the number of voxels with "true" label "1" to the total number of voxels in any 3D image that contains the region of interest (ROI) which is segmented, assuming that all considered 3D images contain more or less the same portion of the human body and are of similar size. This prior knowledge is acquired during a training phase. In this regard, we require a "true" segmentation of the ROI in any image from any patient provided during the training phase. We anticipate two options for acquiring the "true" segmentation. The first and ideal one is to have a highly ranked expert rater segment the ROI, while the second less favorable one is to have multiple less experienced raters segment the ROI and use the mean segmentation as the "true" segmentation. In both cases, we denote the "true" segmentation with ρ_t. The corresponding probability of the initial degree of belief against the event "$T = 1$", i.e., $p(T = 0)$, is set as $p(T = 0) = 1 - p(T = 1)$. The likelihoods $p(D = \mathcal{D}(\rho_*, x) \mid T = 1)$ and $p(D = \mathcal{D}(\rho_*, x) \mid T = 0)$ are estimated using kernel density estimation (KDE)(Breiman et al. (1977); Terrell and Scott (1992)) during the testing phase each based on a different finite data sample acquired during the training phase. In this regard, we require at least one training rater segmentation ρ of the ROI in the image to which ρ_t corresponds. Let $\left(d_1^{(1)}, d_2^{(1)}, ..., d_{\nu_1}^{(1)}\right)$ be a univariate independent and identically distributed sample from the distribution over the signed distances (with respect to rater segmentations of the ROI) of all voxels that have been labeled "1" in the respective "true" segmentation, which is generated based on all available above-described training rater segmentations and ρ_t. Analogous to this, let $\left(d_1^{(0)}, d_2^{(0)}, ..., d_{\nu_0}^{(0)}\right)$ be a sample of rater signed distance of voxels that have been labeled "0" in ρ_t, which is also generated based on all available above-described training rater segmentations. Then, the likelihood $p(D = \mathcal{D}(\rho_*, x) \mid T = 1)$ is estimated as follows:

$$p(D = \mathcal{D}(\rho_*, x) \mid T = 1) = \frac{1}{\nu_1 \eta} \sum_{i=1}^{\nu_1} \Gamma\left(\frac{\mathcal{D}(\rho_*, x) - d_i^{(1)}}{\eta}\right), \tag{4.10}$$

where Γ is a non-negative kernel smoothing function and $\eta > 0$ is a bandwidth that acts as a smoothing parameter, i.e., controls the smoothness of the resulting density curve. Analogous to this, the likelihood $p(D = \mathcal{D}(\rho_*, x) \mid T = 0)$ is estimated as follows:

$$p(D = \mathcal{D}(\rho_*, x) \mid T = 0) = \frac{1}{\nu_0 \eta} \sum_{i=1}^{\nu_0} \Gamma\left(\frac{\mathcal{D}(\rho_*, x) - d_i^{(0)}}{\eta}\right). \tag{4.11}$$

4.4.4 Parametric Model-based Probabilities

In this subsection, we present a method that, similar to the method presented in Subsection 4.4.3, aims at estimating realistic marginal label probabilities $\theta(x)$ based on a single test rater segmentation. The probability $\theta(x)$ is formulated as the probability

that a real-valued random variable D_b will take a value less than or equal to $\mathcal{D}(\rho_*, x)$,. Formally, we formulate it as

$$\begin{aligned}\theta(x) &= p(D_b \leq \mathcal{D}(\rho_*, x)) \\ &= F_{D_b}(\mathcal{D}(\rho_*, x)),\end{aligned} \tag{4.12}$$

where D_b is a real-valued random variable which takes as values the signed distance values (w.r.t. the "true" segmentation) of all pixels/voxels on the boundary of all rater segmentations and F_{D_b} is the cumulative distribution function of the random variable D_b. Note that, we assume that the signed distance is positive inside the segmented ROI and negative outside with zero on the segmentation boundary. Furthermore, we treat the pixels/voxels on the boundary as being inside the ROI. To learn the probability distribution of the random variable D_b, we require a training step. In this regard, we use the same training data as the one used for the method presented in Subsection 4.4.3, i.e., we use the "true" segmentation ρ_t and at least one training rater segmentation ρ as defined in Subsection 4.4.3. We use this training data to generate a univariate independent and identically distributed sample from the probability distribution of the random variable D_b and fit an appropriate parametric probability distribution (e.g., Gaussian) model on this data. Note that, we make the assumption that the provided test rater segmentation ρ_* is close to the unknown true segmentation. Furthermore, the estimated marginal label probability $\theta(x)$ is in fact the probability that imaginary raters (with similar level of expertise with those from the training phase) would assign the label "1" to the pixel/voxel with world position x in mm during the testing phase.

4.5 Efficient Sampling

As mentioned above, the stochastic level set function ϕ is modeled as a GP, i.e., $\phi \sim \mathcal{GP}(m, k)$, with the covariance function being stationary. The general problem we address in this section is how to efficiently draw stationary GP samples. In practice, one can only draw discretized function samples. The GP is therefore discretized at the index locations. In general and in our case in particular, this is not a limitation as one is usually interested in the values of the function samples at discrete locations (e.g., in our case we are interested in the values of the level set function samples at the world positions (in mm) of the pixels/voxels of the image that is segmented into background/-foreground). In order to sample from the discretized GP, we need to evaluate the mean and covariance functions of the GP at the index locations. This results in a mean vector μ and a covariance matrix K. Then, the task of drawing discretized GP samples boils down to drawing multivariate normal samples $u \sim \mathcal{N}(\mu, K)$. We can easily generate standard normal random (column) vector samples $v \sim \mathcal{N}(\mathbf{0}, \mathcal{I})$, by multiple separate

calls to a zero-mean unit-variance scalar Gaussian generator (readily available in many programming environments) (Rasmussen and Williams (2006)). Note that $\mathbf{0}$ is a (column) vector of all 0's (of length n, where n is the number of index set locations), and $\mathcal{I}$ is the identity matrix (of size n). Then, each generated v sample could be affinely transformed $\mathbf{A}v + \mathbf{b}$ (see Appendix 4.A) into u with the desired/target multivariate normal distribution. Therefore, we need to identify $\mathbf{A}$ and $\mathbf{b}$, such that the generated u samples are multivariate normal samples with mean μ and covariance matrix K. In this regard, the derivation of $\mathbf{b}$ is as follows:

$$\begin{aligned} \mathbb{E}\left[\mathbf{A}v + \mathbf{b}\right] &= \mu \\ \mathbf{A}\mathbf{0} + \mathbf{b} &= \mu && \text{See Eq. (4.A.1.1).} \\ \mathbf{b} &= \mu \end{aligned} \tag{4.13}$$

while for $\mathbf{A}$ we have the following:

$$\begin{aligned} \mathrm{cov}(\mathbf{A}v + \mathbf{b}) &= K \\ \mathbf{A}\mathcal{I}\mathbf{A}^T &= K && \text{See Eq. (4.A.1.2).} \\ \mathbf{A}\mathbf{A}^T &= K \end{aligned} \tag{4.14}$$

where $\mathbb{E}\left[a\right]$ is the expected value of any a and $\mathrm{cov}(w)$ is the covariance matrix of any random vector w. Hence, we need to identify a matrix $\mathbf{A}$ that fulfills Eq. (4.14). The standard approach in the literature is to deploy Cholesky factorization/decomposition (known as the "matrix square root") of the positive definite symmetric covariance matrix K to factor it as $K = LL^T$(Rasmussen and Williams (2006)), where L is a lower triangular matrix. In principle, we could set $\mathbf{A} = L$. However, the Cholesky factorization/decomposition turns out to be ill-conditioned and computationally expensive when the number of index set locations is large (Lê et al. (2016, 2015)). In regard to the computational complexity of the Cholesky factorization, it takes $\frac{n^3}{6}$ operations (Rasmussen and Williams (2006)). Furthermore, note that Rasmussen and Williams (2006) suggested to add a small multiple of the identity matrix to the covariance matrix for numerical reasons and claimed that without this stabilization the Cholesky factorization fails.

We propose an alternative factorization $K = SS$. As K is symmetric, S is also symmetric and therefore it also fulfills $K = SS^T$. Therefore, if we set $\mathbf{A} = S$, $\mathbf{A}$ will satisfy Eq. (4.14) as desired. At the same time, we assume that S is constructable by evaluating a stationary function $s(x, x')$ at the index locations, which means that, for example in 2D, S also fulfills the following:

$$\left[SSv\right]_i = \mathrm{VEC}(R^{m,n}\left(R^{m,n}\left(\mathbb{S} * \mathbb{S} * \mathbb{v}\right)\right))_i \,, \tag{4.15}$$

where VEC($\mathbb{A}$) reshapes/converts the 2D/3D array $\mathbb{A}$ into a 1D array, $R^{a,b}(\mathbb{A})$ is a 2D array shift operator which is defined as $\left[R^{a,b}(\mathbb{A})\right]_{u,v} = \mathbb{A}_{u+a,v+b}$, $\mathbb{S}$ is defined, in 2D, as $\mathbb{S}_{u,v} \doteq s(\mathbb{c}_{u,v}, \mathbb{c}_{m,n})$ with $\mathbb{c}$ being a 2D array of the (world) positions (in mm) of the pixels of the 2D image that is segmented and (m, n) being the center in pixel coordinates of the 2D image that gets segmented (note that we do not identify s, we only assume that it exists, and compute $\mathbb{S}$ based on an analogous discretization of the covariance function $k(x, x')$), $\mathbb{v} = \text{VEC}^{-1}(v)$, and $*$ is the convolution operator. The proof of Eq. (4.15) is provided in Appendix 4.C (see Eq. (4.C.3)). As K is also constructed by evaluating a stationary covariance function at the index locations, it fulfills, for example in 2D, the following:

$$[Kv]_i = \text{VEC}(R^{m,n}(\mathbb{K} * \mathbb{v}))_i \,, \tag{4.16}$$

where $\mathbb{K}$, in 2D, is a 2D array (currently of the same size as the 2D image that is segmented) with its values defined as follows:

$$\mathbb{K}_{u,v} \doteq k(\mathbb{c}_{u,v}, \mathbb{c}_{m,n}) \tag{4.17}$$

The proof of Eq. (4.16) is provided in Appendix 4.C (see Eq. (4.C.2)). As $S = \text{MAT}(\mathbb{S})$, where MAT($\mathbb{A}$) is defined, in 2D, as $[\text{MAT}(\mathbb{A})]_{i,j} = \mathbb{A}_{q-s+m,r-t+n}$ with $h(q, r) = i$, $h(s, t) = j$, and h, in 2D, being a function which maps from pixel coordinates to 1D array indices (see also Eq. (4.B.3)), we proceed with deriving $\mathbb{S}$ based on $\mathbb{K}$, in 2D, as follows:

$$\begin{aligned}
SS = K \iff \\
[SSv]_i = [Kv]_i \iff \\
\text{VEC}(R^{m,n}(R^{m,n}(\mathbb{S} * \mathbb{S} * \mathbb{v})))_i = \text{VEC}(R^{m,n}(\mathbb{K} * \mathbb{v}))_i \iff \quad &\text{See (4.15) and (4.16).} \\
R^{m,n}(R^{m,n}(\mathbb{S} * \mathbb{S} * \mathbb{v})) = R^{m,n}(\mathbb{K} * \mathbb{v}) \iff \\
R^{m,n}(\mathbb{S} * \mathbb{S} * \mathbb{v}) = \mathbb{K} * \mathbb{v}
\end{aligned} \tag{4.18}$$

At this point, inspired by Higdon (2002), we guess $\mathbb{S}$ as follows:

$$\mathbb{S} = R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F(\mathbb{K})}\right)\right) \,, \tag{4.19}$$

where $F(\mathbb{A})$ is the multidimensional Fourier transform of the 2D/3D array $\mathbb{A}$, $\sqrt[\odot]{\mathbb{A}}$ is the element-wise square root of the 2D/3D array $\mathbb{A}$, and $F^{-1}(\mathbb{A})$ is the multidimensional discrete inverse Fourier transform of the 2D/3D array $\mathbb{A}$. Then, we substitute the guessed expression for $\mathbb{S}$ in Eq. (4.18):

$$R^{m,n}\left(R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F(\mathbb{K})}\right)\right) * R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F(\mathbb{K})}\right)\right) * \mathbb{v}\right) = \mathbb{K} * \mathbb{v}$$

and prove that the equality holds. First, we make use of the interchange property of the 2D array shift operator (see Appendix 4.D for a proof of it):

$$R^{a,b}(\mathbb{A}) * \mathbb{B} = R^{a,b}(\mathbb{A} * \mathbb{B}) \,, \tag{4.20}$$

where $\mathbb{A}$ and $\mathbb{B}$ are two 2D arrays, to turn the above equation into:

$$R^{m,n}\left(R^{-\frac{m}{2},-\frac{n}{2}}\left(R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*\mathbb{v}\right)\right)\right)=\mathbb{K}*\mathbb{v}$$

then we combine the three 2D array shift operations into one 2D array shift operation as follows:

$$R^{m-\frac{m}{2}-\frac{m}{2},n-\frac{n}{2}-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*\mathbb{v}\right)=\mathbb{K}*\mathbb{v}$$

which is equivalent to:

$$R^{0,0}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*\mathbb{v}\right)=\mathbb{K}*\mathbb{v}$$

and therefore finally, we prove that the following equality holds:

$$\begin{aligned}
F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*\mathbb{v}=\mathbb{K}*\mathbb{v} &\iff\\
F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)*F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)=\mathbb{K} &\iff\\
F^{-1}\left(F\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)\right)\odot F\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)\right)\right)=\mathbb{K} &\iff\\
F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\odot\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)=\mathbb{K} &\iff\\
F^{-1}\left(F\left(\mathbb{K}\right)\right)=\mathbb{K} &\iff\\
\mathbb{K}=\mathbb{K}\,, &
\end{aligned}\tag{4.21}$$

where $\odot$ is the Hadamard product operator. This completes the proof that, when the guessed expression for $\mathbb{S}$ from Eq. (4.19) is substituted in Eq. (4.18), the respective equality holds. Note that in Eq. (4.21), we make use of the discrete version of the convolution theorem (Bracewell and Bracewell (1986)).

In addition, we provide an alternative derivation of $\mathbb{S}$, which does not involve guessing. We start with Eq. (4.18) and algebraically step-by-step derive $\mathbb{S}$, in 2D, as follows:

$$
\begin{aligned}
R^{m,n}\left(\mathbb{S} * \mathbb{S} * \mathbb{v}\right) &= \mathbb{K} * \mathbb{v} \iff \\
R^{\frac{m}{2},\frac{n}{2}}\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S} * \mathbb{S} * \mathbb{v}\right)\right) &= \mathbb{K} * \mathbb{v} \iff \\
R^{\frac{m}{2},\frac{n}{2}}\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) * \mathbb{S} * \mathbb{v}\right) &= \mathbb{K} * \mathbb{v} \iff && \text{Acc. 2D array shift operator interchange property. See Eq. (4.20).} \\
R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S} * R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) * \mathbb{v}\right) &= \mathbb{K} * \mathbb{v} \iff && \text{Acc. commutativity property of convolution.} \\
R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) * R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) * \mathbb{v} &= \mathbb{K} * \mathbb{v} \iff && \text{Acc. 2D array shift operator interchange property. See Eq. (4.20).} \\
R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) * R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) &= \mathbb{K} \iff \\
F\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right)\right) \odot F\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right)\right) &= F\left(\mathbb{K}\right) \iff && \text{Acc. discrete version of the convolution theorem (Bracewell and Bracewell (1986)).} \\
F\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right)\right) &= \sqrt[\odot]{F\left(\mathbb{K}\right)} \iff
\end{aligned}
$$

$$
\begin{aligned}
R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right) &= F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right) \iff \\
R^{-\frac{m}{2},-\frac{n}{2}}\left(R^{\frac{m}{2},\frac{n}{2}}\left(\mathbb{S}\right)\right) &= R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)\right) \iff \\
\mathbb{S} &= R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\right)\right)
\end{aligned} \tag{4.22}
$$

This completes the alternative derivation of $\mathbb{S}$. As can be observed, the result in Eq. (4.22) is the same as our guess in Eq. (4.19).

Having derived $\mathbb{S}$ in terms of $\mathbb{K}$, we can now express S in terms of $\mathbb{S}$, as mentioned above, as follows $S = \text{MAT}(\mathbb{S})$ and draw u samples in the following way:

$$
u = \mathbf{A}v + \mathbf{b} = Sv + \mu \tag{4.23}
$$

Then, each (column) vector sample u would need to be converted into a 2D/3D array sample $\mathbb{u} = \text{VEC}^{-1}(u)$, that can be thresholded to compute a 2D/3D image segmentation sample. It turns out that we can directly compute $\mathbb{u}$ based on $\mathbb{K}$, $\mathbb{v}$, and $\text{VEC}^{-1}(\mu)$

as follows:

$$
\begin{aligned}
\mathbb{u} &= \mathrm{VEC}^{-1}(u) \\
&= \mathrm{VEC}^{-1}(Sv + \mu) && \text{See Eq. (4.23).} \\
&= \mathrm{VEC}^{-1}(Sv) + \mathrm{VEC}^{-1}(\mu) && \text{Here, we use the fact that } \mathrm{VEC}^{-1}(a) \text{ is a linear map.} \\
&= \mathrm{VEC}^{-1}(\mathrm{VEC}(R^{m,n}(\mathbb{S} * \mathbb{v}))) + \mathrm{VEC}^{-1}(\mu) && \text{Acc. Eq. (4.C.2), as } S, \text{ as mentioned above, is constructable by evaluating a stationary function.} \\
&= R^{m,n}(\mathbb{S} * \mathbb{v}) + \mathrm{VEC}^{-1}(\mu) \\
&= R^{m,n}\left(R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})}\right)\right) * \mathbb{v}\right) && \text{See Eqs. (4.19) and (4.22).} \\
&\quad + \mathrm{VEC}^{-1}(\mu) \\
&= R^{m,n}\left(R^{-\frac{m}{2},-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})}\right) * \mathbb{v}\right)\right) && \text{Acc. 2D array shift operator interchange property. See Eq. (4.20).} \\
&\quad + \mathrm{VEC}^{-1}(\mu) \\
&= R^{m-\frac{m}{2},n-\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})}\right) * \mathbb{v}\right) + \mathrm{VEC}^{-1}(\mu) \\
&= R^{\frac{m}{2},\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})}\right) * \mathbb{v}\right) + \mathrm{VEC}^{-1}(\mu) \\
&= R^{\frac{m}{2},\frac{n}{2}}\left(F^{-1}\left(F\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})}\right)\right) \odot F(\mathbb{v})\right)\right) && \text{Acc. discrete version of the convolution theorem (Bracewell and Bracewell (1986)).} \\
&\quad + \mathrm{VEC}^{-1}(\mu) \\
&= R^{\frac{m}{2},\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F(\mathbb{K})} \odot F(\mathbb{v})\right)\right) + \mathrm{VEC}^{-1}(\mu) && (4.24)
\end{aligned}
$$

Note that, $F(\mathbb{v})$, in Eq. (4.24), is distributed the same as $\mathbb{v}$, i.e., we can replace $F(\mathbb{v})$ with $\mathbb{v}$. As can be observed in Eq. (4.24), in order to draw a $\mathbb{u}$ sample, we only need access to $\mathbb{K}$, $\mathbb{v}$, and $\mathrm{VEC}^{-1}(\mu)$. In other words, we do not need to explicitly compute S or $\mathbb{S}$.

4.6 Results and Discussion

In this section, we apply the proposed methods on multi-rater segmentations datasets for which a corresponding radiotherapy dose plan is available. The clinical goal is to

quantify the uncertainty in computed dose metrics (e.g., generalized equivalent uniform dose (gEUD) (Niemierko (1999))), which originates from the uncertainty in the segmentation(s).

4.6.1 Multi-rater segmentations of an esophagus

In this subsection, we perform experiments on a multi-rater segmentations dataset of an esophagus composed of 4 different rater 3D segmentations. These segmentations are shown in Figures 4.1, 4.2, and 4.3. We first estimate MLPs using the proposed

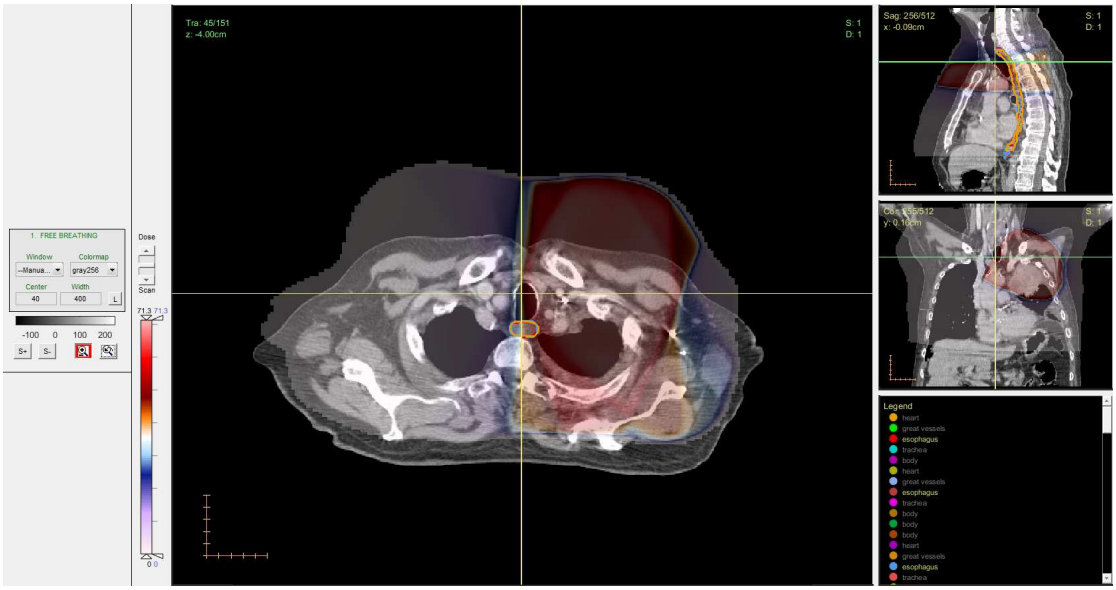

FIGURE 4.1: Multi-rater segmentations of an esophagus. (This figure has been created using the Computational Environment for Radiotherapy Research (CERR) (Deasy et al. (2003)).)

methods in Section 4.4. The α-smoothed empirical MLPs, computed using Eq. (4.5), are shown in Figure 4.4. In Figure 4.4, we can observe that, because the test dataset is relatively small ($n_{rs_*} = 4$), neighboring voxels are not segmented similarly enough. To alleviate this problem, we perform a Gaussian blur on the 3D array which holds the α-smoothed empirical MLPs according to Eq. (4.6). The resulting blurred α-smoothed empirical MLPs are shown in Figure 4.5. As can be observed in Figure 4.5, neighboring voxels are segmented much more similarly than in Figure 4.4. As it is more likely in practice to have a single rater segmentation during testing, we proceed with estimating MLPs using the Bayes' theorem-based model (see Subsection 4.4.3) and the parametric model (see Subsection 4.4.4). The input multi-rater segmentations of the esophagus, shown in Figures 4.1, 4.2, and 4.3, are used as training data for these two models. In this regard, the "true" segmentation ρ_t is set equal to the mean input multi-rater segmentation of the esophagus. The four input multi-rater segmentations of the esophagus

