Information Fusion for Prostate Brachytherapy Planning

by

Saman Nouranian

B.Sc., Amirkabir University of Technology, 1999
M.Sc., University of Tehran, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

June 2016

Abstract

Low-dose-rate prostate brachytherapy is a minimally invasive treatment approach for localized prostate cancer. It takes place in one session by permanent implantation of several small radio-active seeds inside and adjacent to the prostate. The current procedure at the majority of institutions requires planning of seed locations prior to implantation from transrectal ultrasound (TRUS) images acquired weeks in advance. The planning is based on a set of contours representing the clinical target volume (CTV). Seeds are manually placed with respect to a planning target volume (PTV), which is an anisotropic dilation of the CTV, followed by dosimetry analysis. The main objective of the plan is to meet clinical guidelines in terms of recommended dosimetry by covering the entire PTV with the placement of seeds. The current planning process is manual, hence highly subjective, and can potentially contribute to the rate and type of treatment related morbidity.

The goal of this thesis is to reduce subjectivity in prostate brachytherapy planning. To this end, we developed and evaluated several frameworks to automate various components of the current prostate brachytherapy planning process. This involved development of techniques with which target volume labels can be automatically delineated from TRUS images. A seed arrangement planning approach was developed by distributing seeds with respect to priors and optimizing the arrangement according to the clinical guidelines. The design of the proposed frameworks involved the introduction and assessment of data fusion techniques that aim to extract joint information in retrospective clinical plans, containing the TRUS volume, the CTV, the PTV and the seed arrangement. We evaluated the proposed techniques using data obtained in a cohort of 590 brachytherapy treatment cases from the Vancouver Cancer Centre, and compare the automation results with the clinical gold-standards and previously delivered plans. Our results demonstrate that data fusion techniques have the potential to enable automatic planning of prostate brachytherapy.

Preface

This thesis is primarily based on two journal publications and five conference papers, resulting from the collaboration between multiple researchers. The author was responsible for development, implementation and evaluation of the method and the production of the manuscripts. All co-authors have contributed to the editing of the manuscripts and providing feedback and comments. The research conducted in this study was undertaken under the approval of the UBC BCCA Research Ethics Board, certificate number H13-01983 titled *"Information Fusion Technology to Optimize Brachytherapy Procedure"*.

The study from Chapter 2 is presented at:

- S. Nouranian, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Multi-atlas-based Automatic 3D Segmentation for Prostate Brachytherapy in Transrectal Ultrasound Images, in Proceedings of SPIE Medical Imaging, 2013, vol. 8671, pp. 86710O1–7.

- S. Nouranian, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, An Automatic Multi-atlas Segmentation of the Prostate in Transrectal Ultrasound Images Using Pairwise Atlas Shape Similarity, in Medical Image Computing and Computer-Assisted Intervention: MICCAI, 2013, vol. 8150, pp. 173-180.

An extended version of the proposed methodology from Chapter 2 validated on a larger dataset has been published in:

- S. Nouranian, S. S. Mahdavi, I. Spadinger, W. Morris, S. Salcudean, and P. Abolmaesumi, A Multi-Atlas-Based Segmentation Framework for Prostate Brachytherapy, IEEE Trans. Med. Imaging, vol. 34, no. 4, pp. 950-961, 2015.

Dr. Mahdavi provided a dataset of previously studied TRUS volumes and segmentation results of the semi-automatic approach part of the current clinical software package. Scientific inputs of Prof. Abolmaesumi and Prof.

Salcudean helped with development and implementation of the methodology. Dr. Spadinger and Dr. Morris provided advice on clinical evaluation approaches. All authors helped to improve the manuscript structure by their valuable comments and feedback.

Studies presented in Chapter 3 are presented at:

- S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Fast Prostate Segmentation for Brachytherapy based on Joint Fusion of Images and Labels, in Proceedings of SPIE Medical Imaging, 2014, vol. 9036, pp. 90361A1-7.

- S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Data Fusion for Planning Target Volume and Isodose Prediction in Prostate Brachytherapy, in Proceedings of SPIE Medical Imaging, vol. 9415. pp. 94151I1-7, 2015.

Dr. Mahdavi provided the dataset used in evaluation of the presented method in these two papers. Prof. Abolmaesumi, Prof. Salcudean and Dr. Ramezani have contributed to the development and improvement of the methodology. Dr. Spadinger and Dr. Morris provided advice on clinical evaluation approaches. All authors helped to improve the manuscript structure by their valuable comments and feedback.

A study described in Chapter 4 has been published in:

- S. Nouranian, M. Ramezani, I. Spadinger, W. Morris, S. Salcudean, and P. Abolmaesumi, Learning-based Multi-label Segmentation of Transrectal Ultrasound Images for Prostate Brachytherapy, IEEE Trans. Med. Imaging, vol. 35, no. 3, pp. 921-932, 2016.

Prof. Abolmaesumi, Prof. Salcudean and Dr. Ramezani provided scientific input for development and improvement of the methodology. Dr. Spadinger and Dr. Morris provided advice on clinical evaluation approaches. All authors helped to improve the manuscript structure by their valuable comments and feedback.

A study described in Chapter 5 is presented at:

- S. Nouranian, M. Ramezani, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Automatic Prostate Brachytherapy Preplanning Using Joint Sparse Analysis, in Medical Image Computing and Computer-Assisted Intervention: MICCAI, 2015, vol. 9350, pp. 415–423.

Prof. Abolmaesumi, Prof. Salcudean and Dr. Ramezani provided scientific input for development and improvement of the methodology. Dr. Spadinger and Dr. Morris provided advice on clinical evaluation approaches. All authors helped to improve the manuscript structure by their valuable comments and feedback.

Table of Contents

List of Tables

List of Figures

Glossary

AAM Active Appearance Model.

ABS American Brachytherapy Society.

BCCA British Columbia Cancer Agency.

CT Computed Tomography.

CTV Clinical Target Volume.

DRE Digital Rectal Examination.

DSC Dice Similarity Coefficient.

GPU Graphical Processing Unit.

HD Hausdorff Distance.

ICA Independent Component Analysis.

ICRU International Commission of Radiation Units.

jICA joint Independent Component Analysis.

K-SVD K-Singular Value Decomposition.

KSVM Kernel Support Vector Machines.

LDR Low Dose Rate.

LOP Logarithmic Opinion Pool.

LWMV Locally Weighted Majority Voting.

MAD Mean Absolute Radial Distance.

MAP Maximum A Posteriori.

MAS Multi-Atlas Segmentation.

MAXD Maximum Absolute Radial Distance.

mPD Minimum Prescribed Dose.

MR Magnetic Resonance.

MSD Mean Surface Distance.

OR Operating Room.

PCA Principal Component Analysis.

PSA Prostate Specific Antigen.

PTV Planning Target Volume.

RBM Restricted Bultzman Machines.

SA Simulated Annealing.

SIMPLE Selective and Iterative Method for Performance Level Estimation.

SSD Sum of Squared Differences.

SSM Statistical Shape Model.

STAPLE Simultaneous Truth and Performance Level Estimation.

TRUS Transrectal Ultrasound.

VCC Vancouver Cancer Centre.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to Professor Purang Abolmaesumi, my advisor, for the opportunities you gave to me. Thank you for your trust, patience and invaluable support that opened this chapter of my life. Thanks for the generosity, advice and friendship you offered me throughout my experience at UBC.

I am very grateful for being associated with Drs. James Morris and Ingrid Spadinger, whose insightful advices helped me throughout the course of this endeavor. Their patience, dedication and sincere advice in every single step of the thesis have been truly crucial. Thanks for being always supportive and enthusiastic.

I would like to thank supervisory committee Professors Salcudean and Rohling for their valuable comments and advices.

I would like to particularly thank Dr. Mahdi Ramezani for his not only scientific support but also friendship and enthusiasm shared with me during years I started at UBC.

I would like to thank my dearest friend Dr. Mehdi Moradi for his unconditional support and friendship since I started my academic journey. My sincerest thanks to all current and past RCL members at UBC, Vickie, Abtin, Farhad, Siavash, Alex, Jeff, Parmida, Shekoofeh, Rohit, Greg, Philip and many others.

I owe immensely my true friend and wife Samira for her invaluable support in all small and big steps I've taken during this endeavor. And deep my heart, I thank my mom, dad and brother who constantly supported and inspired me to start this journey. You will never be forgotten.

Dedication

I dedicate this to my father who left fingerprints of grace on my life; my mother and my wife without whom none of my success would be possible ...

Chapter 1

Introduction

1.1 Brachytherapy

Prostate cancer is one of the frequently diagnosed types of cancer in North America. According to American and Canadian Cancer Societies, it has affected the life of at least one in six men in North America in 2013 [2, 3]. Recent advances in diagnosis and treatment methods proposed for prostate cancer have shown a significant increase in the rate of survival; e.g. the Canadian Cancer Society reported five-year relative survival ratio of more than 95% for patients diagnosed between 2006 and 2008 comparing to the 86% for the range of 1992 to 1994 [3]. Prostate cancer is routinely detected by Digital Rectal Examination (DRE) and measurement of the Prostate Specific Antigen (PSA). Although PSA level is often elevated in the presence of cancer, reliable diagnosis leading to staging of the cancer is usually made by pathological analysis of the sample tissues collected from the prostate core using biopsy needles under guidance of transrectal ultrasound (TRUS) images.

Several treatment methods have been proposed to eliminate or remove the cancerous tissue or the entire gland, while minimizing unnecessary morbidity and better preserving patients' quality of life (see Figure 1.1). Low-dose-rate (LDR) prostate brachytherapy is an effective treatment option for localized prostate cancer with low and intermediate risk disease (based on National Comprehensive Cancer Network[1] guidelines). It is associated with a good long-term survival rate [4, 5], requires less patient hospitalization when compared to radical prostatectomy, and is delivered in one session as opposed to alternative methods such as external beam therapy that need to be administered over a period of multiple sessions.

In LDR brachytherapy, small radioactive seeds are loaded as multi-seed "trains" into several needles in predefined patterns. They are permanently implanted through the perineum into the prostate and its adjacent tissue using a grid template (see Figure 1.4). Throughout this thesis, we refer to

[1] http://www.nccn.org

the LDR brachytherapy treatment approach that has been developed in the British Columbia Cancer Agency (BCCA) [6]. This program is founded in 1997 by radiation oncologists and medical physicists at the BCCA. Treatment protocol and guidelines are a combination of an in-house manual seed planning algorithm and evolved version of the Seattle preplanning experience [7]. According to the BCCA guidelines, a typical brachytherapy procedure can be divided into four main steps, usually taking place days apart; *volume study*, *seed planning*, *implantation*, and *post-implant analysis*.

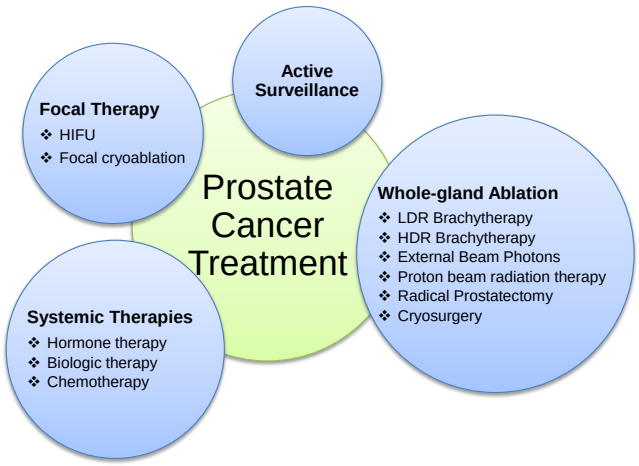


Figure 1.1: Prostate cancer treatment methods.

The volume study step involves determination of the Clinical Target Volume (CTV) or the target eradicable anatomy boundary. Standard-of-care in this process is to use an ultrasound machine and a transrectal ultrasound probe to acquire a sparse volumetric representation of the gland by a set of 5 mm apart 2-D parallel images from the base (bladder) to the apex (pelvic floor). Subsequently, all these 2-D images are contoured by an expert radiation oncologist to provide the CTV, which closely but not necessarily exactly follows the prostate boundary (Figure 1.2). CTV cannot be accurately defined for an individual patient; but is meant to contain all cancerous anatomy that requires to be adequately treated.

In seed planning step, the number and distribution of seeds and needles is determined based on some dosimetry analysis. First, margins are added to the CTV, to obtain the Planning Target Volume (PTV). PTV is determined to account for uncertainties of delivering the planned dose to the CTV. Generally, the dilation parameters are suggested in clinical guide-

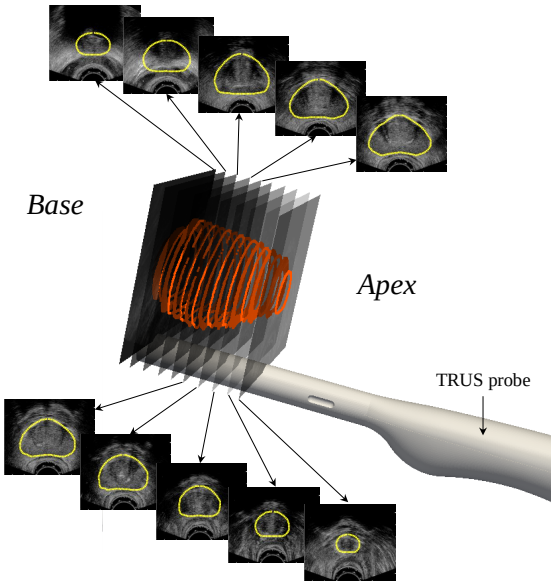


Figure 1.2: Volume study in brachytherapy requires prostate boundary delineation in a series of equally-spaced parallel 2-D transrectal ultrasound images.

lines. Figure 1.3 illustrates the relation between CTV and PTV based on the guidelines used at the Vancouver Cancer Centre (VCC) and the BCCA. A predefined set of margins are used to spare the CTV volume by 0.5 cm superiorly, inferiorly and laterally, 0.3 cm anteriorly and 0.0 cm posteriorly. After this anisotropic dilation of the CTV, the shape and size of the PTV may then be further altered slice by slice as needed to improve the compromise between dose coverage and dose conformality. These guidelines recommend PTV determination, seed placement and needle configurations to be symmetric across the patient mid-lobe.

In the procedure day, patient is anaesthetised and positioned to the same posture and condition as the volume study day. Radiation oncologist follows the preplanned seed map to perform implantation and to deliver minimum prescribed dose (mPD) to the CTV accurately.

1.2 Uncertainties in LDR Brachytherapy

Brachytherapy procedure outcome and morbidity rate is assessed by patient follow up records and recurrence state, months and years after treatment. Like other prostate cancer treatment methods, adjacency of urinary

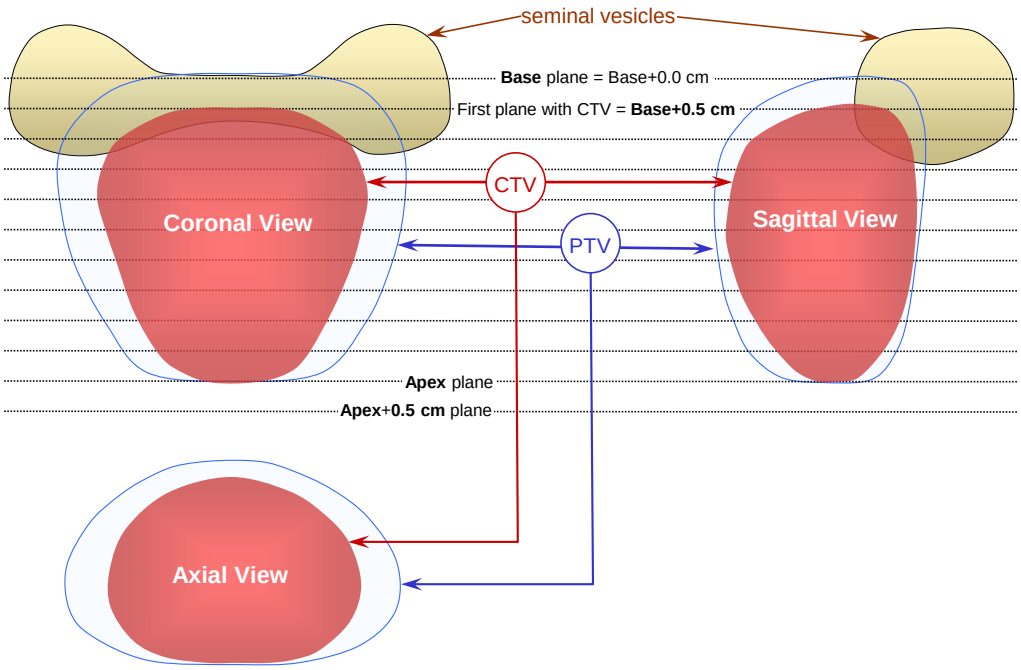


Figure 1.3: Planning target volume (PTV) is anisotropic dilation of clinical target volume (CTV) to account for implantation variability.

The significance of the intra-operative implantation uncertainty has led the American Brachytherapy Society (ABS) to recommend post-implant analysis for every patient [8, 9]. CT imaging is unanimously recommended as the standard protocol to detect seed positions in order to reconstruct the dose [10]. It is recommended that CT imaging should take place at approximately one month post-implant, at which time post-operative swelling has largely resolved [11, 12]. Alternatively, CT imaging may be acquired immediately after the procedure, when there may be significant edema, in order to obtain more timely feedback on implant quality. As part of the standard-of-care at BCCA, day-0 CT is acquired right after the patient post-anaesthetic recovery.

Implantation quality is routinely assessed on the basis of the post-implant D_{90} and V_{100} parameters. The D_{90} parameter is defined as the minimum dose in Gy received by 90% of the prostate volume and the V_{100} parameter is

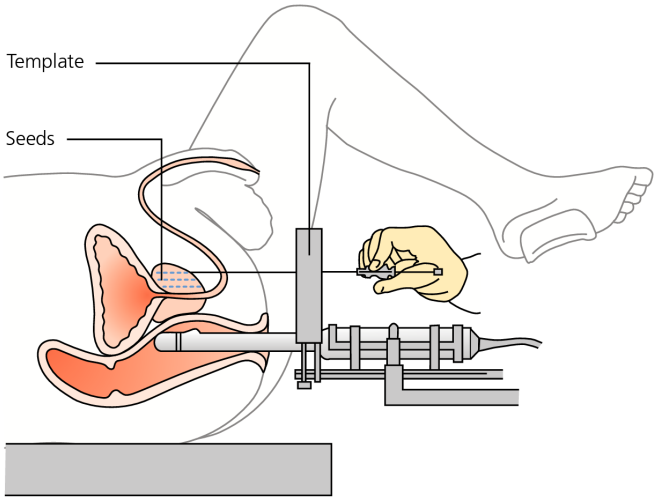


Figure 1.4: Schematic of a prostate brachytherapy implantation procedure[2].

percentage of the prostate volume receiving at least the minimum prescribed dose. Based on the criteria defined for these two parameters, implantation quality is classified as *excellent*, *good* or *suboptimal*.

In addition to the post-implant analysis based on CT, brachytherapy treatment outcome is assessed by periodic monitoring of PSA and testosterone levels over the course of several years after the procedure. Common side effects such as rectal wall and urinary toxicity are also monitored.
and erectile nerve bundles to the cancerous tissue causes prevalent post-procedural complications such as urinary incontinence, erectile dysfunction and rectal wall toxicity. However, this rate of morbidity varies from patient to patient.

Although some gene-related and physiological reasons play a key role in developing recurrent disease [13], brachytherapy outcome is affected by several obstacles to accurately deliver the intended dose to the target anatomy. Treatment-related sources of uncertainty can be divided into two groups: 1) uncertainties associated with pre-operative tasks in determining the target volume and the seed plan (preplanning); and 2) uncertainties stemming from intra-operative tasks to deliver the plan.

Prostate brachytherapy preplanning, as defined at the BCCA, involves

[2]Anatomic schematic is retrieved from the book "Fast Facts: Prostate Cancer" (7^th edition) by Kirby, Roger.

TRUS volume collection, CTV and PTV delineation, and determination of the seed arrangement w.r.t. the PTV. Accurate determination of the CTV is desirable in order to deliver sufficient radiation dose to the prostate while minimizing dose to the urethra and the surrounding tissues, such as bladder and rectum. CTV segmentation error, in combination with other inherent sources of error (e.g. seed delivery error), can produce unnecessary morbidity, such as rectal wall toxicity, if the CTV is overestimated posteriorly, or lead to under-treatment, if the CTV is underestimated.

CTV delineation is a cumbersome task that requires drawing of a contour for each 2-D TRUS image that mainly but not necessarily follows the prostate boundaries. Typically, a volume study image set consists of 7-14 ultrasound images at 5 mm spaced axial planes. These images are often affected by speckles, shadowing and reverberation artifacts, and the boundary of the prostate is not reliably visible, especially in the base and the apex. These characteristics of the TRUS images, in conjunction with the current CTV delineation process which is either manual or semi-automatic, make delineation of the CTV a tedious task and vulnerable to subjective errors [1]. Eliminating user interaction will also provide the means for real-time dosimetry, where planning the seed placement can be corrected in the operating room during the brachytherapy procedure to account for anatomical changes due to patient positioning, internal organ fillings and edema as well as seed placement error.

Definition of PTV is important because it is used as a destination where seeds are usually distributed. PTV is defined to account for intra-operative errors by anisotropic dilation of the CTV. Although the dilation parameters are generally suggested in clinical guidelines, a radiation oncologist normally modifies the contours in different anatomical regions to account for some implicit characteristics of the patient review chart, and his/her own expertise. The standard-of-care at the VCC is to use a manual or semi-automatic approach in contouring the target volumes, which highly depends on the expertise of a radiation oncologist.

The need for CTV and PTV delineation is not limited to the preplanning application. Ideally, segmentation correction or plan refinement just prior to or during implantation facilitates the treatment objective: Prostate shape and size varies due to bladder and rectum fillings, cancer evolution rate and characteristics, patient conditions, the effect of general anesthesia, prostate deformation and the development of edema during the procedure. Hence, there is a high demand for real-time target anatomy segmentation in prostate brachytherapy for clinical applications such as intraoperative preplanning, where the plan is developed in the operating room immediately

prior to implantation, or intraoperative planning in which plan is corrected as the needles are inserted. According to the ABS recommendations for prostate dosimetry, *"Ideally, one should strive for on-line, real-time intra-operative dosimetry to allow for adjustment in seed placement to achieve the intended dose"* [10]. However, the acknowledged limitations of the current target volume delineation process make intraoperative real-time segmentation challenging. A fast and automatic delineation algorithm can potentially lead to a more efficient and less user-dependent target segmentation, which in turn will produce more consistent dosimetric treatment plans while facilitating potential solutions for real-time dosimetry.

Inevitably, operating room (OR) related user variability introduces a significant amount of uncertainty to the whole treatment procedure, since volume study and preplanning take place weeks before the implantation. Deviation between planned and delivered dose mainly rises from three major sources:

- Prostate visibility must be the same as the volume study day. Therefore, patient, brachytherapy grid template, and the probe setup need to be replicated from the day TRUS volume is acquired. However, internal organ condition is not necessarily the same. Pelvic internal muscle tension changes when patient is under general anaesthesia. Moreover, deformation and movement forced on the anatomy by the TRUS probe is not reproducible completely.

- There are some technical challenges associated with delivery of the plan due to needle bending, friction, and deformation of the anatomy by the force applied through needle insertion.

- The prostate boundary is not necessarily the same as the volume study day because of the organ boundary changes. These changes are due to cancer evolution or some medications taken meanwhile. Bladder and rectum fillings may also affect the anatomy on the operation day, although patients are instructed to empty both.

Given the aforementioned sources of variability, brachytherapy treatment outcome and morbidity can benefit from optimization approaches that reduce or eliminate sources of uncertainty from parts or the whole procedure. For instance, automation of the delineation processes eliminates user interaction, however it needs to be capable of incorporating implicit knowledge of radiation oncologists as well as clinical guidelines. A solution to such optimization is to model the brachytherapy procedure as an integrated system

of interacting processes. This model can be used to optimally predict elements of each individual procedure. The proposed model can be generated from previous successful treatment records and applied to future records.

1.3 Background

Review of the prior art in optimization (automation) of the prostate brachytherapy procedure can be divided into two sections: 1) segmentation of the target volume in TRUS images; and 2) automation of the seed planning inside the PTV.

1.3.1 Target Volume Segmentation

Although there has been a lot of work on segmentation of the prostate in TRUS images, they are mostly not directly applicable to segmentation of the CTV and the PTV. This is further discussed in Chapter 4. To the best of our knowledge, there is no work on automatic delineation of the PTV from TRUS images. Hence, in this section we review the closest works, i.e. prostate segmentation for TRUS images.

Several groups have proposed semi-automatic and automatic techniques to alleviate the manual prostate segmentation process in TRUS images. It has been shown that pure texture features of the TRUS images can aid classification of the pixels and delineation of the prostate boundary [14, 15]. Active contours and snakes have also been utilized by several groups to delineate the prostate boundary in 2-D [16, 17] and 3-D [18–22] TRUS images.

Another group of algorithms take advantage of *a priori* knowledge of the prostate shape and TRUS image statistics to improve segmentation results. For example, geometrical model-based approaches using super-ellipses have been recognized as a suitable modelling approach for determining the prostate boundary [1, 23, 24]. One of the key advantages of the super-ellipses-based segmentation approaches is the strong shape regularization that compensates for the relatively low resolution and sparse characteristics of the brachytherapy TRUS volumes. In addition, this approach meets the requirements defined in the BCCA treatment protocol about symmetric CTVs. However, the proposed approaches are semi-automatic, hence, vulnerable to subjective errors and initial conditions. Although fusion of different imaging modalities has been proposed to automate the super-ellipses-based segmentation pipeline [25], it is still far from being integrated

with current standards of care. Moreover, prior information of TRUS images, such as texture features and prostate shape variability in the form of atlas or Statistical Shape Models (SSM) [26–31] have been used to improve robustness of segmentation. For example, Kernel Support Vector Machines (KSVM) [32, 33], Active Appearance Models (AAM) [34] and level sets [35, 36] are some of different statistical modeling approaches proposed to incorporate shape and intensity priors into a robust to noise segmentation algorithm. In an atlas-based segmentation approach, a statistical model, i.e. atlas, is usually transformed into a common coordinate system with the target image. Subsequently, prior knowledge is propagated from the atlas towards the target image. Multi-atlas-based methods are based on intensity image registration and label fusion techniques and generally have not been used in the context of TRUS images due to the poor ultrasound image registration quality. This problem also affects process of selection and fusion of atlases that keeps it challenging for the TRUS images. Therefore, single or multiple atlas-based segmentation approaches are proposed mainly for segmentation of the prostate in CT and MR images [37–41].

Except [1, 24–26], the above mentioned TRUS segmentation methods are not directly applicable to prostate brachytherapy without significant changes to the underlying methods or clinical workflow, because: 1) The delineated CTV in clinical images does not follow the anatomical prostate boundary everywhere; 2) TRUS volumes are sparsely acquired with a slice distance of 5 mm. This sparse nature of brachytherapy TRUS images in addition to the intrinsic artifacts of the ultrasound images, makes the segmentation process even more challenging. Mahdavi *et al.* [25] combined information from another modality, i.e. elastography maps, to aid the segmentation process, however it requires the introduction of additional hardware or the need for special ultrasound machines for elasticity imaging. Ghose *et al.* [26] proposed a solution for automatic prostate segmentation that appears to be applicable for prostate brachytherapy. However, the approach is based on three independent 2-D active shape and appearance models that are generated only for central, apex and base zones; hence, segmentation is inherently 2-D and does not guarantee a smooth CTV delineation that considers the 3-D shape of the prostate.