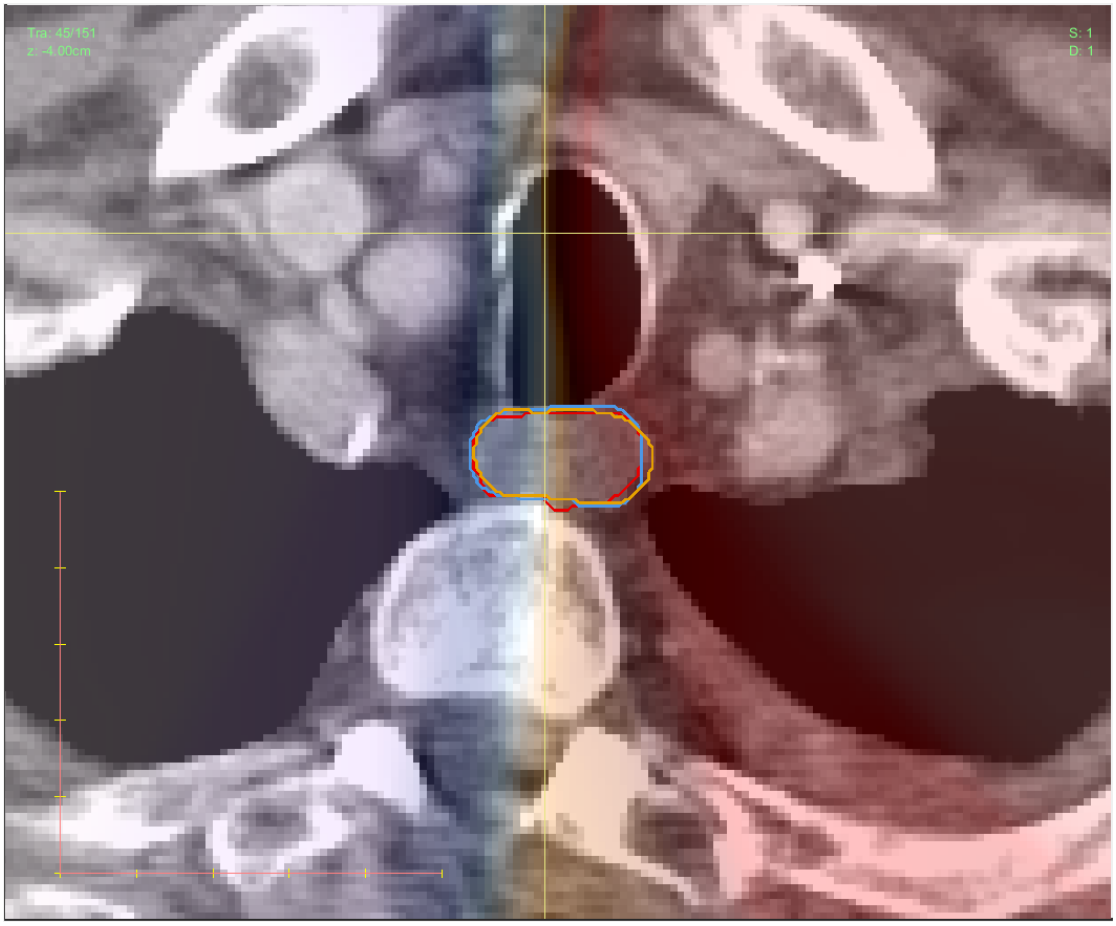

FIGURE 4.2: Multi-rater segmentations of an esophagus (zoomed in, transverse/axial slice 45). (This figure has been created using CERR (Deasy et al. (2003)).)

are used as four different training rater segmentations ρ_i. To demonstrate the different parts of the Bayes' theorem-based model (see Subsection 4.4.3), we proceed with estimating the density functions $p(D = d \mid T = 0)$ and $p(D = d \mid T = 1)$ using Eqs. (4.11) and (4.10) respectively. The corresponding results, for different values of the bandwidth η, are shown in Figures 4.6, 4.7, 4.8 4.9, 4.10, and 4.11. Having estimated the functions $p(D = d \mid T = 0)$ and $p(D = d \mid T = 1)$, we are able to estimate the density function $p(D = d)$ using Eq. (4.9). The respective result, for different values of the bandwidth η, is shown in Figures 4.12, 4.13, and 4.14 Note that $p(T = 0)$ and $p(T = 1)$ are estimated to be equal to 0.999649 and 0.000351 respectively. We then plug $p(D = d \mid T = 1)$, $p(T = 1)$, and $p(D = d)$ in Eq. (4.8) and estimate the desired function $p(T = 1 \mid D = d)$. The result, for different values of the bandwidth η, is shown in Figures (4.15), 4.16, and 4.17. In practice, we estimate $\theta(x)$ (according to Eq. (4.7)) by evaluating $p(T = 1 \mid D = d)$ at $\mathcal{D}(\rho_*, x)$. Example results are shown in Figures 4.18 and 4.19. In regard to the parametric model (see Subsection 4.4.4), we start with learning the probability distribution of the random variable D_b, which is required in

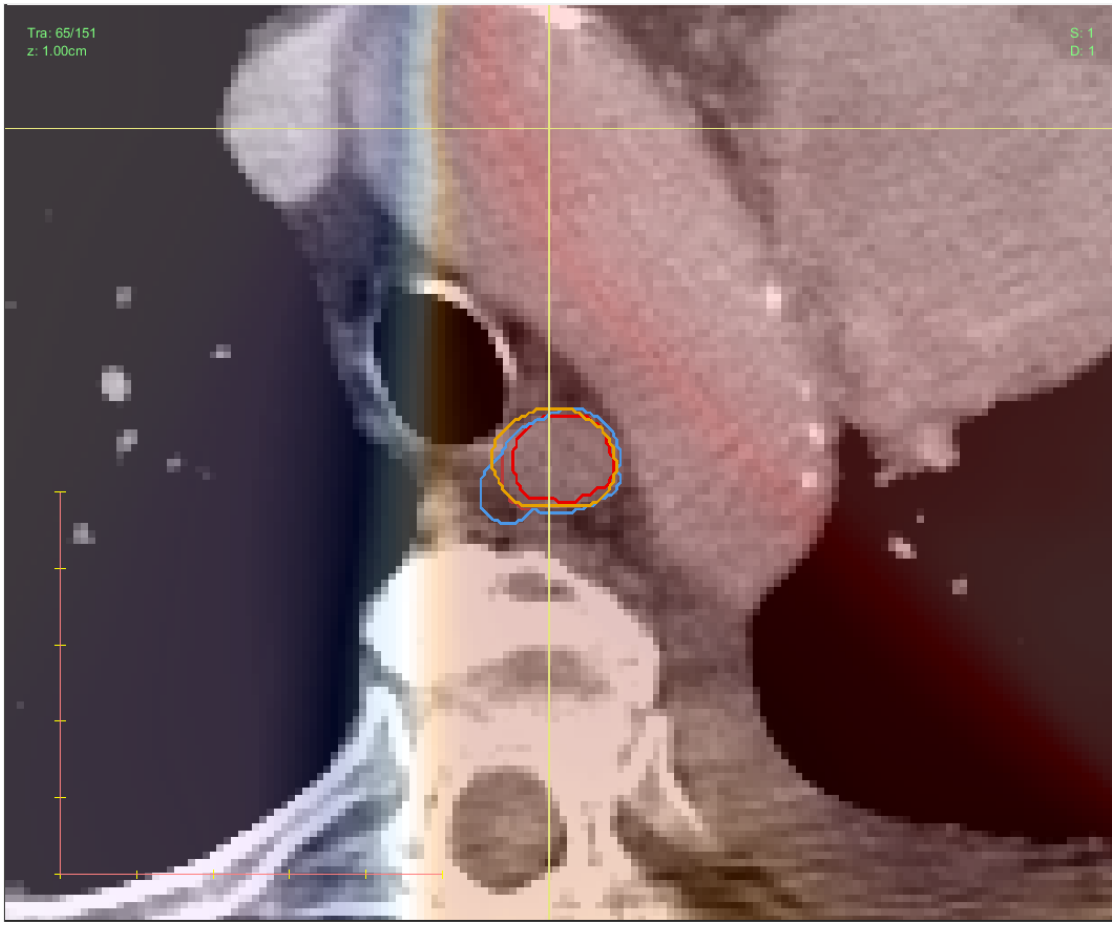

FIGURE 4.3: Multi-rater segmentations of an esophagus (zoomed in, transverse/axial slice 65). (This figure has been created using CERR (Deasy et al. (2003)).)

Eq. (4.12). All training rater segmentations ρ_i are used to generate a univariate independent and identically distributed sample from the unknown probability distribution of the random variable D_b. We then fit a normal probability distribution to this data. The cumulative distribution function of the normal distribution with the fitted mean and standard deviation and the corresponding empirical cumulative distribution function are shown in Figure 4.20. We then set the test rater segmentation ρ_* equal to ρ_t and based on the fitted cumulative distribution function of the random variable D_b, we estimate parametric model-based MLPs using Eq. (4.12). The resulting parametric model-based MLPs are shown in Figure 4.21. The rest of the results in this subsection are based on the estimated parametric model-based MLPs in Figure 4.21. In this regard, we first compute the mean function, discretized at the voxel positions in mm, using Eq. (4.2). The respective result is shown in Figure 4.22. An alternative visualization of the same cross section of the computed discretized mean function is shown in Figure 4.23. In principle, in Figures 4.22 and 4.23, we can observe a cross section of the component $\mathrm{VEC}^{-1}(\mu)$ in Eq. (4.24). The same cross section of an example of the other component $R^{\frac{m}{2},\frac{n}{2}}\left(F^{-1}\left(\sqrt[\odot]{F\left(\mathbb{K}\right)}\odot F\left(\mathbb{v}\right)\right)\right)$ in Eq. (4.24) is shown in Figure 4.24. The

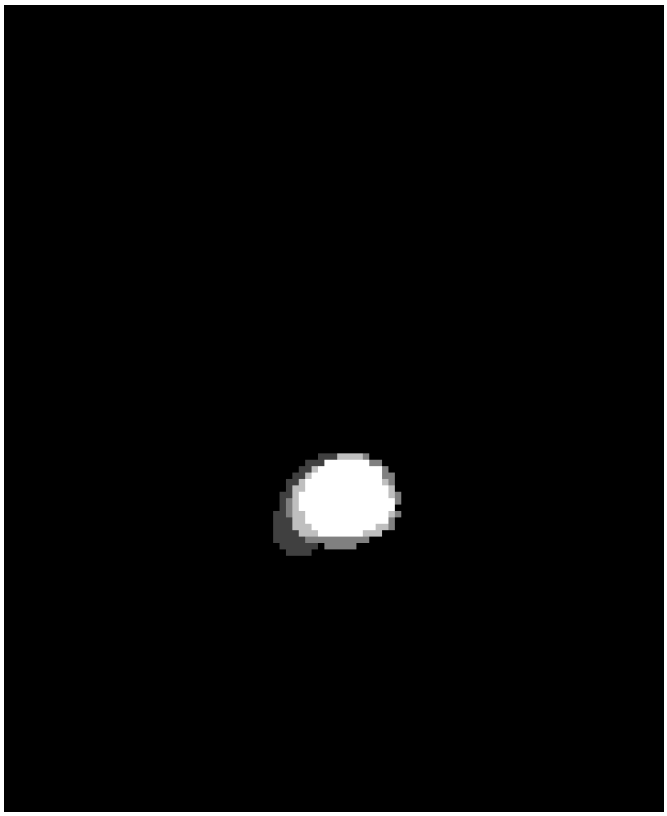

FIGURE 4.4: Esophagus α-smoothed ($\alpha = 10^{-6}$) empirical MLPs (cropped transverse/axial slice 65).

corresponding cross section of the thresholded, according to Eq. (4.1), sample $\mathfrak{u}$ (see Eq. (4.24)), corresponding to Figures 4.23 and 4.24, is shown in Figure 4.25. In simple words, Figure 4.25 shows a 2D cross section of an esophagus image segmentation sample that has been generated using our proposed approach. A larger collection of esophagus image segmentation samples, corresponding to Figure 4.23, are shown in Figures 4.26 and 4.27. We draw many such esophagus 3D image segmentation samples and for each compute the respective gEUD dose metric. An example boxplot of computed gEUD dose metrics in this way is shown in Figure 4.28. The boxplot conveys the uncertainty in the computed dose metric, which originates from the uncertainty in the segmentation. We believe that providing such information will be useful in clinical practice.

4.7 Conclusion and Future Work

We presented a novel method for automatic and efficient generation of plausible image segmentation samples that are close to a single or multiple rater segmentations or in agreement with the output label probabilities of a probabilistic segmentation. By modeling the stochastic level set function as a GP with a suitable covariance function, we were able to draw plausible zero level sets (i.e., segmentation boundaries). The novel

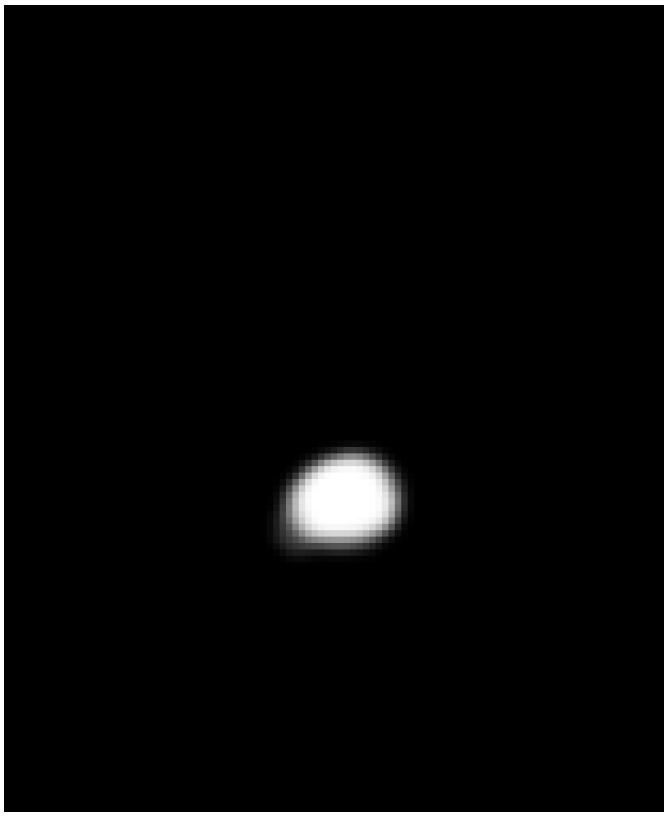

FIGURE 4.5: Esophagus blurred (with a 3D Gaussian smoothing kernel with standard deviation of 1.0) α-smoothed ($\alpha = 10^{-6}$) empirical MLPs (cropped transverse/axial slice 65).

setup of the GP mean function based on MLPs further generalized a current state-of-the-art method for image segmentation sampling and hence broadened the applicability of the proposed method. The proposed novel methods for estimating MLPs based on a single or multiple (e.g., rater) input segmentations allowed to apply the proposed method also in cases when a probabilistic segmentation output is not available. By deploying process convolution and performing the blur of 2D/3D white noise images in Fourier space, we were able to efficiently generate image segmentation samples. The application of the devised methods on multi-rater segmentation datasets for which a corresponding radiotherapy dose plan was available revealed that there is significant uncertainty in the computed dose metrics which we believe may be useful information if considered in clinical practice.

In future work, we plan to explore more sophisticated algorithms (Warfield et al. (2004)) for estimating the true segmentation based on a collection of segmentations generated by human raters or automated segmentation algorithms and analyze the impact on our estimated MLPs. In this regard, we would also like to compare our MLPs with those estimated by the STAPLE algorithm (Warfield et al. (2004)) for example. Furthermore, we would like to perform a larger study on multi-rater segmentation datasets of various

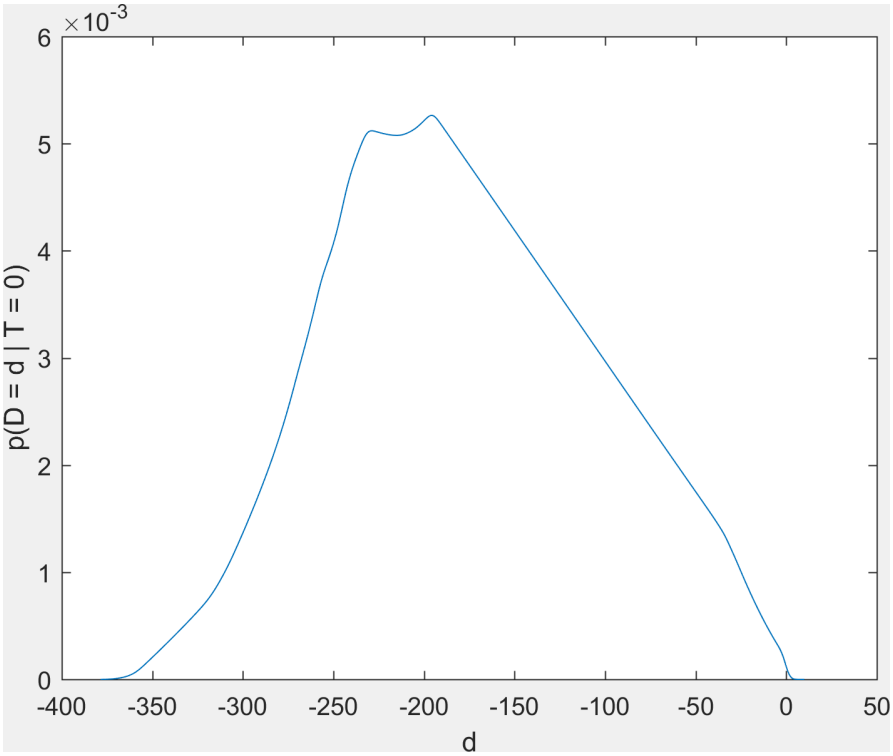

FIGURE 4.6: The graph of the density function $p(D = d \mid T = 0)$ with the bandwidth η in Eq. (4.11) being set to 1.9.

organs and tumors and estimate the uncertainty in numerous dose metrics and analyze thoroughly how this information may influence the outcome of the radiotherapy in clinical practice.

Acknowledgments

This work was supported by the DFG Creative Unit "Intra-Operative Information: What Surgeons Need, When They Need It" under grant number M5 CU IOI, and the NIH grants P41 EB015902, U24 CA180918, and P41 EB015898.

Appendix 4.A Affine Transformation of a Multivariate Normal Random Vector

If one applies an affine transformation on a multivariate normal random vector, the resulting random vector has also a multivariate normal distribution. This is formally defined in the proposition below:

Proposition 4.A.1. Let $\mathbf{y}$ be a $p \times 1$ multivariate normal random vector with mean $\overline{\mathbf{y}}$ and a covariance matrix Σ, i.e., $\mathbf{y} \sim \mathcal{N}(\overline{\mathbf{y}}, \Sigma)$. Let $\mathbf{A}$ be a $p' \times p$ real matrix of rank p',

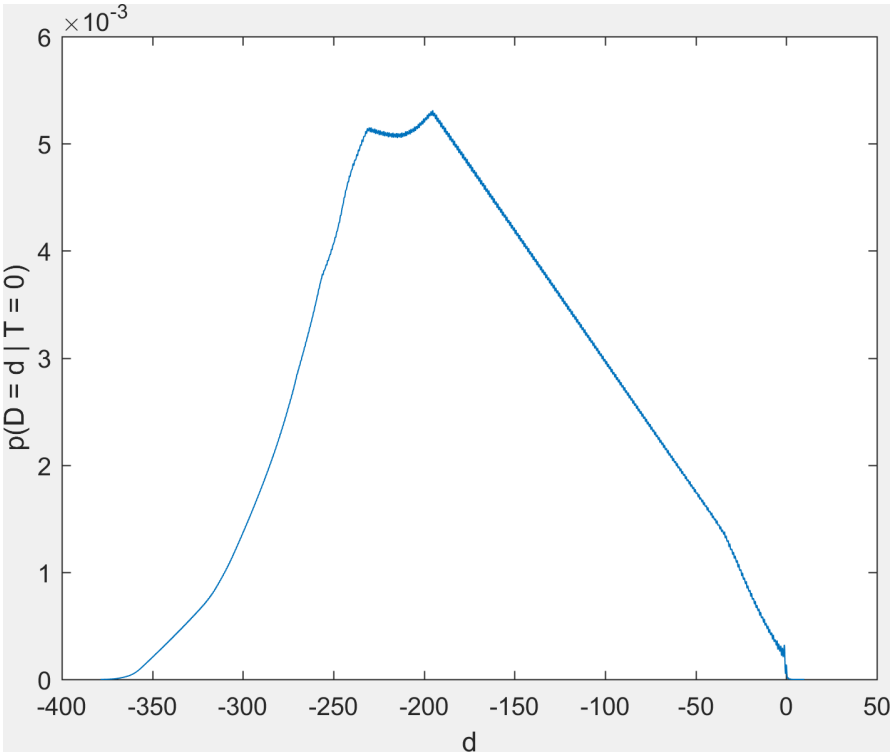

FIGURE 4.7: The graph of the density function $p(D = d \mid T = 0)$ with the bandwidth η in Eq. (4.11) being set to 0.25.

where $p' \leq p$, and $\mathbf{b}$ be a $p' \times 1$ real vector. Then, the affine transformation resultant $p' \times 1$ random vector $\mathbf{z} = \mathbf{A}\mathbf{y} + \mathbf{b}$ has a multivariate normal distribution with a mean:

$$\mathbb{E}[\mathbf{z}] = \mathbf{A}\overline{\mathbf{y}} + \mathbf{b} \tag{4.A.1.1}$$

and a covariance matrix:

$$\text{cov}(\mathbf{z}) = \mathbf{A}\Sigma\mathbf{A}^T \tag{4.A.1.2}$$

For more details, we refer the reader to the book by Rencher (2003).