A more general approach using atlases is the multi-atlas segmentation method that is primarily used for segmentation of MR brain images [42–44], and more recently for CT and MR images of prostate [37, 38, 41]. Multi-atlas-based methods are based on intensity image registration and label fusion techniques and generally have not been used in the context of TRUS images due to the poor ultrasound image registration quality. This problem

also affects process of selection and fusion of atlases that keeps it challenging for the TRUS images.

1.3.2 Seed Arrangement Planning

The main objective in seed planning is to target the PTV with minimum prescribed dose (mPD). Seed arrangements can be planned either prior to, right before or during the implantation procedure. The BCCA program is based on prostate brachytherapy *preplanning*, which requires pre-operative planning of seed arrangements, usually weeks before the implantation procedure.

Over the last two decades, several solutions have been proposed to automate seed planning [45–52]. A common approach for this automation is to define an inverse problem based on dose constraints applied to parameters that characterize features such as target coverage, dose uniformity, and dose to organs at risk. In this regard, various optimization techniques are employed including simulated annealing [47], genetic algorithms [50, 52] and branch-and-bound solution for mixed-integer model [45, 48, 51, 53]. The initialization for these techniques is normally with a random seed arrangement. Considering the very large size of the search space for optimal seed configuration, these solutions are highly subject to local minima and sensitive to initial seed configurations.

1.4 Proposed Solution

A typical brachytherapy preplanning procedure can be subdivided into different processes as shown in Figure 1.5. Obviously, uncertainty propagates throughout the cascading processes and would affect quality of the treatment procedure, i.e. increasing rate of morbidity. Hence, current prostate brachytherapy procedure can substantially benefit from reducing the need for user interaction by automation of parts or the whole planning processes. In this thesis, we aim to facilitate pre-operative processes, where the current procedure is performed manually or semi-automatically.

Current preplanning procedure is sequential and takes place in a geometrical extent of a patient's TRUS volume. Hereafter, we refer to inputs and outputs of processes, i.e. TRUS intensity volume, the CTV and the PTV, as *volumetric information*. Since each process depends on the preceding processes and observed volumetric information, we propose to fuse all available volumetric information and learn their joint relation to devise

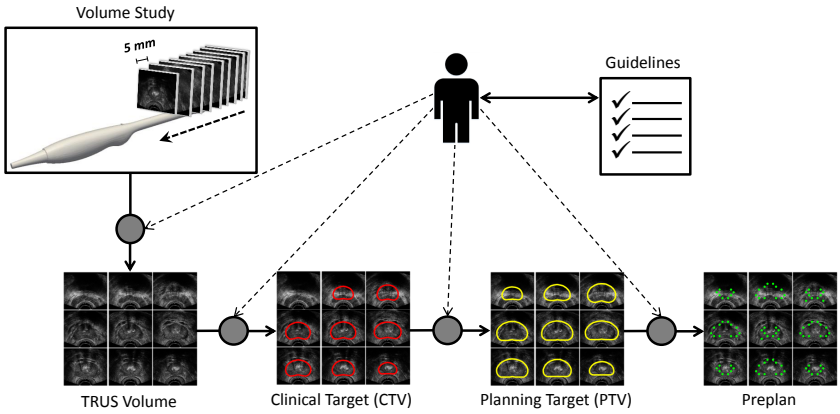


Figure 1.5: Current prostate brachytherapy preplanning requires user intervention as shown by circles.

an automation solution. Our approach is to combine and capture the interrelation between volumetric information generated from processes in the form of mathematical and statistical models.

Architecture and parameters of such models are tuned and optimized using previous treatment knowledge. We take advantage of the previous excellent outcome treatment records at the VCC in our analysis. We propose to represent joint relation between observations of the volumetric information by joint statistical analysis of these treatment records. To this end, we investigate some statistical modelling approaches such as independent component and sparsity analyses, all aiming to manifest the joint relation between volumetric information in a limited set of complex joint patterns.

1.4.1 Objective

The global objective of the thesis is to reduce the need for user interaction in planning prostate brachytherapy treatment. To this end, we investigate automation approaches in preplanning, namely clinical/planning target volume delineation in TRUS images and seed arrangement. We investigate a fusion system that automatically delineates the CTV with accuracy comparable to manual or semi-automatic methods currently used in clinics. Furthermore, we propose a fusion framework that extracts joint patterns between volumetric information representations of the preplanning elements, i.e. CTV and PTV contours. In this thesis, we investigate the feasibility of automatic seed planning by combining the joint learning-based approach with a novel

in-house optimization algorithm. Throughout this thesis, the joint analysis framework is generated and evaluated on a dataset of previously delivered successful treatment records obtained from the VCC.

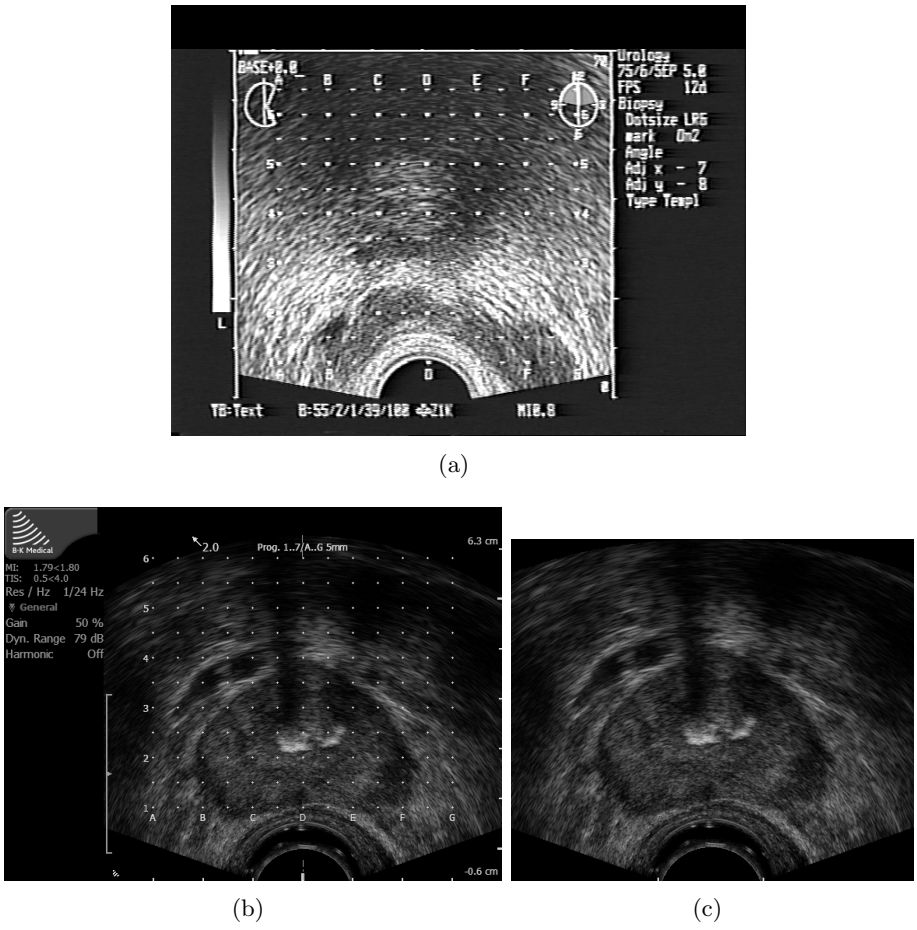
1.4.2 Contributions

The present thesis is an attempt to develop a fusion-based framework to automate prostate brachytherapy preplanning, according to the BCCA treatment program. The proposed model encapsulates variability associated with different elements of determining CTV and PTV as well as seed arrangement using existing prior knowledge in a dataset of treatment records. In the course of achieving this objective, the following contributions were made:

- Proposing a multi-atlas-based segmentation approach for automatic delineation of the CTV from TRUS images.
- Proposing an approach for fast and reliable estimation of the CTV, the PTV and the mPD contours using joint independent component analysis (jICA).
- Proposing a framework for simultaneous estimation of the CTV and the PTV using joint sparse analysis.
- Proposing a seed arrangement optimizer to enforce the BCCA prostate brachytherapy preplanning clinical guidelines.
- Proposing a novel learning-based seed planning framework.

1.5 Materials

Materials used in this dissertation are collected from the VCC under the approval of the University of British Columbia and BCCA Research Ethics Board, certificate number H13-01983. The target population for this study is composed of male patients who have undergone brachytherapy treatment. This research requires a retrospective dataset of brachytherapy patient records which includes the transrectal ultrasound images of the prostate before implantation of seeds, corresponding clinical target contours, planning target contours, seed placement plan and plan dosimetry parameters such as seed activation unit, radial and geometrical dose distribution functions. We looked at the anonymized and de-identified data records of patients who underwent treatment between May 1, 2006 and May 1, 2012. In total, 2,087 patient charts were reviewed.



(a)

(b) (c)

Figure 1.6: Brachytherapy grid was superimposed over the original ultrasound images. a) A sample case captured from a video monitoring device excluded from this study due to lack of an ultrasound recovery solution; and b) a sample image collected from BK machine before ultrasound image recovery; and c) after ultrasound image recovery.

Patient brachytherapy charts had been managed and archived using VariSeed^{TM}LDR Treatment Planning System. All ultrasound images were altered by the ultrasound machine software through superimposing a brachytherapy grid and annotations. As a result, a clean up process was necessary through image processing software to recover the original ultrasound data. This was a challenging task, since majority of the cases in the original cohort were captured from a video monitoring screen (some analog), hence,

had substantially degraded image quality. We focused on a smaller cohort of 590 cases where a BK ultrasound machine was used for imaging. Images from BK machines had sufficient quality to develop the recovery algorithm, and further, they have been recently widely used all across centres in British Columbia after 2010. For the BK machine images, a consistent procedure was implemented to remove superimposed information and recover the original ultrasound images with minimum loss. Figure 1.6 shows some of the sample images from the cohort of this study before and after ultrasound image recovery.

1.6 Thesis Outline

The rest of this thesis is subdivided into four chapters as outlined below:

Chapter 2: Atlas-based Fusion for Brachytherapy Preplanning

In this chapter, we introduce a multi-atlas fusion framework to automatically delineate the clinical target volume in ultrasound images. A dataset of a priori segmented ultrasound images, i.e. atlases, is registered to a target TRUS image. We introduce a pairwise atlas agreement factor that combines an image-similarity metric and similarity between a priori segmented contours. This factor is used in an atlas selection algorithm to prune the dataset before combining the atlas contours to produce a consensus segmentation. The proposed method produces segmentation results that are within the range of observer variability when compared to a semi-automatic segmentation technique that is routinely used at the VCC.

Chapter 3: ICA-based Fusion for Brachytherapy Preplanning

In search of a faster technique for fusion and estimation, in this chapter we introduce a joint ICA-based estimator for two different applications:

- Delineation of the CTV: Real-time dosimetry and intra-operative plan correction requires a fast CTV segmentation algorithm. In this chapter, we propose a computationally inexpensive and fully automatic segmentation approach that takes advantage of previously segmented images to form a joint space of images and their segmentations. We utilize joint independent component analysis method to generate a model which is further employed to produce a probability map of the target segmentation. We show that the proposed approach is fast with

comparable accuracy and precision to those found in previous studies on TRUS segmentation.

- Delineation of the PTV and the mPD simultaneously: We introduce a jICA-based model that enables joint determination of PTV and the minimum prescribed isodose (mPD) map by capturing the correlation between different volumetric information elements consisting of transrectal ultrasound (TRUS) volumes, PTV and isodose contours. Taking advantage of the same jICA technique, we obtain a set of joint components that optimally describe such correlation. We also perform a component stability analysis to generate a model with stable parameters that predicts the PTV and isodose contours solely based on a new patient TRUS volume.

Chapter 4: Sparsity-based Fusion for Brachytherapy Preplanning

Independent component analysis can be seen as a special case of a sparse modeling. Hence, in this chapter, we extend the ICA-based framework by targeting simultaneous delineation of the preplanning elements, i.e. the CTV and the PTV using sparse analyses. In this chapter, we aim to reduce the segmentation variability and planning time by proposing an efficient learning-based multi-label segmentation framework. We incorporate a sparse representation approach in our methodology to learn a dictionary of sparse joint elements consisting of images, and clinical and planning target volume segmentation. The generated dictionary inherently captures the relationships among elements, which also incorporates the institutional clinical guidelines. We show simultaneous segmentation results compared with the currently used clinical algorithm for both target volumes.

Chapter 5: Automatic Seed Plan Estimation

In this chapter, we aim to reduce the preplanning variability by automating the seed arrangement process. We propose a novel framework which uses a retrospective treatment dataset to extract common radioactive seed patterns. The framework captures the inter-relation between the treatment volume delineation and seed arrangements through a joint sparse representation of retrospective data. This representation is used to estimate an initial seed arrangement for a new treatment volume, followed by a novel optimization process which captures the clinical guidelines, to fine-tune the seed arrangement.

Chapter 6: Conclusion and Future Work

This chapter includes a short summary of the thesis followed by suggestions for future work.

Chapter 2

Atlas-based Fusion for Brachytherapy Preplanning

2.1 Introduction

In low-dose-rate brachytherapy, reliable CTV delineation in TRUS images is the foundation of generating appropriate seed plan to deliver sufficient radiation dose to the cancerous tissue. Current standard-of-care at the VCC requires user intervention in determination of the CTV contours. A state-of-the-art algorithm proposed by Mahdavi *et al.* [1] has been integrated into the current workflow that initiates the contours in 2-D planes. These contours are further manipulated based on expert's knowledge to ensure cancerous target is covered. This chapter is an effort to automate the CTV (approximately the prostate boundary) delineation. The main objective is to take advantage of previously contoured TRUS volumes and the state-of-the-art atlas-based segmentation approaches to facilitate prostate/CTV segmentation process.

In this chapter, we introduce an algorithm for 3-D segmentation of the CTV based on multiple atlases that include pairs of images and their segmented labels. The general approach of the multi-atlas segmentation (MAS) method is to transform atlases in an existing dataset to the coordinates of a target image. Subsequently, a consensus segmentation of the target image is produced by fusion of the atlas labels. This approach has been previously used for segmentation of brain MR images [42–44], and more recently for CT and MR images of the prostate [37, 38, 41, 55]. The application of MAS for TRUS segmentation is not straightforward, since it relies on robust registration of images, which is a challenging task in TRUS data. To improve the robustness to registration inaccuracy, several groups have proposed solutions to identify optimal weights for fusion, mainly by applying the fusion

This chapter is adapted from [54]: S. Nouranian, S. S. Mahdavi, I. Spadinger, W. Morris, S. Salcudean, and P. Abolmaesumi, A Multi-Atlas-Based Segmentation Framework for Prostate Brachytherapy, IEEE Trans. Med. Imaging, vol. 34, no. 4, pp. 950-961, 2015.

in a neighbourhood around each target voxel, and by introducing intensity priors to the fusion algorithms, e.g. locally weighted majority voting [56], local maximum a posteriori STAPLE [57] and local logarithmic opinion pool STAPLE [58]. Our observation from both intensity images and their corresponding label maps shows a weak correlation between intensity similarity metrics and label overlap measures in the context of sparse TRUS images. Hence, some recently introduced methods that incorporate intensity based priors in fusion of atlases may not perform as well in the context of our study. An alternative approach proposed by Langerak *et al.* [59] uses a preregistration atlas selection for multi-atlas segmentation. The approach maintains/improves the quality of input atlases based on a heuristic selection of atlases prior to the registration process. The authors observe slightly but not statistically significantly less accuracy for the target segmentation; however, the computational time was significantly reduced. In this chapter, we propose an alternative approach, where we perform atlas selection prior to and after the registration: 1) we use a target specific approach for atlas selection prior to the registration and fusion steps; 2) we introduce an atlas pruning technique to incorporate shape deformation agreement among atlases along with the registration performance. We incorporate a pairwise atlas agreement factor to select an appropriate subset of atlases for fusion, and subsequently produce the consensus segmentation. To satisfy CTV requirements according to the treatment protocol at our institution, consensus segmentation generated by the fusion process is smoothed and made symmetric by fitting tapered and warped ellipses. The method is evaluated on a clinical dataset of 280 prostate TRUS volumes. We compare our proposed configuration for the framework against the manually segmented gold-standard, and show it performs within a range of variability same as a state-of-the-art semi-automatic segmentation approach [23] that is part of the standard-of-care at the Vancouver Cancer Centre (VCC). The initial results of our approach was reported in [60, 61]; here, we provide further details of the approach, and validate it with a much larger dataset. Moreover, we propose to improve the robustness of the approach by introducing the atlas agreement factor. We also compare our proposed multi-atlas segmentation framework with *local* and intensity-driven label fusion techniques.

2.2 Methods

Figure 2.1 shows a schematic block diagram of the proposed MAS framework. In this framework, the term atlas refers to an intensity image and

its corresponding morphometric properties, such as its binary segmentation, which we hereafter refer to as the label image. Generally, MAS approaches assume a strong presence of corresponding structures in the intensity and the label image of an atlas. They use an image registration approach to align a group of atlases with a target image, so that the corresponding labels of the atlases are propagated to the coordinates of the target image. A probabilistic map of the consensus segmentation is achieved by fusing the transformed labels of the atlases.

The key assumption in MAS is that a perfect registration between an atlas and the target image leads to a perfect labeling of the target image by propagating the same transformations from the intensity image to the label image. Where there is imperfect registration, fusion of the information from multiple atlases reduces the sensitivity of the consensus segmentation to the registration quality and the error associated with each individual atlas label. In the context of TRUS images obtained for brachytherapy, registration is a particularly challenging problem, which could lead to relatively poor and spatially inconsistent results, due to the wide range of variability in shape and size of the prostate. Furthermore, fusion of atlas labels based on intensity criteria is highly affected by these specific characteristics of TRUS images which results in an inaccurate alignment of the the atlases and hence poor fusion performance. This implies a requirement to prune atlases before and after registration prior to the fusion step.

Hereafter, we refer to the target volume with V, and each atlas with pair $\{I_i, S_i\}$, $i \in \{1, 2, ..., M\}$, which represents the intensity volume I_i and the corresponding segmentation binary volume $S_i : \mathbb{R}^d \rightarrow \{0, 1\}$ for the i^{th} atlas from total number of M atlases in a d dimensional space.

2.2.1 Target-specific Atlas Selection

In the first step, we reduce the size of the atlas dataset using a target-specific pruning approach, where a smaller set of atlases that are similar to the target image are selected. We follow an iterative procedure to generate a mean atlas [39]. This method does not require inverting nonrigid coordinate transformations. Subsequently, we utilize the mean atlas to bring the target image, along with all atlas intensity images, to a common coordinate system. Afterwards, a similarity metric between the transformed target image and each individually transformed atlas is computed to generate a list of candidate atlases.

By computing the image similarity metric between each pair of transformed atlases, $\{\widetilde{I}_i^n, \widetilde{S}_i^n\}$, and the transformed target image, $\widetilde{V}$, a fixed num-

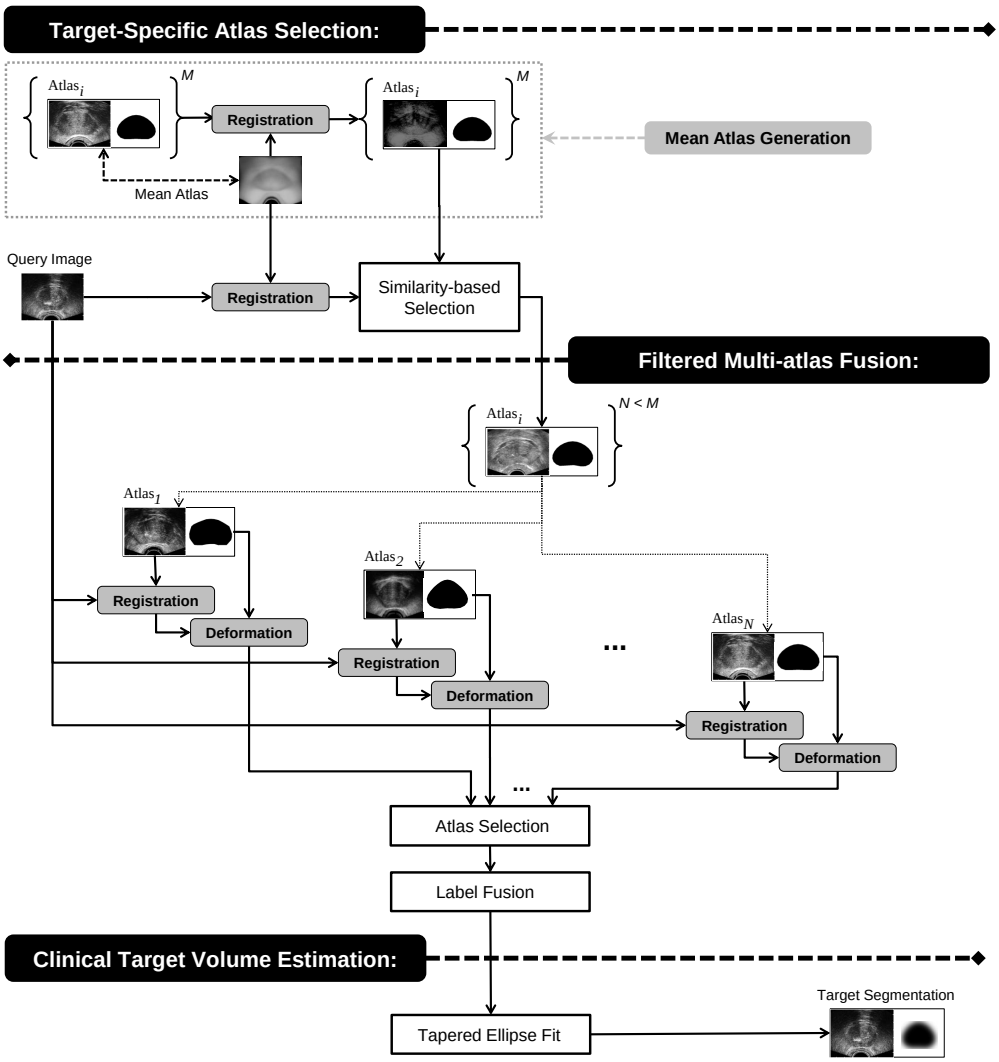


Figure 2.1: Block diagram of the proposed MAS-based TRUS segmentation framework.

ber of N atlases expected to have a higher resemblance to the target volume is selected. Notably, this process does not add a significant computational burden to the segmentation pipeline as the mean atlas, $\widetilde{I}^n$, and all transformed atlases are computed, once, prior to the segmentation process.

2.2.2 Filtered Multi-atlas Fusion

The second block of the proposed MAS framework consists of three cascade processes; namely, registration, atlas filtration and label fusion. For the registration, we use the diffeomorphic Demons registration algorithm [62–64], which has been shown to perform well in the context of ultrasound images [65, 66].

For the atlas-selection, a general consensus in the MAS literature [42, 44, 67, 68] is to use the image similarity metric as a measure to rank atlases. While this methodology may be effective in CT and MR images, in TRUS, due to speckles, shadowing and low signal-to-noise ratio, the sensitivity of the image similarity metric to anatomical variability in terms of shapes and sizes of the prostate is relatively low. To alleviate this issue, we propose a new similarity metric that jointly encodes not only the image similarity in the intensity domain, but also the prostate shape and size similarity in the label domain. We define a pair atlas agreement factor, $f_{i,j}$, between atlas i and j as the harmonic mean of the intensity and label distances:

$$f_{i,j} = \frac{2d^I_{i,j}d^S_{i,j}}{d^I_{i,j} + d^S_{i,j}}, \tag{2.1}$$

where the distance between intensity of the atlases is represented by Sum of Squared Differences (SSD), denoted by $d^I_{i,j}$. Also, the volume error, percentage of non-overlapping volume, is used for shape dissimilarity $d^S_{i,j}$. For each atlas, $\{\widetilde{I}_i, \widetilde{S}_i\}$, we define the pairwise atlas agreement vector $\mathbf{F}_i$ by calculating the agreement factors $f_{i,j}, j \in \{1, ..., N\}$ to all other transformed atlases. We follow a consistent indexing of atlases in all calculations:

$$\mathbf{F}_i = (f_{i,1}, f_{i,2}, ..., f_{i,N})\,. \tag{2.2}$$

The feature vectors $\mathbf{F}_i, i \in \{1, ..., N\}$ are divided into K clusters $\mathbf{G} \in \{G_1, ..., G_K\}$ using the k-means clustering algorithm. Intuitively, members of each cluster share a similar distance pattern with the other members. We utilize the cosine similarity between pairwise atlas agreement vectors as the distance metric for the clustering algorithm. The k-means algorithm minimizes the within cluster distance between members $\mathbf{F}_i$ and the centroid of each cluster $\mathbf{C}_1, ..., \mathbf{C}_K$:

$$\arg\min_{\mathbf{G}} \sum_{i=1}^{K} \sum_{\mathbf{F}_j \in G_i} (1 - \frac{\mathbf{F}_j.\mathbf{C}_k}{\|\mathbf{F}_j\|\|\mathbf{C}_k\|}). \tag{2.3}$$

To minimize the effect of initialization on the clustering algorithm, centroids of the clusters are randomly initiated many times and the best clustering result with the minimum summation of the distance to the centroids is selected. Subsequently, clusters are ranked based on their average image intensity similarity to the target image. We define the ranking parameter r_{G_i} for cluster G_i as the mean SSD between members of cluster Gi and the target image:

$$r_{G_i} = \frac{1}{|G_i|} \sum_{\widetilde{I} \in G_i} \sum_{\mathbf{x} \in \widetilde{I}} \left(\widetilde{I}(\mathbf{x}) - V(\mathbf{x}) \right)^2, \tag{2.4}$$

where $\mathbf{x}$ represents a voxel in the transformed atlas image $\widetilde{I}$.

Next, members of the top ranked cluster, G_T, are selected for the label fusion step. Transformations obtained from the registration step are propagated to the corresponding atlas labels, producing $\widetilde{S}_i, i \in \{1, ..., P\}$. This group of transformed labels are combined together to provide a consensus segmentation $\widehat{L}$ for the target volume V.

There are several label fusion strategies introduced to determine the optimal weight for each atlas label. A group of algorithms adapt a deterministic voting strategy, such as majority voting and globally weighted majority voting, and assign equal or adjusted weights based on a global atlas quality measure (e.g., the registration performance metrics). Another group of algorithms include a stochastic model within a statistical fusion strategy. Simultaneous Truth and Performance Level Estimation (STAPLE) [69] is one of the widely used statistical fusion approaches, which has been shown to perform well for MAS-based segmentation [44, 70]. STAPLE was originally developed to account for intra-rater variability by estimating an underlying ground-truth from multiple segmentations of the same object. To account for spatially inconsistent registration quality, locally weighted variations of the voting [56, 67] and local STAPLE algorithms [57, 58, 71] have been proposed. In our experiments, we evaluated different label fusion techniques using a fixed number of input atlases to produce a consensus segmentation.

2.2.3 Clinical Target Volume Estimation

According to the VCC treatment protocol [6], the CTV contours are desired to be symmetric and smooth. The treatment protocol instructs to draw symmetric contours with respect to the mid-sagittal plane. This is to minimize the effect of intra-operative image reproducibility error, that causes deviation from the pre-planned dose distribution. To accomplish this, we

follow the approach proposed by Mahdavi *et al.* [1] and fit a tapered ellipse to unwarped contours. The algorithm comprises the following steps: 1) obtaining the unwarping parameters based on anatomical landmarks, 2) unwarping the contours w.r.t. the unwarping parameter, 3) fitting a tapered ellipse to unwarped contours, and 4) warping the fitted tapered-ellipse back to the target coordinates. The required steps are shown in Figure 2.2. Three initialization landmarks are identified automatically for each slice from the segmentation result (i.e. STAPLE algorithm output) defined as 1) P_1: probe center, 2) P_4: mid-posterior and 3) P_2: lower posterior-lateral. The tapered-ellipse fit process is expressed as an optimization problem using the *l-bfgs* optimization algorithm [72]. The explained pipeline will generate a smooth and symmetric contour with respect to the sagittal plane for all 2-D TRUS planes.