Appendix 4.B Equivalence between K and $\mathbb{K}$

Let c be a 1D array of the (world) positions (in mm) of the pixels/voxels of the 2D/3D image that is segmented. Note that, $c = \text{VEC}(\mathbb{c})$, where $\mathbb{c}$ is a 2D/3D array of the (world) positions (in mm) of the pixels/voxels of the 2D/3D image that is segmented. In 2D, the position in mm of the pixel with coordinates (u, v), i.e., $\mathbb{c}_{u,v}$, is defined as follows:

$$\mathbb{c}_{u,v} = \mathbf{M}\begin{bmatrix} u \\ v \end{bmatrix} + \mathbf{t}\ , \tag{4.B.1}$$

where $\mathbf{M}$ is linear transformation and $\mathbf{t}$ is a translation/displacement vector. Next, we present the proof, in 2D, about the equivalence between a matrix that is constructed by

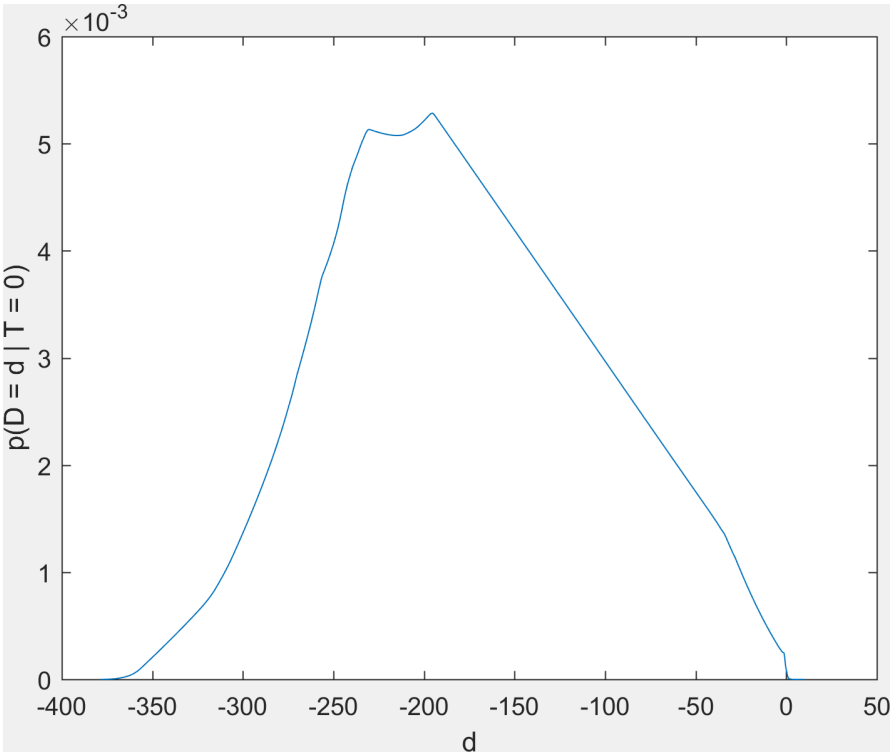

FIGURE 4.8: The graph of the density function $p(D = d \mid T = 0)$ with the bandwidth η in Eq. (4.11) being set to 0.5.

evaluating a stationary (covariance) function, e.g., $k(x, x')$, at the index locations, i.e., e.g., the matrix K, and a 2D array whose entries hold the values of the same stationary (covariance) function evaluated at the positions (in mm) of the respective entry and the

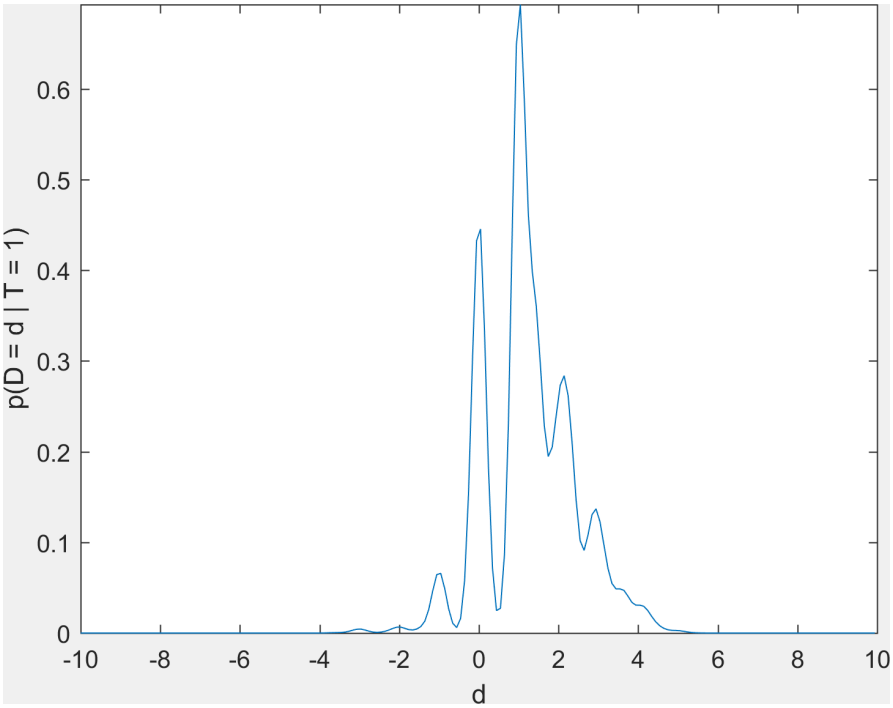

FIGURE 4.9: The graph of the density function $p(D = d \mid T = 1)$ with the bandwidth η in Eq. (4.10) being set to 0.18.

center of the 2D array, e.g., the 2D array $\mathbb{K}$:

$$
\begin{aligned}
K_{ij} &= k\left(c_i, c_j\right) \\
&= k\left(\mathbb{c}_{h^{-1}(i)}, \mathbb{c}_{h^{-1}(j)}\right) && \text{Let } (q,r) = h^{-1}(i). \\
& && \text{Let } (s,t) = h^{-1}(j). \\
&= k\left(\mathbb{c}_{q,r}, \mathbb{c}_{s,t}\right) \\
&= k\left(\mathbb{c}_{q,r} - \mathbb{c}_{s,t} + \mathbb{c}_{m,n}, \mathbb{c}_{s,t}\right. && \text{Here, we use the property} \\
&\quad \left. - \mathbb{c}_{s,t} + \mathbb{c}_{m,n}\right) && \text{that the deployed covariance} \\
& && \text{function is stationary.} \\
&= k\left(\mathbf{M}\begin{bmatrix} q \\ r \end{bmatrix} + \mathbf{t} - \mathbf{M}\begin{bmatrix} s \\ t \end{bmatrix}\right. && \text{See Eq. (4.B.1).} \\
&\quad \left. - \mathbf{t} + \mathbf{M}\begin{bmatrix} m \\ n \end{bmatrix} + \mathbf{t}, \mathbb{c}_{m,n}\right) \\
&= k\left(\mathbf{M}\begin{bmatrix} q-s+m \\ r-t+n \end{bmatrix} + \mathbf{t}, \mathbb{c}_{m,n}\right) && \text{Use the fact that } \mathbf{M} \text{ is a} \\
& && \text{linear map.} \\
&= k\left(\mathbb{c}_{q-s+m,r-t+n}, \mathbb{c}_{m,n}\right) && \text{See Eq. (4.B.1).} \\
&= \mathbb{K}_{q-s+m,r-t+n} && \text{By definition of } \mathbb{K}. \text{ See Eq. (4.17).} \quad (4.B.2) \\
&= \mathrm{MAT}(\mathbb{K})_{i,j} && \text{Acc. definition of MAT}(\mathbb{A}). \quad (4.B.3)
\end{aligned}
$$

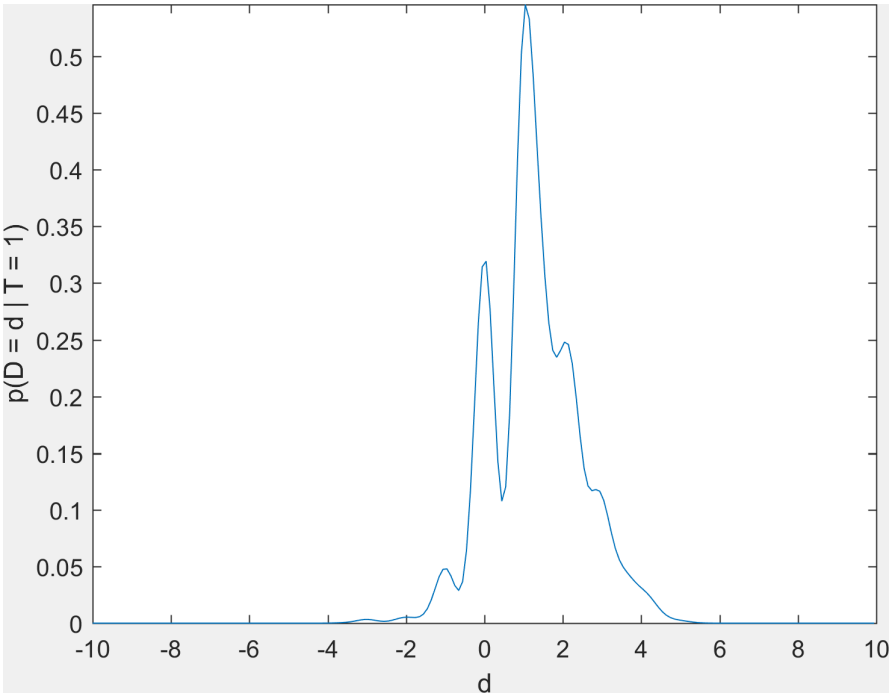

FIGURE 4.10: The graph of the density function $p(D = d \mid T = 1)$ with the bandwidth η in Eq. (4.10) being set to 0.25.

This completes the proof about the equivalence between K and $\mathbb{K}$. Note that the same equivalence holds for S and $\mathbb{S}$, i.e., $S = \text{MAT}(\mathbb{S})$.

Appendix 4.C Equivalence between Matrix Multiplication and Convolution

In this section, we first provide the proof, in 2D, about the equivalence between matrix-vector multiplication with a matrix that is constructed by evaluating a stationary (covariance) function, e.g., $k(x, x')$, at the index locations, i.e., e.g., with the matrix K, and a 2D array convolution with a 2D array convolution kernel whose entries hold the values of the same stationary (covariance) function evaluated at the positions (in mm) of the respective entry and the center of the 2D array, e.g., with the 2D array $\mathbb{K}$ as a convolution kernel. In this regard, the convolution of a 2D array $\mathbb{A}$ with another 2D array $\mathbb{B}$ results in another 2D array $\mathbb{C} = \mathbb{A} * \mathbb{B}$ with its entries defined as follows:

$$\mathbb{C}_{i,j} = [\mathbb{A} * \mathbb{B}]_{i,j} = \sum_{u,v} \mathbb{A}_{u,v} \mathbb{B}_{i-u,j-v} \,, \tag{4.C.1}$$

where u and v range over all legal subscripts for $\mathbb{A}_{u,v}$ and $\mathbb{B}_{i-u,j-v}$. In fact, we provide the proof of Eq. (4.16). However, the same equivalence holds also between a matrix-vector multiplication with the matrix S and a 2D array convolution with the 2D array

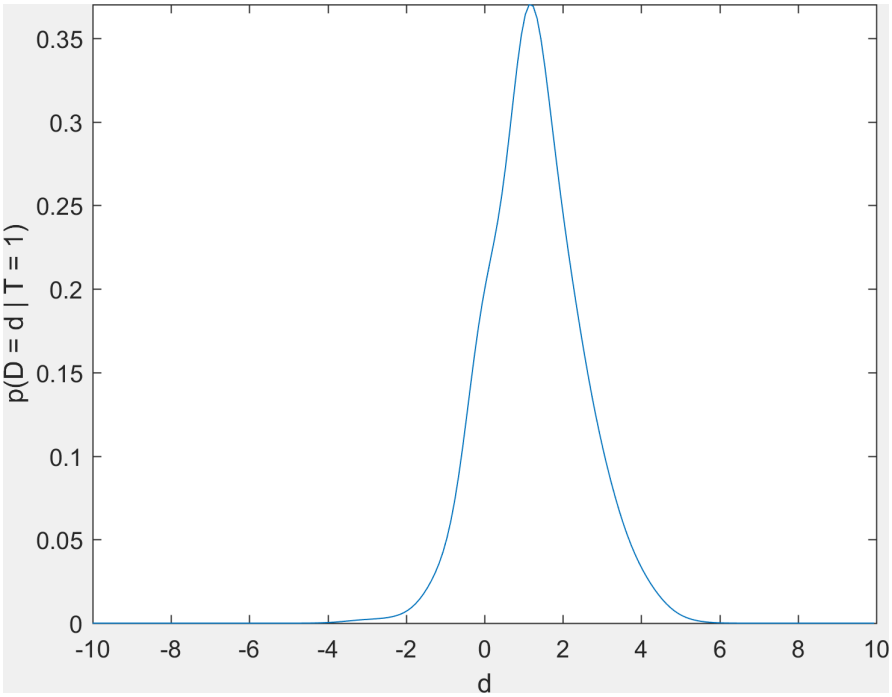

FIGURE 4.11: The graph of the density function $p(D = d \mid T = 1)$ with the bandwidth η in Eq. (4.10) being set to 0.5.

$\mathbb{S}$ as a convolution kernel. The proof of Eq. (4.16) is then as follows:

$$
\begin{aligned}
[Kv]_i &= \sum_j K_{ij} v_j \\
&= \sum_j k\left(c_i, c_j\right) v_j \\
&= \sum_{h^{-1}(j)} k\left(\mathbb{c}_{h^{-1}(i)}, \mathbb{c}_{h^{-1}(j)}\right) \mathbb{v}_{h^{-1}(j)} \\
&\qquad\qquad \text{Let } (q, r) = h^{-1}(i). \\
&\qquad\qquad \text{Let } (s, t) = h^{-1}(j). \\
&= \sum_{s,t} k\left(\mathbb{c}_{q,r}, \mathbb{c}_{s,t}\right) \mathbb{v}_{s,t} \\
&= \sum_{s,t} k\left(\mathbb{c}_{q,r} - \mathbb{c}_{s,t} + \mathbb{c}_{m,n}, \mathbb{c}_{s,t}\right. \qquad \text{Here, we use the property} \\
&\qquad \left. - \mathbb{c}_{s,t} + \mathbb{c}_{m,n}\right) \mathbb{v}_{s,t} \qquad \text{that the deployed covariance} \\
&\qquad\qquad \text{function is stationary.} \\
&= \sum_{s,t} k\left(\mathbf{M}\begin{bmatrix} q \\ r \end{bmatrix} + \mathbf{t} - \mathbf{M}\begin{bmatrix} s \\ t \end{bmatrix}\right. \qquad \text{See Eq. (4.B.1).} \\
&\qquad \left. - \mathbf{t} + \mathbf{M}\begin{bmatrix} m \\ n \end{bmatrix} + \mathbf{t}, \mathbb{c}_{m,n}\right) \mathbb{v}_{s,t}
\end{aligned}
$$

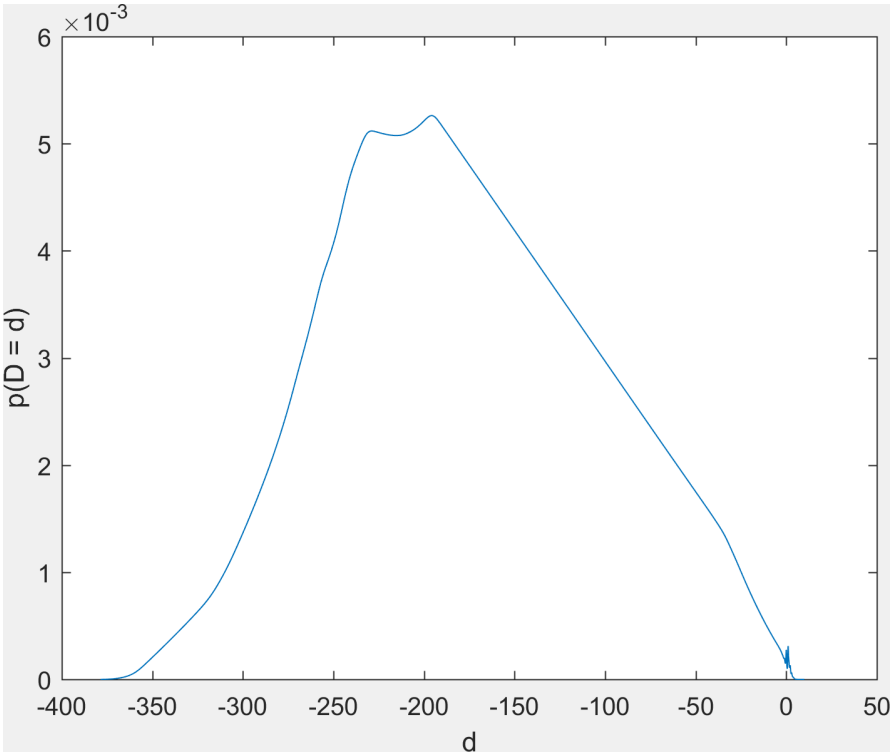

FIGURE 4.12: The graph of the density function $p(D = d)$ with the bandwidth η in Eqs. (4.10) and (4.11) being set to 0.18 and 1.9 respectively.

$$= \sum_{s,t} k\left(\mathbf{M}\begin{bmatrix} q-s+m \\ r-t+n \end{bmatrix} + \mathbf{t}, \mathbb{c}_{m,n}\right) \mathbb{v}_{s,t} \quad \text{Use the fact that } \mathbf{M} \text{ is a linear map.}$$

$$= \sum_{s,t} k\left(\mathbb{c}_{q-s+m,r-t+n}, \mathbb{c}_{m,n}\right) \mathbb{v}_{s,t} \quad \text{See Eq. (4.B.1).}$$

$$= \sum_{s,t} \mathbb{K}_{q-s+m,r-t+n} \mathbb{v}_{s,t} \quad \text{By definition of } \mathbb{K}. \text{ See Eq. (4.17).}$$

$$= \sum_{s,t} \mathbb{v}_{s,t} \mathbb{K}_{q+m-s,r+n-t} \quad \text{Rearranging.}$$

$$= [\mathbb{v} * \mathbb{K}]_{q+m,r+n} \quad \text{Acc. definition of 2D array convolution. See Eq. (4.C.1).}$$

$$= [\mathbb{K} * \mathbb{v}]_{q+m,r+n} \quad \text{Acc. commutativity property of convolution.}$$

$$= \left[R^{m,n}\left(\mathbb{K} * \mathbb{v}\right)\right]_{q,r} \quad \text{Acc. definition of the 2D array shift operator } R^{a,b}\left(\mathbb{A}\right).$$

$$= \text{VEC}(R^{m,n}\left(\mathbb{K} * \mathbb{v}\right))_{h(q,r)} \quad \text{Acc. definition of the operator VEC}(\mathbb{A}).$$

$$= \text{VEC}(R^{m,n}\left(\mathbb{K} * \mathbb{v}\right))_i \tag{4.C.2}$$

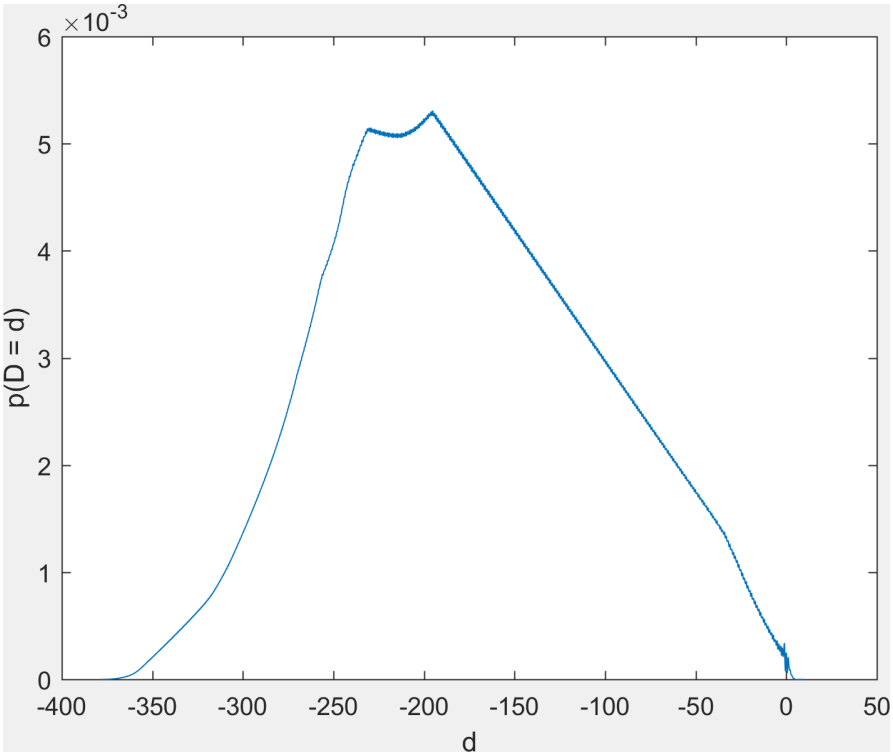

FIGURE 4.13: The graph of the density function $p(D = d)$ with the bandwidth η in both Eq. (4.10) and Eq. (4.11) being set to 0.25.

This completes the proof of Eq. (4.16). Next, we provide the proof, in 2D, about the equivalence between matrix-matrix-vector multiplication with a matrix that is constructed by evaluating a stationary (covariance) function, e.g., $k(x, x')$, at the index locations, i.e., e.g., with the matrix K, and a 2D array double convolution with a 2D array convolution kernel whose entries hold the values of the same stationary (covariance) function evaluated at the positions (in mm) of the respective entry and the center

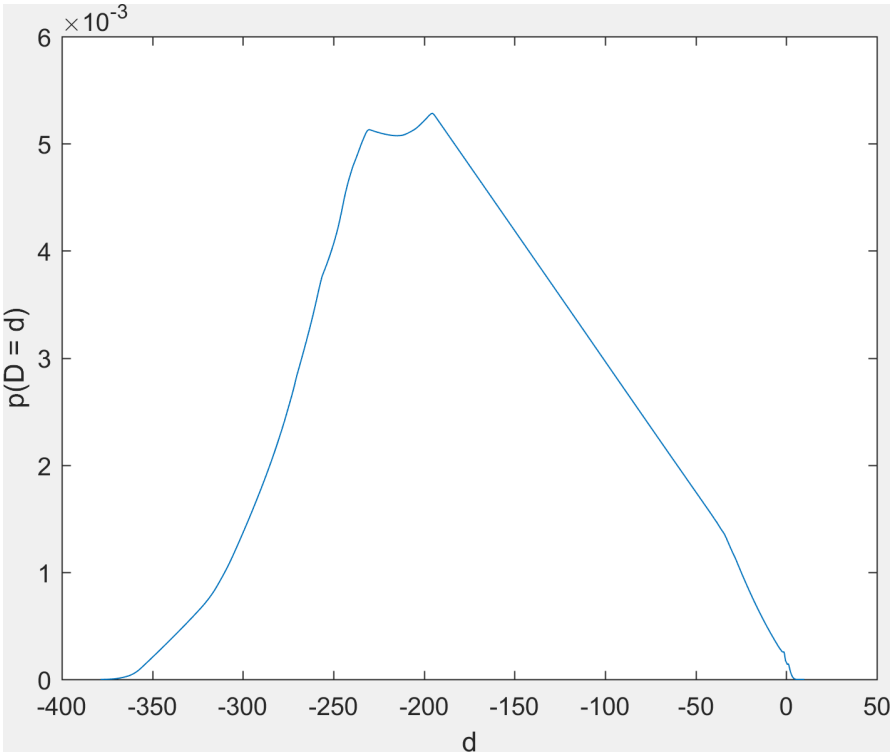

FIGURE 4.14: The graph of the density function $p(D = d)$ with the bandwidth η in both Eq. (4.10) and Eq. (4.11) being set to 0.5.

of the 2D array, e.g., with the 2D array $\mathbb{K}$ as a convolution kernel:

$$
\begin{aligned}
[KKv]_i &= [Kw]_i && \text{Let } w = Kv. \\
&= \mathrm{VEC}(R^{m,n}\left(\mathbb{K} * \mathbb{w}\right))_i && \text{Follows directly from Eq. (4.C.2).} \\
& && \text{Note that, } \mathbb{w} = \mathrm{VEC}^{-1}(w), \\
& && \text{analogous to } \mathbb{v}. \\
&= \mathrm{VEC}(R^{m,n}\left(\mathbb{K} * R^{m,n}\left(\mathbb{K} * \mathbb{v}\right)\right))_i && \text{This holds because} \\
& && \mathbb{w} = \mathrm{VEC}^{-1}(w) = \mathrm{VEC}^{-1}(Kv) = \\
& && \mathrm{VEC}^{-1}(\mathrm{VEC}(R^{m,n}\left(\mathbb{K} * \mathbb{v}\right))) = \\
& && R^{m,n}\left(\mathbb{K} * \mathbb{v}\right) \\
&= \mathrm{VEC}(R^{m,n}\left(R^{m,n}\left(\mathbb{K} * \mathbb{K} * \mathbb{v}\right)\right))_i && \text{This follows from the interchange} \\
& && \text{property of the 2D array shift} \\
& && \text{operator. See Eq (4.20).}
\end{aligned}
\tag{4.C.3}
$$

Note that the same equivalence holds also between a matrix-matrix-vector multiplication with the matrix S and a 2D array double convolution with the 2D array $\mathbb{S}$ as a convolution kernel. I.e., the above proof also serves as a proof of Eq. (4.15).

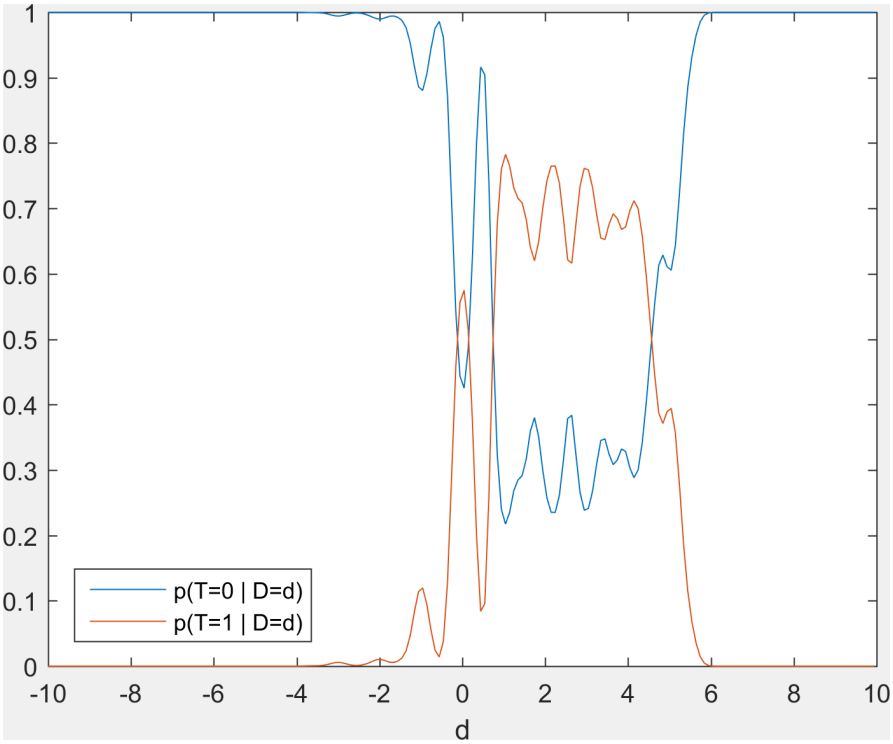

FIGURE 4.15: The graph of the functions $p(T = 1 \mid D = d)$ and $p(T = 0 \mid D = d)$ with the bandwidth η in Eqs. (4.10) and (4.11) being set to 0.18 and 1.9 respectively.