2.3 Materials

A dataset of 280 TRUS volumes of brachytherapy patients is used for evaluation of the proposed segmentation algorithm. All images were acquired at the VCC, a few weeks before the treatment day. Data acquisition and planning procedures were carried out in accordance to the local treatment protocol [6]. Instructions in the protocol standardize all TRUS volumes to be trimmed from prostate base to apex, while the gland is visually located in the middle of each axial B-mode image. As mentioned earlier, each TRUS volume consists of 7 to 14 parallel equally spaced (5 *mm* apart) axial B-mode images of the prostate which are captured using a side-firing transrectal probe. For each B-mode image, the prostate gland is delineated using a software. This software is part of the standard of care, and works based on a state-of-the-art semi-automatic prostate segmentation method presented by Mahdavi *et al.* [1]. These contours are manually corrected afterwards by an expert clinician, and subsequently are used as the CTV reference contours for evaluation of the proposed algorithm.

Images are all preprocessed to remove overlaid device labels and marks, and are down-sampled by a factor of three (136×165 pixels) to accelerate the computational processes. The intensity range is normalized across all atlases w.r.t the target image by histogram matching prior to the atlas registration process. Subsequently, all atlases are geometrically aligned, and if necessary, resampled (with the nearest-neighbour technique) w.r.t. the target image extent. The extent of each individual volume is limited to the base and apex axially, and is standardized in two other axes (sagittal and coronal) among

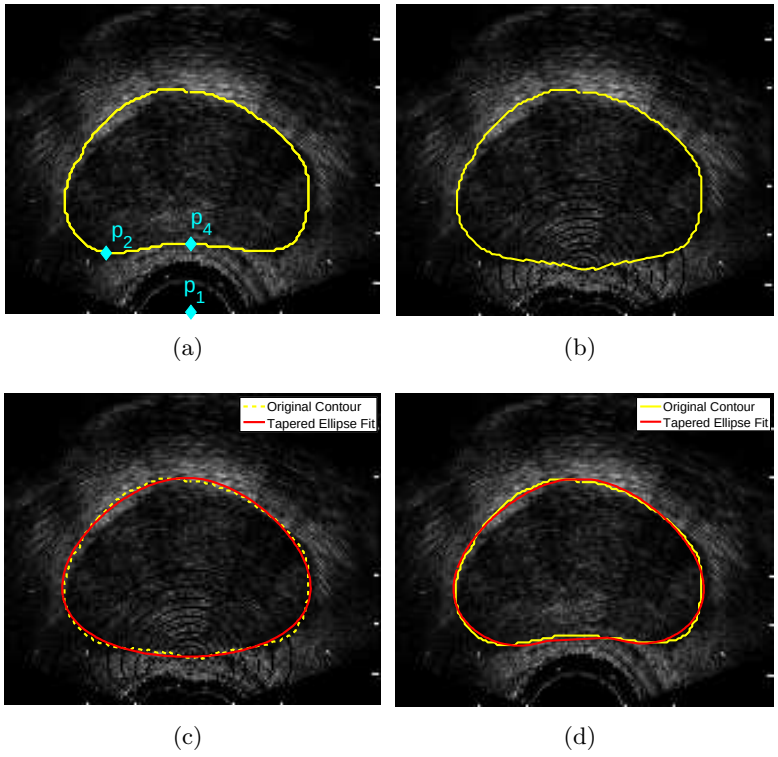


Figure 2.2: Process of making contours smooth and symmetric. a) initial point required for unwarping parameter estimation, b) unwarped image and contour, c) tapered ellipse fit, and d) final warped tapered ellipse and the original contour. All contours are generated for a sample mid-gland 2-D TRUS image.

all atlases.

Hyper-parameter K, which refers to the optimal number of clusters, is tuned by performing validation on a subset of the dataset. Statistics of the segmentation outcome on the validation dataset are obtained for different values of K, and the optimum number of clusters is determined. Potentially, this parameter exposes a characteristic of the training dataset, e.g. variation in anatomy coverage, which is intuitively expected to be the same between validation and test datasets. We adapt a systematic sampling of the dataset to divide it into 60% train, 20% validation, and 20% test cases. The partitioning process is repeated 10 times in order to evaluate robustness of the proposed segmentation algorithm. While none of the test cases are included in training and validation sets within a partitioning process, each test atlas might appear at most twice in test datasets among partitioning attempts. In each trial, the target volume is registered to the mean atlas as explained in Section 2.2.1 in order to prune the training dataset into a final set of 100 atlases, which are further transformed into the coordinates of the target image.

2.3.1 Segmentation Evaluation

For every case, segmentation performance is evaluated against the clinical gold-standard as well as the semi-automatic segmentation algorithm of Mahdavi *et al.* [1]. We perform statistical non-inferiority analysis using paired t-test. We define the null hypothesis as both semi-automatic and proposed algorithms produce results with the same mean value for the evaluation metrics (errors). Contours of the semi-automatic algorithm are directly obtained from the software used as part of the standard of the care at the VCC along with the gold-standard contours. We carry out computation of volumetric and surface-based distance measures to evaluate the segmentation performance. We use the Dice Similarity Coefficient (DSC) to quantify shape similarity between two binary volumes, i.e. the segmentation outcome and the gold-standard. We also denote the volume error by V_{err} and define it as $100 - DSC$. If we denote $\widehat{L}$ and L as the estimated and true binary CTV volumes of the target image, we compute the DSC as

$$DSC = \frac{2\left|\widehat{L} \cap L\right|}{\left|\widehat{L}\right| + |L|} \times 100. \tag{2.5}$$

In brachytherapy, volume is a key parameter in treatment planning and dosimetry analysis, as there is a high correlation between the number of

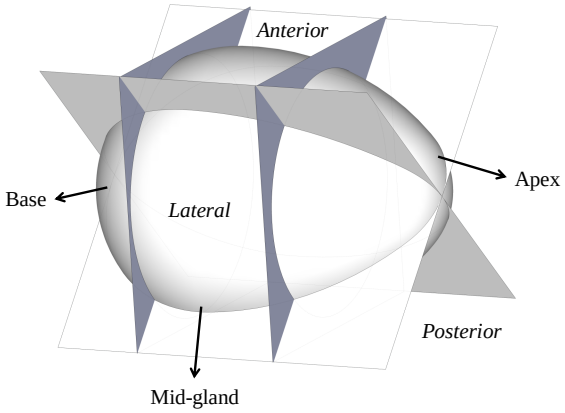


Figure 2.3: Prostate sectioning for segmentation evaluation

seeds and needles required and the volume size. We quantize this difference by calculating volume difference coefficient as

$$V_{diff} = \frac{(\left|\widehat{L}\right| - |L|)}{|L|} \times 100, \tag{2.6}$$

which is a signed coefficient representing size difference between estimated and the gold-standard CTV binary volumes.

For the mid-gland slice contours, we represent the distance by calculating Mean Absolute Radial Distance (MAD), and Maximum Absolute Radial Distance (MAXD) in between. The mid-gland slice is expected to provide the clearest view of the prostate anatomy for the delineation process. We also quantify the overall surface distance by calculating the mean surface (MSD) and Hausdorff distance (HD) between two contour sets.

Multi-atlas segmentation is highly dependent to the registration performance, i.e. global SSD between the registered atlases and the target images. Rohlfing [73] criticized tissue overlapping metrics and their correlation with registration accuracy. He showed that a good global overlapping metric does not necessarily correspond to a good registration result. To account for the weakness of global metrics, we break down computation of the similarity metrics (both volumetric and surface-based) into several anatomical regions similar to Mahdavi *et al.* [1]. From a clinical perspective, each of the anatomical regions are subject to a different expected accuracy and importance with respect to the treatment planning protocol. These regions are defined by partitioning the whole gland into six sectors. Axial partitioning

is made by dividing w.r.t. 0.3, 0.4 and 0.3 of the total prostate length from base to apex. Each partition is further divided into lateral (left and right), posterior and anterior sections by two perpendicular sagittal planes passing through the axial line of the prostate, producing nine general anatomical sectors as shown in Figure 2.3.

2.4 Experiments and Results

TRUS images provide a field of view that encompasses some of the adjacent tissues, including portions of the bladder and rectum. Obviously, the global registration metrics do not represent the registration performance related to the region of interest. Therefore, we define a mask volume M on an area where there is a disagreement between transformed atlas labels:

$$M(\mathbf{x}) = \begin{cases} 1, & \exists(i,j), \widetilde{S}_i(\mathbf{x}) \neq \widetilde{S}_j(\mathbf{x}) \\ 0, & otherwise \end{cases} \tag{2.7}$$

We individually obtain and use a volumetric mask M for each new target image to include pixels with higher uncertainty in calculation of the registration quality metric.

All atlas and target registrations are performed in 3-D on the whole TRUS volume. We use the Demons registration algorithm [62–64] with a Gaussian kernel size of 5 mm in the axial direction.

A linear correlation analysis of the intensity and label structure matching is illustrated in Figure 2.4. We use SSD for quantifying the intensity matching, and DSC for binary label shape similarity. We obtain a low correlation coefficient of 0.29 with an arbitrary distribution of the observed paired metrics (SSD, DSC) for all 10 experiments on partitioned dataset, which points to the necessity of an atlas pruning technique prior to and after registration.

We choose $N = 100$ atlases to form the MAS framework dataset prior to the registration process. After the registration step, we apply the clustering on pairwise atlas agreement factor vectors and choose the highest ranked cluster w.r.t the average global SSD of its members. All members of the winning cluster are fused using the STAPLE algorithm to produce the probability segmentation map of the target image. Calculation of intensity-based similarity metrics is confined within the area of the generated mask from all transformed atlases, as explained above.

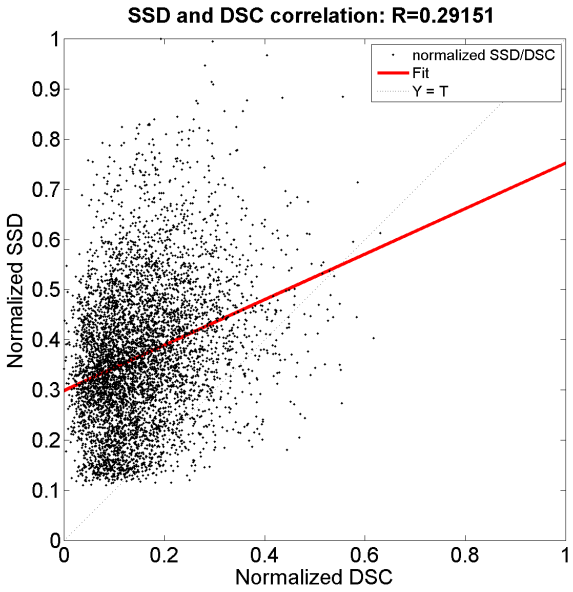


Figure 2.4: Correlation analysis between SSD (registration performance) and DSC (segmentation overlap) performed on test cases. This analysis has been achieved by running the experiments on partitioned dataset for 10 times.

2.4.1 Mean Atlas Generation

Figure 2.5.c shows the mid-gland slice of the average atlas generated from the intensity volumes of atlases in a training dataset. Blurriness of the mean atlas relates to the performance of the registration algorithm in the training dataset, whereas in the current context, it is highly affected by intrinsic characteristics of the TRUS images. We use the Demons registration algorithm, and apply the mean atlas generation algorithm for 10 iterations. We found this number sufficient to converge to a consistent average volume. Figures 2.5.a and 2.5.b illustrate the mid-gland slice result of a registration process between a sample TRUS volume and the mean atlas.

2.4.2 Global and Sector-based Performance Measures

Table 2.1 shows the error statistics calculated for all 10 trials, in which a total number of 560 bootstrapped cases is included in the error calculation. We report mean and standard deviation of the error metrics to compare

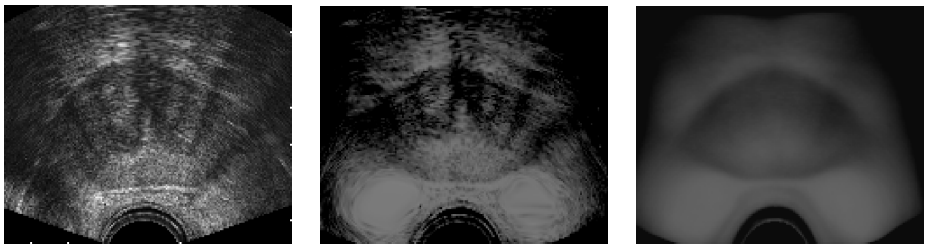


Figure 2.5: For the mid-gland slice: a) a sample atlas, b) registered to the mean atlas, and c) the generated mean atlas at the same slice.

Table 2.1: Volumetric, contour and surface error statistics (mean±standard deviation) calculated from 10 random dataset partitionings are shown for the proposed automatic segmentation approach and compared with the approach of Mahdavi *et al.* [1].

	Volumetric Error (%)		**Mid-gland Contour Error** (mm)		**Surface Error** (mm)	
	V_{err}	V_{diff}	MAD	$MAXD$	MSD	HD
Automatic	10.25±3.74	-2.37±13.73	1.77±0.98	4.52±2.04	1.33±0.60	5.48±1.71
Semi-auto.	8.81±5.60	3.56±14.66	1.20±1.21	3.36±2.56	1.11±0.81	5.57±2.28

accuracy and precision of the proposed automatic segmentation method. Since the gold-standard segmentation is based on manual initialization of the semi-automatic segmentation of Mahdavi *et al.* [1] in the mid-gland, our mid-gland contour metrics show the most significant error amongst the metrics.

In Tables 2.2 and 2.3, we show evaluation metrics (i.e. DSC and MSD) obtained from 10 different random partitionings of the dataset ($M = 280$) into 60-20-20% training-validation-test sets per sector. To provide detailed analysis of accuracy distribution, a volumetric measure of DSC is shown in Figure 2.6. In the posterior region, the automatic algorithm provides a higher DSC value in comparison to the semi-automatic algorithm. In contrast, at the anterior region there is less accuracy obtained compared to the semi-automatic algorithm. In the rest of the regions both methods perform with similar accuracy. Noticeably, there are cases for which the semi-automatic algorithm presents the highest DSC value, meaning there has been no need to modify the generated contours.

Two typical segmentation results are shown in Figure 2.9, each overlaid by the gold-standard and semi-automatic based contours.

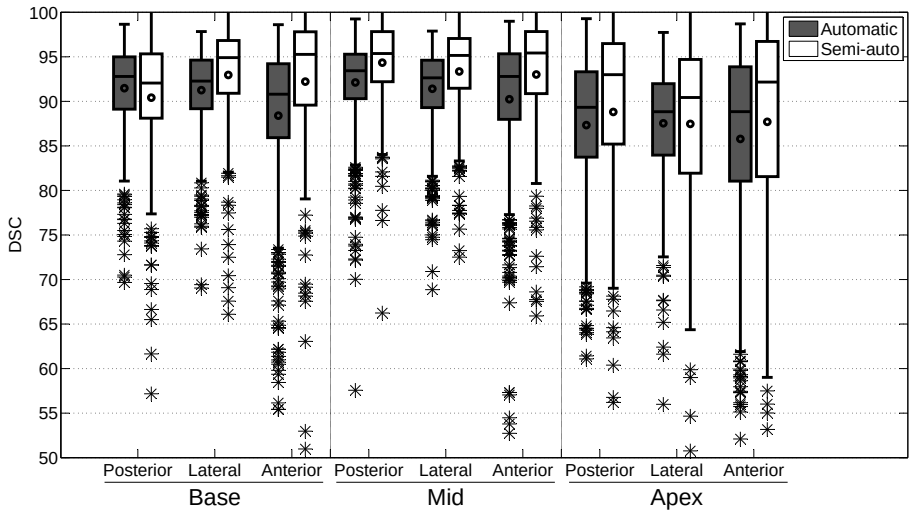


Figure 2.6: DSC values computed for nine anatomical sectors of the prostate and for two segmentation methods: 1) the proposed automatic approach, and 2) the semi-automatic method. Statistics are computed from 10 random partitionings of the dataset into 60-20-20% train-validation-test sets, i.e. 560 test cases.

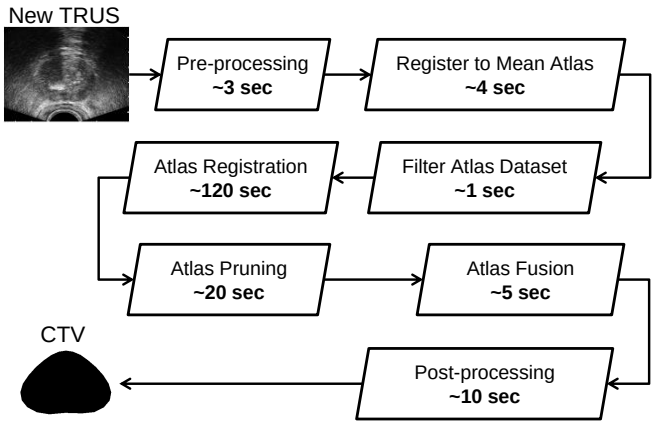


Figure 2.7: Average operating computation time measured for each individual process of the proposed segmentation pipeline and its configuration with total number of 100 atlases.

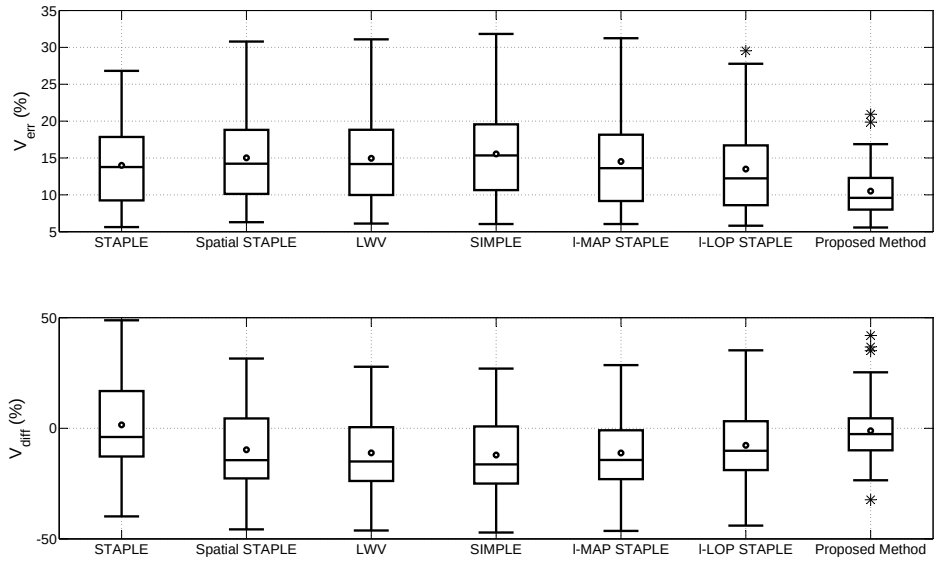


Figure 2.8: Comparison between different fusion algorithms on all pre-selected atlases and the proposed algorithm. V_{err} and V_{diff} represent shape and size dissimilarity between the estimation result and the gold-standard for all test cases in a random dataset partitioning.

Table 2.2: Grand mean and pooled standard deviation of Dice (DSC) calculated for 10 different trials of randomly dividing data into 60-20-20% training-validation-test sets.

	Automatic			
DSC %	Base	Mid	Apex	**Total**
Ant.	88.40±8.50	90.24±7.84	85.78±10.78	89.75±3.75
Lat.	91.26±4.62	91.42±4.56	87.54±6.04	
Post.	91.48±4.87	92.12±4.75	87.33±8.13	
	Semi-automatic			
DSC %	Base	Mid	Apex	**Total**
Ant.	92.22±8.69	93.01±7.20	87.70±12.74	91.19±5.60
Lat.	92.96±6.02	93.35±5.57	87.47±10.17	
Post.	90.42±7.63	94.35±4.75	88.81±11.94	

2.4.3 Computational Cost

Standard-of-care at the VCC, where the TRUS data for this study were collected, is to conduct the planning process a few weeks before the scheduled implantation date, a time interval during which the treatment plan is generated and the seeds and needles ordered from the manufacturer. Currently, $3 \sim 4$ min of the planning procedure is dedicated to the segmentation process [1], and no time constraints are applied to the segmentation process as it is a very early step in the planning procedure. However, because it is amongst the first of several steps in the planning pipeline, it should be performed as quickly and efficiently as possible.

We measure execution time of the proposed automatic segmentation approach to ensure there is no burden added to the standard of the care at the cost of automating the segmentation procedure. Computation time is measured on a standard PC (Intel Core i7, 2.93GHz, 8GB RAM). All the main functions are implemented in C++ and MATLAB®.

The mean atlas generation process takes place only once. The mean atlas and all atlases transformed to the coordinates of the mean atlas are stored for the subsequent steps. Fig 2.7 provides an overview of the computation time of individual processes of the proposed algorithm. Reported numbers are for the non-optimized code employed in two different programming interfaces (i.e. C++ and MATLAB®). The proposed method runs under 3 minutes according to the maximum execution time measured for the evaluation dataset. Therefore, the automatic approach is fast enough that allows the physicians to perform more tasks such as contour modification while it

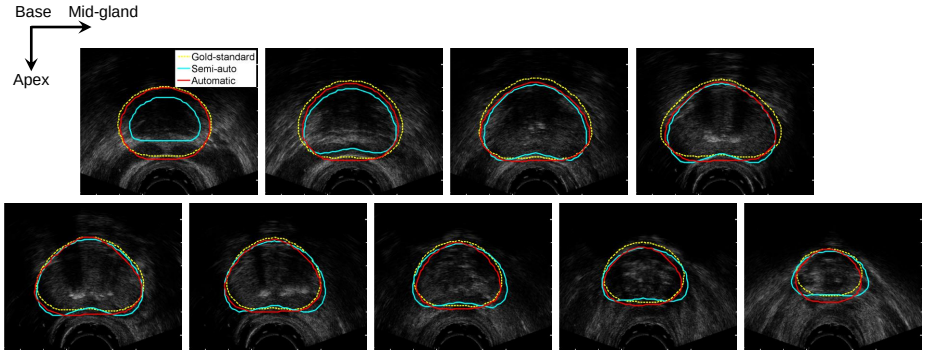
Table 2.3: Grand mean and pooled standard deviation of mean surface distance (MSD) calculated for 10 different trials of randomly dividing data into 60-20-20% training-validation-test sets.

	Automatic			
MSD *mm*	Base	Mid	Apex	**Total**
Ant.	0.96±0.78	0.82±0.70	0.65±0.66	1.33±0.60
Lat.	0.92±0.56	0.96±0.56	0.69±0.47	
Post.	0.63±0.38	0.56±0.34	0.53±0.44	
	Semi-automatic			
MSD *mm*	Base	Mid	Apex	**Total**
Ant.	0.58±0.69	0.63±0.60	0.51±0.64	1.11±0.81
Lat.	0.68±0.66	0.82±0.71	0.58±0.55	
Post.	0.59±0.52	0.37±0.30	0.45±0.60	

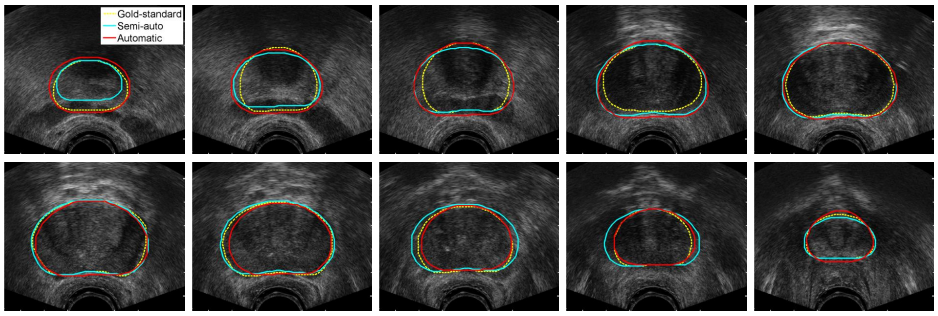
generates contours for a patient.

2.4.4 Comparison of Different Fusion Methods

Figure 2.8 shows volume and shape distance metrics obtained by the compared fusion algorithms. A total of 100 atlases, preselected for the multi-atlas registration and fusion, are processed by: 1) the conventional STAPLE algorithm [69]; 2) spatial STAPLE [71]; 3) Locally Weighted Majority Voting [56, 67] (LWMV); 4) Selective and Iterative Method for Performance Level Estimation (SIMPLE) [68]; 5) local Maximum A Posteriori STAPLE (l-MAP STAPLE) [57]; and 6) local Logarithmic Opinion Pool STAPLE [58] (l-LOP STAPLE). We have used implementation of the local LOP and the local MAP STAPLE from http://crl.med.harvard.edu/, and the SIMPLE and spatial STAPLE from http://www.nitrc.org/projects/masi-fusion/. The parameters used were those recommended by the authors of the original algorithms. In *local* variation of the algorithms, we limit neighbourhood window size to 1.4 *mm* in 2-D axial planes. Statistical significance analysis of the measured V_{err} using paired t-test shows significant improvement by the proposed method when compared against the evaluated fusion algorithms (p-value< 0.01). Among the other fusion approaches, local LOP STAPLE provides the most accurate segmentation results both in terms of shape and size, as measured by V_{err} and V_{diff}.



(a)



(b)

Figure 2.9: Two examples of the automatic segmentation result compared to the semi-automatic approach of Mahdavi *et al.* [1]. The gold standard contour is shown by dashed line.

2.5 Discussion and Conclusion

We presented an automatic 3-D segmentation algorithm for delineation of the CTV in prostate brachytherapy. The algorithm uses a non-parametric deformable registration method to align a set of *a priori* atlas volumes with a target volume. The CTVs delineated for all atlas volumes are then combined in a STAPLE framework to provide a probability map of the target image CTV. We proposed post-processing steps by adapting a tapered-ellipse fit algorithm to obtain the consensus segmentation of the CTV. The proposed algorithm can potentially reduce the inter- and intra-observer variability of the CTV contouring by eliminating the need for user intervention. However, modification of the automatically generated contours by clinicians is inevitable for seamless translation of the approach to clinic.

To reduce the computational cost and maintain the efficiency of the algorithm, we take advantage of a forward mean atlas generation process prior to the registration of atlases. We filter a large dataset of atlases into a smaller set of similar atlases using the generated mean atlas, hence reducing the chance of participation for those atlases with lower similarity between their intensity volumes and the target volume.

Our observation from the registration performance concludes that participation of all atlases in the fusion procedure decreases performance and accuracy of the segmentation; namely, reduces the mean DSC and raises the standard deviation. One reason can be the fact that registration accuracy is not equally spread over different anatomical regions within the region of interest. Moreover, there is no strong evidence behind the assumption of *the lower SSD in intensity volumes, the higher DSC between binary volumes is expected*. For the ultrasound modality, deformation fields obtained from registration algorithms are highly affected by the presence of speckles and other artifacts that appear as registration noise in the boundary region of the prostate. Therefore, propagation of the deformations from the intensity domain to the label's binary domain is subject to some distortions. Eventually, filtering the registered atlases instead of including all atlases reduces the registration error.

In a major departure from earlier prostate segmentation approaches, the proposed pipeline is applicable to the sparse ultrasound volumes collected in brachytherapy. Intuitively, we expect a better registration quality for dense volumes, because more anatomical features contribute to the registration and fusion processes. However, this will increase computational cost of the segmentation pipeline.