Appendix 4.D 2D Array Shift Operator Interchange Property

In this section, we provide a proof of Eq. (4.20), i.e., a proof of the interchange property of the 2D array shift operator. Let $\mathbb{A}$ and $\mathbb{B}$ be two 2D arrays. Then, the proof is as follows:

$$\left[R^{a,b}\left(\mathbb{A}\right) * \mathbb{B}\right]_{i,j} \stackrel{?}{=} \left[R^{a,b}\left(\mathbb{A} * \mathbb{B}\right)\right]_{i,j} \iff \tag{4.D.1}$$

$$\sum_{u,v}\left[R^{a,b}\left(\mathbb{A}\right)\right]_{u,v}\mathbb{B}_{i-u,j-v} \stackrel{?}{=} \left[\mathbb{A} * \mathbb{B}\right]_{i+a,j+b} \iff \tag{4.D.2}$$

$$\sum_{u,v}\mathbb{A}_{u+a,v+b}\mathbb{B}_{i-u,j-v} \stackrel{?}{=} \sum_{u,v}\mathbb{A}_{u,v}\mathbb{B}_{i+a-u,j+b-v} \iff \tag{4.D.3}$$

$$\sum_{u,v}\mathbb{A}_{u+a,v+b}\mathbb{B}_{i-u,j-v} \stackrel{?}{=} \sum_{u,v}\mathbb{A}_{u+a,v+b}\mathbb{B}_{i+a-(u+a),j+b-(v+b)} \iff \tag{4.D.4}$$

$$\sum_{u,v}\mathbb{A}_{u+a,v+b}\mathbb{B}_{i-u,j-v} = \sum_{u,v}\mathbb{A}_{u+a,v+b}\mathbb{B}_{i-u,j-v} \tag{4.D.5}$$

This completes the proof of Eq. (4.20). Note that the left-hand sides of Eqs. (4.D.1) and (4.D.2) are equal according to Eq. (4.C.1), while their right-hand sides are equal according to the defintion of the 2D array shift operator $R^{a,b}\left(\mathbb{A}\right)$. In an analogous fashion, we make the transition from Eq. (4.D.2) to Eq. (4.D.3). Furthermore, the

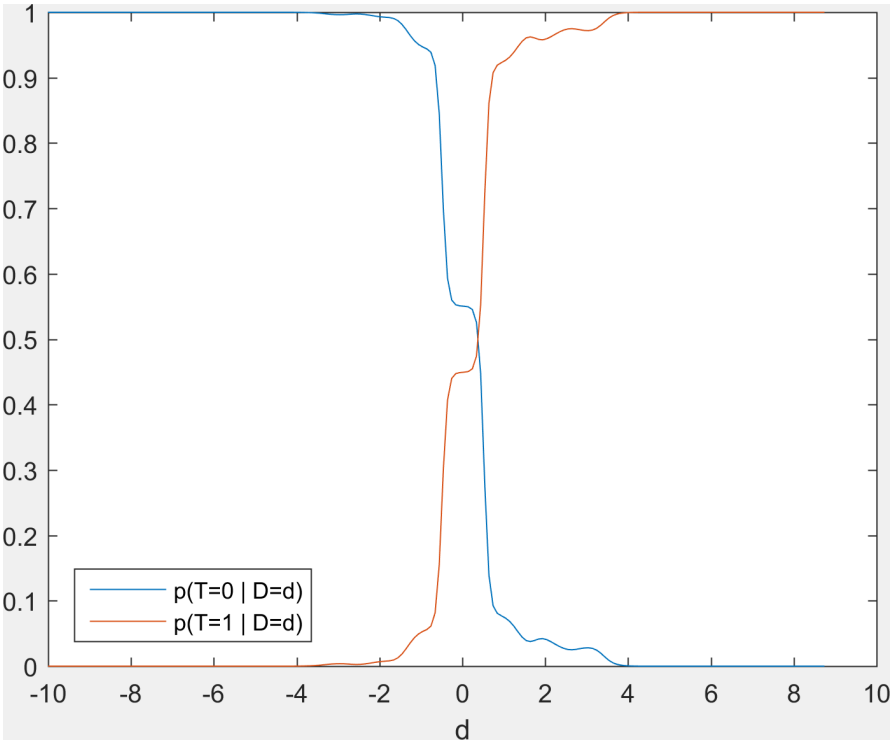

FIGURE 4.16: The graph of the functions $p(T = 1 \mid D = d)$ and $p(T = 0 \mid D = d)$ with the bandwidth η in both Eq. (4.10) and Eq. (4.11) being set to 0.25.

substitutions $u \leftarrow u + a$ and $v \leftarrow v + b$ on the right-hand side of Eq. (4.D.4) are possible because u and v in the sum on the right-hand side of Eq. (4.D.3) range over all legal subscripts for $\mathbb{A}_{u,v}$ and $\mathbb{B}_{i+a-u,j+b-v}$. Note that also the following holds:

$$
\begin{aligned}
R^{a,b}(\mathbb{A}) * \mathbb{B} &= R^{a,b}(\mathbb{A} * \mathbb{B}) && \text{See Eq. (4.20).} \\
&= R^{a,b}(\mathbb{B} * \mathbb{A}) && \text{Acc. commutativity property} \\
& && \text{of convolution.} \\
&= R^{a,b}(\mathbb{B}) * \mathbb{A} && \text{See Eq. (4.20).} \\
&= \mathbb{A} * R^{a,b}(\mathbb{B}) && \text{Acc. commutativity property} \\
& && \text{of convolution.}
\end{aligned} \tag{4.D.6}
$$

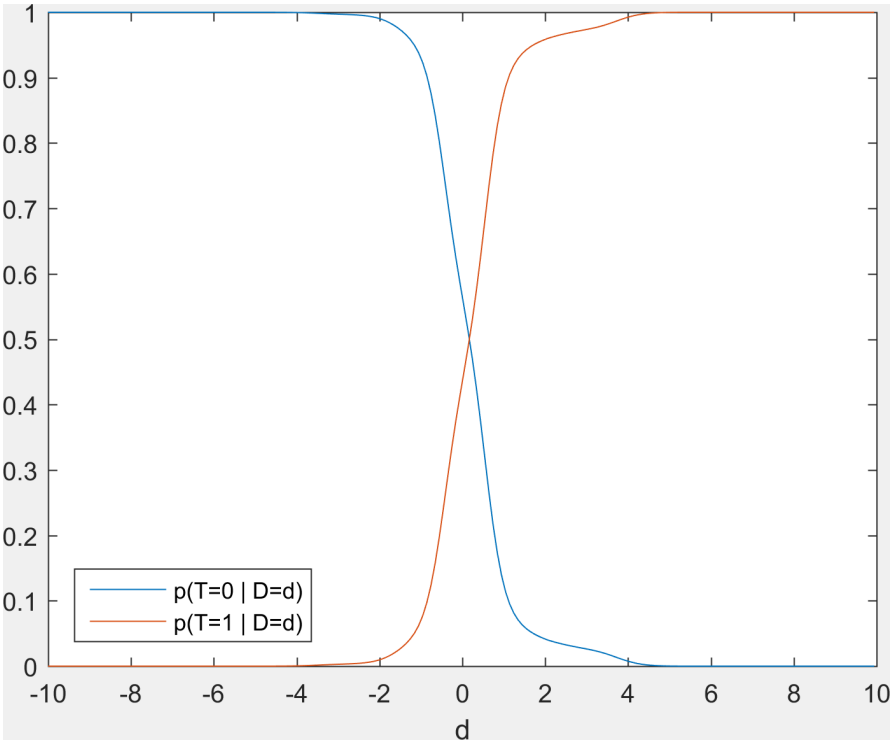

FIGURE 4.17: The graph of the functions $p(T = 1 \mid D = d)$ and $p(T = 0 \mid D = d)$ with the bandwidth η in both Eq. (4.10) and Eq. (4.11) being set to 0.5.

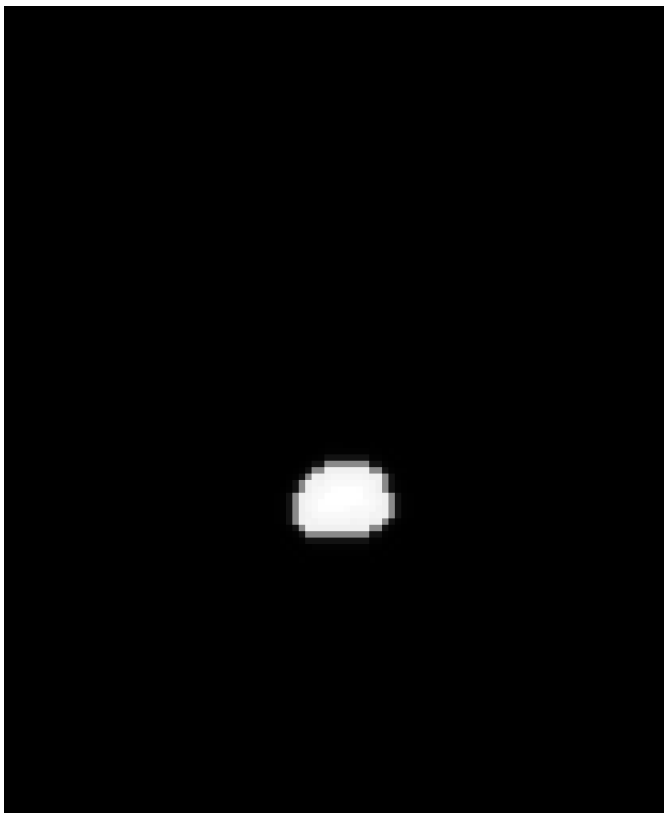

FIGURE 4.18: Esophagus Bayesian MLPs (cropped transverse/axial slice 65) with the bandwidth η in Eqs. (4.10) and (4.11) being set to 0.18 and 0.25 respectively.

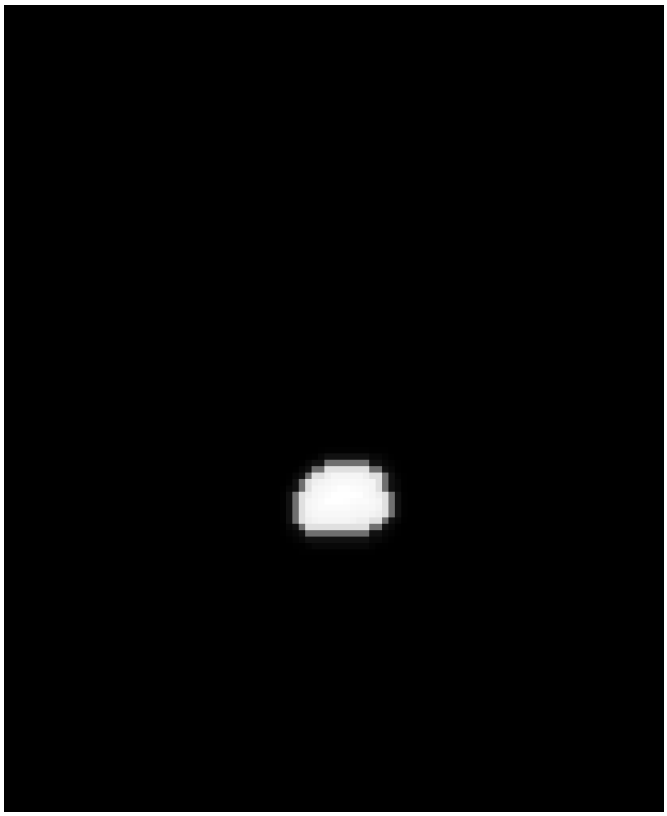

FIGURE 4.19: Esophagus Bayesian MLPs (cropped transverse/axial slice 65) with the bandwidth η in both Eqs. (4.10) and (4.11) being set to 0.5.

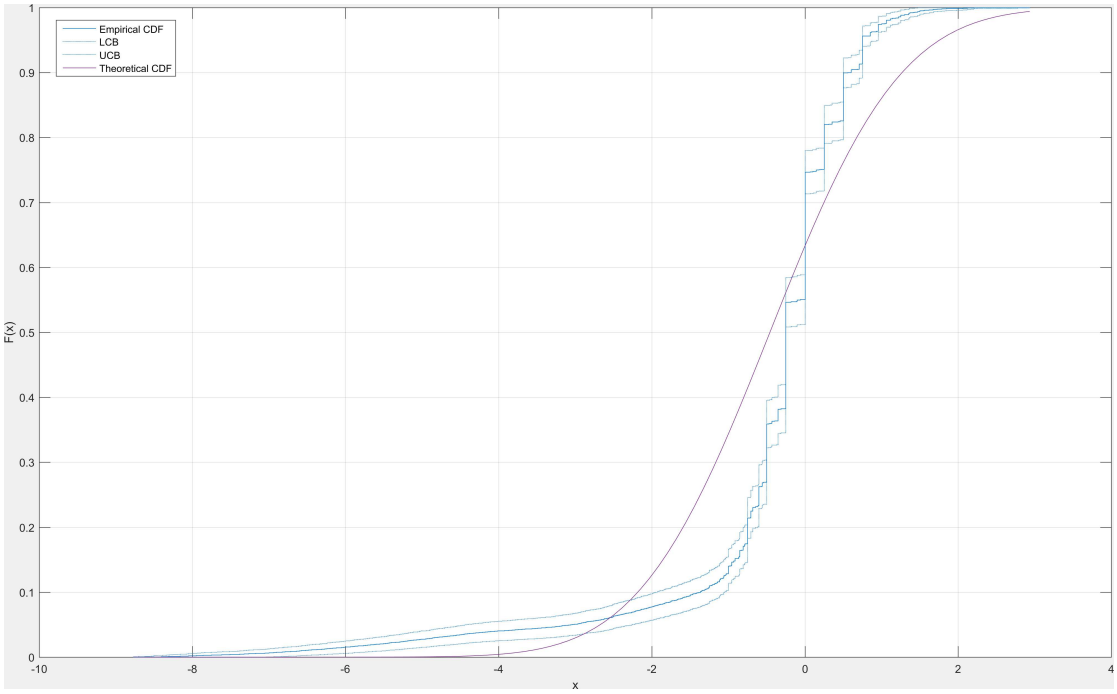

FIGURE 4.20: Empirical cumulative distribution function and the corresponding fitted cumulative distribution function of the normal distribution to the sample from the unknown probability distribution of the random variable D_b.

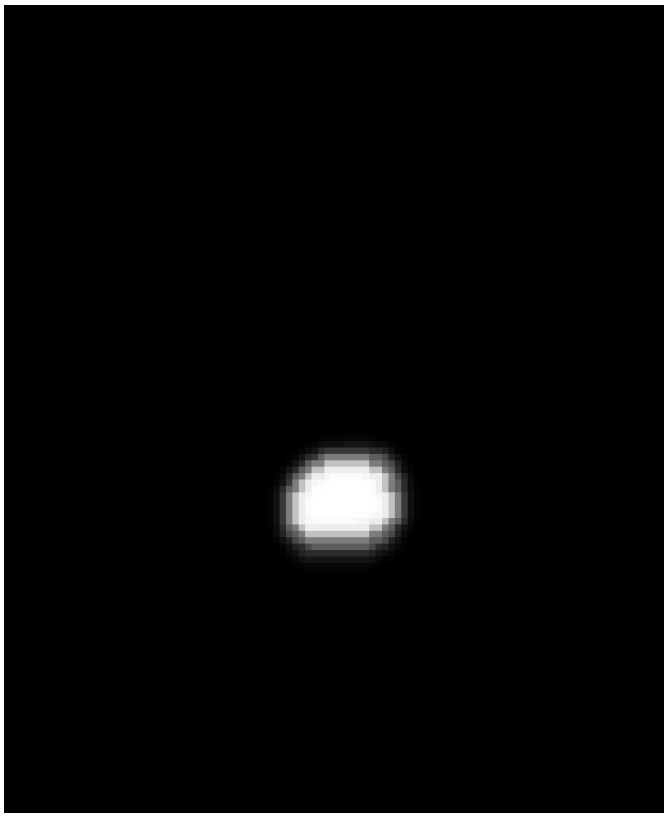

FIGURE 4.21: Esophagus parametric model-based MLPs (cropped transverse/axial slice 65).

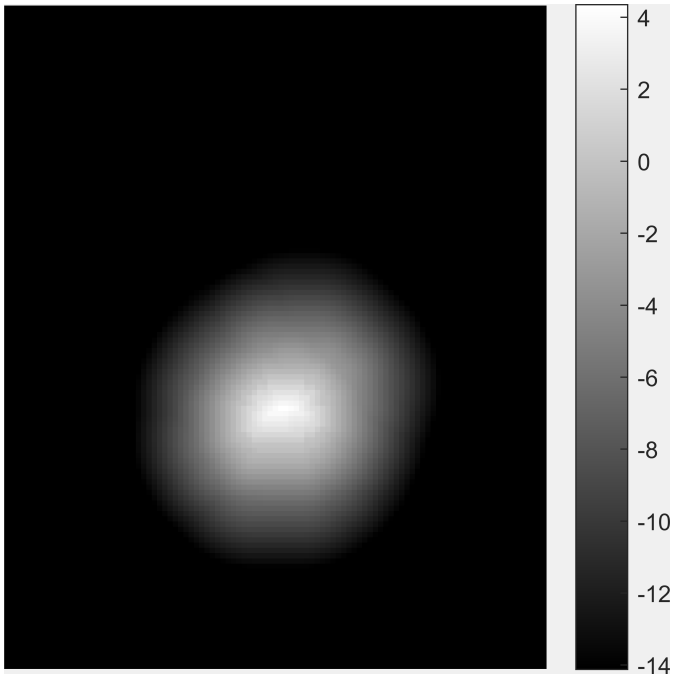

FIGURE 4.22: Esophagus parametric model-based MLPs-based mean function (cropped transverse/axial slice 65). Note that for visualization purposes the mean function values that are equal to $\pm\infty$ are replaced with the maximum/minimum finite mean function value.

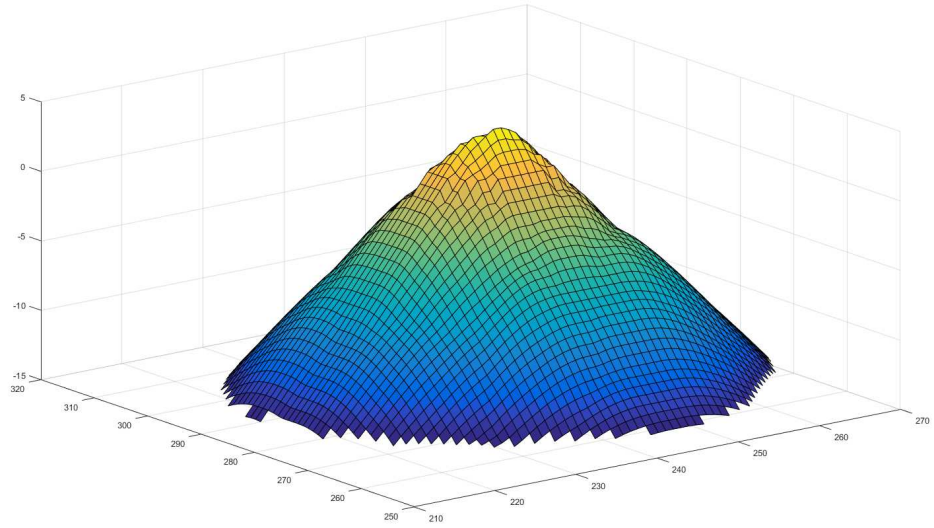

FIGURE 4.23: Alternative visualization of the esophagus parametric model-based MLPs-based mean function (transverse/axial slice 65). This is a cross section of the component $\text{VEC}^{-1}(\mu)$ in Eq. (4.24). Note that for visualization purposes the mean function values that are equal to $\pm\infty$ are ignored.

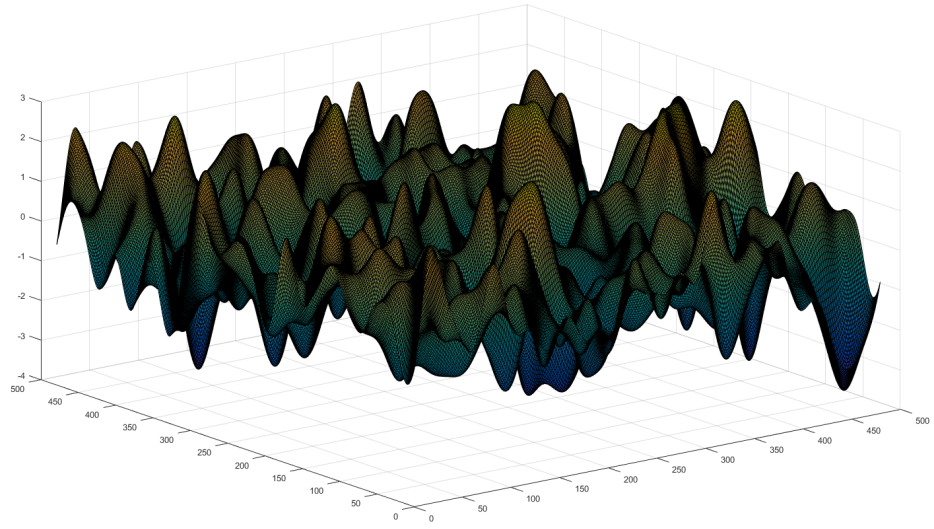

FIGURE 4.24: Cross section (transverse/axial slice 65) of an example of the component $R^{\frac{m}{2},\frac{n}{2}}\left(F^{-1}\left(\sqrt[\circledcirc]{F\left(\mathbb{K}\right)}\odot F\left(\mathbb{v}\right)\right)\right)$ in Eq. (4.24). In this example, the characteristic length-scale (hyper-) parameter l is set equal to 15.

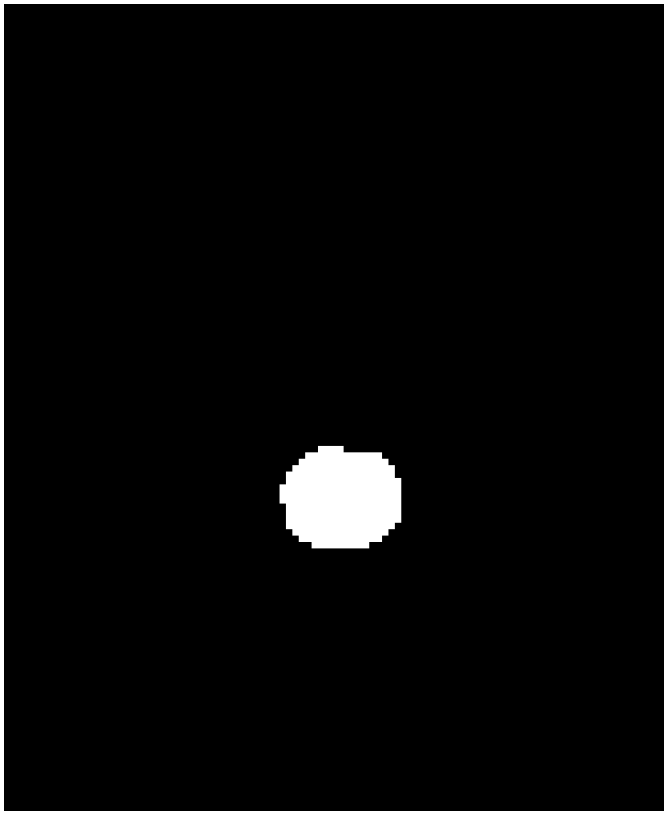

FIGURE 4.25: Cross section (transverse/axial slice 65) of the thresholded sample $\mathfrak{u}$ (see Eq. (4.24)), corresponding to Figures 4.23 and 4.24.

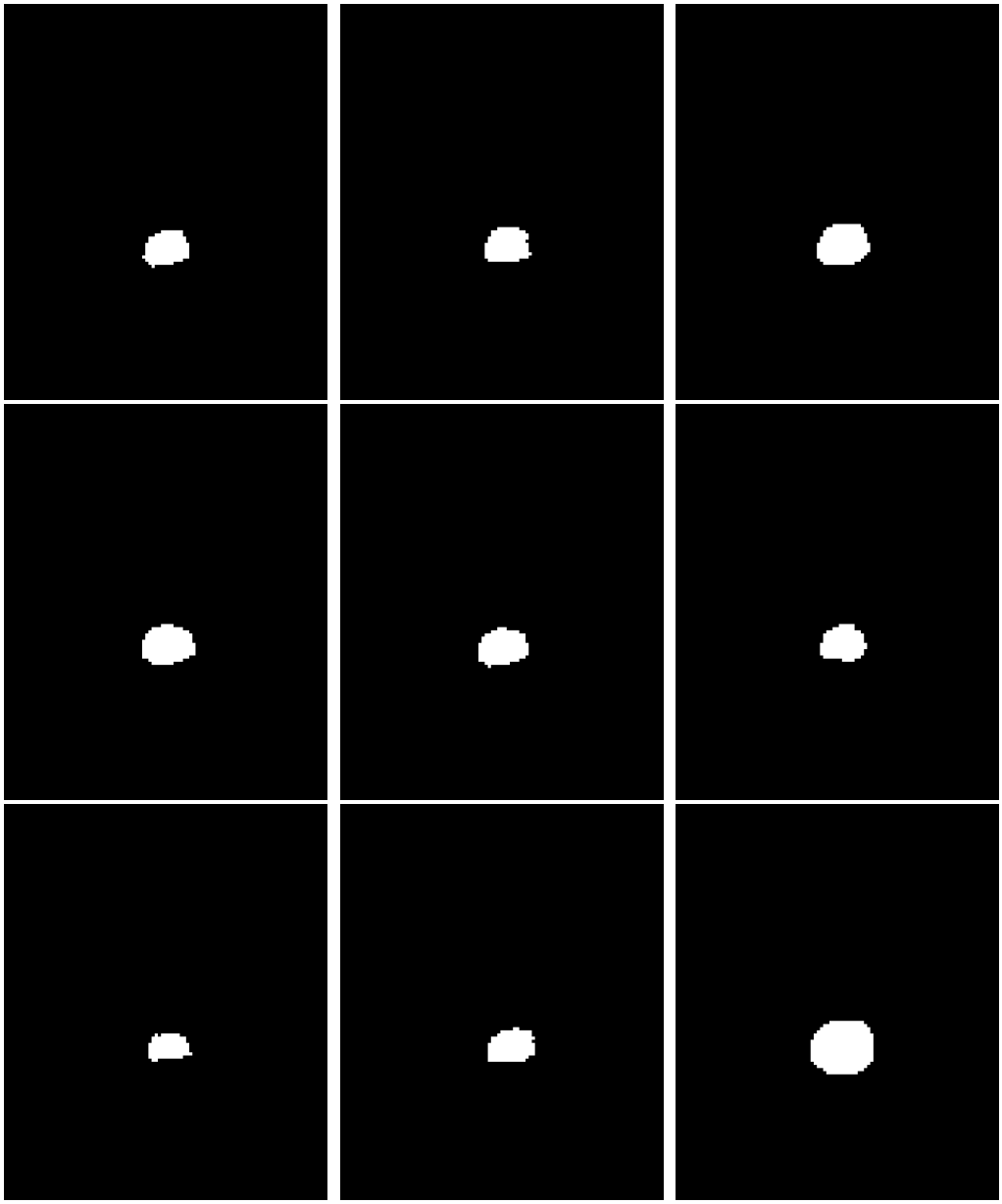

FIGURE 4.26: Esophagus image segmentation samples with their (transverse/axial slice 65) 2D cross section being shown, corresponding to Figure 4.23.

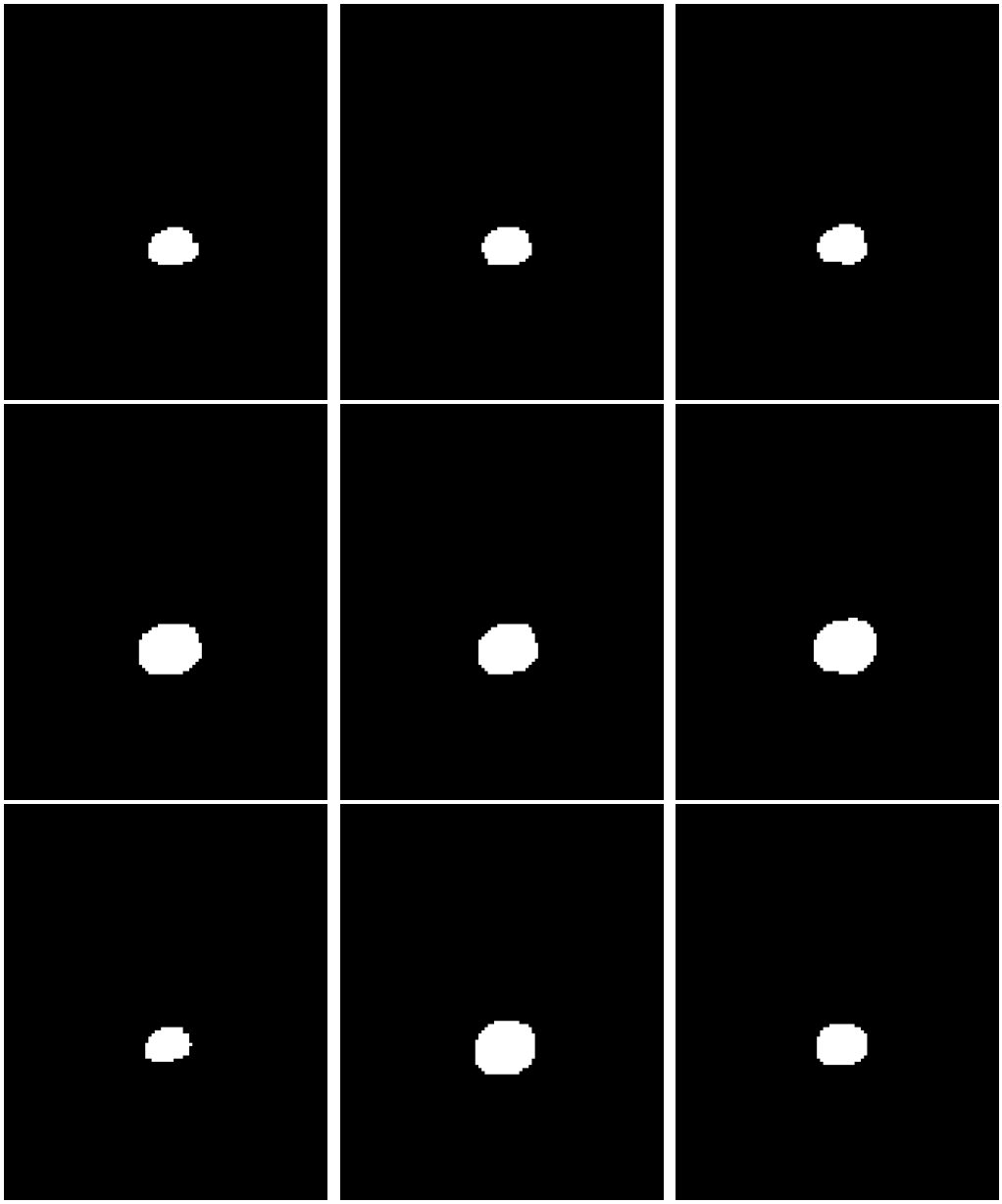

FIGURE 4.27: Esophagus image segmentation samples with their (transverse/axial slice 65) 2D cross section being shown, corresponding to Figure 4.23.

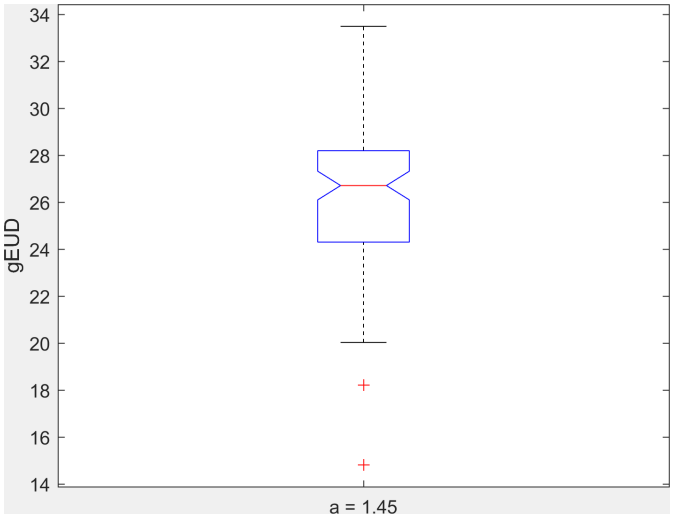

FIGURE 4.28: Boxplot of computed gEUD dose metrics for example drawn esophagus image segmentation samples. The parameter a of the gEUD dose metric has been set to 1.45 (Fogliata et al. (2018); Luxton et al. (2007); Belderbos et al. (2005)).

Chapter 5

Information Fusion for Real-time Motion Estimation in Image-guided Breast Biopsy Navigation[1]

Bojan Kocev[1,2], Joachim Georgii[1], Lars Linsen[2], Horst Karl Hahn[1,2]

[1]Fraunhofer MEVIS, Bremen, Germany

[2]Jacobs University, Bremen, Germany

Abstract

A real-time delivery of accurate soft-tissue intervention navigation information is one of the most crucial aspects for accepting the soft-tissue navigation systems for intra-operative use. Currently, soft-tissue navigation systems face some obstacles in terms of registration of the virtual navigation information on the deformable soft-tissue organ. Most of them perform a rigid registration between the virtual data and the organ, and then provide the surgeon with all navigation information. However, they suffer from the disadvantage that the virtual information is not correctly registered to the deformable organ. In order to enable a real-time non-linear registration between the virtual navigation information and the deformable organ, we incorporate different means for tracking

[1]This material originally appeared in Kocev et al. (2014).

the soft-tissue internal and on-surface local motion. Furthermore, we introduce an intelligent information fusion engine for combining the various soft-tissue local motion tracking information into a global motion information channel. The fusion engine is the interface to the motion measurements, a motion dynamics model, and static shape information, which are combined to compute the a posteriori estimate of the current state of the deformed shape. The dynamics model is realized as a finite-element deformation simulation. In order to test the feasibility of our devised information fusion engine, we have employed it for capturing the global motion of a breast phantom during an image-guided biopsy. The biopsy planning navigation data, in the form of a prior diagnostic MRI, is continuously updated over time according to the a posteriori estimate of the global motion. As a result, the real-time changes in the shape of the breast are always reflected in the biopsy navigation information.

5.1 Introduction

Smart information processing and interactive visualization systems have advanced significantly over the past years, while fascinating us in many different ways. One of the most prominent fields in which the incorporation of such systems is very beneficial, is the field of modern medical technology for computer-assisted intervention planning and navigation. Currently, modern medical technologies employing such systems are widely accepted for pre-operative planning and diagnosis. However, their intra-operative use is still facing some obstacles in terms of intervention navigation and surgeon's interaction with the virtual information system. The problem in terms of intervention navigation is twofold:

1. there are *inaccuracies in the registration* of the navigation information on the deformable soft-tissue organ (Meinzer et al. (2008)), and
2. there is *an increase in the cognitive load* of the intervention specialist during the transfer of the navigation information from the system to the intervention situs.

The inaccuracies in the registration of the navigation information on the deformable soft-tissue organ are, in general, due to the rigid registration between the virtual data and the soft-tissue organ or due to incomplete and erroneous soft-tissue motion tracking. In other words, the deformation component of the motion, which the soft-tissue organ undergoes over time, is not properly taken into account.
The obstacles in terms of surgeon's interaction with the virtual information are related to the fact that the *navigation system is usually placed far away* from the surgeon which

inhibits direct interaction.

Projector-based soft-tissue navigation and surgeon-computer natural interaction mechanisms (Kocev et al. (2011, 2013)) are very promising in circumventing the obstacles concerning the transfer of the virtual navigation information and the difficulties in the interaction therewith. *However, the overall acceptance of the navigation systems as beneficial for the intra-operative medical interventions is still confronting big challenges, mainly due to the inaccuracies in the registration of the navigation information on the deformable soft-tissue organ.*

5.1.1 Information Fusion System

In order to increase the accuracy of the above-described registration task, we devised an intelligent information fusion engine for real-time estimation of the motion that a tracked instance undergoes over time. The engine fuses, in real time, three information sources: motion dynamics, motion measurements, and shape information of the tracked instance whose motion is being estimated.

The tracked instance is represented in a discrete fashion (see Section 5.3), and its motion state at any discrete time k is directly defined by the positions of all points which constitute its discrete representation. The information fusion is performed at the points which constitute the discrete representation of the tracked instance. In accordance with this, the information sources are estimated on a per-point basis, at the points which constitute the state of the tracked instance. Our navigation information visualization engine is then able to update the virtual navigation information, in real time, according to the output of our intelligent information fusion engine, such that the changes in the pose and shape of the tracked instance are reflected in the displayed virtual information.

The general concept for the above-described information fusion, with an accompanying navigation information visualization engine, is depicted in Figure 5.1. In this regard, the work of Comaniciu et al. (2004) was influential to us in our formulation of the general information fusion problem in this manner.

In this work, we actually present the solution of a specific information fusion problem instance, depicted in Figure 5.2.

The motion dynamics modeling is embedded in a finite element model (see Section 5.5), which is set with parameters specific to the instance whose motion is being estimated. The finite element model is then used to simulate in real time the motion dynamics of the tracked instance, which serves as a prediction of the motion which the tracked instance undergoes over a specified time period. Furthermore, the finite element-based model is actually extracted from a segmented volumetric scan of the instance whose motion

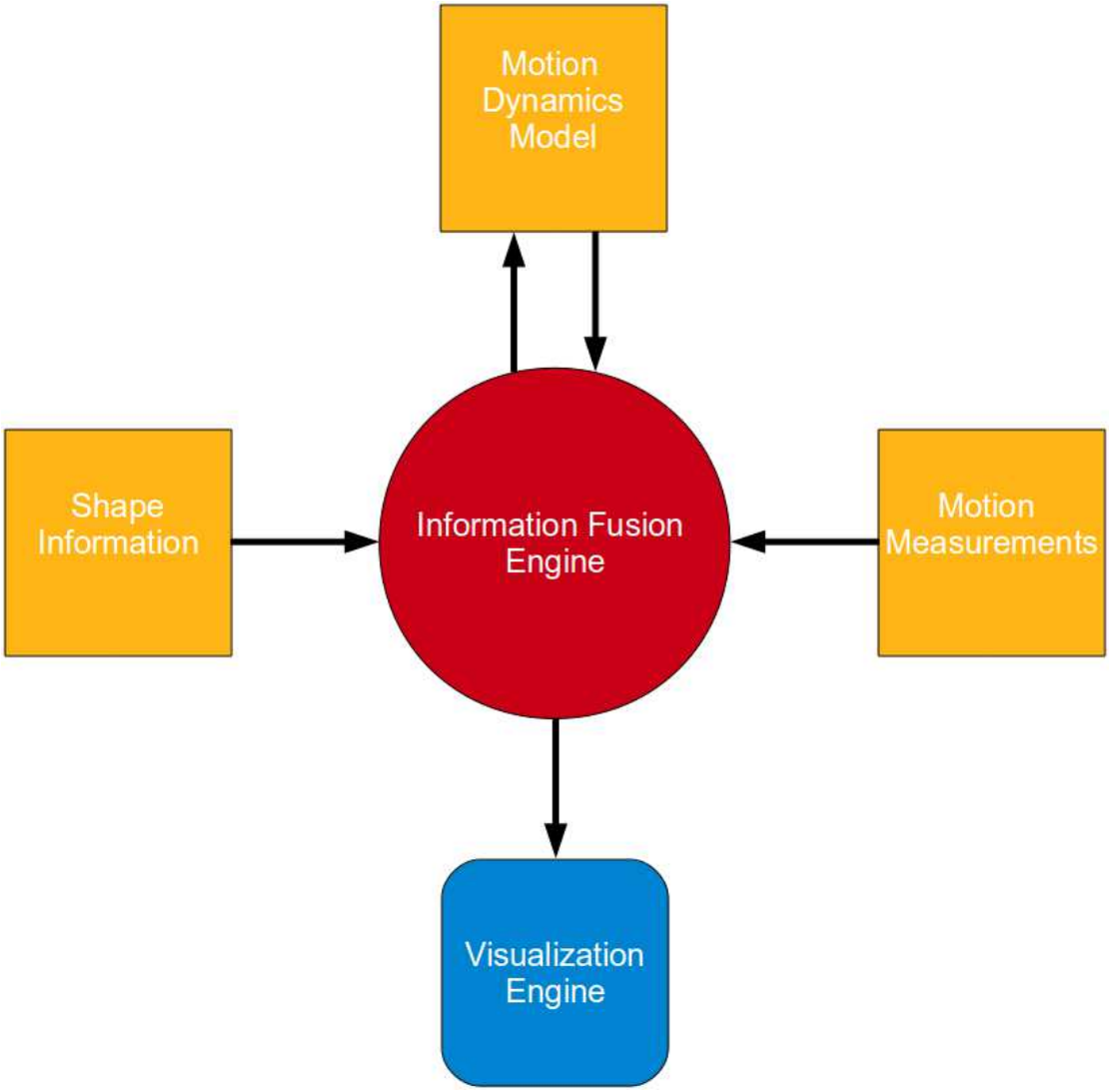

FIGURE 5.1: General Information Fusion Concept. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

is being estimated. In other words, the shape information is also embedded into the finite element-based model. Therefore, the prediction $\mathcal{E}_k$ (see Figure 5.2) by our finite element-based nonlinear motion prediction model is actually the fusion of the motion system dynamics and the shape information sources.

The motion measurements information source is composed of real-time surface (electromagnetic-based) and volumetric (ultrasound-based) tracking data (see Section 5.4), which captures the motion of the tracked instance at points which do not necessarily correspond in number as well as physically to the points constituting the state of the tracked instance. Therefore, the displacement vectors, which describe the change in the position of all tracked points from time $k-1$ to k, are interpolated at the discrete points which constitute the state of the tracked instance. In that way, we obtain the approximation $\mathcal{I}_k$ (see Figure 5.2) of the measurement displacement vectors information source at the

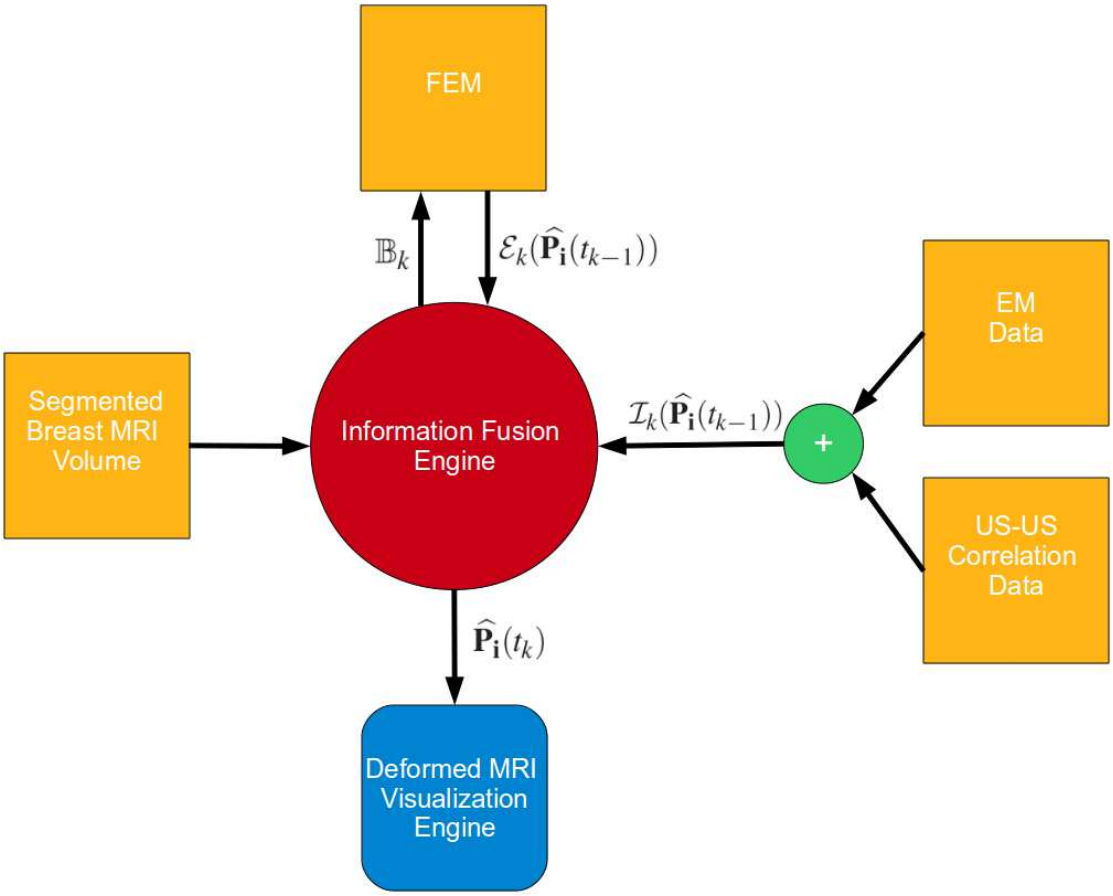

FIGURE 5.2: Information fusion problem instance in the context of breast biopsy navigation. $\mathcal{E}_k$ is the a priori displacement estimate (taking into account the set of boundary conditions $\mathbb{B}_k$) and $\mathcal{I}_k$ is the approximation of the measurement displacement vectors information source at a specific state point, at a discrete time index k. $\widehat{\mathbf{P}}_\mathbf{i}(t_k)$ is the a posteriori estimate of the position of state point $\mathbf{P}_\mathbf{i}$, at time t_k. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

state points.

The overall information fusion and estimation of the a posteriori state $\widehat{\mathbf{P}}$ (see Figure 5.2) are described in Section 5.6. The virtual navigation information, in the form of a prior diagnostic MRI, is updated in real time according to the output of our intelligent information fusion engine (see Section 5.7), such that it always reflects the estimated current shape and internal structure of the tracked instance. Since the overall information fusion and navigation information update is performed in real time, we need to exploit concurrency and parallelism efficiently and correctly (see Section 5.8). Results are presented and discussed in Section 5.9.

5.1.2 Contribution

Our main contribution to the state of the art can be summarized as follows:

1. real-time information fusion engine for motion measurements, motion dynamics, and shape information to estimate the a posteriori motion state and visualize the deformed shape;
2. computation and integration of displacements from measurements and models;
3. application to real-time image-guided breast-biopsy navigation.

5.2 Related Work

Baumhauer et al. (2008) pointed out that probably the greatest challenge, in the field of computer assisted navigation for endoscopic soft tissue interventions, relates to the intraoperative measurement and modeling of organ shift and tissue deformation of "unconstrained" organs in thoracic and abdominal cavities. Furthermore, they noted that as navigation is performed over certain period of time, a continuous correction of tissue motion and deformation would be required for constant and reliable navigation accuracy. In this section, we will discuss a list of selected related work of others about: measurement of the organ shift and tissue deformations, soft tissue deformation modeling, and information fusion of these and related information sources.

Zhang et al. (2006) employed magnetically tracked needles and biomechanical models, while compensating the liver respiratory motion. They have implemented and extended the so-called paired-point, sensor orientation-based, and needle-based (needles are implanted in the soft-tissue organ and sampled both in the electromagnetic space as well as in the CT image space) registration methods. Furthermore, they rely on the affine transform proposed by Horn (1987) for simulating the small-range deformation.