A non-optimized implementation of our proposed algorithm runs under

4 minutes in a standard PC, which is comparable with the current segmentation speed in a typical brachytherapy volume study process [1]. However, individual atlas registration to the target volume is an independent process, and a parallel implementation on multiple CPU cores and possibly GPU can reduce the computation time further to the order of seconds. Therefore, our method has the potential to handle the requirement of real-time dosimetry for brachytherapy with intra-operative planning.

2.5.1 Comparison with Other Methods

A comparison analysis with other advanced fusion algorithms shown in Figure 2.8, indicates that the low correlation between structures in intensity and label domains can mislead the local or intensity-driven weight assignment performance, and hence may undermine the fusion results. This observation implies that selecting the appropriate atlases prior to fusion can significantly improve the accuracy of consensus segmentation in the context of TRUS images. Our framework can potentially benefit from other atlas selection methods, such as the approach recently proposed by Asman *et al.* [74] which has been evaluated on MR images of cervical vertebrae.

We retain clinically comparable results in terms of accuracy and precision against the gold-standard CTVs on a dataset of sparse TRUS volumes obtained from 280 patients. The inter- and intra-observer variability of the prostate segmentation in the context of brachytherapy has been documented before [1]. Specifically, the inter-observer variability of manually segmented contours has been reported to be on the order of $7.25 \pm 0.39\%$ and $6.64 \pm 2.36\%$ for V_{err} (i.e $100 - DSC$) and volume difference (D_{diff}), respectively. Moreover, the intra-observer variability of manually segmented contours has been reported to be on the order of $5.95 \pm 1.59\%$. Our proposed fully-automatic method improves robustness and consistency of the segmentation, and produces results within the accepted range of variability.

We also compare segmentation results to some reported results in the literature; however, direct comparison is not possible. Tutar *et al.* [75] reported volume overlap of $82.8 \pm 6.2\%$ (corresponding to DSC value of 90.59%) as disagreement between three different experts in manual boundary delineation of 30 post-implant TRUS prostate volumes. Akbari [32] and Yang [33] obtained DSC values of $88.1 \pm 1.44\%$ and $90.7 \pm 2.5\%$, respectively, for a small dataset of five dense TRUS volumes. Heimann *et al.* [34] reported overlap error of between $16.7 \pm 5.2\%$ to $17.6 \pm 6.8\%$ (corresponding to DSC value of 90.35% and 90.89%) for successful segmentation of 25-33 out of 35 dense TRUS volumes. Gong *et al.* [23] reported inter-observer variability

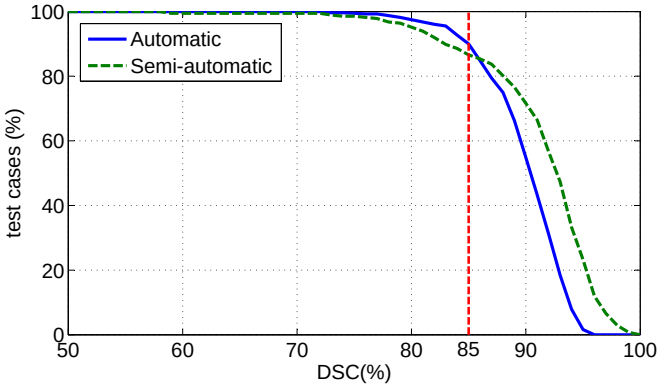


Figure 2.10: Performance comparison between the proposed automatic approach and the semi-automatic method presented as percentage of test cases segmented with respect to DSC values. The proposed method achieves greater success rate for DSC values less than 86% compared to the semi-automatic method.

of 1.82 ± 1.44 mm for the contour distance in all planes obtained from five experts and out of 125 images of 16 patients.

Figure 2.10 provides a comparison chart for segmentation success rate, in terms of DSC with the gold-standard, on a 10 random dataset partitioning. Herein, we define the accuracy threshold of DSC as a parameter to compute the success rate, i.e. percentage of results with greater DSC value. It can be observed that for the accuracy threshold of 85% the proposed MAS approach is successful for 90% of the test volumes (i.e. 560 cases), which is comparable to the success rate of the semi-automatic algorithm [1] which is 87%. However, the semi-automatic approach outperforms in terms of the success rate for higher segmentation accuracy threshold values.

Detailed evaluations on nine anatomical sections shows maximum surface error in the posterior-apex section (see Figure2.6). Although average DSC obtained from the semi-automatic algorithm is higher than the proposed MAS framework, less standard deviation is seen in most of the regions. Highest DSC values are obtained in the mid-gland section in both algorithms.

2.5.2 Dataset Dependency

Definition of CTV and subsequently the Planning Target Volume (PTV) is a subjective process which varies in between brachytherapy clinics. The

multi-atlas segmentation approach takes advantage of existing segmentations by projecting them onto the target TRUS volume. This makes the proposed MAS approach a unique method that combines intensity-based segmentation methods with a learning-based technique that is capable of learning delineation style in the atlas dataset. In contrast to those segmentation methods which rely on boundary region detection (e.g., edge-based methods or active shape models), the multi-atlas approach projects the existing pattern in the atlas dataset onto the target image. This feature makes the approach a potential solution for problems similar to the brachytherapy CTV delineation, in which the clinical label does not necessarily follow the real target anatomy boundary in all regions.

A drawback of the proposed algorithm is the dependency on the dataset, which affects the accuracy of the final segmentation. However, it might be desirable to form a generic atlas dataset that includes atlases representing different prostate shapes and sizes.

In summary, the proposed multi-atlas segmentation approach is a clinically adaptable algorithm that can be integrated with the current brachytherapy volume study procedure. Although the segmentation execution time is not a design criteria for the brachytherapy preplanning, the proposed method can be accelerated and optimized for real-time clinical diagnostic or treatment applications such as targeted biopsy and real-time dosimetry. A potential solution is to replace the Demons registration method with the GPU-accelerated version of *NiftyReg* [76].

Chapter 3

ICA-based Fusion for Brachytherapy Preplanning

In search for a machine learning fusion algorithm that can efficiently extract joint variations between different volumetric information representations of the brachytherapy preplanning elements, we investigate joint independent component analysis (jICA) in this chapter. This chapter is subdivided into two parts aiming to utilize jICA methodology in estimation of contours; and prediction of both planning target volumes and the mPD isodose contours.

3.1 Clinical Target Volume Estimation from TRUS Images

3.1.1 Introduction

In prostate brachytherapy, usually TRUS volumes are axially sparse. Considering some intrinsic characteristics of the ultrasound images and poor visibility of the prostate near the base and the apex, CTV delineation remains a challenging problem. Moreover, there is a high demand for an efficient and fast prostate segmentation approach that can be used intraoperatively. This potentially allows the physicians to correct the plan w.r.t. the operating room conditions to ensure CTV is properly targeted. However, most of the existing segmentation approaches are limited in terms of the speed and performance to fulfil intraoperative requirements including our approach on using multiple atlases explained in Chapter 2 [54, 60, 61]. Therefore, a fast segmentation algorithm with minimum interaction with the operator facilitates potential solutions for the real-time dosimetry.

For the first time in the context of prostate segmentation, we jointly analyze *a priori* knowledge of the CTV and TRUS images by generating

This section is adapted from [77]: S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Fast Prostate Segmentation for Brachytherapy based on Joint Fusion of Images and Labels, in Proceedings of SPIE Medical Imaging, 2014, vol. 9036, p. 90361A1-7.

a model using joint Independent Component Analysis (jICA). We use a dataset of TRUS images and their delineated CTVs determined by expert clinicians that comprises a large variety of prostate shapes. We form a joint data matrix including intensity images and their corresponding labels in order to seek statistically independent basis functions, e.g. joint image-label sources, within the existing dataset.

The model generated from jICA can be potentially used in a real-time segmentation application. The target image is projected onto the joint basis vectors in order to obtain the optimum coefficients for reconstruction. We assume the target label can be reconstructed by a linear combination of the obtained basis vectors. Accordingly, same coefficients are employed to reconstruct a probability map of the target label. The target label is estimated from this probability map using a globally optimum threshold followed by a contour smoothing process. We evaluate this method against the manually segmented prostate contours that were used by clinicians to plan brachytherapy procedures for 60 patients. We show clinically acceptable results that satisfies speed requirement for a real-time clinical application.

3.1.2 Methods

Model Generation based on joint ICA

Independent Component Analysis (ICA) is a statistical approach to decompose a set of observations, $\mathbf{x}$, into components which are maximally independent:

$$\mathbf{x} = \mathbf{a}^T \mathbf{C} \tag{3.1}$$

where $\mathbf{a}$ is the mixing vector and $\mathbf{C} = [\mathbf{c}_1 \mathbf{c}_2 ... \mathbf{c}_K]^T$ is the component matrix. ICA works with higher order statistics of the observation data and assumes that hidden sources are statistically independent, non-Gaussian with linear mixing process. For the context of this study, ICA is performed on a reduced dimensionality observation matrix using Principal Component Analysis (PCA).

Joint Independent Component Analysis (jICA) [78], is an extension of ICA used as a data fusion technique to combine information from multiple modalities. jICA has been successfully applied in cognitive and brain studies [79–81]. We employ jICA to learn the correlation between intensity images and their binary labels. We obtain a set of joint components consist of the intensity and the corresponding label maps.

For each patient data, we form a joint observation vector $\mathbf{x}_i$ that is described as 1-D representation of TRUS volume and the corresponding seg-

mentation. A joint observation matrix $\mathbf{X}$ is defined as a stack of observation vectors:

$$\mathbf{X} = [\mathbf{x}_1\mathbf{x}_2...\mathbf{x}_N]^T \quad where \quad \mathbf{x}_i = \begin{bmatrix} \mathbf{x}_i^v \\ \mathbf{x}_i^l \end{bmatrix}, \qquad i = 1, 2, ..., N. \tag{3.2}$$

Elements of the vector $\mathbf{x}_i^v$ are scalar values of a particular voxel inside the TRUS volume, and $\mathbf{x}_i^l$ is a binary vector representing the label volumes.

The proposed segmentation pipeline consists of two main parts, modelling and evaluation. The modelling process involves decomposition of the observation matrix $\mathbf{X}$ into K number of joint independent components $\mathbf{C} = [\mathbf{c}_1\mathbf{c}_2...\mathbf{c}_K]^T$. The evaluation process uses the joint basis vectors on a test TRUS volume. The test volume is partially projected onto the subcomponents $\mathbf{c}_j^v$ to obtain the mixing vector $\mathbf{a}$. Same mixing vector is further used to combine the second part of the joint components, $\mathbf{c}_j^l$ to generate a segmentation probability map. A heuristic threshold value is used to convert the label probability map into a binary map and its corresponding edge contours. The generated contours are smoothed to produce the final estimation of the target segmentation. Figure 3.1 shows block diagram of the proposed segmentation pipeline.

Label Map Reconstruction and Smoothing

A label map is produced from a linear combination of the label basis vectors $\mathbf{c}_j$, using same coefficients obtained from the intensity image projection $\mathbf{a}$. In a post-processing step, we convert the produced label map into the final estimated binary volume using an empirical threshold value ($\tau = 0.4$ for a normalized label map). Edge contours of the estimated binary volume are smoothed slice by slice. Contours are represented in a polar coordinate system assuming their geometrical center as the origin and further transformed into the Euler coordinate system to perform B-spline curve fitting. Figure 3.2 shows smoothing process for a single contour of a sample result. Coordinates of the contour points are mapped to B-spline curve-fits in two vertical (x) and horizontal (y) directions separately.

Experiments

The proposed method is evaluated on a dataset of 60 TRUS volume images of patients undergoing brachytherapy using leave-one-out cross validation approach. Each TRUS volume consists of 7 to 14 transverse plane 2-D ultrasound images axially acquired from base to apex while the prostate gland

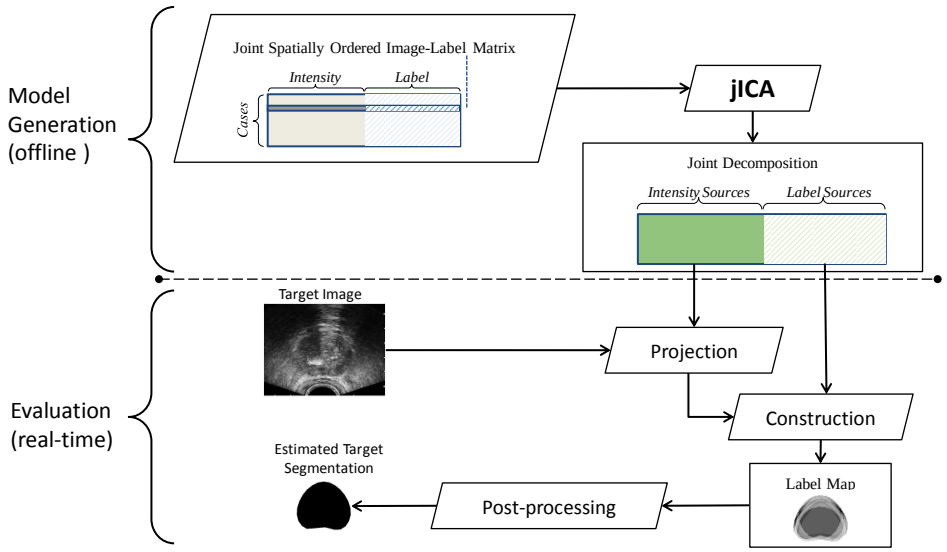


Figure 3.1: Block diagram of the proposed segmentation approach. Model generation is performed off-line and based on jICA. Real-time application utilizes the generated model to estimate for the target CTV delineation.

is visually aligned in the middle of all ultrasound images. For each TRUS volume, the delineated prostate gland in all 2-D ultrasound images is provided by an expert radiation oncologist as the gold-standard in this study. The accuracy and precision of the algorithm is evaluated using 1) volumetric measures such as dice similarity coefficient (DSC) and percent volume difference, V_{diff} [1]; 2) contour distance measures such as the Mean and Maximum Absolute radial Distances (MAD and MAXD) and 3) surface dis-

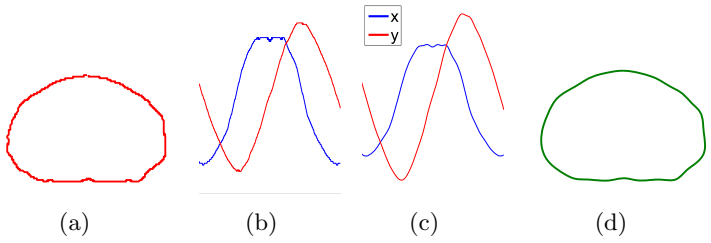


Figure 3.2: Contour smoothing process, (a) a sample contour, (b) x and y coordinate coefficients, (c) B-spline curve fitting, (d) smoothed contour.

tance measures such as Hausdorff Distance (HD) and Mean Surface Distance (MSD) against the gold standard determined by the expert clinician.

In order to generate the spatially ordered joint image-label matrix, all prostate image and label volumes are re-sampled towards the size of the axially longest TRUS volume. To speed up the process, 2-D ultrasound images of each volume are down-sampled by a factor of three. All intensity images are normalized using grand mean and standard deviation computed from dataset members. Joint image-label vectors are formed using a unique spatial order and stacked to form the joint data matrix. Data dimensionality is reduced to the order of 35 using PCA in order to retain over 95% of the information from eigenvectors. Followed by a whitening process, i.e. subtraction of the mean data, jICA decomposition is performed to obtain 35 joint basis vectors.

3.1.3 Results and Discussion

Statistics of the volumetric, contour and surface distance evaluation metrics are compared with the gold standard segmentation and shown in Table 3.1. A direct comparison with the reported results in the literature is not possible. However, our proposed method produces results with comparable volumetric measures against other TRUS image segmentation reports. The presented results have to be considered within the context that the gold-standard itself is subject to inter- and intra-observer variability on top of the variability associated with the segmentation method. Our gold-standard segmentations are obtained from actual delivered treatment plans which are produced by one expert clinician. Since our comparison is with respect to manual segmentations of one clinician, we refer to the literature and review the clinical significance of the obtained results. For segmenting the prostate boundary for planning brachytherapy [1], the inter- and intra-observer variability of manually segmented contours in terms of the volume error $(1 - DSC)$ has been reported to be in the order of $4.65 \pm 0.77\%$ and $5.95 \pm 1.59\%$, respectively. In another study [75], volume overlap of $82.8 \pm 6.2\%$ is measured when estimating for the inter-observer variability of manual prostate boundary delineation which corresponds to 9.41% volume error. The value of $1.82 \pm 1.44\ mm$ [23] is also reported for the contour distance inter-observer variability in a 2-D segmentation study.

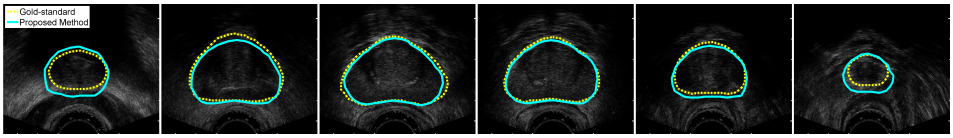
Figure 3.3 shows generated contours in six slices from base to apex for two average segmentation outcomes based on the calculated DSC value. Volumetric (DSC) and surface distance (MSD) evaluation metric histograms are shown in Figure 3.4 as well as TRUS images of an outlier with lowest

Table 3.1: Comparison of various evaluation metrics between the proposed segmentation approach and the manually segmented contours determined by an expert clinician.

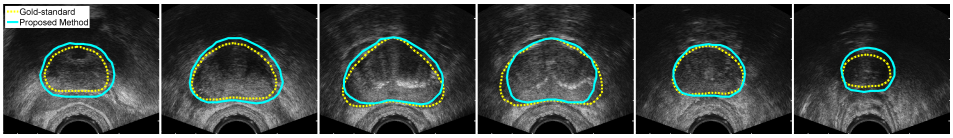
Volumetric Measures (%)		Contour Errors (mm)		Surface Errors (mm)	
DSC	V_{diff}	MAD	MAXD	MSD	HD
90.28 ± 3.07	-1.69 ± 11.85	1.84 ± 0.93	4.52 ± 1.77	1.30 ± 0.48	5.39 ± 1.46

DSC and highest MSD value. Due to strong appearance of the ultrasound artifacts, delineation of the prostate boundary is very difficult for this case in most of the slices. In addition, size of the prostate is largest among the dataset and most likely is not learned by the modelling approach which justifies failure of the proposed automatic segmentation approach for this case.

The proposed segmentation procedure requires minimum computational cost to produce the label map, since projection of the target intensity image is formulated as a least squares problem. We measured execution time from a non-optimized code on a standard PC (Intel Core i7, 2.93 GHz, 8 GB RAM) to be 645 ± 24 ms. Our smoothing approach consists of coordinate system transformation and B-spline fit which adds 1158 ± 432 ms to the segmentation pipeline. We conclude to execution time of less than two seconds on a standard PC which is sufficient for enabling real-time dosimetry planning.

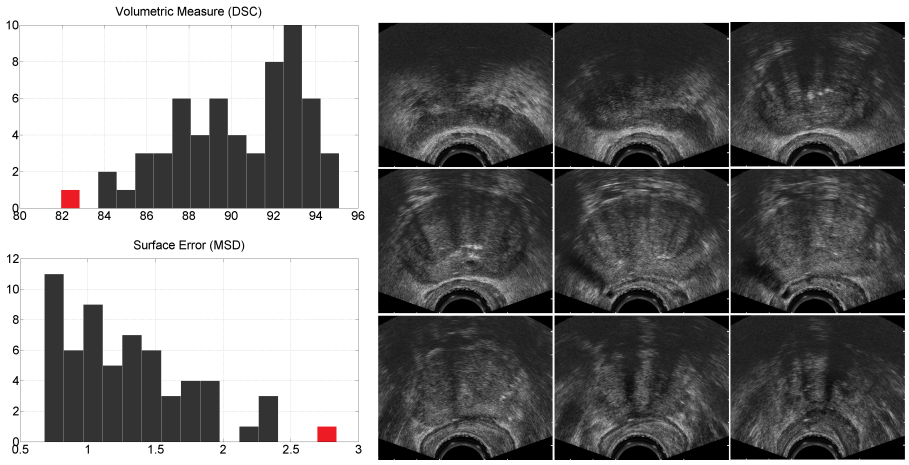


(a)



(b)

Figure 3.3: Contours for six equally-spaced slices (base to apex) of two typical segmentation results.



(a) (b)

Figure 3.4: a) Histogram of the volumetric error and surface distance measure. b) TRUS images of the marked outlier (in red color) from base (top left) to apex (bottom right). The prostate boundary near base and apex is not visible.

3.1.4 Conclusion

By using joint independent component analysis on a spatially ordered joint data matrix of TRUS images and their corresponding CTV labels, we have introduced a fast segmentation approach that can be potentially used in real-time brachytherapy dosimetry applications. The proposed segmentation approach generates a set of joint image-label basis vectors. Coefficients obtained from partial projection of the intensity image onto these basis vectors is further used to combine the label basis vectors to produce a label map for the target image. Evaluation of this method on a clinical brachytherapy dataset shows results with clinically acceptable accuracy and repeatability while it performs in less than two seconds. Accordingly, an optimized implementation of the algorithm with a larger dataset has the potential to handle the requirement of real-time dosimetry for brachytherapy with intra-operative planning.

3.2 Planning Target Volume and Minimum Prescribed Isodose Estimation from TRUS Images

3.2.1 Introduction

Essentially, PTV is the reference anatomical target in the preplanning of prostate brachytherapy [5, 6, 83]. As explained in Chapter 1, the PTV is obtained by inhomogeneous expansion of the CTV according to the institutional guidelines and the prostate anatomical regions. Subsequently, the number and pattern of seeds and needles is determined. Seed planning requires adherence to selected control parameters, such as the percentage of PTV receiving the minimum Prescribed Dose value (mPD).

Here, we aim to incorporate prior knowledge in a dataset of treatment records including TRUS volumes, PTVs and the seed coordinates to simultaneously predict the PTV and the isodose contour for a new patient, solely based on the TRUS volume. The isodose contour of the prescribed dose can be considered as one of the plan representation parameters, as it reflects the mPD distribution inside and outside of the PTV. We calculate the isodose contours based on the seed coordinates and adapt the joint Independent Component Analysis (jICA) [78–81], as a reconstruction based modeling approach to capture relation between TRUS intensity volumes, PTV and the prescribed isodose map. We propose a model with a set of joint components that are identified using jICA. The PTV and the isodose map of a new patient TRUS volume are represented as a weighted sum of the identified joint components in the derived model.

A major shortcoming of the ICA-based modeling is the instability of the identified sources, since majority of the mathematical solutions proposed for ICA are sensitive to initial conditions. To alleviate this issue, we repeat the ICA decomposition process multiple times. Stable components are then chosen from the centroids of the clusters representing compactness and separability of the decomposed space. In this chapter, we present a stability analysis for the jICA approach. Subsequently, we follow a holdout validation approach to evaluate our generated model from fusion of TRUS, PTV and the prescribed isodose. We report the result of our analysis on data obtained from 120 patients.

This section is adapted from [82]: S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Data fusion for planning target volume and isodose prediction in prostate brachytherapy, Proc. SPIE, vol. 9415. p. 94151I-94151I-7, 2015.

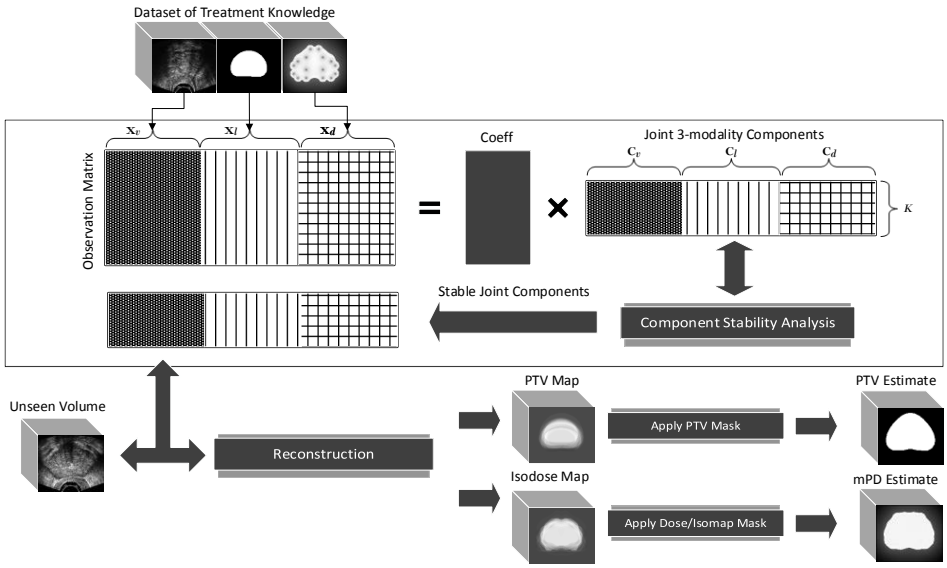


Figure 3.5: Isodose prediction pipeline consists of a modeling (offline) process and a prediction (real-time) process. Prediction process simultaneously estimates PTV and isodose contours solely based on an unseen TRUS volume.

3.2.2 Method

The block diagram of the proposed method is shown in Figure 3.5. A retrospective dataset of treatment records is used to identify joint components of the ICA model. The dataset consists of three information modalities; namely, the TRUS volume, PTV binary volume and the isodose binary volume which is obtained by performing dosimetry calculations on the seed plan. The modeling procedure takes place in an offline mode. The generated model encodes the information within the dataset into a set of joint components by performing the joint component stability analysis. In the prediction phase, an unseen TRUS volume is only partially projected onto the TRUS part of the joint components to obtain the appropriate coefficients that minimize the reconstruction error. Same coefficients are adapted to combine the other modality parts of the joint components and generate a probability map of the corresponding PTV and isodose. We apply a mask to convert these maps into the target binary PTV and isodose estimations.

We denote a TRUS volume as $V(p) \in \mathbb{R}$, the associated PTV volume as $S(p) \in \{0, 1\}$, and the corresponding mPD volume as $D(p) \in \{0, 1\}$ in a

discrete space $p \in \Omega$, where $\Omega \subset \mathbb{R}^3$. A unique spatial pattern $\mathbf{u}$, in space Ω, is used to convert the original volumes into row vectors $\mathbf{x}_v$, $\mathbf{x}_l$, and $\mathbf{x}_d$, respectively, so that each voxel is represented with the same index in all three vectorized modalities. We form a joint observation vector $[\mathbf{x}_v\mathbf{x}_l\mathbf{x}_d]$ by concatenating these data vectors, and hereafter refer to i^{th} patient record with $\mathbf{x}^{(i)}$.

mPD Volume Generation

According to the guidelines used at our local cancer centre, we use the updated dosimetry formalism proposed by Task Group #43 [84] in order to obtain the target isodose contours from seed locations. We follow the point source (1-D) formalism for the radio-active seeds. The general 1-D formalism introduces minimal error and simplifies the dose calculation by ignoring the seed orientation. The maximum initial dose $\dot{D}$ is expressed as

$$\dot{D}(r) = S_K.\Lambda.\left(\frac{r_0}{r}\right)^2.g_p(r).\phi_{an}(r), \tag{3.3}$$

where S_K represents air-kerma strength of the radio-active seed (in μGym^2h^{-1}, aka U), Λ is dose-rate constant in water (in $cGyh^{-1}U^{-1}$), r_0 denotes the reference distance which is 1 cm from the seed center, g_p accounts for photon attenuation and scatter effects and ϕ_{an} accounts for anisotropies due to seed construction. Total dose D deposited by a single point model source is calculated as

$$D(r) = \dot{D}(r)\int_{t=0}^{t=T} e^{-\lambda t}dt, \tag{3.4}$$

where T is the time of radiation (in h). The isodose contour of the prescribed dose (usually set at 144 Gy) outlines the tissue volume that receives the prescribed mPD or greater. Maximum PTV coverage and conformity of this dose distribution are important factors considered in devising a treatment plan for a new patient.

Model Generation

A joint three-modality observation matrix $\mathbf{X}$ is formed by stacking all patient records $\mathbf{x}^{(i)}, i = 1, 2, ..., N$, where N is the size of training dataset. We generate a set of K joint independent components $\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, ..., \mathbf{c}^{(K)}$ using

the InfoMax approach [85], so that

$$\mathbf{C} = \mathbf{W}\hat{\mathbf{X}}, \quad where \;\; \mathbf{C} = \begin{bmatrix} \mathbf{c}_v^{(1)} & \mathbf{c}_l^{(1)} & \mathbf{c}_d^{(1)} \\ \mathbf{c}_v^{(2)} & \mathbf{c}_l^{(2)} & \mathbf{c}_d^{(2)} \\ \vdots & \vdots & \vdots \\ \mathbf{c}_v^{(K)} & \mathbf{c}_l^{(K)} & \mathbf{c}_d^{(K)} \end{bmatrix}, \tag{3.5}$$

where $\mathbf{W}$ is the demixing matrix and each component $\mathbf{c}^{(j)}$ is divided into three modality parts $[\mathbf{c}_v^{(j)}\mathbf{c}_l^{(j)}\mathbf{c}_d^{(j)}]$, and $\hat{\mathbf{X}}$ is the whitened observation matrix. The ICA algorithm is sensitive to the whitening process which involves centering the data vectors and transforming them so that the covariance matrix is close to identity. The whitening process decorrelates the data vectors prior to the decomposition step. In general, ICA works with higher order statistics of the observation data and assumes that the hidden components are statistically independent, non-Gaussian with linear mixing process.

Component Stability

The ICA decomposition algorithms do not necessarily identify stable and reliable components from an existing dataset. This uncertainty correlates with the inherent properties of the ICA contrast function and the optimization techniques undertaken. Moreover, the observation dataset is mostly limited in size and may induce statistical errors in the component estimation process. One of the approaches to overcome this problem is to repeat the decomposition process multiple times and introduce an average estimation of the results. Similar to the approach proposed in [86], we repeat the joint ICA decomposition m times in order to obtain the K components. Demixing matrices of $\mathbf{W}_i$ are stacked into a matrix $\dot{\mathbf{W}}$ in order to calculate the mutual correlation coefficient from

$$\mathbf{R} = \dot{\mathbf{W}}\hat{\mathbf{X}}\hat{\mathbf{X}}^T\dot{\mathbf{W}}^T, \tag{3.6}$$

where an element of matrix $\mathbf{R}$ represented by r_{ij} refers to the mutual correlation between components $\mathbf{c}^{(x)}$ and $\mathbf{c}^{(y)}$ in two different decomposition attempts. A hierarchical clustering on the dissimilarity coefficients $1 - |r_{ij}|$ groups the corresponding components in different decomposition attempts into clusters. We introduce centroids of the clusters as the stable joint components of the model.

A key parameter in our modeling approach is the number of joint components, K. Ideally, minimum reconstruction error is achieved when the

maximum possible number of components is considered in the ICA algorithm. However, some of the generated components will reflect the systematic noise and the variability enforced by the joint decomposition algorithm. A trade off between noise cancellation and reconstruction error minimization is targeted by a hold-out validation approach. We randomly separate M patient records for the validation dataset. The validation set is never used in the prediction experiments. We choose the optimal number of components by observing the reconstruction error in three modalities on the validation dataset.

Prediction

For a new unseen TRUS volume $\mathbf{x}_v^t$, we solve a least squares problem to find the projection coefficients $\mathbf{a}$, so that

$$\mathbf{x}_v^t \approx \mathbf{a}^T \begin{bmatrix} \mathbf{c}_v^{(1)} \\ \mathbf{c}_v^{(2)} \\ \vdots \\ \mathbf{c}_v^{(K)} \end{bmatrix}. \tag{3.7}$$

According to the modeling constraint, same coefficients can be adapted to reconstruct the other two modalities, the PTV and the mPD volume. However, a linear weighted combination of the PTV and isodose parts of the components produces probabilistic estimates $\tilde{\mathbf{x}}_l^t$ and $\tilde{\mathbf{x}}_l^t$ for the target binary volumes. We propose to generate a threshold mask from the training dataset used for modeling. For each 2-D axial plane of the sparse volume out of S planes, we calculate optimal plane threshold $\hat{\tau}_s, s = 1, 2, ..., S$:

$$\hat{\tau}_s = \underset{\tau}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} |\mathbf{x}_l^{(i)} - b(\mathbf{a}_i^T \mathbf{C}_l, \tau)|, \quad 0 < \tau < 1, \tag{3.8}$$

where $b(\mathbf{v}^T, \tau)$ is a function that converts the normalized scalar volume $\mathbf{v}$ into a binary volume by applying the threshold value τ. For any test volume, after obtaining the probability segmentation map, we apply $\hat{\tau}_s$ to plane s, in order to estimate the target segmentation contour in that plane. Furthermore, edge contours of the estimated binary volume are smoothed slice by slice to produce the final estimation of the target segmentation. Contours are represented in a polar coordinate system assuming their geometrical center as the origin and further transformed into the Euler coordinate system to perform B-spline curve fitting. The generated contours are smoothed to produce the final estimation of the target segmentation.

3.2.3 Experiments

Records from a cohort of 120 patients who had undergone brachytherapy treatment were used in the proposed modeling approach. We only included TRUS volumes of axially approximately equal length prostates in the cohort of this study. We enforce this constraint to perform the feasibility study while satisfying the approximate anatomy alignment condition. For each TRUS volume, the PTV and seed coordinates are provided by expert radiation oncology practitioners. We calculate the dose distribution map and subsequently, the mPD volume of 144 Gy based on a point source model described in Section 3.2.2. TRUS volumes consist of nine equally spaced (5 mm) 2D ultrasound images, collected from base to apex of the prostate in the transverse plane and in accordance with the VCC protocol.

In all analyses, we randomly use 80% of the cases in our training dataset and evaluate the hypothesis on the remaining 20%. To avoid bias in the modeling process, we repeat the jICA decomposition with stable components for 10 times.

Predicted PTV and mPD volumes are evaluated against the gold-standard. We evaluate PTV with volume error (V_{err}) and volume difference (V_{diff}) as shape and size dissimilarity metrics [1, 54, 60]. In addition to the volumetric measures, we validate the predicted mPD volume based on the dosimetry parameter V100. The V100 indicates the percentage of the PTV volume receiving 100% of the prescribed dose.

3.2.4 Results and Discussion

To determine the optimum number of components in our modeling approach, we calculate the reconstruction error for PTV and mPD volume using different numbers of components as shown in Figure 3.6. Intuitively, the number of components in our model represents the complexity of the data vector space as well as the power of the model in capturing the correlation between modalities. The smaller the number of components, the larger the reconstruction error observed due to the fact that the model is unable to capture the correlation between modalities. As the number of components grows, the reconstruction error is reduced; however, there is an increase when the number of components is close to the total number of cases included in the training dataset. This can be due to the stronger presence of components that reflect noise in the training dataset. Our linear decomposition approach equivalently optimizes the weights for all components, hence is prone to miss-guidance by the these components. We choose the optimum

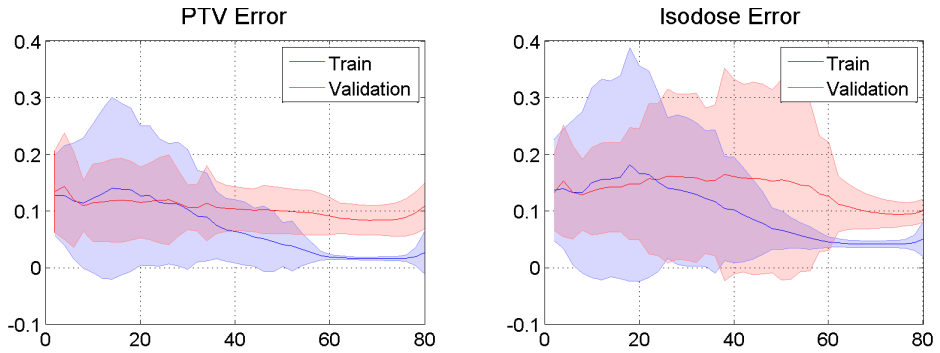


Figure 3.6: Reconstruction error V_{err} calculated for PTV and Isodose on training and validation datasets.

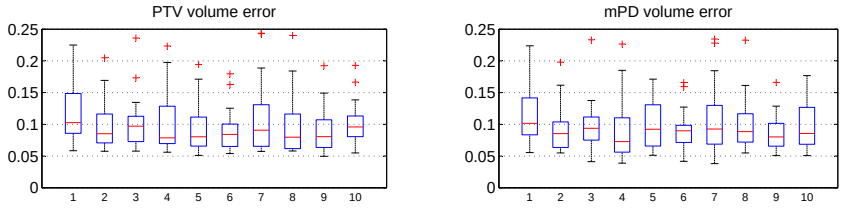


Figure 3.7: V_{err} statistics shown with boxplots from 10 random dataset partitioning into train and test sets.

number of 40 components as a compromise on lower reconstruction error among the three modalities.

Figure 3.7 shows the statistics of the V_{err} in 10 random dataset partitioning. The obtained results show that PTV and mPD volume prediction errors are not statistically different. It can be inferred that the ICA-based modeling approach is able to equally capture the correlation between both PTV and isodose volumes and the original TRUS volume.

The calculated grand mean and pooled standard deviation of the metrics from 10 trials are shown in Table 3.2. We calculated $V100$ with respect to the gold standard PTV. At the Vancouver Cancer Centre, clinical guidelines recommend a value greater than 98% for the dosimetry parameter of $V100$. If we include the predicted PTV in $V100$ calculation, we observe that the proposed modeling approach always meets the planning standard $V100$ measure, i.e. $98.46 \pm 1.09\% \approx 98\%$ compared to the gold standard $V100$ measured at $97.51 \pm 0.56\%$. Paired t-test statistical significance analysis fails to show any difference in the mean value of the calculated $V100$ for

the proposed method and the gold-standard with p-value< 0.05. Five 2D TRUS images (from base to apex 1 cm apart) of a randomly chosen patient record are shown in Figure 3.8. Top row shows the treatment gold standard compared to the bottom row as the predicted PTV and mPD volume mask overlaid on the TRUS images.

Table 3.2: Grand mean and pooled standard deviation of V_{err}, V_{diff} and $V100$ calculated for the estimated PTV and mPD volume in 10 different random training/testing datasets.

PTV		Isodose		
V_{err}	V_{diff}	V_{err}	V_{diff}	V100
$10.02 \pm 4.5\%$	$7.55 \pm 14.77\%$	$9.74 \pm 4.23\%$	$9.21 \pm 14.17\%$	$97 \pm 3.55\%$

3.2.5 Conclusion

We proposed a reconstruction based modeling approach to find the correlation between TRUS volume, PTV and the mPD volume. We employed the jICA algorithm on a set of treatment records in order to introduce joint independent components representing the correlation between these modalities. Instability of the components caused by the jICA algorithm was reviewed by clustering the components based on the correlation between jICA demixing coefficients in different trials. We also introduced parameter tuning procedures for the number of components, and the binary conversion mask for the mixed components.

The proposed approach enables us to project the prior knowledge in a set of brachytherapy treatment records onto an unseen TRUS volume. Consequently, it can lead to reduced inter- and intra-observer variability associated with the brachytherapy PTV delineation and seed planning processes; as the dose plan and expert's knowledge can contribute to the treatment outcome and morbidity rate [87, 88]. The proposed modeling approach is applicable in risk stratification and toxicity/morbidity prediction. It runs in less than a couple of seconds on a typical PC and learns the correlation pattern inside the dataset regardless of the prostate boundary and texture features, therefore, is an applicable approach to capture institution/clinician specific knowledge. Future work will include seeds coordinates determination in addition to the 150% isodose contours which are not contiguous in all planes.

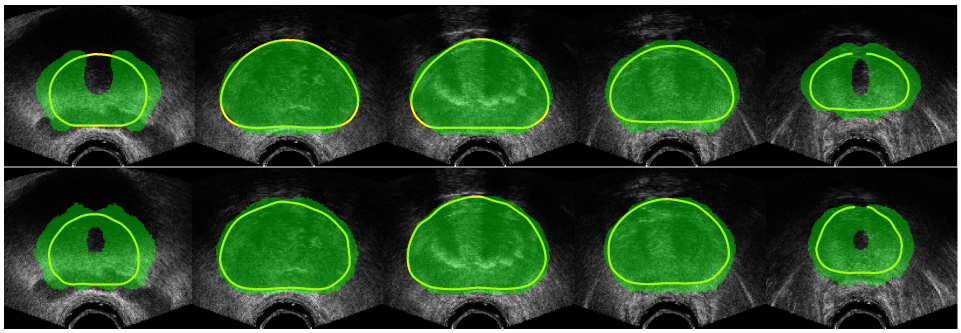


Figure 3.8: Gold standard (top) and predicted (bottom) TRUS, PTV and mPD volume (base to apex, from left to right, 1 cm apart) of a randomly chosen test case.

Chapter 4

Sparsity-based Fusion for Brachytherapy Preplanning

4.1 Introduction

There is no evidence supporting the idea that joint patterns between volumetric information of brachytherapy such as TRUS and CTV volumes are statistically independent. Although, this assumption has led the modeling approach to look for joint patterns that are spatially locally distributed in the extent of the joint space. Hence, sparse pattens can potentially be a new perspective in defining the joint patterns in the fused information space. Moreover, sparsity can be seen as a generalization of the ICA-based modeling that produces joint components that are sparse instead of statistically independent. Hence, here in this chapter we investigate sparse analysis and its strength in representing the joint patterns in the fusion framework.

In this chapter we aim to propose a fusion algorithm that automates delineation of both CTV and PTV contours. Although there are some clinical guidelines but definition of the CTV and the PTV is not definite and varies from case to case. Hence, most of the previous TRUS segmentation approaches are not directly applicable to CTV and PTV segmentation. We have visualized two sample actual treatment cases in Figure 4.1 to emphasize on the difference between prostate boundary and the two target volumes. Figure 4.1 shows some anatomical structures in the anterior region of the mid-gland that are outside the prostate boundary but inside the CTV contours. In general, CTV is not necessarily the real prostate boundary and is defined for each patient to maximize the likelihood of receiving prescribed radiation dose for the target anatomy while minimizing the irradiation to the surrounding tissue. According to International Commission of Radiation Units and Measurements (ICRU Report 50) and ABS,

This chapter is adapted from [89]: S. Nouranian, M. Ramezani, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Learning-based Multi-label Segmentation of Transrectal Ultrasound Images for Prostate Brachytherapy, IEEE Trans. Med. Imaging, vol. 35, no. 3, pp. 921-932, 2016.

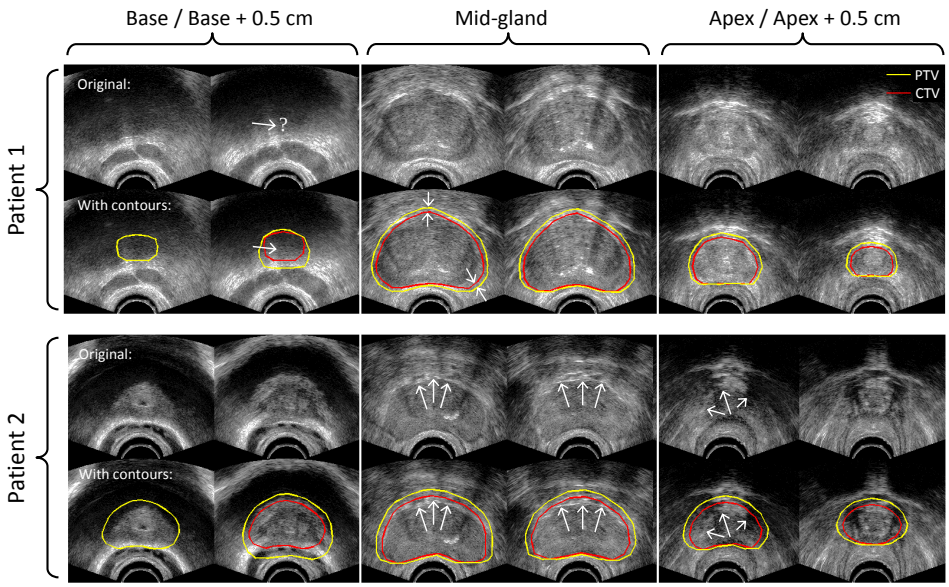


Figure 4.1: Sample CTV and PTV contours used for treatment of two patients. CTV is not necessarily the prostate boundary: 1) the prostate boundary is not visible in near base and apex views (patient 1 base views), and 2) in mid-gland views some surrounding tissue is included in CTV (patient 2 mid-gland and apex views). PTV-CTV margin varies within and between patients (patient 1 mid-gland views); hence, PTV is anisotropic dilation of CTV per patient based on CTV, guidelines and expertise of radiation therapist.

CTV is an anatomical-clinical concept that contains cancerous tissue and sub-clinical microscopic malignant disease, and may contain adjacent tissues such as extra-capsular extensions of the prostate. Although there are some guidelines for delineation of the PTV based on CTV, PTV is also determined based on some patient specific clinical parameters (prostate length, size, shape and visibility) and more specifically the clinician's expertise. A visual comparison of the two TRUS images in Figure 4.1 shows that PTV contours are not a predefined dilation of the CTV and size of the margin in different regions of the prostate varies from case to case. Moreover, the prostate boundary is not usually visible close to the base and the apex of the prostate in TRUS axial views. Furthermore, most of the previous TRUS segmentation reports either need user interaction or require access to dense TRUS volumes. Also our approach for multi-atlas-based segmentation (see Chapter 2) shows clinically acceptable results [54, 60, 61], but none of these methods meet requirements for a real-time, automatic segmentation application such as speed which is highly demanded for intraoperative clinical applications. Moreover, none of the methods provide two target anatomical labels for TRUS voxels, i.e. CTV and PTV, and therefore require two independent, unconstrained and sequential segmentation attempts.

In this work, we aim to incorporate prior knowledge in a large dataset of treatment record observations including sparse TRUS volumes and the CTV/PTV labels to introduce a fast and automatic multi-label segmentation method. We introduce our solution in the form of a joint model that captures the relation between the three observations. The proposed model is used to simultaneously estimate the CTV and the PTV labels for a new patient, solely based on the TRUS volume. We propose a sparse representation of joint observations and train a set of joint sparse dictionary elements, aka *atoms*. Each atom is a non-linear model of a particular joint pattern between three volumetric information spaces. Consequently, projection of a partial observation (i.e. TRUS volume) onto the trained dictionary elements (i.e. joint atoms) estimates a set of coefficients suitable to combine other parts of the atoms that correspond to CTV and PTV labels.

To the best of our knowledge, this work is the first framework proposed for fast and automatic prostate brachytherapy treatment target labeling of the TRUS images. In summary, key contributions of the proposed method are: 1) We constrain the dictionary learning algorithm to not only satisfy the sparsity but also to force equal contributions for each volumetric information space of the joint atoms. Hence, we obtain an optimum set of coefficients from one volumetric information space (i.e. TRUS) and apply it to the other expected volumetric information spaces (i.e. CTV and PTV).

2) In contrast to other statistical shape modeling approaches, this method is not iterative (fast) and does not rely on local shape and appearance characteristics and incorporates all voxel values. Architecture of the proposed framework including choice of the pre-processing and the proposed thresholding algorithm to convert estimated probability maps of CTV and PTV into binary labels are specifically tailored to the context of this study.

We evaluate the proposed method on 590 patient records and compare estimated contours against the manually segmented treatment target contours that were used by clinicians to plan the seed arrangements. A 5-fold cross validation shows clinically acceptable results that also satisfies speed requirements for a real-time clinical application. We compare performance of the proposed multi-label model with single-label models generated from joint TRUS/CTV and joint TRUS/PTV labels, and show advantage of multi-label model over single-label models. We also show an added value of the proposed method: it can adapt to different clinics, clinicians or treatment groups, as it inherently captures the expertise and implicit guidelines used in different centres in the form of a dataset specific sparse dictionary.

4.2 Methods

Figure 4.2 illustrates the proposed pipeline for multi-label segmentation of TRUS images. We propose a joint sparse dictionary learning approach in a one-time offline process of *learning* performed on a set of observations from TRUS, CTV and PTV. Subsequently, the generated model is used in a real-time process of *estimation*. In the following sections, we introduce observation space, modeling and its successive predictors.

4.2.1 Problem Definition and Observation Space Formation

We represent patient volumetric information in a unique discrete Cartesian space of $\Omega \subset \mathbb{R}^3$. We assume all patients' volumetric information datasets are coregistered in this Cartesian space for further analysis. We define three volumetric information spaces of target anatomy within the discrete space Ω as TRUS intensity and corresponding binary labels for the CTV and the PTV. We assume there exist certain observable patterns for each voxel w.r.t. the other voxels within Ω in and among three volumetric information spaces. Our goal is to learn joint patterns from a dataset of observations by including multiple volumetric information spaces in one unique learning algorithm. Our approach is to generate a joint sparse dictionary of elements [90], i.e. atoms, so that each volumetric information space can be

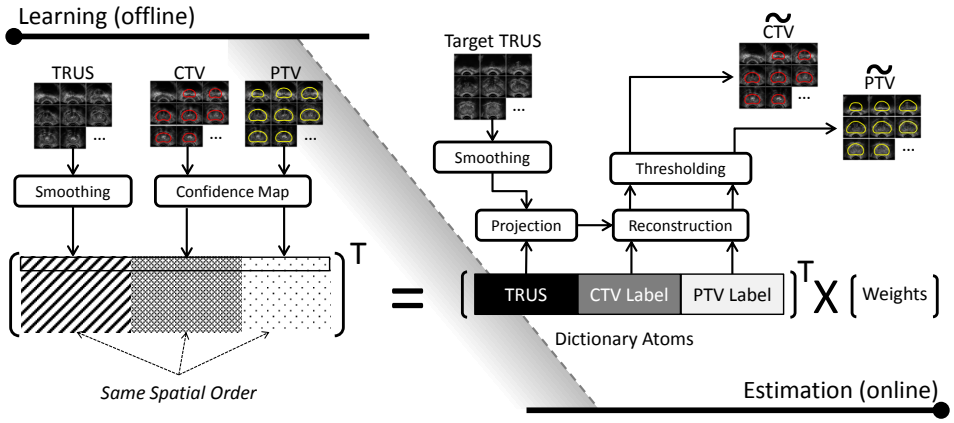


Figure 4.2: Illustration of the proposed system for simultaneous automatic labeling of the TRUS volumes.

represented by a weighted linear combination of these atoms. For each patient $i \in \{1, 2, ..., N\}$ in a dataset of N patient treatment records, we use a global spatial indexing order in the space Ω to generate three observation vectors of $\mathbf{v}^{(i)} \in \mathbb{R}^{d\times 1}$, $\mathbf{l}_{ctv}^{(i)} \in \{0,1\}^{d\times 1}$ and $\mathbf{l}_{ptv}^{(i)} \in \{0,1\}^{d\times 1}$, representing TRUS intensities, CTV and PTV labels, respectively, where d is the number of discrete coordinates in Ω. Therefore, each voxel is represented with the same index in all vectorized triplets of volumetric information spaces. We introduce a full joint observation vector as concatenation of the three observations $[{\mathbf{v}^{(i)}}^T, \mathbf{l}_{ctv}^{(i)T}, \mathbf{l}_{ptv}^{(i)T}]^T$.

Presence of noise in observation vectors can potentially decrease performance of the training process and consequently make the atoms less likely to capture the correlation among input volumetric information spaces. This is important when it comes to the noisy and speckled TRUS images. We aim to reduce the input noise in TRUS intensities by adapting a Gaussian smoothing kernel to convolve with 2-D axial images. The filtered images are then used to generate smoothed TRUS observation vector of $\hat{\mathbf{v}}^{(i)}$.

CTV and PTV delineations are also affected by both intrinsic noise and observer uncertainty. Hence, we adapt an approach [91] to provide confidence measures from binary segmentations. We aim to assign a confidence degree to voxels based on their distance to the segmentation boundary. Original CTV and PTV binary volumes are converted into signed distance maps, and vectorized to produce signed distance vectors $\mathbf{r}^{(i)}$. Then we apply a sigmoid function to calculate voxel's likelihood to be inside the target label

binary volume:

$$\hat{\mathbf{r}}^{(i)} = P(\mathbf{r}^{(i)}) = \frac{1}{1 + \exp(-\alpha \mathbf{r}^{(i)})}, i \in 1, 2, ..., N. \tag{4.1}$$

This transformation produces a likelihood map, in which the points in the neighborhood of the target boundary are associated with lower confidence degree than the points sitting at the boundary. Parameter α, the slope of the sigmoid function, determines the neighborhood size for confidence degree calculation. This parameters is empirically set to 0.8 to represent ~ 3 mm margin of less than 95% confidence in label assignment.

Finally, we generate the joint observation matrix $\hat{\mathbf{X}}$ by stacking processed joint observation vectors of $\mathbf{x}^{(i)} = [\hat{\mathbf{v}}^{(i)T}, \hat{\mathbf{r}}_{ctv}^{(i)T}, \hat{\mathbf{r}}_{ptv}^{(i)T}]^T$. Each voxel can be considered as a feature in the learning process; however, due to the stationarity characteristics of the observations from voxels, features in a neighborhood reflect a high correlation. We decorrelate the observation features by whitening the input observation matrix. Hence, we train the model on the whitened observation matrix $\tilde{X}$ to obtain a joint sparse dictionary representing global joint patterns in Ω for the three volumetric information spaces.

4.2.2 Joint Sparse Dictionary Learning

In our joint learning approach, we aim to introduce $\mathbf{D} \in \mathbb{R}^{3d \times K}$ as a dictionary of K atoms that satisfies $\|\mathbf{x}^{(i)} - \mathbf{D}\mathbf{c}^{(i)}\| \leq \epsilon$, where $\mathbf{c}^{(i)} \in \mathbb{R}^{K \times 1}$ is a vector of weights obtained for joint observation vector $\mathbf{x}^{(i)}$ with fewest number of nonzero elements.

Since we introduce the joint observation vector $\mathbf{x}^{(i)}$ to the learning algorithm, the generated dictionary $\mathbf{D}$ contains atoms that jointly represent the observations in three parts, i.e. $\mathbf{D} = [\mathbf{D}^{vT}, \mathbf{D}^{l_{ctv}{}^T}, \mathbf{D}^{l_{ptv}{}^T}]^T$. Hence, atom coefficients are constrained to be equal for all three volumetric information spaces while satisfying the sparsity and optimal reconstruction conditions throughout the learning process.

We reformulate the training process for the whitened joint observation matrix $\tilde{\mathbf{X}} \in \mathbb{R}^{3d \times N}$ and express the joint dictionary learning as $\tilde{\mathbf{X}} \approx \mathbf{D}\mathbf{C}$, where $\mathbf{C} \in \mathbb{R}^{K \times N}$ refers to the coefficient matrix [92].