Kocev et al. (2013) incorporated an algorithm for creating, in real time, a virtual point-based representation of the deformable surface of a tracked instance lying on a surgical table. They first create a virtual point-based model of the whole scene (viewed by a Kinect camera), and then segment the surface of the tracked instance in the acquired point cloud. In this way, they are able to sample the global motion signal of the deformable tracked instance over time on the surface spatial domain. However, they do not sample on the tracked instance's interior spatial domain and their method may suffer from occlusion problems whenever parts of the tracked region are not seen by the Kinect camera.

Cash et al. (2005) employed a range scan point cloud acquired from the exposed soft-tissue organ surface, which is then rigidly aligned to a preoperative (predeformed) complete three-dimensional surface of the organ. They account for the deformation by using a linearly elastic FEM, which is implemented by using an incremental framework to

resolve geometric nonlinearities. The boundary conditions for the incremental formulation are generated from the intraoperatively acquired range scan surfaces of the exposed soft-tissue organ surface. However, they do not measure the organ interior local deformation, but rely solely on the FEM constrained with the surface boundary conditions.
Several research groups have developed methods for brain shift compensation. vSkrinjar et al. (2002) proposed a biomechanical-model-based approach for brain shift comensation, which is guided by limited intraoperative (exposed brain) surface data. Dumpuri et al. (2007, 2010) computed an atlas of model deformations based on different loading conditions preoperatively, and used it with a constrained linear inverse model to predict the intra- and post-operative distributed brain shift. Miller et al. (2012) employed the so-called Meshless Total Lagrangian Explicit Dynamics Method (MTLED), for computing brain deformations during surgery. The problem geometry is based on patient-specific MRI data, while the nodes are distributed automatically through the domain. They reported a Hausdorff distance difference between previously validated Finite Element results and their meshless results of less than 0.2 mm. However, in the context of real-time breast motion estimation, handling the motion deformation component might be more challenging and result in higher Hausdorff distance differences.
A relatively recent overview over ongoing research in the field of physically based deformation modeling is given by Nealen et al. (2005). For an overview over ongoing research in the field of breast biomechanics modeling, in the context of information fusion from different imaging modalities, we refer to the article by Rajagopal et al. (2010).
The Kalman filter, in its information filter form (Anderson and Moore (1979)), is the simplest and most well-known example of fusion. It performs the fusion of the measurements and the system dynamics information sources. The measurements and the system dynamics predictions have independent uncertainty distributions which are expressed with mean vectors and covariance matrices. Most of the other information fusion algorithms employ the original idea of Kalman for fusing the two above-mentioned information sources. However, it is often necessary to adapt or extend the Kalman fusion framework to handle additional information sources, such that more complicated motions can be estimated in an optimal and rigorous fashion.
Zhou et al. (2005) continued the work of Comaniciu et al. (2004), and thoroughly presented their complete information fusion framework for robust shape tracking, in a rigorous fashion. They follow the basic idea in treating the measurement, the shape model, and the prediction as noisy measurements with covariance matrices and fuse all the information in an optimal way. More specifically, they apply the subspace fusion on the Kalman fusion (measurement + prediction based on system dynamics) result and a subspace source (shape model), and in this way they combine all the available knowledge in the information space. They employ a *strongly adapted-PCA (SA-PCA)* model (Comaniciu et al. (2004)) to augment the statistical generic shape model with information

specific to the currently tracked shape (e.g., the initial contour of the tracked heart of a specific patient), and in this way obtain a *deformation model of the current case*. The *SA-PCA* model is then actually fused with the above-specified Kalman fusion result, such that in the end they actually fuse four information sources: the system dynamics, measurement, subspace model, and the tracked instance-specific information. In our solution, we can interpret the *finite element-based model*, extracted from the segmented volumetric scan of tracked instance, as the above-described *deformation model of the current case*. By setting the finite element-based model with tracked instance-specific material properties (density and elastic modulus), we actually incorporate the system dynamics in the finite element-based model. Therefore, the prediction of our finite element-based nonlinear motion prediction model is actually the fusion of the system dynamics and the *deformation model of the current case*.

5.3 Tracked Instance I^n State Representation

The tracked instance which undergoes some form of motion is denoted as I^n, where n is the number of points which constitute its discrete representation. The overall state of the tracked instance I^n, at time t, is represented by the state vector function:

$$\mathbf{S}(t) = \begin{bmatrix} \mathbf{P}_0(t) & \mathbf{P}_1(t) & \dots & \mathbf{P}_{n-1}(t) \end{bmatrix} \tag{5.1}$$

where $\mathbf{P}_i(t)$ is a vector variable function defined as

$$\mathbf{P}_i(t) = \begin{bmatrix} x_i(t) \\ y_i(t) \\ z_i(t) \end{bmatrix},$$

where $x_i(t)$, $y_i(t)$, and $z_i(t)$ are three scalar variables which define the 3D position, in a Cartesian coordinate system in Euclidean space, of the *i-th* point at time t. The n points P_i constitute the discrete representation of the tracked instance I^n, as shown in Figure 5.3.

The global motion of the tracked instance, from any discrete time point $k-1$ to k, is described by the displacement function:

$$u : \Omega \rightarrow \mathbb{R}^3 \tag{5.2}$$

where $\Omega \subset \mathbb{R}^3$ spans the 3D subspace of all possible 3D position values for the points which constitute the state representation (5.1) of the tracked instance I^n.

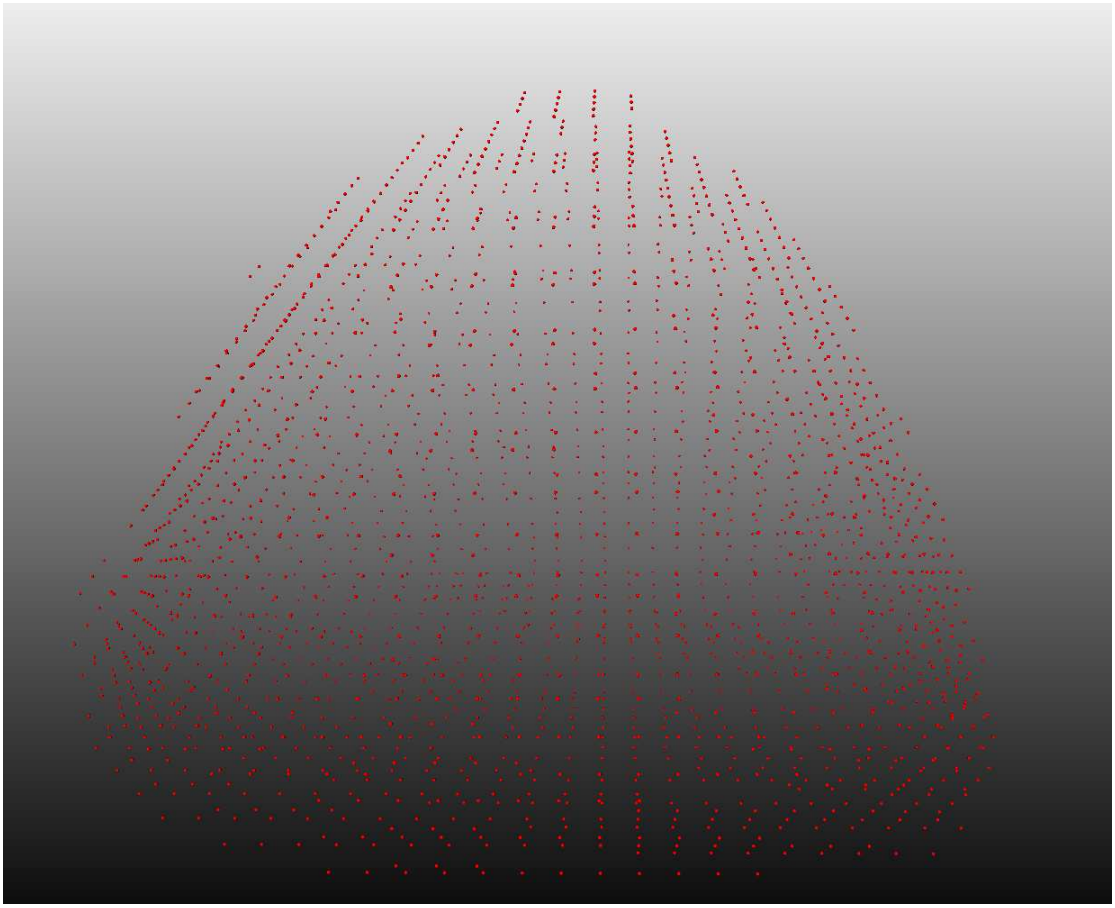

FIGURE 5.3: Discrete representation of the breast phantom. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

5.4 Motion Measurements Information Source

In general, all tracking data captures some form of motion. The measured tracking data at a discrete time point k, for an instance I^n which undergoes some form of motion, is a discrete set of points:

$$\mathbb{M}_k = \left\{ \mathbf{M_i}(t_k) : \begin{array}{c} \mathbf{M_i}(t_k) \; is \; the \; measurement \; of \\ the \; i-th \; point \; at \; time \; t_k \end{array} \right\} \tag{5.3}$$

whose cardinality theoretically may range from 1 to an arbitrarily large number. $\mathbf{M_i}(t_k)$ is a vector variable measurement function which gives the vector value of the observable or tracked i-th point at time t_k. In practical examples, the arbitrarily large number is sufficiently big, while the expected minimum cardinality of the set $\mathbb{M}_k$ is imposed by the minimum amount of information necessary for updating the overall state $\mathbf{S}(t_k)$ of the tracked instance correctly.

If $|\mathbb{M}_k| = n$ and $\forall \mathbf{M} \in \mathbb{M_k} \; \exists \mathbf{P_i}$ such that in reality they both correspond to the same

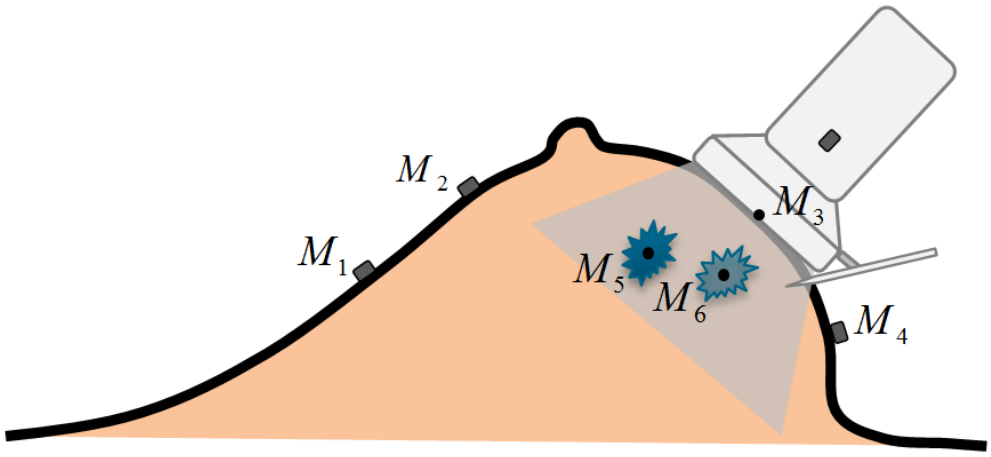

FIGURE 5.4: Breast skin surface and interior tracking data. In this example, the discrete set of points $\{\mathbf{M_1}, \mathbf{M_2}, \mathbf{M_3}, \mathbf{M_4}, \mathbf{M_5}, \mathbf{M_6}\}$ constitute the set $\mathbb{M}_k$ from Eq. 5.3. The four dark gray rectangular bodies, among which three are attached directly on the skin surface and one on the utlrasound probe, are the tracked electromagnetic bodies/sensors. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

physical point, then the state $\mathbf{S}(t_k)$ of the tracked instance I^n, at time t_k, is completely defined directly by the measurements. However, in real scenarios we have $|\mathbb{M}_k| < n$. Furthermore, the cardinality of the set $\mathbb{M}_k$, in general, varies and not always the same physical points are tracked over time, such that the overall spatial sampling density and coverage of the global motion signal varies from one time point to another.

Furthermore, the sets of measurements $\mathbb{M}_k$ and $\mathbb{M}_{k-1}$ define the set of samples, denoted by $\mathbb{S}_k$, from the output of the function (5.2) at any discrete time point k. The spatial domain Ω (defined in Eq. (5.2)) of the displacement function which captures the motion of the breast (treated as an I^n instance), is then the 3D subspace encompassing the breast interior and bounded by the breast skin surface. Therefore, the location of each sample $s \in \mathbb{S}_k$ could be either somewhere on the breast skin surface or in the breast interior.

In order to sample the output of the displacement function on the breast surface spatial domain, we attached small electromagnetic bodies/sensors to the breast skin surface, as illustrated in Figure 5.4. The 3D position of each attached electromagnetic body/sensor defines the location of a different sample from the output of the above-described displacement function, on the breast skin surface spatial domain. Furthermore, the sample values at these locations, at a discrete time k, are based on the electromagnetic tracking data, for the corresponding bodies/sensors, at the discrete time points $k-1$ and k.

In order to sample the output of the displacement function on the breast interior spatial domain, we incorporated a real-time ultrasound imaging device which is localized in the electromagnetic tracking space. This is achieved by attaching an electromagnetic

body/sensor on the ultrasound probe, as illustrated in Figure 5.4, which enables us to localize the probe in the electromagnetic tracking space. Furthermore, we compute the position (at a discrete time k) of the contact point between the ultrasound probe and the breast skin surface, based on the tracking data (at a discrete time k) for the sensor attached on the probe and on a prior calibration of the displacement vector from the sensor location to the bottom mid point on the ultrasound probe. Moreover, we are able to use the position of this contact point as an additional sample location on the breast skin surface. Having this in hand and using a prior information about the definition of the ultrasound image space with respect to a 3D local frame positioned at the above-described contact point, we obtain the transformation between the ultrasound image space and the electromagnetic tracking space. The real-time ultrasound imaging device then captures the interior of the breast by acquiring 2D images over time, as illustrated in Figure 5.4. The 2D ultrasound images reveal distinctive breast soft-tissue structures, which are identified and tracked over time by a digital image correlation (DIC) variant of the real-time capable algorithms by Isard and Blake (1998) and Zhang et al. (2010). The position of every tracked point in the ultrasound image space, within the identified and tracked breast soft-tissue structures, is localized in the electromagnetic tracking 3D space, as described above. The 3D position of every tracked point, localized in the electromagnetic tracking space, then defines the location of a different sample from the output of the above-described displacement function on the breast interior spatial domain. The sample values at these locations, at a discrete time k, are then based on the DIC tracking data, for the corresponding tracked soft-tissue points, at the discrete time points $k-1$ and k.

5.5 Motion Dynamics and Shape Information Sources

The motion, which the tracked instance I^n undergoes between two discrete time points $k-1$ and k, in general, contains two components: a rigid-body displacement and a deformation. The rigid-body displacement is composed of a rotation and a translation component and it preserves the shape and size of the tracked instance I^n, where the change in shape and size is measured against the initial or undeformed state $\mathbf{S}(t_0)$. On the other hand, the change in the deformation component of the motion is responsible for any change in shape and size which the tracked instance I^n undergoes between the two discrete time points. The deformation component of the motion at time k has changed with respect to the one at time *k-1*, if there is a nonzero relative displacement between all or some of the points which constitute the discrete representation of the tracked instance I^n.

In our case, we need to model the dynamics of the motion which the breast undergoes over time. Soft-tissue deformation modeling is a challenging task, because it involves a major deformation component which is difficult to model when the material properties of the tracked instance I^n are not well known. Furthermore, we need an appropriate nonlinear motion model, which shall be able to predict, in real time, the change in the motion deformation component from time *k-1* to *k*. This kind of nonlinear models in addition impose the challenge of extracting appropriate motion heuristics from the global motion signal sampling data $\mathbb{S}_k$, which are then used as boundary conditions by the nonlinear motion fitting model. In other words, the nonlinear model depends both on the time t_k and on a number of carefully selected boundary conditions to which the model is fitted, when predicting the motion deformation component.

In order to achieve a realistic soft-tissue deformation modeling, one needs to employ a physical model of the tracked instance I^n. As pointed by Georgii and Dick (2012), physics-based modeling is preferred over simplified models, because in a physics-based prediction model the accuracy is much higher, at least, from a theoretical point of view, which is especially important for our target medical applications. Furthermore, finite elements are a well-known mathematical tool for accurate modeling of the behavior of deformable objects based on the theory of elasticity. When compared to finite difference methods, which consider only the values at discrete samples, the finite element methods take the continuum within an element into account by providing a well defined interpolation function. In this manner, a higher accuracy is guaranteed. The degree of freedom of an element is defined by the number of "free" vertices which constitute the element, and the data values are only given at these vertices. Without any loss in the generality of our state formulation (5.1), we assume that the points which constitute the state representation of the tracked instance I^n and all the finite element vertices have 1-to-1 correspondence, and each pair of corresponding points contain information for the same physical point. In other words, the data values at all the finite elements' vertices will contain the predicted complete motion state information at time *k*.

Having these arguments in hand for the suitability of this nonlinear model for our particular problem, we decided to incorporate the framework by Georgii and Westermann (2008, 2005) for physical simulation of deformable volumetric bodies in real time, which is built upon the physical laws of continuum mechanics. Their framework is based on an implicit finite element method and it employs a multigrid approach for the efficient numerical simulation of elastic materials. Their proposed approach enabled us to do efficient realistic and numerically stable simulation of heterogeneous bodies (described by tetrahedral or hexahedral grids).

Regarding the set of boundary conditions at time *k*, denoted as $\mathbb{B}_k$, we check if the influence measure of the nearest-neighbor sample $s \in \mathbb{S}_k$ of every state point $\mathbf{P}_i$ is above some threshold, and only then compute $\mathcal{I}_k(\mathbf{P}_i)$ and add the pair $(\mathbf{P}_i,\ \mathcal{I}_k(\mathbf{P}_i))$ to the

set $\mathbb{B}_k$. The influence measure is a function of the Euclidean distance between the nearest-neighbor sample $s \in \mathbb{S}_k$ and the respective state point $\mathbf{P}_i$. $\mathcal{I}_k$ is the identified best interpolant on $\mathbb{S}_k$, which currently employs a simple nearest-neighbor interpolation strategy.

An example of a finite element model of the CIRS triple modality (X-ray Mammography, MRI, Ultrasound) breast biopsy training phantom, is shown in Figure 5.5 right.

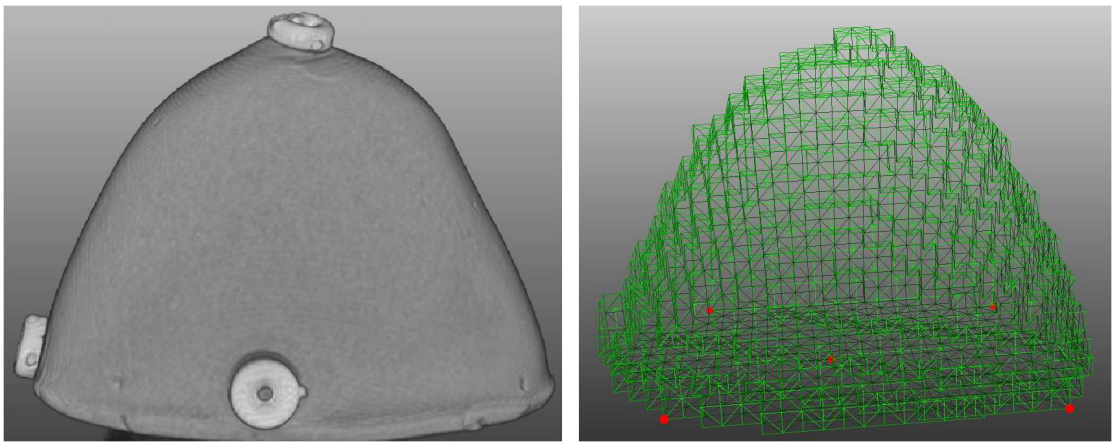

FIGURE 5.5: Breast phantom. Left to right: MRI scan with 4 markers (one on the back side); the finite element-based model composed of tetrahedra elements which are extracted from the MRI scan data on the left. The points in red are fixated, i.e., the FEM-based nonlinear motion prediction model considers these vertices as not moving. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

The fixation points are chosen such that they reflect the expected realistic behavior of the motion which the breast phantom, shown in Figure 5.9, can undergo. The points $\mathbf{P}_i$, which constitute the state representation of the breast, have 1-to-1 correspondence with the tetrahedral finite element vertices, as depicted in Figure 5.6. Furthermore, the set of boundary conditions $\mathbb{B}_k$ are applied on the tetrahedral finite element model on a per-vertex basis. We then compute the a priori displacement estimate $\mathcal{E}_k(\mathbf{P}_i)$ for all state points $\mathbf{P}_i$ for which a boundary condition is not provided, while taking into account the provided boundary conditions $\mathbb{B}_k$. In this regard, we employ a geometric multigrid solver on the tetrahedral grid to efficiently solve the resulting system of linear equations (Georgii et al. (2010)). As the finite element-based model is extracted from the MRI scan data, shown in Figure 5.5 left, and its material properties are set to match those of the CIRS phantom, the a priori displacement estimate $\mathcal{E}_k(\mathbf{P}_i)$ is actually the fusion of the motion dynamics and shape information sources at the state point $\mathbf{P}_i$. Furthermore, the a priori displacement estimate $\mathcal{E}_k(\mathbf{P}_i)$ serves as a prediction of the global motion signal at the state point $\mathbf{P}_i$, at time k.

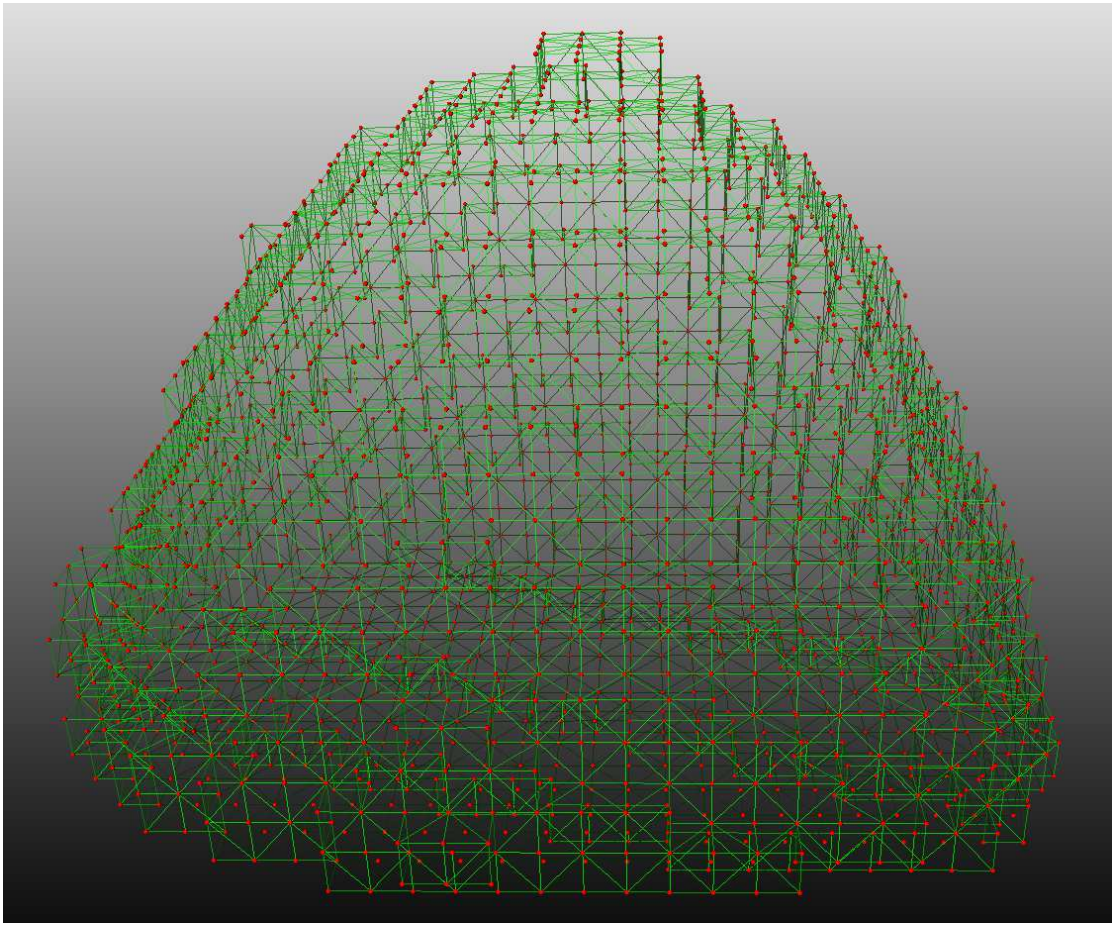

FIGURE 5.6: 1-to-1 correspondence between the points which constitute the state representation of the breast on one hand, and the tetrahedral finite element vertices on the other. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

5.6 A Posteriori State Estimation

While computing the a posteriori position estimate of each point $\mathbf{P}_i$, part of the state representation (5.1) of the tracked instance I^n (e.g., the breast phantom), we consider whether a boundary condition is provided for that state point or not. For all state points for which a boundary condition is provided in $\mathbb{B}_k$, we compute the a posteriori estimate as follows:

$$\widehat{\mathbf{P_i}}(t_k) = \widehat{\mathbf{P_i}}(t_{k-1}) + \mathcal{I}_k(\widehat{\mathbf{P_i}}(t_{k-1})) \tag{5.4}$$

while for those for which it is not provided as:

$$\widehat{\mathbf{P_i}}(t_k) = \widehat{\mathbf{P_i}}(t_{k-1}) + \mathcal{E}_k(\widehat{\mathbf{P_i}}(t_{k-1})) \tag{5.5}$$

where $\widehat{\mathbf{P_i}}(t_{k-1})$ is the previous a posteriori position estimate of the state point $\mathbf{P}_i$. In other words, there are state points, on one hand, whose nearest-neighbor motion signal sampling point has a strong enough influence on them, while on the other hand there are

state points whose nearest-neighbor motion signal sampling point does not have such a strong influence on them. Therefore the a posteriori position estimates of the former are computed using (5.4), while of the latter using (5.5).