We use the K-singular value decomposition (K-SVD) method [93] to train the joint model and generate the dictionary matrix $\mathbf{D}$. K-SVD is a generalization of the K-means clustering algorithm for vector quantization,

and can be defined as the following optimization problem:

$$\min_{\mathbf{D},\mathbf{C}} \left\{ \|\tilde{\mathbf{X}} - \mathbf{DC}\|_F^2 \right\} \text{ subject to } \forall_i, \|\mathbf{c}^{(i)}\|_0 \leq \mathbf{T}_0, \tag{4.2}$$

where $\|.\|_F$ is the Frobenius matrix norm and $\mathbf{T}_0$ is the sparsity level or the number of nonzero coefficients. K-SVD as a minimization problem solver consists of two steps: 1) assuming $\mathbf{D}$ is fixed, an orthogonal mapping pursuit algorithm [94] is used for sparse coding, i.e. obtaining $\mathbf{c}^{(i)}$ by decoupling the optimization problem into N individual sub-problems. 2) Dictionary atoms are updated one by one. Assuming $\mathbf{D}$ and $\mathbf{C}$ are fixed except one column (joint atom) in the dictionary matrix, i.e. $\mathbf{d}_j = [\mathbf{d}_j^{v^T} \mathbf{d}_j^{{l_{ctv}}^T} \mathbf{d}_j^{{l_{ptv}}^T}]^T$, an optimization problem is defined with a penalty term as:

$$\|\tilde{\mathbf{X}} - \mathbf{DC}\|_F^2 = \|\mathbf{E}_j - \mathbf{d}_j \mathbf{c}_j\|_F^2, \tag{4.3}$$

where $\mathbf{E}_j$ is the summation of the error for all N observations when jth atom is removed and $\mathbf{c}_j$ is a vector of coefficients of that joint atom for all observations. Then, representation error $\hat{\mathbf{E}}_j$ is calculated from those error columns corresponding to observation that use atom $\mathbf{d}_j$ (with nonzero coefficient). Eventually, atoms and their corresponding sparse weights are sequentially updated using the singular value decomposition:

$$\hat{\mathbf{E}}_j = \mathbf{U}\boldsymbol{\Delta}\mathbf{V}^{\mathbf{T}} \tag{4.4}$$

Solution for $\mathbf{d}_j$ is then defined as the first column of $\mathbf{U}$ and coefficient vector $\mathbf{c}_j$ as first column of $\mathbf{V}$ multiplied by $\boldsymbol{\Delta}(1,1)$.

Since the K-SVD approach is sensitive to the initial conditions, i.e. the initial dictionary, we reduce this sensitivity by conducting independent component analysis on training dataset. We adopt the InfoMax approach [85] to obtain K initial dictionary atoms for the joint sparse dictionary learning algorithm. We let the K-SVD algorithm iterate through both steps for 30 times to generate a sparse joint dictionary $\mathbf{D}$ in our offline learning process.

4.2.3 Multi-label Estimation

In a real-time *estimation* process (see Figure 4.2), a new unseen TRUS image set $\mathbf{v}^t$ is smoothed by the same Gaussian kernel to provide an incomplete observation vector of $\hat{\mathbf{v}}^t$. We solve a least squares problem, i.e. the projection step, to find the optimal set of dictionary coefficients $\mathbf{c}^t$ as

$$\tilde{\mathbf{c}}^t = \underset{\mathbf{c}^t}{\operatorname{argmin}} \|\hat{\mathbf{v}}^{\mathbf{t}} - \mathbf{D}^v \mathbf{c}^t\|_2^2, \tag{4.5}$$

where $\|.\|_2^2$ denotes the Euclidean norm.

We use the obtained coefficients from the projection to linearly combine the other two parts of the atoms (related to CTV and PTV observations) and calculate $P(\mathbf{l}_{ctv}^t) = \mathbf{D}^{l_{ctv}}\mathbf{c}^t$ and $P(\mathbf{l}_{ptv}^t) = \mathbf{D}^{l_{ptv}}\mathbf{c}^t$. However, $P(\mathbf{l}_{ctv}^t)$ and $P(\mathbf{l}_{ptv}^t)$ are probabilistic estimates of the CTV and the PTV for the test case. We propose to generate a threshold mask from the observation matrix $\hat{\mathbf{X}}$ to convert the probability estimations into a binary estimation of the CTV and the PTV. We calculate optimal plane-wise threshold $\hat{\tau_q}, q = 1, 2, ..., Q$ for CTV and PTV labels separately:

$$\hat{\tau_q} = \underset{\tau}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^{N} |\mathbf{l}^{(i)} - \delta(\mathbf{D}^l \mathbf{c}^{(i)}, \tau)|, \quad 0 < \tau < 1, \tag{4.6}$$

where Q is the number of axial planes acquired for the volume study in space Ω, and $\delta(\mathbf{y}, \tau)$ is a threshold function that converts the normalized scalar observation vector of $\mathbf{y}$ into a binary vector by applying the threshold value τ. We apply $\hat{\tau_q}$ to plane q for the test case, in order to estimate the CTV and the PTV labels per plane.

4.3 Materials

A retrospective cohort of 590 patients who underwent prostate brachytherapy at the Vancouver Cancer Centre (VCC) is used to evaluate the proposed method. For each patient, a complete set of observations consisting of TRUS images, CTV and PTV contours is collected. Volume study routine in VCC involves TRUS image collection from the base to the apex of the prostate captured every 5 mm in the axial direction using a side-firing endorectal probe, hence the resulting image set is a sparse representation of the volume of the prostate and the adjacent tissue. According to the size of the prostate, total 7-14 TRUS images of size 415×490 pixels (physical spacing of 0.1557 mm by 0.1560 mm) were collected per patient. The institutional clinical volume study protocol obliges the prostate to be maintained in the middle of the axial plane, making the anatomy appear symmetric or nearly symmetric in all ultrasound images. These series of 2-D images are then contoured semi-automatically [1] and further modified to produce CTV contours. Following the clinical guidelines and radiation oncologist expertise, the PTV contours are drawn. We convert both contour sets into binary volumes for the purpose of our study. According to the aforementioned workflow, TRUS volume, CTV and PTV are all confined axially within the base to the apex

of the prostate for each patient. Therefore, target anatomy is implicitly co-registered among all patients. No re-sampling of the images was necessary, since, all TRUS volumes were acquired with a consistent depth parameter and the same transducer type; hence, all pixel observations reside on the unified coordinate system of Ω.

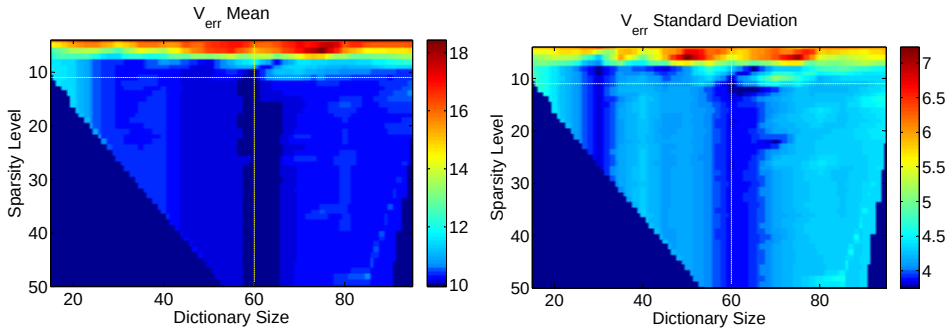
We evaluate the performance of the multi-label estimator and compare it with contours of the actual delivered treatment plans, by calculating volume shape and size similarity coefficients; namely, V_{err} and V_{diff} in different anatomical sectors of the prostate [24]. We also measure distance between actual and estimated CTV and PTV surfaces measures in terms of Hausdorff Distance (HD) and Mean Surface Distance (MSD).

4.4 Experiments and Results

A 5-fold cross validation approach was implemented to measure accuracy of the multi-label estimation. We performed some experiments to tune the hyper-parameters of the proposed model. Subsequently, we validated the model on all the 590 cases to review its performance in general. Additionally, we divided the cohort into physician specific sub-groups in order to investigate adaptability of the method. In all experiments we implemented a 5-fold cross validation approach after hyper-parameter tuning.

4.4.1 Model Parameter Tuning

We randomly partitioned the cohort into 80% training and 20% validation groups, to tune the hyper-parameters of the proposed model, namely, number of atoms and level of sparsity. Partitioning for parameter tuning is performed once to obtain an estimation of the complexity of the joint information in the dataset. Although tuning can be based on multiple observations from repeated random partitioning, we did not find performance of the proposed algorithm be very sensitive to these parameters. We characterized performance of the model by measuring volume errors for estimated CTVs V_{err}^{ctv} on the validation dataset. An exhaustive grid search is carried out to determine complexity of the patterns in terms of the minimum number of atoms and sparsity level to estimate CTV labels from TRUS volumes. Figure 4.3 depicts mean and standard deviation of CTV V_{err} calculated for different values of sparsity level and dictionary size. Cold region (shown in dark blue) in both heat maps represents combination of the two hyper-parameters that produces results with both lower mean and lower standard deviation of the error. It can be observed that increasing dictionary size



(a) (b)

Figure 4.3: Hyper-parameter tuning by grid search on sparsity level and dictionary size; a) the mean, and b) the standard deviation of CTV V_{err} calculated for validation group, i.e. random partitioning of the dataset into 80% training and 20% validation. Dashed line shows the optimum parameters used in this study.

and sparsity levels includes larger number of atoms that represent noise in the system. Moreover, very small sparsity level fails to learn patterns in the dataset. As a trade-off between estimation performance and fewer number of atoms and coefficients, we chose a dictionary size of 60 atoms with sparsity level of 11, as we observed lower validation error variation in estimating the CTV.

4.4.2 Multi-label Estimation

Figure 4.4 illustrates some of the patterns captured by the proposed joint sparse dictionary learning. It is observed that atoms jointly represent spatially distributed patterns as is highlighted in Figure 4.4. Each atom shown in this figure represents observation variations in different regions, e.g. base-anterior, posterior, lateral and apex-anterior. This figure also shows how joint patterns are captured, as the PTV patterns are highly correlated with CTV patterns per each atom.

General statistics of the estimation performance characterized by volumetric measures in different anatomical sectors of the prostate are shown in Table 4.1 and Figure 4.5. We also measured MSD and HD as 0.95 ± 0.41 mm and 5.33 ± 1.47 mm for CTV, and 1.19 ± 0.48 mm and 5.48 ± 1.51 mm for PTV.

It can be observed that a lower estimation error is obtained in mid-

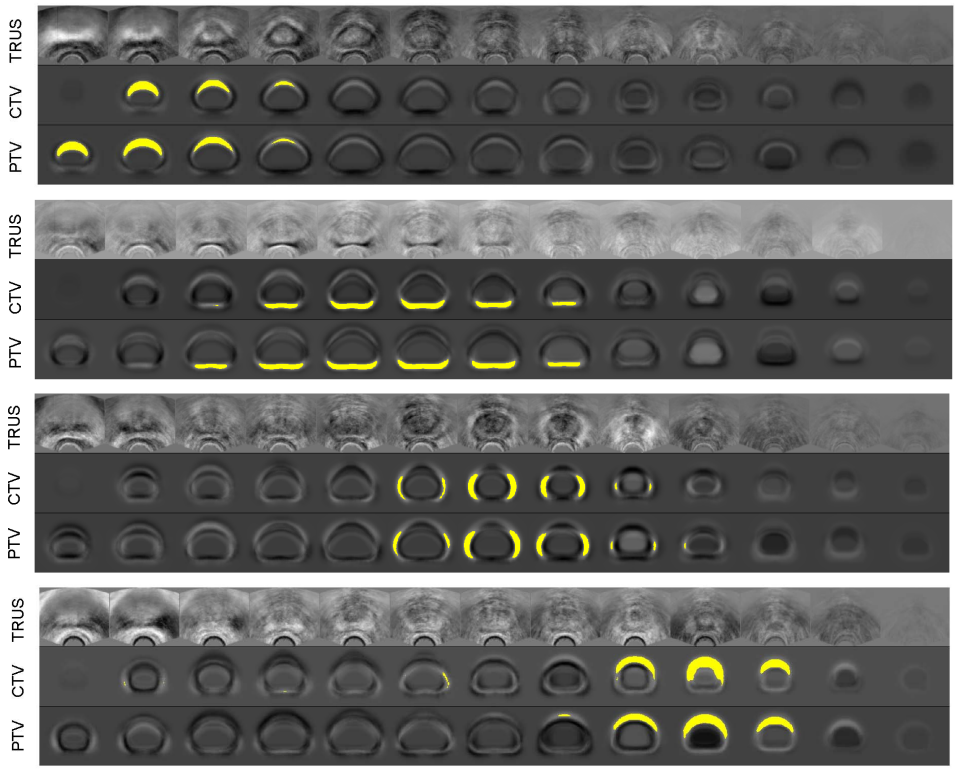


Figure 4.4: Learned sparse dictionary atoms. Each joint atom represents region-specific modulation in base-anterior, posterior, lateral and apex-anterior regions for all three volumetric information spaces. Highlighted area shows coordinates with more than 50% modulation value within the same volumetric information space.

Table 4.1: Mean and standard deviation of V_{err} calculated for nine anatomical sectors of the prostate over CTV and PTV.

V_{err}%	**Base**		**Mid**		**Apex**		**Total**	
	CTV	PTV	CTV	PTV	CTV	PTV	CTV	PTV
Ant.	11.33±8.53	10.79±7.63	8.75±6.91	8.38±6.36	13.81±9.46	12.69±8.83	9.92±3.51	8.84±3.13
Lat.	8.42±4.19	7.40±3.32	7.41±3.53	6.49±3.15	11.35±5.48	9.59±4.76		
Post.	8.87±6.79	7.91±6.06	7.44±6.31	7.06±5.97	12.15±9.05	11.21±8.41		

posterior and mid-lateral sectors of the prostate. However, in anterior sectors and apex region a larger deviation from the actual PTV is measured. It is reasonable to consider that the prostate boundary is poorly visible anteriorly and particularly in close to apex and base TRUS images; hence,

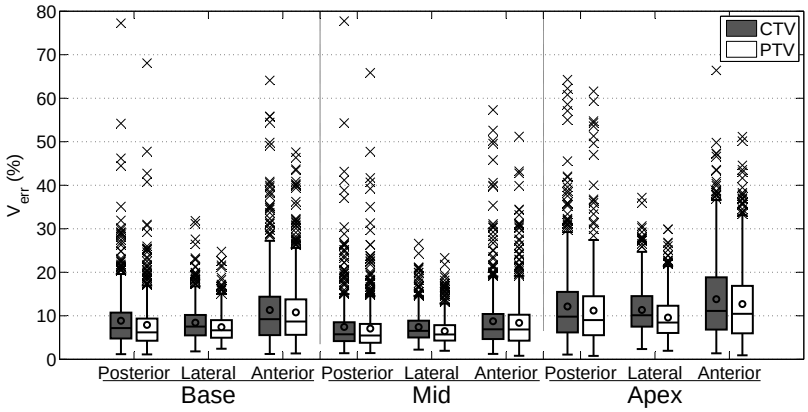


Figure 4.5: Label estimation errors in terms of V_{err} in different anatomical sectors of the prostate. Circles represent mean and the horizontal line the median of the calculated error in each sector.

a greater observer variability is expected. Figure 4.5 also shows that less number of outliers are seen in lateral sectors. In general, we observe lower volumetric errors for the estimated PTV in all sectors as well as the entire volume. Two typical multi-label estimation results are shown in Figure 4.6 per slice from base to apex of the prostate.

4.4.3 Dataset Adaptability Analysis

We extract three sub-groups from the cohort of 590 patient records based on physician in charge of the treatment planning. We selected the three physicians with the largest number of assigned cases, hereafter referred to as physician A, B and C. To review dataset adaptability of the proposed multi-label segmentation approach, we train the model for groups A, B and C independently and perform 5-fold cross validation within each group (internal validation). Subsequently, we use the trained model obtained for each fold to validate on the remainder of the entire 590 patient cohort (external validation).

Hyper-parameter tuning for group A (174 cases), B (156 cases) and C (135 cases) is performed as explained in Sec. 4.4.1. General statistics of the volumetric error measure V_{err} for the three groups are shown in Figure 4.7.

Table 4.2: Comparison to two other brachytherapy CTV segmentation methods for sparse TRUS volumes. "⋆" and "•" indicate a statistically significant difference between mean and standard deviation of the error measured by the proposed multi-label approach and each of the reported methods using paired t-test and Levene's test, respectively.

	Volumetric Errors (%)		**Surface Errors (mm)**		**Speed (sec)**	**Automatic**
	V_{err}	V_{diff}	MSD	HD		
Mahdavi et al. [1]	8.81±5.60 ⋆ •	3.56±14.66 ⋆	1.11±0.81 ⋆ •	5.57±2.28 •	$\sim 120 - 240$	No
Nouranian et al. [54]	10.25±3.74	-2.37±13.73	1.33±0.60 ⋆ •	5.48±1.71 •	~ 180	Yes
Proposed Method	9.95±3.53	-3.25±11.42	0.98±0.39	5.4±1.38	~ 0.3	Yes

4.4.4 Comparison of the CTV Labeling with Other Methods

The proposed multi-label segmentation approach is not directly a prostate segmentation solution. Since to the best of our knowledge, there are no prior work on automatic PTV estimation from TRUS, we compare the performance of the CTV estimations we obtain with two of the recently published works: 1) a semi-automatic approach proposed by Mahdavi et al. [1] which has been used at VCC for more than five years; and 2) our previous work on CTV segmentation using label fusion from multi-atlas approach [54].

Current planning procedure at VCC is based on a semi-automatic approach that requires manually locating some anatomical landmarks in the mid-gland [1]. Contours obtained from this state-of-the-art method are further modified by clinicians to provide CTV. Then a dilated PTV volume is generated and manually corrected based on clinician's expertise and knowledge to make sure clinical guidelines are met.

Table 4.2 shows statistics of the estimated CTV error from 1) the semi-automatic algorithm, 2) the automatic multi-atlas approach, and 3) the proposed multi-label algorithm as well as the execution time. We used the same dataset of 280 cases that previous works are evaluated on. We obtained 35 atoms and sparsity level of 10 by optimizing hyper-parameters of our model on this dataset, and evaluated its performance using 5-fold cross validation as explained in Sec. 4.4.

A statistical significance analysis has been performed on all reported error measures between the proposed multi-label approach evaluated on the dataset of 280 cases, and each of the two CTV segmentation methods using paired t-test and Levene's test. As for paired t-test, we define the null hypothesis as equivalence between mean (t-test) and variance (Levene's test) of observations from the same evaluation measure for the proposed method and

each of the other two methods. Results as indicated by "⋆" and "•" in Table 4.2 show statistically significant improvement for MSD when compared to the other two methods (p-value< 0.01 for both tests). Hausdorff distance is also significantly improved compared to the semi-automatic algorithm, by analyzing variance of the distributions using Levene's test. However, the semi-automatic algorithm outperforms in terms of V_{err} when compared to the proposed method. Although, according to the reported observer variability of CTV contouring, they all satisfy the clinical acceptance requirements. It is also observed that the proposed method is at least two orders of magnitude faster than other compared methods for CTV delineation. Figure 4.8 compares the proposed method with the other two CTV segmentation approaches [1, 54] in terms of percentage of cases estimated with different threshold values for V_{err}. The overall observer variability is shown by a vertical dashed line as the summation of reported inter- and intra-observer variabilities (see[95, 96]). The proposed multi-label method can generate a higher percentage of successfully labeled cases w.r.t. to the overall observer variability threshold when compared to the previous work [54], but still has lower rate when compared to the semi-automatic standard-of-the-care algorithm proposed by Mahdavi et al. [1].

4.4.5 Multi-label Model Comparison with Single-label Model

Joint dictionary atoms of the proposed method represent relation among three volumetric information spaces of the treatment record, i.e. TRUS, CTV and PTV. Learning process is constrained to produce equal weights for all three information spaces of the joint atoms. We further review significance of the relation between TRUS/CTV and TRUS/PTV by training models that generate single label from TRUS images to verify performance of the proposed multi-label segmentation algorithm. Three single-label joint dictionaries are trained for TRUS/CTV, TRUS/PTV and CTV/PTV to measure performance of the model in capturing the relation between two volumetric components. We obtained optimum hyper-parameters of each single-label model independently by performing exhaustive grid search. We used 55 and 60 atoms for TRUS/CTV and TRUS/PTV models respectively, while sparsity level remained in the same range of 10. Performance of each single-label model is compared with the proposed multi-label approach using 5-fold cross validation. Figure 4.9 shows the difference between two single-label models and the unified multi-label model in terms of CTV and PTV V_{err}. Both single-label models are trained using the proposed joint

sparse dictionary learning algorithm. Results show that joint patterns between TRUS and CTV are better captured when compared to joint patterns between TRUS and PTV. This is expected as there is more variability in determining PTV rather than CTV since PTV is determined based on CTV and the original TRUS image as well as some clinical guidelines and clinician's expertise. However, if we constrain the three volumetric information in a unified multi-label model, we achieve a higher accuracy for PTV at the price of slightly increased error for CTV. Gray area in Figure 4.9 right, represents the multi-label CTV and PTV estimation error per case. It is sorted to compare how the single-label model performs for the same case as the multi-label model. It is observed that CTV estimation is more persistent. It can be concluded that the dictionary training constrained by PTV has minimum impact on the estimation of CTV, while in contrast, PTV estimation benefits from being constrained by CTV components of the observation matrix. Our experiment on measuring performance of a two-part model to capture correlation between CTV and PTV shows $16.28 \pm 2.39\%$ V_{err} in estimation of PTV from CTV. This shows a better performance when compared to the single-label model (TRUS/PTV) and still significantly higher error (p-value< 0.01) when compared to the multi-label approach.

4.4.6 Execution Time

The proposed method includes two independent groups of computations, one to train the model and obtain the sparse dictionary atoms, and the other to estimate segmentation for an unseen TRUS volume. Assuming that the input TRUS volume is confined axially within base and apex of the prostate and is in adherence with the clinical protocols, the mathematical computation cost of the estimator is trivial. The proposed method has been implemented in MATLAB® environment (Mathworks, MA, USA) using C++ and MATLAB scripts. The computational burden of the proposed estimator is on solving the least squares problem for TRUS projection onto the joint atom space, which is facilitated by taking advantage of efficient implementation of the QR decomposition method in MATLAB environment. Therefore, we obtained very fast segmentation results from our CPU implementation on a standard PC (Intel Core i7, 2.93GHz, 8GB RAM) with execution time less than 300 ms including smoothing pre-processing and thresholding post-processing steps. This characteristic of the proposed method makes it an appropriate choice for some real-time segmentation applications such as intraoperative preplanning.

4.5 Discussion and Conclusion

In this work, a fast, automatic multi-label simultaneous segmentation method is introduced that is capable of incorporating previous treatment knowledge in the form of a joint linear model. We proposed a joint multiple volumetric information modeling approach that encodes relationships among multiple volumetric information spaces, i.e. TRUS, CTV and PTV, in a set of sparse dictionary atoms. The proposed method takes advantage of stationarity of the observations from the anatomy jointly and in multiple volumetric information spaces. The data driven approach looks for certain joint patterns constrained by the same projection coefficients in the form of a set of joint sources.

Given the TRUS volumes are constrained with the clinical protocols, i.e. Vancouver Cancer Centre guidelines, this approach can automatically and simultaneously provide CTV and PTV contours solely from an unseen TRUS volume. The proposed method is resilient to ultrasound artifacts as it incorporates voxel information w.r.t. all other voxels within the same observation case and amongst all other observation cases. This automatic approach can potentially be used in brachytherapy intraoperative preplanning, where the seed plans are constructed prior to the implantation procedure. This data driven multi-label segmentation approach provides both CTV and PTV contours simultaneously with maximum fidelity to the observed previous target anatomy labels, hence can produce contours that require minimum further manipulation in a time costly scenario such as intraoperative preplanning.

4.5.1 Clinical Significance

The observer variability involved in locating the base and the apex axial views of TRUS volumes as well as contouring CTV and PTV, reflects itself in the overall labeling accuracy. It is difficult to dissociate sources of variability from other factors such as inter- and intra-observer variability of labeling, since the true prostate shape, and the base and apex planes are unknown. In the lack of a true gold-standard, we argue that a clinically acceptable method should achieve accuracy within the overall observer variability determined from planning. Such analysis has been performed at VCC with several clinicians planning for several patients [1], and inter- and intra-observer variability of manual contouring of the CTV has been reported to be in the order of $4.65 \pm 0.77\%$ and $5.95 \pm 1.59\%$ for V_{err}, respectively [1]. Hence, the overall observer variability can be estimated as the summation of these two variabilities (see [95, 96]). We argue that this range of the overall observer

variability (approximately 10%) is clinically acceptable, since a statistical analysis of 10 years clinical outcome for VCC has shown excellent outcome and low rate of cancer recurrence [5]. A greater variability is expected for delineation of the PTV, since more user interaction is involved on top of CTV labeling to determine the PTV. This potentially significant variability is mainly due to the low signal-to-noise ratio of the target anatomy in the sparse TRUS images, especially close to the base and the apex.

4.5.2 Computational Cost

This learning-based segmentation approach provides a set of joint volumetric information dictionary elements and hence reduces the dimension of observation variations to the order of number of sparse coefficients associated with the atoms. Hence, the computational complexity of the multi-label segmentation is reduced dramatically to the order of a least squares problem solution. Hence, this fast simultaneous segmentation can potentially facilitate some clinical applications such as intraoperative preplanning and targeted biopsy.

4.5.3 Dataset Adaptability

An added advantage of the proposed data-driven multi-label segmentation approach is the adaptability of the model to different datasets. Our experiments on three physician-based sub-groups within the studied cohort of patients shows a high fidelity to the existing knowledge in the training dataset, as the trained model, i.e. joint sparse dictionary, performs superiorly on prediction of both segmentations for the same physician when compared to the other physicians. Although in our study all physicians use the same clinical protocols within VCC, our results show that the proposed method captures clinician's planning style in addition to the enforced guidelines in an institution. Hence, this method can be used in different institutions with maximum adaptability to their current segmentation routine and guidelines. However, preplanning protocols used in this study are an evolved and extended version of Seattle protocol [7] for prostate brachytherapy which is widely used in North America since 2009. Hence, extent of the volumetric information spaces as well as the orientation of the prostate in axial views are common beyond the collected dataset in this study. These guidelines are recommended to facilitate easy registration during patient setup; hence reduce deviation of the delivered treatments from preplans.

Our proposed joint intensity and labels approach has the following unique

key features: 1) In contrast to the most statistical models that incorporate shape priors, the proposed method is not iterative and is super fast (computation time in the order of milliseconds); 2) The information encoded in the dictionary is constrained by sparsity which provides a unique mapping that captures spatially distributed joint patterns instead of wholistic modes of variations. Since the variability of CTV and PTV are larger in the base and the apex in comparison to the mid-gland, we expect that spatially distributed joint patterns generated by our approach can capture such local variations with sufficient accuracy. 3) Prior intensity-shape-based models (such as [35, 36]), have only focused on prostate segmentation, hence their approach is not directly applicable to delineate PTV that is an anisotropic patient-specific dilation of the CTV.

4.5.4 Limitations

In the present work, we assume volumetric information spaces of all patients are pre-aligned according to the clinical protocols at VCC. Ideally, accurate alignment of the volumetric spaces can lead to a more accurate model (dictionary) or a lower number of joint components to represent the variations among treatment cases. In the context of our study, there are two main limitations that would affect the registration process: 1) Prostate shapes and sizes vary considerably; 2) TRUS image quality is low, especially in the base and the apex of the prostate. Hence, performing rigid registration as a pre-processing requirement would introduce a potential failure point in the overall automatic labeling workflow. Given that our results on data obtained at VCC are clinically acceptable without such pre-processing step [5], we conclude that enforcing a unified clinical workflow for data acquisition within a center can provide TRUS data that are aligned with accuracies sufficient for our method, considering all observer variabilities.

In summary, we proposed a multi-label segmentation approach evaluated on a retrospective dataset of treatment records at VCC. Results of the proposed models are compared against the actual CTV and PTV, and we have shown clinically promising results that can reduce the variability associated with segmentation of the two anatomical targets. Our previous CTV segmentation work was not directly applicable to PTV labeling since there is low correlation between TRUS intensity and PTV labels without incorporating CTV labels [54]. A comparison of our proposed approach to other CTV segmentation methods show significant improvement in clinical labeling error measures such as mean surface distance and V_{diff}, and speed-up of computation time in two orders of magnitude. This method can poten-

tially aid the clinicians by providing simultaneous estimation of CTV and PTV instantly, solely based on the collected TRUS images. The proposed system can be used for training, decision support or some real-time clinical applications such as intraoperative preplanning and targeted biopsy. It can be adapted to different institutions, physicians or generally the training dataset, and is potentially capable of predicting multiple contours that are biased towards the implicit knowledge and expertise hidden in a dataset which may not be easily translated into clinical guidelines.

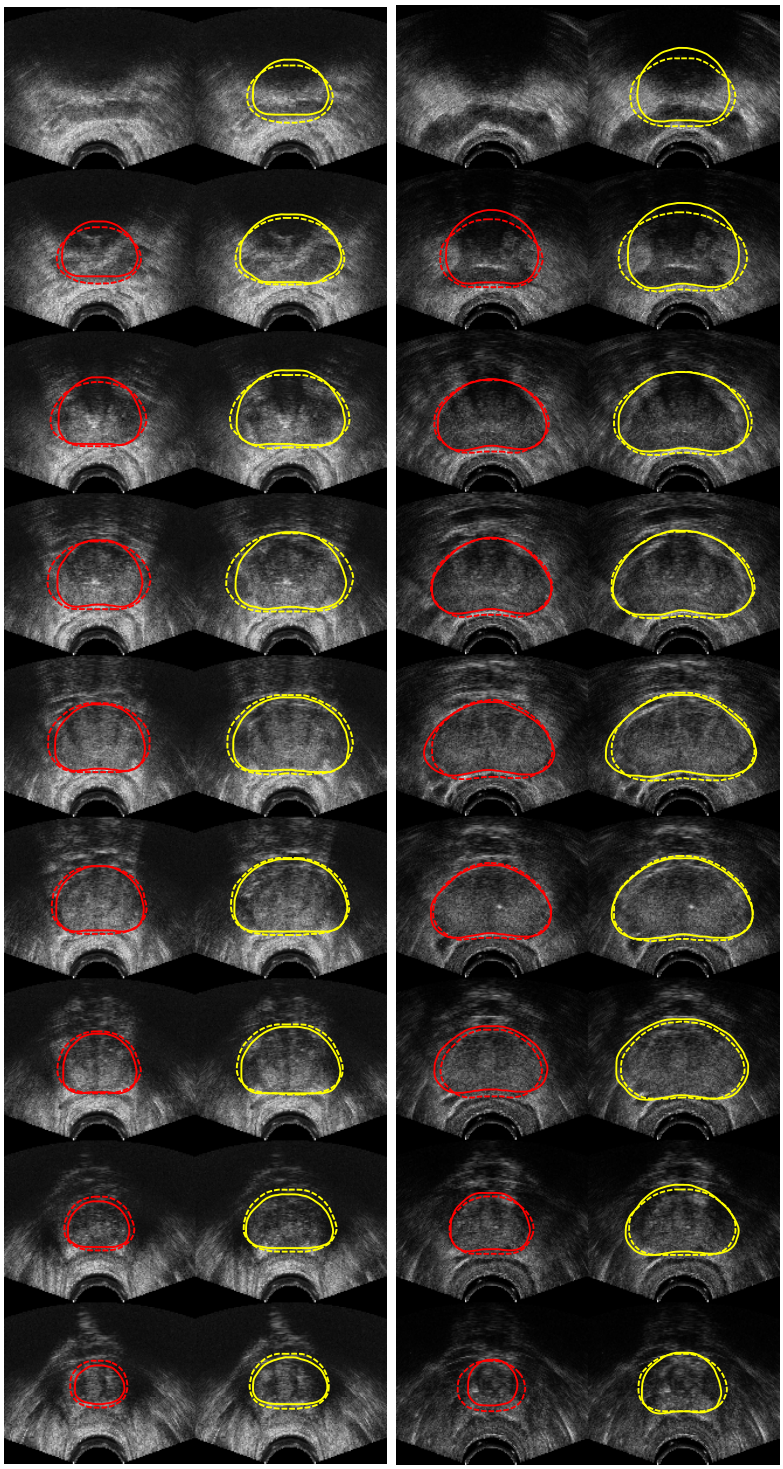
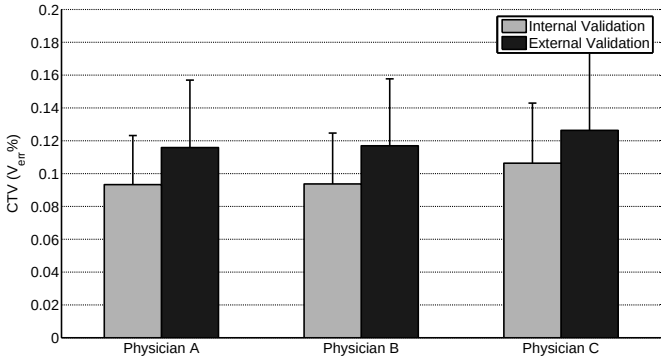
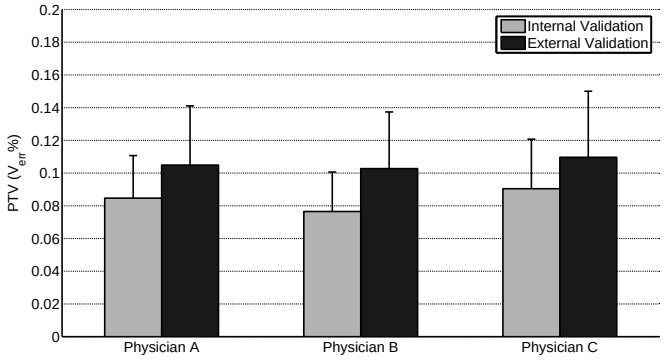


Figure 4.6: Two typical results of multi-label estimation from TRUS volume. Left column in each case shows the actual and estimated (dashed line) CTV and right column the actual and estimated (dashed line) PTV from base to apex (top to bottom).



(a)



(b)

Figure 4.7: Learning-based multi-label segmentation adaptability analysis. For three different groups a model is trained and evaluated internally and externally on the remainder of the cases.

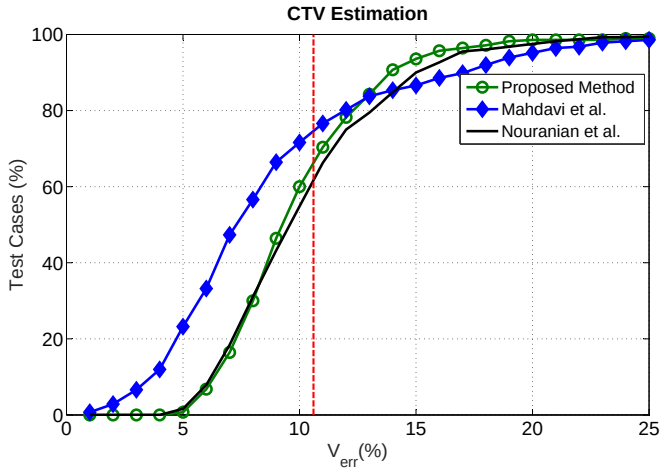


Figure 4.8: Cumulative histogram of the CTV estimation V_{err} comparing the proposed multi-label approach with the two CTV segmentation algorithms: 1) A semi-automatic standard of care algorithm [1]; and 2) an automatic multi-atlas label fusion algorithm [54]. Red vertical dashed line represents overall observer variability as a threshold to compare performance of the methods.

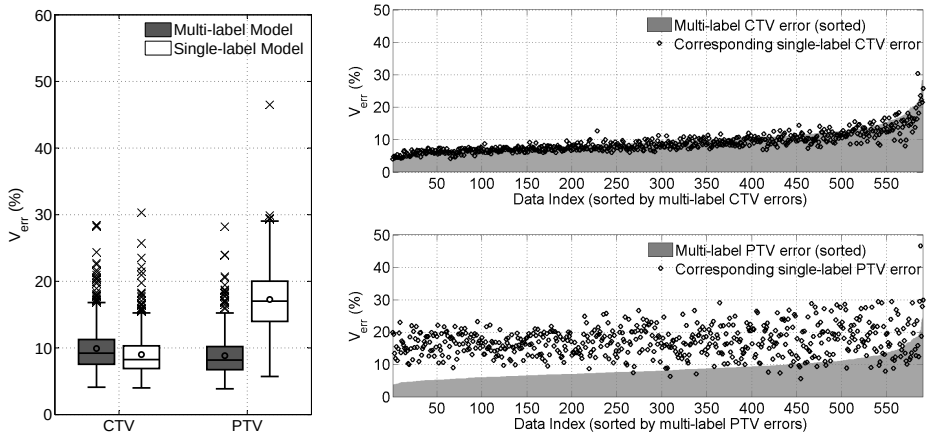


Figure 4.9: Comparison of single-label models generated for TRUS/CTV and TRUS/PTV estimation with the multi-label approach using the joint sparse dictionary learning algorithm. CTV single-label model performs better while PTV single-label model fails to capture TRUS/Label correlation. Label estimation error per case (right) shows that constraining three volumetric information all together in the joint dictionary learning model has minimum impact on CTV estimation (top) but a significant impact on improving PTV estimation when compared with single-label models.

Chapter 5

Automatic Seed Plan Estimation

5.1 Introduction

In LDR prostate brachytherapy, seed arrangements can be planned either prior to, right before or during implantation procedure. According to the BCCA prostate brachytherapy program, *preplanning* is a treatment approach, which requires pre-operative planning of seed arrangements, usually days before the implantation procedure. Seed planning takes place in an anisotropically dilated volume around the prostate boundary, known as Planning Target Volume (PTV). Standard-of-care for prostate brachytherapy preplanning and delivery is to use a standard brachytherapy grid template. This template is used for intra-operative delivery of the treatment, where needles are inserted according to the preplan, and confines needle placements to the grid points. Although clinical guidelines may vary somewhat between institutions, the preplanning treatment approach has shown excellent outcomes in the past decade [5].

Currently the standard-of-care in the majority of clinics is to create the plan manually, which is subject to a certain degree of observer variability. Treatment variability also stems from large set of possible solutions that all meet the clinical guidelines. However, a trained expert selects the recommended plan from this large set based on his/her expertise. Preplanning automation according to clinical guidelines and based on previous successful plans, can mitigate this problem.

In this work, we propose a new framework to automate the prostate brachytherapy preplanning. It consists of two components: 1) a plan estimator that represents sparse joint patterns between PTV and seed plans, and 2) a stochastic optimizer that encodes clinical standards to fine-tune

This chapter is adapted from [97]: S. Nouranian, M. Ramezani, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, Automatic Prostate Brachytherapy Preplanning Using Joint Sparse Analysis, in Medical Image Computing and Computer-Assisted Intervention: MICCAI, 2015, vol. 9350, pp. 415-423.

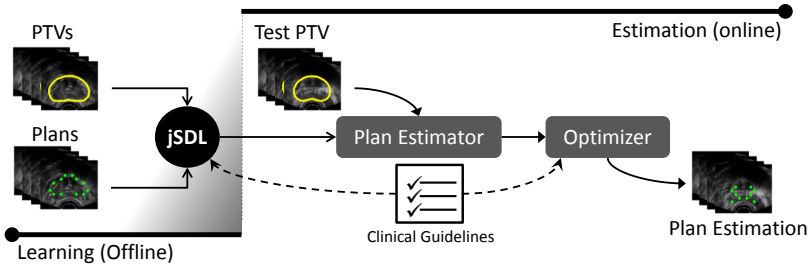


Figure 5.1: Block diagram of the proposed framework. Initial estimation of a plan is optimized, based on a novel cost function, to ensure that clinical guidelines are met.

the outcome of the plan estimator. We show that the estimator aids the optimizer to converge to a solution that meets the clinical standards. We build the plan estimator by employing a sparse dictionary learning algorithm which encodes the joint patterns between PTV and seed plans in a large retrospective dataset, by a set of sparse joint dictionary elements. The estimated plan is used as an initial seed configuration for the optimizer. The stochastic optimizer is designed based on simulated annealing, while enforcing clinical requirements through a novel cost function that captures those requirements. The proposed pipeline is evaluated on a dataset of 590 patients who underwent brachytherapy treatment at the VCC. We demonstrate an accuracy of 86%, defined on the basis of the clinical standards and previous treatment plans.

5.2 Methods

Figure 5.1 shows the proposed framework which consists of two parts: 1) a one-time offline process of generating a model, based on the joint information between PTV and seed configurations in retrospective patient data. We use a joint sparse dictionary learning approach for training the model; and, 2) an online process of plan estimation and optimization. The optimizer rearranges the seed configuration in the neighborhood of the initial plan estimation.

5.2.1 Joint Sparse Dictionary Learning

We represent all patients' volumetric information, consisting of the PTV and seed coordinates, in a unique discrete Cartesian space of $\Omega \subset \mathbb{R}^3$, and

a brachytherapy grid template space of $\Psi \subset \Omega$ with equal spatial spacing (e.g. 5 mm). We assume that this information is coregistered in the template space for further analysis. We define two information modalities as observation of two sets of labels assigned to each point in the discrete space of Ψ, i.e. labels w.r.t. the PTV volume and labels w.r.t. the seed distribution in Ψ. Specifically, for each patient $i \in \{1, 2, ..., N\}$ in a dataset of N patient treatment records, we use a unique spatial pattern in the space Ψ to generate $\mathbf{l}^{(i)} \in \mathbb{R}^{d\times 1}$ and $\mathbf{s}^{(i)} \in \{0, 1\}^{d\times 1}$, corresponding to PTV labels and seed placement labels in brachytherapy grid space, where d is the number of discrete coordinates in Ψ. We introduce joint observation vector $\mathbf{x}^{(i)} = [\mathbf{l}^{(i)^T}, \mathbf{s}^{(i)^T}]^T$ to train a dictionary of sparse joint patterns.

Our main objective is to introduce $\mathbf{D} \in \mathbb{R}^{2d\times K}$ as a dictionary of K atoms that satisfies $\mathbf{x}^{(i)} = \mathbf{D}\mathbf{c}^{(i)}$, where $\mathbf{c}^{(i)} \in \mathbb{R}^{K\times 1}$ is a vector of coefficients obtained for observation vector $i \in 1, 2, ..., N$.

We reformulate the dictionary training process by stacking the joint observation vectors $\mathbf{x}^{(i)}$ into a matrix of joint observations $\mathbf{X} \in \mathbb{R}^{2d\times N}$ expressing the joint dictionary learning as $\mathbf{X} = \mathbf{D}\mathbf{C}$, where $\mathbf{C} \in \mathbb{R}^{K\times N}$ refers to the coefficient matrix, and $\mathbf{D}$ contains atoms that jointly represent the observations in two parts, i.e. $\mathbf{D} = [\mathbf{D}^{l^T}, \mathbf{D}^{s^T}]^T$. Hence, atom coefficients are constrained to be equal for both modalities while satisfying the sparsity and efficient reconstruction conditions throughout the learning process. We use K-singular value decomposition (K-SVD) method [93] to train the model and generate $\mathbf{D}$. K-SVD is a generalization of the K-means clustering algorithm for vector quantization. This is an iterative approach that consists of two steps, 1) updating the atoms and their weights sequentially for the columns of $\mathbf{D}$ using singular value decomposition, and 2) to approximate sparsity using orthogonal mapping pursuit [94].

5.2.2 Seed Plan Estimation

We perform joint sparse dictionary learning on observation vectors which are defined on the brachytherapy grid space, i.e. $\Psi \subset \Omega$. Coordinates of the planning grid are obtained individually for each patient according to the clinical guidelines. Subsequently a randomly chosen case is used as reference to align all other planning grids. After transformation of PTVs and seed plans in space Ψ, we obtain aligned observation vectors of $\hat{\mathbf{l}}$ and $\hat{\mathbf{s}}$ for PTV and seed plan, respectively. We generate the observation matrix $\hat{\mathbf{x}}^{(i)} = [\hat{\mathbf{l}}^{(i)}, \hat{\mathbf{s}}^{(i)}]$ from the actual plan and the PTV to generate the joint sparse dictionary of $\mathbf{D}$, where $\hat{\mathbf{X}} = \mathbf{D}\mathbf{C}$ and $\mathbf{D} = [\mathbf{D}^{l^T}, \mathbf{D}^{s^T}]$. The dictionary $\mathbf{D}$ forms the model that captures the joint relationship between PTV labels

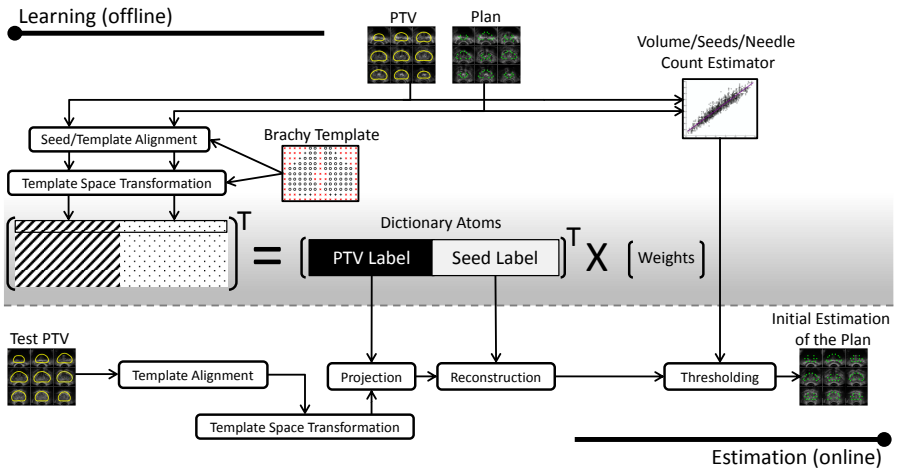


Figure 5.2: Initial plan estimation using the joint sparse dictionary learning algorithm.

and the corresponding seed arrangement.

An online *estimation* process includes an aligned PTV of a new case to form the observation vector $\mathbf{l}^t \in \{0,1\}^{d\times 1}$. Partial projection coefficients are obtained by solving a least squares problem $\tilde{\mathbf{c}}^t = \text{argmin}_{\mathbf{c}^t} \|\hat{\mathbf{l}}^{\mathbf{t}} - \mathbf{D}^{\mathbf{l}}\mathbf{c}^t\|_2^2$, where $\|.\|_2^2$ denotes the Euclidean norm. Subsequently, a probability map of seed arrangement is calculated by linear weighted combination of the joint atoms, $P(\mathbf{s}^t) = \mathbf{D}^s\mathbf{c}^t$.

We propose to use a dynamic threshold value to convert the reconstructed $P(\mathbf{s}^t)$ into an initial estimation of the seed arrangement, $\tilde{\mathbf{s}}^t$. We take advantage of the high correlation between number of seeds and the size of the PTV, and generate a polynomial regression model from the training dataset. This model is then used to estimate the number of required seeds (according to a confidence interval), for a test case based on its PTV size.

5.2.3 Plan Optimization

The estimated seed arrangement from learning-based algorithm is not constrained by planning quality measures, and only captures the arrangement patterns w.r.t. the PTV coordinates; hence does not necessarily meet the clinical guidelines. We propose an optimization process to rearrange estimated seed configurations towards maximizing adaptation to the clinical standards. Table 5.1 lists some of the clinical standards that we encode into

Table 5.1: Clinical guidelines at our local cancer centre:

Condition	Goal
Seed distribution w.r.t the grid	symmetric
Seed distribution w.r.t the grid	follows the mask
150% isodose	contiguous
Seed spacing	$> 7\ mm$
V_{100}	$> 98\%$
V_{150}	between 50% and 60%
V_{200}	$< 21\%$
Number of adjacent needles (row)	< 3
Number of adjacent needles (column)	< 4

the optimization algorithm. Given obtaining optimal seed arrangement is an inverse problem solution, we develop a simulated annealing (SA) algorithm that the large search space is confined in the neighborhood of the estimated seed arrangement $\tilde{\mathbf{s}}^t$.

We follow the formalism proposed by Task Group #43 [84] to calculate the dose distribution map in the discrete space of Ω. We propose an optimization algorithm based on SA architecture as a stochastic optimizer for a PTV volume, $\mathbf{l}_{ptv} \subset \Omega$ and the corresponding estimated seed arrangement $\mathbf{s} \subset \Psi$, which we hereafter refer to as a 2-tuple $\mathbf{y} =< \mathbf{l}_{ptv}, \mathbf{s} >$. We define plan quality indicators as:

$V_p(\mathbf{y})$: % of the PTV volume that is reached by at least $p\%$ or the mPD.

$D_p(\mathbf{y})$: The minimum dose that reaches $p\%$ of the PTV volume.

$C_{needle}(\mathbf{s})$: Total needles required to deliver the plan.

Architecture of the optimization algorithm is illustrated in Figure 5.3. We introduce a cost function that is optimized through an annealing process followed by a guideline-based correction algorithm. Plan is iteratively updated if acceptance condition is met within that iteration. A global parameter, aka temperature, asymptotically approaches zero as iterations progress. We use the temperature parameter in both annealing and correction processes. We define the objective function as

$$J(\mathbf{y}) = \alpha J_{V_{100}}(\mathbf{y}) + \beta J_{V_{150}}(\mathbf{y}) + \gamma J_{V_{200}}(\mathbf{y}) + \delta J_{needles}(\mathbf{y}) + \Gamma(\mathbf{s}), \qquad (5.1)$$

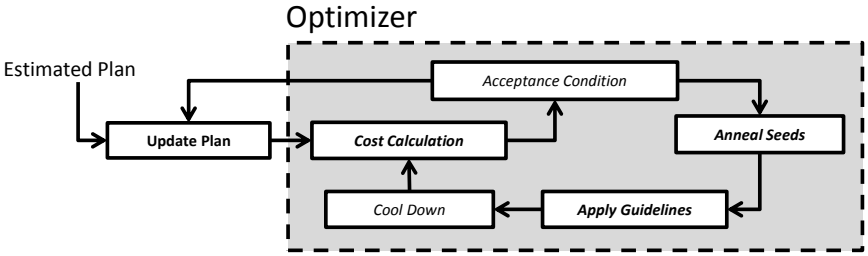


Figure 5.3: Architecture of proposed optimization algorithm based on simulated annealing.

where α, β, γ and δ are relative importance weights assigned to different terms of the objective function. Each term represents clinical quality indicators. Function $\Gamma(\mathbf{y})$ is defined to account for contiguity of the 150% isodose by counting the non-contiguous planes and penalizing the cost function. $J_{V_{100}}$, $J_{V_{150}}$ and $J_{V_{200}}$ are introduced to calculate deviation of the dose map from the boundary conditions:

$$J_{V_{100}}(\mathbf{y}) = 1 - V_{100}(\mathbf{y}), \tag{5.2}$$

$$J_{V_{150}}(\mathbf{y}) = u(V_{150}(\mathbf{y}) - 0.6)(V_{150}(\mathbf{y}) - 0.6) + u(0.5 - V_{150}(\mathbf{y}))(0.5 - V_{150}(\mathbf{y})), \tag{5.3}$$

$$J_{V_{200}}(\mathbf{y}) = u(V_{200}(\mathbf{y}) - 0.21)(V_{200}(\mathbf{y}) - 0.21), \tag{5.4}$$

$$J_{needles}(\mathbf{y}) = u(\mathcal{N}(\mathbf{s}) - R(\mathbf{y}))\frac{(\mathcal{N}(\mathbf{s}) - R(\mathbf{y}))}{R(\mathbf{y})} + \zeta(\mathbf{s}), \tag{5.5}$$

where $u(x)$ is a Heaviside step function that returns zero for $x < 0$ and one for $x >= 0$. $\mathcal{N}(\mathbf{s})$ represents number of needles required for the seed configuration $\mathbf{s}$, and $R(\mathbf{y})$ is an estimation of the total needles expected for the seed configuration in $\mathbf{y}$. We use a high order polynomial regression model to provide priors for needle count and PTV size relationship. $\zeta(\mathbf{s})$ penalizes the cost function by returning 1, if there are more than two adjacent needles in a row or three in a column.

The SA algorithm iteratively anneals the seed arrangement $\mathbf{s}$, as is explained in Algo. 1. Number of seeds k and range of movement r are linearly modified in each iteration based on the temperature cool down function. In

each iteration, seeds are ranked by assigning a movement probability value P_M, which is calculated by alternating between observations of two random variables: 1) ratio of locally non-overlapping volume between isodose volume and PTV, and 2) ratio of seed density over total number of seeds. Calculation of P_M is obtained in a neighborhood of size L which is also a linear function of the SA temperature. We propose the alternating approach for calculating P_M to push the annealing towards filling the uncovered areas, while making the seed arrangement more evenly distributed inside the PTV. Eventually, top k number of candidate seeds are chosen based on P_M.

Algorithm 1 Seeds annealing

```
1: procedure MARKING CANDIDATE SEEDS FOR ANNEALING(per iteration)
2:     for all seeds do
3:         movePrb ← calculate seed move probability P_M in neighborhood of size L
4:     candidateSeeds ← top k number of seeds with highest P_M
5:
6: procedure SEED ANNEALING
7:     for a randomly chosen candidate seed in candidateSeeds do
8:         availableSpots ← find all spots on Ψ in a neighborhood of r
9:         for all spots do
10:            if spot is a forbidden spot on the grid then
11:                availableSpots ← availableSpots - spot
12:
13:        if availableSpots is empty then
14:            continue the loop.
15:        else
16:            seed ← update with a randomly chosen spot in availableSpots
17:
18: call Seeds arrangement correction procedure Algo. 2
```

Annealing procedure is followed by a seed arrangement correction procedure that is explained in Algo. 2. Seed locations, rearranged in the annealing process, are marked in accordance to their adherence to the clinical guidelines, including seed spacing, brachytherapy grid mask and smaller number of needles. Those seeds that do not satisfy the clinical guidelines are moved in a neighborhood of size r until all seeds comply with the clinical guidelines. In our proposed SA-based optimization algorithm an exponentially dropping temperature function with re-annealing scheme is adapted, that gradually decreases the seed displacement range r as well as total number

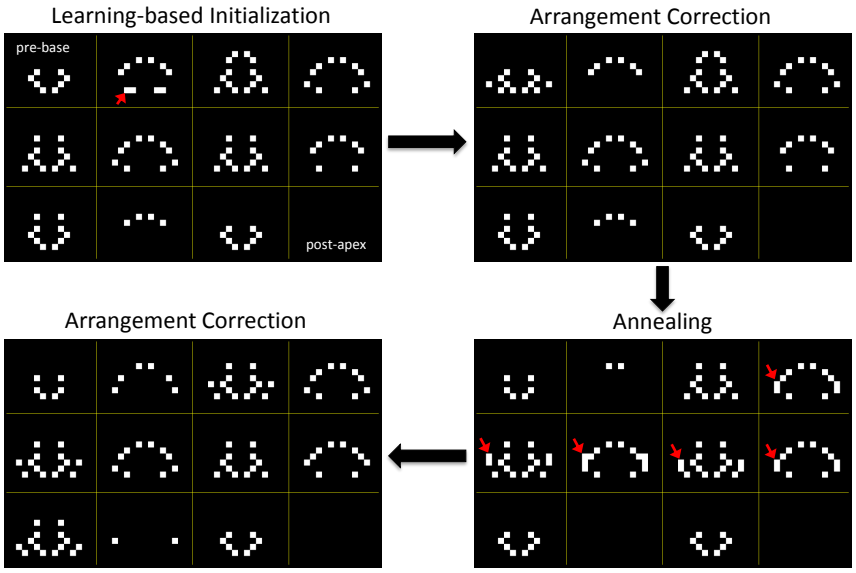


Figure 5.4: Seed arrangement correction and annealing process for one iteration of the optimization algorithm. Arrows point to a location where seed arrangement needs to be corrected based on clinical guidelines.