5.7 Real-time Virtual Navigation Information Update

The virtual navigation information, e.g., a prior diagnostic MRI, shall be updated in real time according to the output of our intelligent information fusion engine, such that it reflects the estimated current shape and internal structure of the tracked breast instance. In our case, we decided to always display the MRI slice which corresponds to the current US 2D image. Therefore, the output of our information fusion engine at a discrete time k, i.e., the estimated state $\mathbf{S}(t_k)$, shall be used to update the prior MRI scan such that the corresponding MRI slice can be correctly sampled.
The prior MRI scan is taken at time t_0, i.e., when the motion state of the tracked breast is $\mathbf{S}(t_0)$. On the other hand, we have the displacement field which describes the motion of each state point $\mathbf{P}_i$ during the transition from state $\mathbf{S}(t_0)$ to state $\mathbf{S}(t_k)$. We could first move the MRI voxels, such that the 3D MRI scan reflects the estimated current shape of the tracked instance, and then sample the MRI volume at the plane corresponding to the US image plane. Alternatively, we could sample the above-described displacement field on the plane defined by the 2D US acquisition image, and then compute the MR image value for each plane pixel by sampling the MRI volume at the voxel which actually moved to the same plane point (based on the previously sampled displacement vector at the same pixel). We choose the latter approach for computing the corresponding MRI slice, because it avoids unnecessary computations. However, the sampling plane and the displacement field are both defined in the EM world coordinate system. Meaning, we also need to transform their descriptions to the MRI world coordinate system, in which the positions of the MRI voxel are described. Therefore, we need to compute the transformation between the EM and the MRI world coordinate systems.
The relationship between the EM and the MRI Cartesian coordinate systems, is found by using pair of measurements of the coordinates of four points in both systems. The measurements of the coordinates in the MRI world coordinate system are given by the 3D positions of the four MR markers (see Figure 5.5) in the MRI world, while their coordinates in the EM world coordiante system are acquired by pointing each of them with an EM-tracked sensor/pointer. The transformation parameters (rotation, translation, and scaling) are then found by employing the closed-form solution, by Umeyama (1991), of the general absolute orientation problem.
Figure 5.7 provides an example of a 2D US image together with its corresponding 2D

MR slice, where the 2D MR slice is sampled, as described above, from the prior MRI scan of the breast phantom.

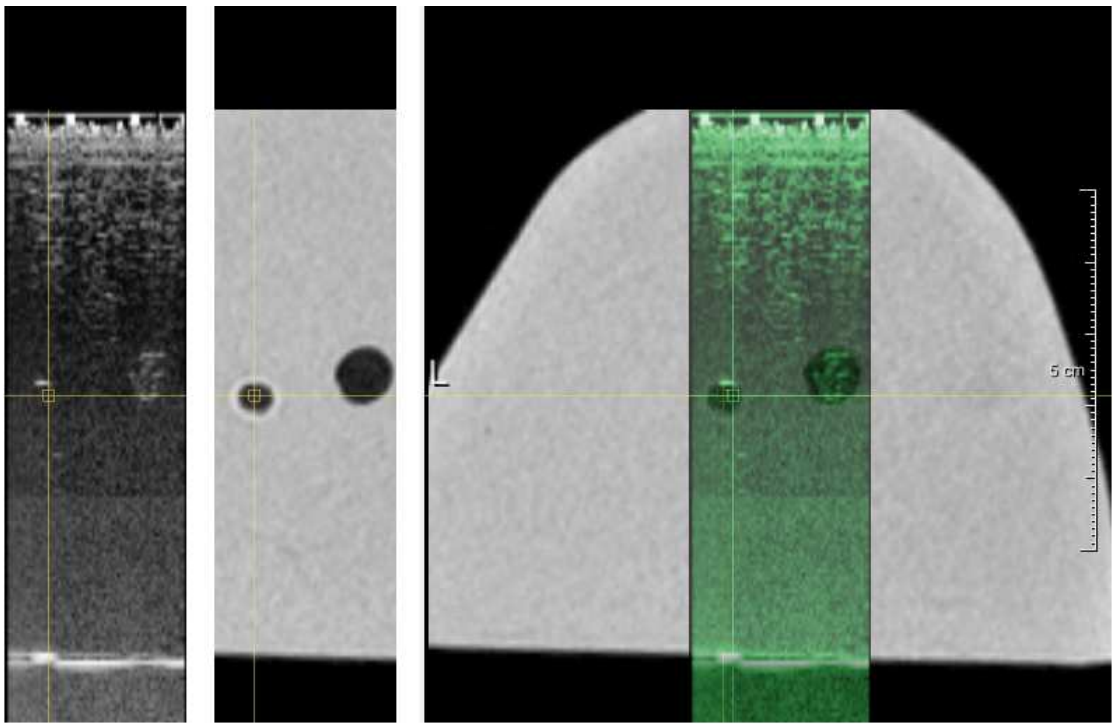

FIGURE 5.7: US-MRI correspondence example. (Left to right) current US 2D image; corresponding MRI slice with the same extents as the current US 2D image; overlay of the current US 2D image on the full corresponding MRI slice. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

5.8 Concurrency and Parallelism Aspects

In order to make the estimation of the motion, a tracked instance I^n undergoes over time, possible in real time, we had to extensively parallelize each information source generation component as well as the overall information fusion engine. When striving for performance, programming in terms of threads (using low-level APIs) can be an inconvenient way to do multi-threaded programming. Logical tasks are a more appropriate choice, because they match better parallelism to available resources, have a faster start-up and shut-down, have a more efficient evaluation order, improved load balancing, and they provide higher-level thinking. Therefore, we have employed the Intel®Threading Building Blocks (Intel®TBB) library (Intel), which supports scalable parallel programming using standard ISO C++ code. We define three different Intel®TBB-based logical tasks which can be described as follows:

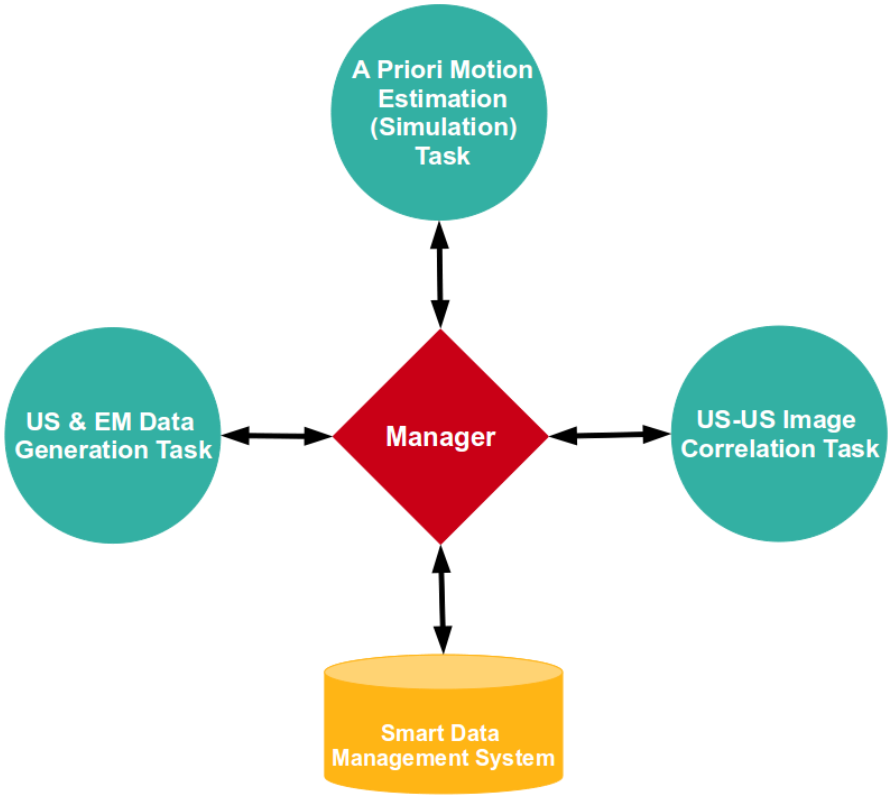

FIGURE 5.8: Depiction of the communication flow in our intelligent information fusion system. (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

1. a task which generates 2D ultrasound (US) images and electromagnetic tracking data in real time;

2. a task which performs the US-US image correlation in real time, and

3. a task which performs the fusion of the motion dynamics and the shape information sources (finite element-based simulation), i.e., the a priori motion estimation, in real time.

The communication and the synchronization between the different logical tasks is handled by a so called *Manager* unit (depicted in Figure 5.8), which we have implemented in addition. Furthermore, we had to devise a smart data management system to handle all data generated by the different tasks in a thread-safe manner.
The sampling of the displacement field (described in section 5.7) on the plane defined by the 2D US acquisition image, is extensively parallelized using OpenMP (OpenMP Architecture Review Board). Furthermore, the 2D sampling of the prior MR image, based on the above-described sampled displacement field, is performed on the graphics processing unit (GPU).

5.9 Results and Discussion

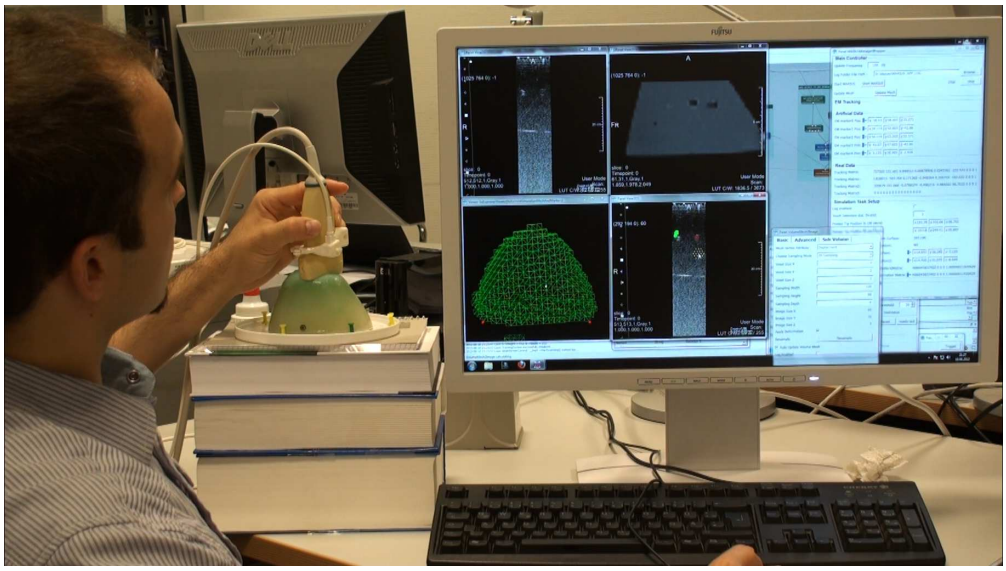

FIGURE 5.9: Hybrid Image-guided Breast Biopsy Navigator. (Please see the accompanying video material.) (This figure originally appeared in Kocev et al. (2014). ©Eurographics Association 2014. Reproduced by kind permission of the Eurographics Association.)

The presented information fusion algorithm has been tested within our devised breast biopsy navigation system, depicted in Figure 5.9. This figure demonstrates the functionality of our information fusion engine, when estimating the global motion of the CIRS triple modality breast biopsy training phantom in real time.

The resolution of the input ultrasound (US) image, as shown in the top-left image viewer on the computer screen in Figure 5.9, is set to 512×512. Furthermore, the ultrasound acquisition depth is set to 120 mm.

The bottom-right image viewer, as depicted on the computer screen in Figure 5.9, shows an overlay of the two displacement vectors (capturing the local motion of the tracked breast lesions in the ultrasound image space) on the input US image. In this example, the global motion signal sampling data set $\mathbb{S}_k$ contains two samples from the output of the displacement function 5.2 on the breast interior spatial domain. Their sample values are set with the values of the overlaying displacement vectors, shown in the bottom-right image viewer on the computer screen in Figure 5.9, transformed to the EM world coordinate system. The set of boundary conditions $\mathbb{B}_k$, at time k, are then extracted from the set $\mathbb{S}_k$ and used as Dirichlet conditions (Georgii et al. (2010)) in the finite element model of the breast phantom.

The breast phantom is fixated with five pins. Two of them can be observed in front, while the remaining three cannot be seen because one is below (in the middle), and the other two behind the breast phantom (see Figure 5.9). One could also observe the virtual representation of these fixation points, rendered as red spheres, in the virtual

scene containing the finite element model of the breast shown in the bottom-left image viewer on the computer screen in Figure 5.9, or more clearly in Figure 5.5.
For the performed tests, we used the following parameter configuration of the breast phantom finite element model:

Parameter	Value/Type
Integration Type	Dynamic Euler
Strain Type	Corotated Cauchy Strain
Stiffness	1000
Poisson Ratio	0.48
Density	1000
Damping	2.5
Time Step	0.033
Number of VCycles	1

Having set the finite element model with the above information, we are able to compute the a priori displacement estimate $\mathcal{E}_k(\mathbf{P}_i)$ for all state points $\mathbf{P}_i$ for which a boundary condition is not given.
Then, we are able to compute the a posteriori position estimate for all state points $\mathbf{P}_i$, as described in Section 5.6.
Furthermore, we are able to continuously update the prior diagnostic image (in this example an MRI image), based on the real-time output of our fusion algorithm, as described in 5.7. In the top-right image viewer on the computer screen in Figure 5.9, one could observe that the updated diagnostic MRI image map reflects the true shape and internal structure of the examined breast (see also Figure 5.7). In this way, we can actually claim that we are also able to provide a hybrid image-guided (in this example US-MRI-guided) biopsy navigation.
We analyzed the real-time performance of our information fusion system on a desktop PC (Intel(R) Core(TM) 2 Quad CPU Q9000 @ 2.00 GHz 2.00 GHz, 4 GB RAM, Windows 7 Professional 64 bit) and obtained that on average we achieve 20-30 updates of the current state per second. The non-constant update rate is mainly due to the concurrent nature of the different components in the information fusion engine.
However, the accuracy of our information fusion system has only been visually inspected in our laboratory setting using the CIRS phantom and not yet for estimating the motion of deformable human organs in an intervention room. In future work, we plan on first validating the results of our information fusion algorithm against simulated ground-truth global motion signal values. One could simulate the ground-truth global motion signal using the finite element model of the soft-tissue tracked instance. Then one would need to sample the ground-truth global motion signal at locations which do not necessarily

correspond to the state points $\mathbf{P}_i$, and use these samples as ground-truth global motion signal samples. These samples would constitute the motion measurements information source which will be fused with the dynamics and the shape information sources by our fusion engine. The result will be then compared against the above-specified ground-truth global motion signal. For an overall system validation, we plan on generating the ground-truth navigation information for a certain number of time points (e.g., by acquiring MRI scans of the deformed soft-tissue tracked instance at the respective time points) and comparing it against the updated virtual navigation information (e.g., the deformed prior MRI scans) based on the output of our information fusion algorithm at the respective time points.

5.10 Conclusion and Future Work

We devised an intelligent information fusion engine for real-time estimation of the motion that a tracked instance undergoes over time. The engine fuses the tracked instance's motion dynamics, motion measurements, and shape information sources. As a result, all available knowledge in the information space is combined.

We incorporated electromagnetic bodies/sensors for sampling the global motion signal on the deformable organ's surface spatial domain. For sampling the global motion signal on the deformable organ's interior spatial domain, we used an appropriate real-time US-US image correlation algorithm which identifies distinctive soft-tissue structures and tracks them over time.

By employing a finite element model, we were able to best model, based on the theory of elasticity, the dynamics of the motion which a deformable organ undergoes over time. Using this nonlinear motion model, we were able to predict, in real time, the change in the motion deformation component from time $k-1$ to k.

We showcased the feasibility of our devised information fusion engine by employing it for capturing the global motion of a breast phantom during an image-guided biopsy. In this way, we enabled the real-time update of the biopsy planning navigation data according to the posteriori estimate of the global motion. In return, the real-time changes in the shape of the breast are always reflected in the navigation information.

As we assumed error-free measurements and modeling, our a posteriori position estimates depend either only on the interpolated displacement or only on the a priori displacement estimate. In other words, currently we do not model the uncertainties neither in the electromagnetic tracking nor in the US-US image correlation data. In future work, we plan on quantifying these uncertainties and including this knowledge in the information space. This will enable us to incorporate the uncertainty information source in the a posteriori state estimation, such that a proper (always assuring consistent

and conservative a posteriori position estimates) uncertainty-aware information fusion is possible.

Acknowledgements

This work was supported by the Fraunhofer Internal Programs under Grant No. MAVO 823 287.

Chapter 6

Conclusion and Future Work

We addressed different aspects of the problem of uncertainty in image-guided therapies. In Chapter 4, we proposed a solution to the problem of automatic and efficient generation of plausible image segmentation samples that are similar to a single or multiple rater segmentations or in agreement with the output class label probabilities of a probabilistic segmentation. By devising a novel stationary GP-based generative segmentation model, we were able to model the uncertainty in image segmentation. The stationary GP served as a model of a stochastic level set function whose zero level sets were used to model the segmentation boundaries, which in return allowed us to treat the boundaries in an uncertainty-aware fashion. By computing our novel GP mean function based on input marginal label probabilities (MLPs), we further generalized a current state-of-the-art method for image segmentation sampling and hence broadened the applicability of the proposed method. That and the setup of the GP with a squared exponential covariance function allowed us to draw plausible zero level sets (i.e., segmentation boundaries). The proposed novel methods for estimating MLPs based on a non-probabilistic single or multiple (e.g., rater) input segmentations allowed us to also apply the proposed method in cases when a probabilistic segmentation output is not available. By devising a fast method for sampling from stationary GPs, we were able to draw the image segmentation samples efficiently. In this regard, we employed convolution to blur 2D/3D white noise images with a 2D/3D image smoothing kernel in Fourier space and in this way we were able to efficiently draw realizations from the stationary GP on a 2D/3D image grid. We also found as secondary results of this research component a computationally more efficient "matrix square root" factorization than the standard Cholesky factorization as well as an efficient matrix-vector multiplication, though they both depend on the assumption of stationarity. The generative segmentation model allowed us to automatically generate many plausible image segmentation samples by simply drawing GP samples and thresholding them, which we then used for estimating and visualizing different aspects of the

uncertainty in image-guided therapies (e.g., the uncertainty in various radiotherapy dose metrics). By inspecting the visualized uncertainties (e.g., conveyed in the boxplot, in Figure 4.28, of computed radiotherapy gEUD dose metrics for example esophagus image segmentation samples), we could conclude that there might be a significant uncertainty in some image-guided therapies and thus the knowledge about this uncertainty may be useful information if considered in clinical practice (e.g., for evaluating the quality of the computed radiotherapy dose plan in terms of possible adverse effects on the patient). The devised generative segmentation model may be potentially further improved in different ways. For example, instead of using a constant characteristic length-scale (hyper-) parameter over the spatial domain for the GP covariance function, in some cases, it may be better to use a spatial location-dependent length-scale. To achieve that, we could, for example, employ the Gibbs' covariance function (Gibbs (1998)), which may potentially allow us to better control the variability of the stochastic level set function samples over the spatial domain. However, the Gibbs' covariance function is non-stationary, i.e., our devised fast method for sampling from stationary GPs would not be applicable in this case. Furthermore, instead of setting the "true" segmentation to a single non-probabilistic segmentation (e.g., performed by a highly ranked expert rater) or to the average of multiple non-probabilistic segmentations (e.g., performed by less experienced raters), we could estimate it in a more sophisticated fashion, for example by employing the STAPLE algorithm (Warfield et al. (2004)), which may potentially enable us to estimate more accurate MLPs.

In Chapter 2, we proposed a solution to the problem of interpolating in an uncertainty-aware fashion randomly non-uniformly spatiotemporally scattered uncertain soft tissue motion measurements and after that, in Chapter 3, we devised a method about propagating the uncertainty from the motion estimation result to a corresponding registered pre-operative/pre-interventional medical image and visualizing it in an uncertainty-aware fashion thereafter. By devising a novel spatiotemporal GP regression model, we were able to perform the interpolation of the motion measurements in an uncertainty-aware optimal unbiased fashion. In principle, the spatiotemporal GP formalism enabled the uncertainty-aware estimation of anatomy displacements at any location, and for any time interval from measured uncertain motions that are sparse in space and time. By estimating the full a posteriori motion field distribution, we also obtained the uncertainty in the resulting MAP estimate of the soft tissue motion at any spatiotemporal location quantified. By using a composite prior GP mean function, we were able to learn the global and local drifts present in the latent process mean function. This and the use of a squared exponential covariance function with an anisotropic distance measure (ADM) enabled our proposed GP model to, at the same time, learn more and yield statistically significantly better predictions than the similar state-of-the-art GP

models. By performing a spatially-varying blur to an image containing a grid, we were able to convey the overall uncertainty in the displayed MAP deformed image estimate in a way consistent with our cognitive models of which perceptual elements contain variability or uncertainty. By estimating POI/ROI (e.g., tumor) MLPs and showing visually mapped 2D cross sections thereof or alternatively visualizing MLP isocontours, we were able to visualize where a POI/ROI (e.g., a tumor) is after deformation. By visually conveying the amount and various forms/types of uncertainty in the registered pre-interventional/pre-operative navigation information, we hope to enable surgeons and intervention specialists to make better informed safety-critical decisions and thus potentially improve the overall outcome of surgeries or interventions. However, the deployed GP regression framework has some limitations. For example, when the number of measurements/observations is considerably large then both storing the covariance (or Gram) matrix and solving the resulting linear systems may become unfeasible on standard modern workstations (Rasmussen and Williams (2006)). Major research has been performed by different groups with the aim to address or deal with this problem (Rasmussen and Williams (2006)), however we believe that there is still space for further improvement. Furthermore, our model might be potentially improved by using a heteroscedastic noise (see the work of Goldberg et al. (1997)) instead of a homoscedastic one. Another potential improvement of our model might be achieved by removing the assumption of independence between the motion signal output dimensions, which would require to use a vector-valued GP instead of a separate scalar-valued GP for each output dimension (see Hein and Bousquet (2004) and Micchelli and Pontil (2005)). In addition, our model might be potentially enhanced by using varying characteristic length-scale (hyper-) parameters over the spatiotemporal domain for the GP covariance function, which can be achieved, for example, by deploying the Gibbs' covariance function (Gibbs (1998)).

In future work, we plan to extend/adapt our proposed sampling method to address the problem of how to efficiently draw samples from non-stationary GPs, which would allow us to use non-stationary covariance functions (e.g., the above-mentioned Gibbs' covariance function). In this regard, Lê et al. (2016) already did some work in this direction that is based on the idea of using supervoxels. Furthermore, we plan to explore more sophisticated algorithms (Warfield et al. (2004)) for estimating the true segmentation based on a collection of non-probabilistic (e.g., rater) segmentations and analyze the impact on our estimated MLPs. In general, we also intend to thoroughly evaluate our estimated MLPs, e.g., by comparing them against those estimated by the STAPLE algorithm (Warfield et al. (2004)). In addition, we would like to perform a larger study on various organs and tumors segmentation datasets and estimate and visualize different aspects of the uncertainty in the respective image-guided therapies and analyze thoroughly how this information may influence the outcome of these therapies in clinical

practice. Regarding our work on human anatomy motion estimation, we plan to identify appropriate formalisms and, if needed, approximation approaches to optimize the conditioning of the proposed GP model on large training datasets. We would also like to apply our proposed motion estimation model for registering real pre-interventional/pre-operative navigation data on deformable soft-tissue organs during a real intervention or surgery. Furthermore, we intend to deploy our proposed GP model for modeling organ deformation in the context of radiotherapy. With respect to the visualization thrust of our research, we would like to apply our proposed method for uncertainty-aware visualization of registered pre-operative/pre-interventional patient images (coupled with planning/navigation data) during a real intervention or surgery.

Acknowledgements

I want to express my enormous gratitude to both my co-supervisors, Prof. Dr. Horst Karl Hahn, and Prof. Dr. Lars Linsen whose continuous support, guidance and great feedback played a great role in creating the work presented in this thesis.

I would also like to thank the other PhD committee members, Prof. Dr. Ron Kikinis, Prof. Dr. William M. Wells, Jun.-Prof. Dr. Christian Hansen, and Prof. Dr. Andreas Birk for their valuable input.

At the same time, I would like to thank Prof. Dr. Robert Cormack for providing us with valuable radiotherapy data without which crucial part of the presented results would not have been possible.

I would like to thank Fraunhofer MEVIS for all the support during my PhD studies.

At the end, my family receive my deepest gratitude and love for inspiring and encouraging me, and fully supporting me throughout my work and life.

Bibliography

Adler, R.J., Taylor, J.E., 2009. Random fields and geometry. Springer Science & Business Media.

Allendes Osorio, R., Brodlie, K.W., 2008. Contouring with uncertainty, in: Theory and Practice of Computer Graphics 2008. Proceedings., Eurographics Association. pp. 59–66.

Anderson, B.D., Moore, J.B., 1979. Optimal Filtering. Prentice-Hall.

Barber, D., 2012. Bayesian reasoning and machine learning. Cambridge University Press.

Baumhauer, M., Feuerstein, M., Meinzer, H.P., Rassweiler, J., 2008. Navigation in endoscopic soft tissue surgery: perspectives and limitations. Journal of endourology 22, 751–766.

Belderbos, J., Heemsbergen, W., Hoogeman, M., Pengel, K., Rossi, M., Lebesque, J., 2005. Acute esophageal toxicity in non-small cell lung cancer patients after high dose conformal radiotherapy. Radiotherapy and oncology 75, 157–164.

Bersvendsen, J., Toews, M., Danudibroto, A., Wells, W.M., Urheim, S., Estépar, R.S.J., Samset, E., 2016. Robust spatio-temporal registration of 4d cardiac ultrasound sequences, in: SPIE Medical Imaging, International Society for Optics and Photonics. pp. 97900F–97900F.

Bishop, C.M., 2006. Pattern recognition. Machine Learning 128.

OpenMP Architecture Review Board, . OpenMP. http://openmp.org.

Bonneau, G.P., Hege, H.C., Johnson, C.R., Oliveira, M.M., Potter, K., Rheingans, P., Schultz, T., 2014. Overview and state-of-the-art of uncertainty visualization, in: Scientific Visualization. Springer, pp. 3–27.

Botchen, R.P., Weiskopf, D., Ertl, T., 2005. Texture-based visualization of uncertainty in flow fields, in: VIS 05. IEEE Visualization, 2005., IEEE. pp. 647–654.

Bracewell, R.N., Bracewell, R.N., 1986. The Fourier transform and its applications. volume 31999. McGraw-Hill New York.

Breiman, L., Meisel, W., Purcell, E., 1977. Variable kernel estimates of multivariate densities. Technometrics 19, 135–144.

Cash, D.M., Miga, M.I., Sinha, T.K., Galloway, R.L., Chapman, W.C., 2005. Compensating for intraoperative soft-tissue deformations using incomplete surface data and finite elements. Medical Imaging, IEEE Transactions on 24, 1479–1491.

Chai, T., Draxler, R.R., 2014. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. Geoscientific Model Development 7, 1247–1250.

Chakrabarti, A., Zickler, T., Freeman, W.T., 2010. Analyzing spatially-varying blur, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE. pp. 2512–2519.

Chan, A.B., 2013. Multivariate generalized gaussian process models. arXiv preprint arXiv:1311.0360 .

Chlan, E.B., Rheingans, P., 2005. Multivariate glyphs for multi-object clusters, in: IEEE Symposium on Information Visualization, 2005. INFOVIS 2005., IEEE. pp. 141–148.

Ciresan, D., Giusti, A., Gambardella, L.M., Schmidhuber, J., 2012. Deep neural networks segment neuronal membranes in electron microscopy images, in: Advances in neural information processing systems, pp. 2843–2851.

Comaniciu, D., Zhou, X.S., Krishnan, S., 2004. Robust real-time myocardial border tracking for echocardiography: an information fusion approach. Medical Imaging, IEEE Transactions on 23, 849–860.

Deasy, J.O., Blanco, A.I., Clark, V.H., 2003. Cerr: a computational environment for radiotherapy research. Medical physics 30, 979–985.

Deitrick, S.A., 2007. Uncertainty visualization and decision making: Does visualizing uncertain information change decisions, in: Proceedings of the XXIII International Cartographic Conference, pp. 4–10.

Djurcilov, S., Kim, K., Lermusiaux, P., Pang, A., 2002. Visualizing scalar volumetric data with uncertainty. Computers & Graphics 26, 239–248.

Djurcilov, S., Kim, K., Lermusiaux, P.F., Pang, A., 2001. Volume rendering data with uncertainty information, in: Data Visualization 2001. Springer, pp. 243–252.

Dumpuri, P., Thompson, R.C., Cao, A., Ding, S., Garg, I., Dawant, B.M., Miga, M.I., 2010. A fast and efficient method to compensate for brain shift for tumor resection therapies measured between preoperative and postoperative tomograms. Biomedical Engineering, IEEE Transactions on 57, 1285–1296.

Dumpuri, P., Thompson, R.C., Dawant, B.M., Cao, A., Miga, M.I., 2007. An atlas-based method to compensate for brain shift: Preliminary results. Medical Image Analysis 11, 128–145.

Fan, A.C., Fisher, J.W., Wells, W.M., Levitt, J.J., Willsky, A.S., 2007. Mcmc curve sampling for image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 477–485.

Fogliata, A., Thompson, S., Stravato, A., Tomatis, S., Scorsetti, M., Cozzi, L., 2018. On the geud biological optimization objective for organs at risk in photon optimizer of eclipse treatment planning system. Journal of applied clinical medical physics 19, 106–114.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2014. Bayesian data analysis. volume 2. Chapman & Hall/CRC Boca Raton, FL, USA.

Georgii, J., Dick, C., 2012. Efficient finite element methods for deformable bodies in medical applications. Critical Reviews^{TM} in Biomedical Engineering 40, 155–172.

Georgii, J., Lagler, D., Dick, C., Westermann, R., 2010. Interactive deformations with multigrid skeletal constraints, in: Proceedings of the 7th Workshop On Virtual Reality Interaction and Physical Simulation, pp. 39–47.

Georgii, J., Westermann, R., 2005. A multigrid framework for real-time simulation of deformable volumes, in: Workshop On Virtual Reality Interaction and Physical Simulation.

Georgii, J., Westermann, R., 2008. Corotated finite elements made fast and stable, in: Proceedings of the 5th Workshop On Virtual Reality Interaction and Physical Simulation, pp. 11–19.

Gerardo-Castro, M.P., Peynot, T., Ramos, F., 2015. Laser-radar data fusion with gaussian process implicit surfaces, in: Field and Service Robotics, Springer. pp. 289–302.

Gerig, T., Shahim, K., Reyes, M., Vetter, T., Lüthi, M., 2014. Spatially varying registration using gaussian processes, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 413–420.

Gibbs, M.N., 1998. Bayesian Gaussian processes for regression and classification. Ph.D. thesis. Citeseer.

Glocker, B., Komodakis, N., Navab, N., Tziritas, G., Paragios, N., 2009. Dense registration with deformation priors, in: International Conference on Information Processing in Medical Imaging, Springer. pp. 540–551.

Goldberg, P.W., Williams, C.K., Bishop, C.M., 1997. Regression with input-dependent noise: A gaussian process treatment. Advances in neural information processing systems 10, 493–499.

Grigoryan, G., Rheingans, P., 2004. Point-based probabilistic surfaces to show surface uncertainty. IEEE Transactions on Visualization and Computer Graphics 10, 564–573.

Harrower, M., 2003. Representing uncertainty: Does it help people make better decisions, in: UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop, Citeseer. pp. 18–20.

Hein, M., Bousquet, O., 2004. Kernels, associated structures and generalizations. Max-Planck-Institut fuer biologische Kybernetik, Technical Report .

Higdon, D., 2002. Space and space-time modeling using process convolutions, in: Quantitative methods for current environmental issues. Springer, pp. 37–56.

Horn, B.K., 1987. Closed-form solution of absolute orientation using unit quaternions. JOSA A 4, 629–642.

Intel, . Intel Threading Building Blocks. https://www.threadingbuildingblocks.org.

Irani, M., Anandan, P., 2000. Factorization with uncertainty, in: Computer Vision-ECCV 2000. Springer, pp. 539–553.

Isard, M., Blake, A., 1998. Condensation-conditional density propagation for visual tracking. International journal of computer vision 29, 5–28.

Kanazawa, Y., Kanatani, K.i., 2001. Do we really have to consider covariance matrices for image features?, in: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, IEEE. pp. 301–306.

Khaleghi, B., Khamis, A., Karray, F.O., Razavi, S.N., 2013. Multisensor data fusion: A review of the state-of-the-art. Information Fusion 14, 28–44.

Knight, J.C., 2002. Safety critical systems: challenges and directions, in: Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on, IEEE. pp. 547–550.

Kniss, J.M., Van Uitert, R., Stephens, A., Li, G.S., Tasdizen, T., Hansen, C., 2005. Statistically quantitative volume visualization, in: VIS 05. IEEE Visualization, 2005., IEEE. pp. 287–294.

Kocev, B., Georgii, J., Linsen, L., Hahn, H.K., 2014. Information Fusion for Real-time Motion Estimation in Image-guided Breast Biopsy Navigation, in: Bender, J., Duriez, C., Jaillet, F., Zachmann, G. (Eds.), Workshop on Virtual Reality Interaction and Physical Simulation, The Eurographics Association. doi:10.2312/vriphys.20141227.

Kocev, B., Hahn, H.K., Linsen, L., Wells, W.M., Kikinis, R., 2019. Uncertainty-aware asynchronous scattered motion interpolation using gaussian process regression. Computerized Medical Imaging and Graphics 72, 1–12.

Kocev, B., Ojdanic, D., Peitgen, H., 2011. An approach for projector-based surgeon-computer interaction using tracked instruments, In: Proc. of GI Workshop: Emerging Technologies for Medical Diagnosis and Therapy.

Kocev, B., Ritter, F., Linsen, L., 2013. Projector-based surgeon-computer interaction on deformable surfaces. International Journal of Computer Assisted Radiology and Surgery , 1–12URL: http://dx.doi.org/10.1007/s11548-013-0928-1, doi:10.1007/s11548-013-0928-1.

Kuss, M., 2006. Gaussian process models for robust regression, classification, and reinforcement learning. Ph.D. thesis. TU Darmstadt.

Lê, M., Unkelbach, J., Ayache, N., Delingette, H., 2015. Gpssi: Gaussian process for sampling segmentations of images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 38–46.

Lê, M., Unkelbach, J., Ayache, N., Delingette, H., 2016. Sampling image segmentations for uncertainty quantification. Medical image analysis 34, 42–51.

Ledesma-Carbayo, M.J., Kybic, J., Desco, M., Santos, A., Suhling, M., Hunziker, P., Unser, M., 2005. Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. IEEE transactions on medical imaging 24, 1113–1126.

Leedan, Y., Meer, P., 2000. Heteroscedastic regression in computer vision: Problems with bilinear constraint. International Journal of Computer Vision 37, 127–150.

Lim, J.H., Ong, S.H., Xiong, W., 2015. Biomedical Image Understanding: Methods and Applications. John Wiley & Sons.

Lodha, S.K., Faaland, N.M., Charaniya, A.P., Varshney, P.K., Mehrotra, K., Mohan, C., 2002. Visualization of uncertain particle movement, in: Proceedings of the Computer Graphics and Imaging Conference, Citeseer. pp. 226–232.

Lodha, S.K., Sheehan, B., Pang, A.T., Wittenbrink, C.M., et al., 1996. Visualizing geometric uncertainty of surface interpolants, in: Graphics Interface, pp. 238–245.

Lundström, C., Ljung, P., Persson, A., Ynnerman, A., 2007. Uncertainty visualization in medical volume rendering using probabilistic animation. IEEE transactions on visualization and computer graphics 13, 1648–1655.

Lüthi, M., Jud, C., Vetter, T., 2011. Using landmarks as a deformation prior for hybrid image registration, in: Pattern Recognition. Springer, pp. 196–205.

Lüthi, M., Jud, C., Vetter, T., 2013. A unified approach to shape model fitting and non-rigid registration, in: International workshop on machine learning in medical imaging, Springer. pp. 66–73.

Luxton, G., Keall, P.J., King, C.R., 2007. A new formula for normal tissue complication probability (ntcp) as a function of equivalent uniform dose (eud). Physics in Medicine & Biology 53, 23.

Meinzer, H.P., Maier-Hein, L., Wegner, I., Baumhauer, M., Wolf, I., 2008. Computer-assisted soft tissue interventions, in: Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on, IEEE. pp. 1391–1394.

Mezger, U., Jendrewski, C., Bartels, M., 2013. Navigation in surgery. Langenbeck's Archives of Surgery 398, 501–514.

Micchelli, C.A., Pontil, M., 2005. On learning vector-valued functions. Neural computation 17, 177–204.

Miller, K., Horton, A., Joldes, G., Wittek, A., 2012. Beyond finite elements: A comprehensive, patient-specific neurosurgical simulation utilizing a meshless method. Journal of biomechanics 45, 2698–2701.

Nealen, A., Müller, M., Keiser, R., Boxerman, E., Carlson, M., 2005. Physically based deformable models in computer graphics, in: Proceedings of Eurographics, pp. 71–94.

Niemierko, A., 1999. A generalized concept of equivalent uniform dose (eud). Med Phys 26, 1100.

Niethammer, M., Pohl, K.M., Janoos, F., Wells III, W.M., 2017. Active mean fields for probabilistic image segmentation: Connections with chan–vese and rudin–osher–fatemi models. SIAM journal on imaging sciences 10, 1069–1103.

O'Hagan, A., Kingman, J.F.C., 1978. Curve fitting and optimal design for prediction. Journal of the Royal Statistical Society. Series B (Methodological) 40, 1–42. URL: http://www.jstor.org/stable/2984861.

Pang, A.T., Wittenbrink, C.M., Lodha, S.K., 1997. Approaches to uncertainty visualization. The Visual Computer 13, 370–390.

Perperidis, D., Mohiaddin, R.H., Rueckert, D., 2005. Spatio-temporal free-form registration of cardiac mr image sequences. Medical image analysis 9, 441–456.

Pohl, K.M., Fisher, J., Bouix, S., Shenton, M., McCarley, R.W., Grimson, W.E.L., Kikinis, R., Wells, W.M., 2007. Using the logarithm of odds to define a vector space on probabilistic atlases. Medical Image Analysis 11, 465–477.

Polak, E., Ribiere, G., 1969. Note sur la convergence de méthodes de directions conjuguées. Revue franccaise d'informatique et de recherche opérationnelle, série rouge 3, 35–43.

Politi, M.C., Han, P.K., Col, N.F., 2007. Communicating the uncertainty of harms and benefits of medical interventions. Medical Decision Making 27, 681–695.

Rajagopal, V., Nielsen, P.M., Nash, M.P., 2010. Modeling breast biomechanics for multi-modal image analysissuccesses and challenges. Wiley Interdisciplinary Reviews: Systems Biology and Medicine 2, 293–304.

Rasmussen, C., Williams, C., 2006. Gaussian Processes for Machine Learning. Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, USA.

Rasmussen, C.E., 2006. Gaussian processes for machine learning .

Rasmussen, C.E., Nickisch, H., 2010. Gaussian processes for machine learning (gpml) toolbox. The Journal of Machine Learning Research 11, 3011–3015.

Rencher, A.C., 2003. Methods of multivariate analysis. volume 492. John Wiley & Sons.

Risholm, P., Pieper, S., Samset, E., Wells III, W.M., 2010. Summarizing and visualizing uncertainty in non-rigid registration, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 554–561.

Rohr, K., Stiehl, H.S., Sprengel, R., Buzug, T.M., Weese, J., Kuhn, M., 2001. Landmark-based elastic registration using approximating thin-plate splines. IEEE Transactions on medical imaging 20, 526–534.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.

Saad, A., Hamarneh, G., Moller, T., 2010a. Exploration and visualization of segmentation uncertainty using shape and appearance prior information. IEEE Transactions on Visualization and Computer Graphics 16, 1366–1375.

Saad, A., Möller, T., Hamarneh, G., 2010b. Probexplorer: Uncertainty-guided exploration and editing of probabilistic medical image segmentation, in: Computer Graphics Forum, Wiley Online Library. pp. 1113–1122.

Sabuncu, M.R., Yeo, B.T., Van Leemput, K., Fischl, B., Golland, P., 2010. A generative model for image segmentation based on label fusion. IEEE transactions on medical imaging 29, 1714–1729.

Sakhaee, E., Entezari, A., 2016. A statistical direct volume rendering framework for visualization of uncertain data. IEEE transactions on visualization and computer graphics 23, 2509–2520.

Sanyal, J., Zhang, S., Bhattacharya, G., Amburn, P., Moorhead, R., 2009. A user study to compare four uncertainty visualization methods for 1d and 2d datasets. IEEE transactions on visualization and computer graphics 15, 1209–1218.

Schlegel, S., Korn, N., Scheuermann, G., 2012. On the interpolation of data with normally distributed uncertainty for visualization. Visualization and Computer Graphics, IEEE Transactions on 18, 2305–2314.

Schütze, H., Manning, C.D., Raghavan, P., 2008. Introduction to information retrieval. volume 39. Cambridge University Press Cambridge.

Shi, W., Jantsch, M., Aljabar, P., Pizarro, L., Bai, W., Wang, H., ORegan, D., Zhuang, X., Rueckert, D., 2013. Temporal sparse free-form deformations. Medical image analysis 17, 779–789.

vSkrinjar, O., Nabavi, A., Duncan, J., 2002. Model-driven brain shift compensation. Medical Image Analysis 6, 361–373.

Sotiras, A., Davatzikos, C., Paragios, N., 2013. Deformable medical image registration: A survey. IEEE transactions on medical imaging 32, 1153–1190.

Stytz, M.R., Parrott, R.W., 1993. Using kriging for 3d medical imaging. Computerized Medical Imaging and Graphics 17, 421–442.

Terrell, G.R., Scott, D.W., 1992. Variable kernel density estimation. The Annals of Statistics , 1236–1265.

Thévenaz, P., Blu, T., Unser, M., 2000a. Image interpolation and resampling. Handbook of medical imaging, processing and analysis , 393–420.

Thévenaz, P., Blu, T., Unser, M., 2000b. Interpolation revisited [medical images application]. IEEE Transactions on medical imaging 19, 739–758.

Torsney-Weir, T., Saad, A., Moller, T., Hege, H.C., Weber, B., Verbavatz, J.M., Bergner, S., 2011. Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration. IEEE Transactions on Visualization and Computer Graphics 17, 1892–1901.

Turk, G., O'brien, J.F., 1999. Variational implicit surfaces. Technical Report. Georgia Institute of Technology.

Umeyama, S., 1991. Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on pattern analysis and machine intelligence 13, 376–380.

Vandemeulebroucke, J., Rit, S., Kybic, J., Clarysse, P., Sarrut, D., 2011. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. Medical physics 38, 166–178.

Wachinger, C., Golland, P., Reuter, M., Wells, W., 2014. Gaussian Process Interpolation for Uncertainty Estimation in Image Registration. Springer International Publishing, Cham. pp. 267–274. doi:10.1007/978-3-319-10404-1_34.

Wang, Y., Georgescu, B., Chen, T., Wu, W., Wang, P., Lu, X., Ionasec, R., Zheng, Y., Comaniciu, D., 2013. Learning-based detection and tracking in medical imaging: a probabilistic approach, in: Deformation Models. Springer, pp. 209–235.

Warfield, S.K., Zou, K.H., Wells, W.M., 2004. Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. IEEE transactions on medical imaging 23, 903–921.

Warfield, S.K., Zou, K.H., Wells, W.M., 2006. Validation of image segmentation by estimating rater bias and variance, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 839–847.

Warfield, S.K., Zou, K.H., Wells, W.M., 2008. Validation of image segmentation by estimating rater bias and variance. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 366, 2361–2375.

Wassermann, D., Toews, M., Niethammer, M., Wells III, W., 2014. Probabilistic diffeomorphic registration: Representing uncertainty, in: International Workshop on Biomedical Image Registration, Springer. pp. 72–82.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biometrics bulletin 1, 80–83.

Williams, O., Fitzgibbon, A., 2007. Gaussian process implicit surfaces. Gaussian Proc. in Practice , 1–4.

Wittenbrink, C.M., Pang, A.T., Lodha, S.K., 1996. Glyphs for visualizing uncertainty in vector fields. IEEE transactions on Visualization and Computer Graphics 2, 266–279.

Wörz, S., Rohr, K., 2008. Physics-based elastic registration using non-radial basis functions and including landmark localization uncertainties. Computer Vision and Image Understanding 111, 263–274.

Zehner, B., Watanabe, N., Kolditz, O., 2010. Visualization of gridded scalar data with uncertainty in geosciences. Computers & Geosciences 36, 1268–1275.

Zhang, H., Banovac, F., Lin, R., Glossop, N., Wood, B.J., Lindisch, D., Levy, E., Cleary, K., 2006. Electromagnetic tracking for abdominal interventions in computer aided surgery. Computer Aided Surgery 11, 127–136.

Zhang, X., Günther, M., Bongers, A., 2010. Real-time organ tracking in ultrasound imaging using active contours and conditional density propagation, in: Medical Imaging and Augmented Reality. Springer, pp. 286–294.

Zhao, Q., Pizer, S., Alterovitz, R., Niethammer, M., Rosenman, J., 2017. Orthotropic thin shell elasticity estimation for surface registration, in: International Conference on Information Processing in Medical Imaging, Springer. pp. 493–504.

Zhou, X.S., Comaniciu, D., Xie, B., Cruceanu, R., Gupta, A., 2004. A unified framework for uncertainty propagation in automatic shape tracking, in: Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, IEEE. pp. I–872.

Zhou, X.S., Gupta, A., Comaniciu, D., 2005. An information fusion framework for robust shape tracking. Pattern Analysis and Machine Intelligence, IEEE Transactions on 27, 115–129.