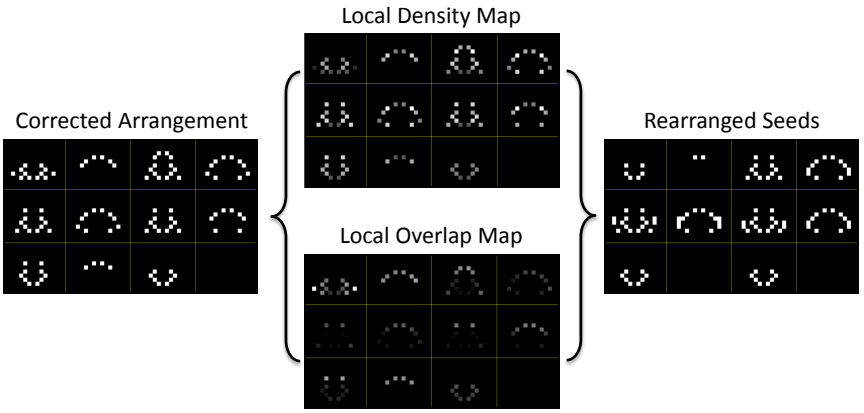


Figure 5.5: Seeds are rearranged (annealed) based on probability values calculated by alternating between local seed density maps and local overlap map between 100% isodose volume and PTV. Local neighborhood size is obtained from temperature parameter of the simulated annealing.

of candidate seeds to be moved k.

Figure 5.4 illustrates the annealing procedure in one iteration of the optimization algorithm. Initial seed configuration is corrected according to the clinical guidelines. Then seeds are moved in a neighborhood based on probability maps illustrated in Figure 5.5. Probability maps alternate between local seed density and local overlap between PTV and 100% isodose volume, and try to fill cold regions while avoid congestion of the seeds in one region.

Algorithm 2 Seeds arrangement correction

```
1: procedure MARKING SEEDS NEED CORRECTION
2:     for all seeds do
3:         if seed is on a forbidden spot on the grid then
4:             markedSeeds ← markedSeeds + seed
5:         if there is at least one < 7 mm distant seed then
6:             markedSeeds ← markedSeeds + seed
7:
8: procedure SEED CORRECTION
9:     r ← default neighborhood size
10:    for a randomly chosen marked seed in markedSeeds do
11:        search:
12:        availableSpots ← find all spots on Ψ in a neighborhood of r
13:        for all spots do
14:            if spot is a forbidden spot on the grid then
15:                availableSpots ← availableSpots - spot
16:            if min distance to the closest seed < 7 mm then
17:                availableSpots ← availableSpots - spot
18:            if no needle exist in the new spot then
19:                availableSpots ← availableSpots - spot
20:
21:        if availableSpots is empty then
22:            r ← r + 1.
23:            goto search.
24:        else
25:            seed ← update with a randomly chosen spot in availableSpots
```

5.3 Experiments and Results

A cohort of 590 patients who underwent prostate brachytherapy at the Vancouver Cancer Centre (VCC), is used to evaluate the proposed automatic preplanning framework. For each patient, PTV contours and actual seed arrangement used for the procedure are collected. All patients were treated with a plan based on the standard monotherapy dose prescription of 144 Gy mPD.

We performed our experiments by dividing the cohort of patients into 80% training and 20% test cases. A 5-fold cross validation approach was implemented to measure performance of the proposed framework. Objective function coefficients were heuristically chosen as $\alpha = 0.6$, $\beta = 0.7$, $\gamma = 1.0$ and $\delta = 0.1$. The rationale behind choosing importance weights is to treat V_{200} as the most significant term and reduce the coefficients for V_{150} and V_{100} relatively. This is due to the fact that V_{200} is closer to the geometrical seeds arrangement in contrast to V_{100} which represents overall dose coverage. We assign lower weight to the term representing the total number of needles, to push the optimizer towards seed reconfiguration rather than minimization of total needle count. Overall performance of the optimizer is dependent on proper weight assignment. Heuristically chosen values produce sub-optimal results.

The performance analysis consists of two parts: 1) plan estimation followed by optimization, and 2) optimization of randomly generated plans. We use the same parameters for the optimizer in the two experiments. Total number of required seeds, σ, is determined from the PTV size as indicated in Sec. 5.2.2.

We repeat our experiments for 95% confidence interval extremes, σ_- and σ_+, to determine the number of required seeds in both analyses; hence, we report performance of the plans for all three runs (i.e., with σ, σ_- and σ_+) for each test case. To tune the hyper-parameters of the proposed joint sparse dictionary model, we randomly partitioned the cohort into 80% training and 20% validation datasets. An exhaustive grid search on dictionary size and level of sparsity resulted in optimal number of 20 atoms with 5 sparse coefficients.

Figure 5.6 shows general statistics of the plan quality indicators V_{100}, V_{150} and V_{200} obtained from the optimization process using estimated plan and randomly generated plan. Hatched area shows the recommended zone for each parameter according to clinical guidelines. Results tabulated in Table 5.2 further show performance of the proposed optimization algorithm in terms of dosimetry parameters and total number of needles. It is seen that

the optimizer can further fine-tune the actual plans towards an improved set of quality indicators.

Table 5.2: Mean and standard deviation of the plan quality metrics calculated for actual, estimated and random plans with and without optimization.

	$V_{100}(\%)$	$V_{150}(\%)$	$V_{200}(\%)$	$D_{90}(Gy)$	#Needles
No Optimization:					
Actual	96.39±1.22	52.60±3.78	19.35±4.40	167.99±3.14	25±3
Proposed Estimator	91.71±5.56	54.99±7.57	25.48±5.85	156.27±14.89	26±4
Randomly Initialized	48.50±15.78	16.00±6.76	7.78±3.03	91.45±18.75	59±5
With Optimization:					
Actual	97.13±1.35	52.15±2.91	17.80±2.25	168.39±3.82	26±3
Proposed Estimator	97.12±1.95	52.32±3.48	18.03±2.54	167.79±6.62	28±4
Randomly Initialized	95.04±2.60	48.94±4.45	17.46±2.76	159.93±7.12	31±5

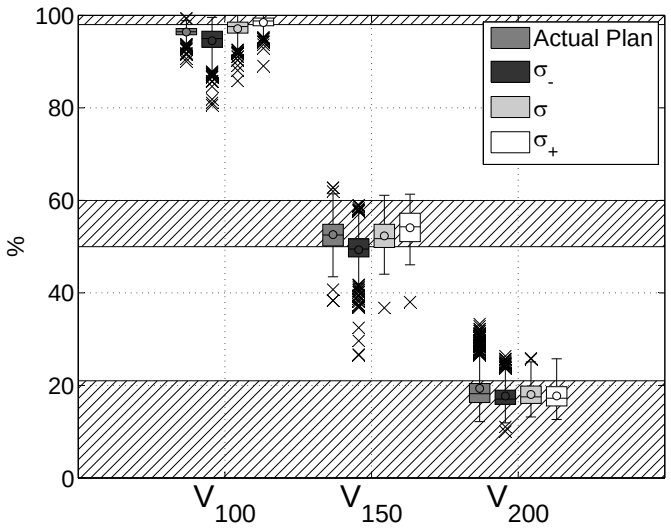
A successful plan needs to meet all or most clinical standards; therefore, we compare the percentage of successful preplans in terms of recommended clinical guidelines and actual delivered plans. Our observation shows 47% and 14% success rate for the optimizer initialized by estimated plan and random plan, respectively. These rates rise up to 86% and 65% when best of the optimization result is chosen for σ, σ_- and σ_+.

We conducted a paired t-test analysis with the null hypothesis indicating whether calculated dosimetry metrics of the plans obtained from different approaches, i.e. learning-based and randomly initialized seed plans, have the same mean values. We observed a statically significant difference between obtained quality indices with p-value< 0.001 indicating that the null hypothesis is rejected. Since mean values for all quality indices for randomly initialized seed plans were below the clinical guidelines, we obtained a significant improvement by initializing seed plans from the learning-based algorithm.

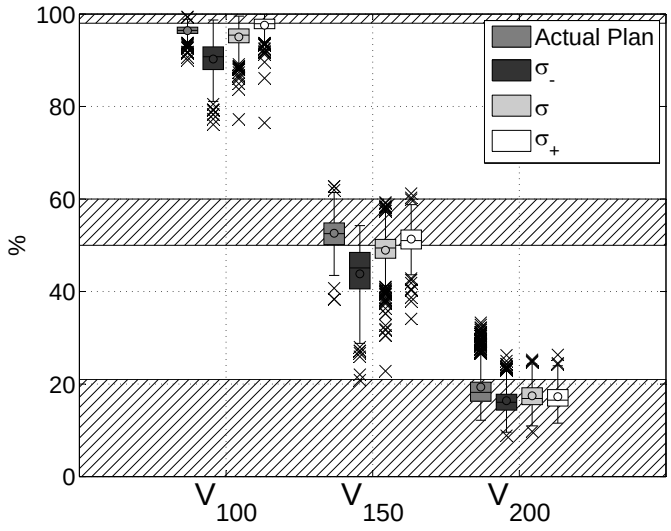
According to the results shown in Table 5.2 the proposed estimator produces plans with quality indicators within the same range of the actual delivered plans; as paired t-test analysis for mean values of all four dosimetry parameters between actual and proposed plans shows equivalence of the mean values with p-value< 0.001.

5.4 Discussion and Conclusion

In this work, an automatic prostate brachytherapy treatment planning framework is introduced that incorporates previous successful treatment knowledge in a model by capturing joint sparse patterns between PTV and seed



(a)



(b)

Figure 5.6: Comparison between statistics of plan quality metrics for a) estimated plan followed by optimization, and b) randomly generated plan followed by optimization. Number of seeds used in both methods are obtained from polynomial regression model.

configuration. Results of the model, i.e. plan estimator, is further optimized by a stochastic optimizer that uses simulated annealing architecture and a novel cost function to apply clinical guidelines and improve the plan quality indicators. Initialization of the optimizer from the estimated plan reduces the very large search space for all possible plans to a neighborhood of the initially estimated plan.

The proposed system is evaluated on a retrospective dataset of treatment records at the Vancouver Cancer Centre. We evaluate performance of the system by reporting the statistics of the plan quality indicators against the actual delivered plan, and show a very high success rate to converge to acceptable plans in terms of clinical guidelines. We show that the automatically generated plans require same range of needles to deliver the plan when compared with the actual plans used for treatment.

Each of the two main components of the proposed framework are attributed to some limitations. In this study, learning-based sparse dictionary generation is applied to learn geometrical distribution of the seeds irrespective to their activation and dose distribution since our training dataset allowed. It is more desired to introduce a mathematical model which is trained for dose distribution patterns, otherwise multiple models for different seed activations may be needed. The proposed optimization algorithm tunes the seed arrangements according to the strict clinical guidelines; however, introduces a large set of parameters to the proposed preplanning framework. Optimizer parameter tuning would benefit from sensitivity analysis.

Our proposed method can potentially aid the clinicians in training, decision support or suggesting plans which can be further modified. The proposed framework can be used as a tool to improve consistency across various centres and clinicians by automatically providing a reference plan that is in compliance with knowledge existed in one unified training dataset. The optimization algorithm can also be used to reduce the number of needles required to deliver the plan.

The framework is computationally efficient, currently running on a commodity personal computer with an unoptimized MATLAB® code within 2 minutes. Hence, it can be further optimized for intra-operative applications. The framework can also adapt to different institutions, requirements or datasets, and is capable of predicting a clinically acceptable plan that captures an implicit knowledge of a clinician or a group of clinicians, so that clinician-specific planning can be performed.

Chapter 6

Conclusion and Future Work

Current flow of the information for prostate brachytherapy preplanning procedure is prone to subjective errors, due to poor visibility of the target anatomy in TRUS images and high dependency to expertise of the physician. Accumulation of the errors in preplanning processes can potentially influence treatment outcome and rate of morbidity. In this thesis, in an effort to reduce the need for user interaction, we have introduced methods and algorithms to automate sequential processes involved in brachytherapy preplanning. Various information fusion algorithms have been introduced to represent joint relation between volumetric information elements of preplanning in the form of mathematical and statistical models. These models are trained based on previous treatment records obtained from the Vancouver Cancer Centre that have shown to result in an excellent treatment outcome.

In Chapter 2, we introduced a CTV delineation algorithm based on fusion of multiple atlases. Prior knowledge in delineation of CTV contours is incorporated in a set of atlases and is propagated through a non-rigid deformable registration algorithm to a new patient's TRUS volume. The proposed method generated clinically acceptable results when compared to the gold standards; however, it lacks the requirements for real-time applications such as intra-operative dosimetry and plan correction, because of the computation expenses. In search for a fast and reliable fusion algorithm, we formulated the segmentation problem with a linear joint model that encapsulates the joint patterns between TRUS intensity volumes and the corresponding contours in a set of joint components. In Chapter 3, we adapted the independent component analysis to estimate joint components and presented its application in estimation of CTV, PTV and mPD contours. In Chapter 4, we extended the proposed methodology by introducing sparsity to the modeling process. Further, we proposed a framework for simultaneous estimations of two planning labels, i.e. CTV and PTV. The proposed approach has been evaluated on a large dataset of 590 treatment case.

In Chapter 5, we proposed a fusion algorithm for estimation of the seed arrangement. We introduced a two-part framework that firstly generates

initial seed distribution utilizing the joint sparse model, and secondly rearranges seeds using a greedy optimization algorithm. Generated plans are compared against actual plans in terms of plan quality indicators. We conducted adaptability analysis to investigate a feature of the proposed methodology in capturing dataset-specific knowledge. We concluded that the proposed framework can generate better plans in terms of dosimetry parameters when trained and evaluated on a dataset of the same physician.

In conclusion, in this thesis, we aimed to automate various components of the preplanning procedure for prostate brachytherapy. We investigated various atlas-based and join statistical modeling approaches, while incorporating clinical guidelines. We achieved our goal by analyzing a large retrospective brachytherapy treatment planning dataset at the BC Cancer Agency. We demonstrated that real-time estimation of CTV and PTV is feasible. Further, we solved the inverse problem of seed placement by using a learning-based initial seed distribution to obtain a plan that closely complies with the clinical guidelines.

The contributions of this thesis are summarized as follows:

- A multi-atlas fusion framework is introduced for automatic delineation of the CTV from TRUS images. A dataset of a priori segmented ultrasound images, i.e. atlases, is registered to a target image. Corresponding labels of the atlases are fused based on pairwise shape similarity. Evaluation of the proposed segmentation approach on a set of transrectal prostate volume studies produces segmentation results that are within the range of observer variability when compared to a semi-automatic segmentation technique that is routinely used at the VCC.

- A fusion criteria is introduced to sort and rank the atlases in the multi-atlas-based segmentation approach. We propose pairwise atlas agreement factor that combines an image-similarity metric and similarity between a priori segmented contours. This factor is used in an atlas selection algorithm to prune the dataset before combining the atlas contours to produce a consensus segmentation.

- A fast and efficient fusion framework based on ICA is introduced for estimation of single or multiple planning contours, i.e. the CTV, the PTV and the mPD. Generally, this linear decomposition algorithm is trained on a set of complete joint observation vectors, and is tested on a single partial observation vector to estimate missing parts of the observation vector. We introduce a method to calculate a global threshold

value from training dataset that converts the estimated probability maps into final binary volumes.

- A sparsity-based fusion framework is introduced for simultaneous estimation of the CTV and the PTV. This framework is a generalization of the ICA-based modeling approach that is presented to automate delineation of CTV and PTV by finding sparse joint patterns extracted from previous treatment records.

- An optimization algorithm is proposed to produce seed arrangement w.r.t. the clinical guidelines at the VCC. This optimizer is designed on top of a simulated annealing architecture and encodes some key guidelines in both objective function and the annealing process.

- We propose a learning-based seed planning framework that combines the proposed optimizer with the sparsity-based fusion framework. The learning-based fusion framework provides seed configuration based on sparse joint patterns in a planning dataset. The initial seed arrangement is further perturbed according to the clinical guidelines using the in-house optimizer. This plan estimation framework has been evaluated on a large dataset of treatment records collected from the VCC and shows acceptable results based on clinically suggested plan quality indicators.

6.1 Future Work

Novel methods have been presented in this thesis for automation of the prostate brachytherapy planning procedure using information fusion and machine learning techniques. A number of interesting areas of research can be suggested as follows:

- The proposed CTV segmentation using multiple atlases is essentially built on top of an intensity-based registration process. Image registration is more challenging in the context of TRUS images mainly because of the poor visibility of the CTV near the base and the apex. Definition of the CTV is also a subjective concept which limits the proposed approach to correctly estimate the target contours for a new case. To alleviate this, a very large dataset is required that includes variety of CTV variations. hence we recommend:

  - Using a larger dataset of atlases. This increases likelihood of having more similar cases in the atlas dataset for a new patient TRUS volume.
  - To regularize intensity-based deformable registration algorithm based on the location proximity to the base or the apex.
  - In case of a very large dataset of atlases, the proposed framework would benefit from pre-registration grouping of atlases w.r.t. a similarity metric. This can be implemented either generally using pair- or group-wise clustering algorithms, or particularly by including new patient's TRUS volume in similarity calculation.
  - Our CPU-based non-optimized implementation of the framework produces segmentation results in few minutes. Since atlas registration processes are performed individually and independently, our implementation can potentially be improved by orders of magnitude if it utilizes graphical processing unit (GPU) for the registration algorithm. An improved implementation can potentially be used in real-time segmentation applications such as intraoperative dosimetry and prostate guided biopsy.

- Our proposed joint CTV and PTV estimation framework takes advantage of sparsity constraints in generating joint fusion-based models. The method is efficient, fast and reliable; however, accuracy of estimation can be improved by considering the following points:

  - A larger dataset and a more generic model that can be applied to different size and shape prostate cases. Bootstrapping and dataset pruning based on some clinical labels such as size, shape, biomarkers or image quality provides one or multiple joint models that represent clinically appealing variations of the patients.
  - We propose a method to convert estimated probability maps of the target volumes into final binary estimation of the target volumes. This approach uses a dataset (training) dependent global threshold value for all the query TRUS volumes. We recommend using a case-specific threshold value that can use priors from the training dataset to estimate optimum threshold value per each query case.
  - The sparsity-based fusion algorithm assumes all the volumetric information representation of the context are roughly pre-aligned. This is guaranteed in case of the VCC dataset, since clinical guidelines oblige the physician to maintain the prostate in the middle

of the TRUS images while moving the probe axially from the base to the apex. We suggest to add a pre-processing step that performs a rigid alignment of the training and the test cases to ensure this assumption is held true independent of the data acquisition protocol.

- The proposed target volume segmentation approaches have to be integrated into the current clinical workflow, e.g. at the VCC. It is recommended to provide an efficient implementation of the algorithms that can communicate with the current software in clinics with minimum change in the existing software infrastructure.

- A definition of confidence for segmentation is highly desirable. This is mostly beneficial for clinical integration, as it provides a confidence level for contours to help clinicians to perform corrections. Confidence can be defined from level of similarity to those cases included in the training set.

- Methods proposed and reviewed in this thesis are based on a linear modeling approach, while complexity of the context is captured via extracted joint patterns. Utilizing some non-linear methods may improve accuracy of the model in capturing the joint patterns in the fused information space.

- In the course of investigating fusion algorithms, we see ICA and sparse analysis as special cases of a general modeling approach. We recommend to review application of some unsupervised techniques that encode observation vectors into natural clusters such as restricted Bultzman machines (RBM) [98], auto-associative networks and random forest [99].

- Although adaptability can be one of the key features of the proposed methodology, a remaining question to be addressed as future work is to determine how applicable this methodology is to other datasets outside VCC.

  – The dataset used in this study is limited to a large cohort of patients imaged by the same brand ultrasound machine which has been widely used across BC. However, any other dataset that provides an extent of the anatomy which observations are reasonably standardized can be used for joint pattern analysis. Imaging

parameters and hardware related artifacts might appear as unwanted complex patterns that may influence performance of the segmentation. It is clear that removing those disruptive patterns by standardizing the imaging protocol or splitting the training dataset provides better resources for training models with more powerful joint patterns expressing the segmentation problem.

  - Joint patterns learned by the model represent the segmentation related information in the training dataset. Therefore, a new patient with incompatible imaging data (e.g. non-axial view, different spacing or unusual imaging parameters) can be detected from reconstruction error after projection onto the joint components. Eventually, an inclusion criteria can be defined to prune the outliers in test time.

- Plan estimation approach presented in this thesis shows feasibility of automatic seed planning using a two component automation framework: 1) A fusion-based learning component that estimated a seed configuration in accordance with the priors in the training dataset; and 2) an optimization algorithm that strictly enforces the clinical guidelines per case. Current approach limits the optimizer to search for a clinically acceptable seed arrangement in the neighborhood of the estimated seed configuration. This approach can be pushed more towards the priors by mixing the learning-based model and optimization algorithm. We recommend to guide the optimization process, per iteration, to follow patterns more compliant with the joint model. This can be implemented in the annealing process, where the seed arrangement is modified slightly towards the clinical guidelines. We suggest to introduce more optimization constraints to the objective function according to the VCC guidelines:

  - Total number of seeds outside PTV is desired to be minimum. Hence, priors of such factor can be used in the objective function to minimize this value. Our presented framework generates plans with greater number of seeds outside PTV for the larger size prostates, while this value is less for the smaller size prostates when compared to the actual plans. A new weighted term can be introduced to the objective function to apply this constraint.
  - The proposed approach generates plans with higher entropy in terms of seed arrangement when compared to the actual plans. This is observable when analyzing plans in the coronal views,

which show less regularity in the peripheral zone. For example, needles with single seeds are not common in actual plans; however, the optimizer is not constrained to avoid use of such needles. We suggest to use a grid template probability map that provides a guide for placement of regularly loaded strands or special loaded needles.

- It is not usually recommended to place three vertical needles in a column in anterior regions. Although limit of three vertical needles has been introduced to the optimizer, it is not constrained location-wise.
- One of the reasons actual plans express more regularity compared with the proposed method is the fact that actual preplanning process starts off by placing regular strands in desired grid locations. Hence, we suggest to generate a separate model for needle configuration based on PTV and utilize it in the annealing process or in initializing seed arrangement.

Bibliography

[1] S. S. Mahdavi, N. Chng, I. Spadinger, W. J. Morris, and S. E. Salcudean, "Semi-automatic segmentation for prostate interventions," *Med. Image Anal.*, vol. 15, no. 2, pp. 226–237, 2011.

[2] C. K. e. Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, Kosary CL, Yu M, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Lewis DR, Chen HS, Feuer EJ, "SEER Cancer Statistics Review, 1975-2010," Bethesda, MD, 2012.

[3] Advisory Committee on Cancer Statistics, "Canadian Cancer Statistics," 2013.

[4] K. A. Hinnen, J. J. Battermann, J. G. H. van Roermund, M. A. Moerland, I. M. Jürgenliemk-Schulz, S. J. Frank, and M. van Vulpen, "Long-term biochemical and survival outcome of 921 patients treated with I-125 permanent prostate brachytherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 76, no. 5, pp. 1433–8, Apr. 2010.

[5] W. J. Morris, M. Keyes, I. Spadinger, W. Kwan, M. Liu, M. McKenzie, H. Pai, T. Pickles, and S. Tyldesley, "Population-based 10-year oncologic outcomes after low-dose-rate brachytherapy for low-risk and intermediate-risk prostate cancer," *Cancer*, vol. 119, no. 8, pp. 1537–46, Apr. 2013.

[6] M. Keyes, W. J. Morris, I. Spadinger, C. Araujo, A. Cheung, N. Chng, J. Crook, R. Halperin, V. Lapointe, S. Miller, H. Pai, and T. Pickles, "Radiation oncology and medical physicists quality assurance in British Columbia Cancer Agency Provincial Prostate Brachytherapy Program," *Brachytherapy*, vol. 12, no. 4, pp. 343–55, 2013.

[7] J. E. Sylvester, P. D. Grimm, S. M. Eulau, R. K. Takamiya, and D. Naidoo, "Permanent prostate brachytherapy preplanned technique: The modern Seattle method step-by-step and dosimetric outcomes," *Brachytherapy*, vol. 8, no. 2, pp. 197–206, 2009.

[8] Y. Yu, L. L. Anderson, Z. Li, D. E. Mellenberg, R. Nath, M. C. Schell, F. M. Waterman, A. Wu, and J. C. Blasko, "Permanent prostate seed implant brachytherapy: report of the American Association of Physicists in Medicine Task Group No. 64," *Med. Phys.*, vol. 26, no. 10, pp. 2054–76, Oct. 1999.

[9] W. L. Smith, C. Lewis, G. Bauman, G. Rodrigues, D. D'Souza, R. Ash, D. Ho, V. Venkatesan, D. Downey, and A. Fenster, "Prostate volume contouring: a 3D analysis of segmentation using 3DTRUS, CT, and MR," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 67, no. 4, pp. 1238–47, Mar. 2007.

[10] S. Nag, W. Bice, K. DeWyngaert, B. Prestidge, R. Stock, and Y. Yu, "The American Brachytherapy Society recommendations for permanent prostate brachytherapy postimplant dosimetric analysis," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 46, no. 1, pp. 221–30, Jan. 2000.

[11] Z. Chen, N. Yue, X. Wang, K. B. Roberts, R. Peschel, and R. Nath, "Dosimetric effects of edema in permanent prostate seed implants: a rigorous solution," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 47, no. 5, pp. 1405–19, Jul. 2000.

[12] N. Yue, Z. Chen, R. Peschel, A. P. Dicker, F. M. Waterman, and R. Nath, "Optimum timing for image-based dose evaluation of 125I and 103PD prostate seed implants," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 45, no. 4, pp. 1063–72, Nov. 1999.

[13] M. Keyes, C. Macaulay, M. Hayes, J. Korbelik, W. J. Morris, and B. Palcic, "DNA ploidy measured on archived pretreatment biopsy material may correlate with prostate-specific antigen recurrence after prostate brachytherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 86, no. 5, pp. 829–34, Aug. 2013.

[14] W. D. Richard and C. G. Keen, "Automated texture-based segmentation of ultrasound images of the prostate," *Comput. Med. Imaging Graph.*, vol. 20, no. 3, pp. 131–40, 1996.

[15] S. D. Pathak, V. Chalana, D. R. Haynor, and Y. Kim, "Edge-guided boundary delineation in prostate ultrasound images," *IEEE Trans. Med. Imaging*, vol. 19, no. 12, pp. 1211–9, Dec. 2000.

[16] H. M. Ladak, F. Mao, Y. Wang, D. B. Downey, D. A. Steinman, and A. Fenster, "Prostate boundary segmentation from 2D ultrasound images," *Med. Phys.*, vol. 27, pp. 1777–88, 2000.

[17] P. Abolmaesumi and M. R. Sirouspour, "An interacting multiple model probabilistic data association filter for cavity boundary extraction from ultrasound images," *IEEE Trans. Med. Imaging*, vol. 23, no. 6, pp. 772–784, 2004.

[18] Y. Wang, H. N. Cardinal, D. B. Downey, and A. Fenster, "Semiautomatic three-dimensional segmentation of the prostate using two-dimensional ultrasound images," *Med. Phys.*, vol. 30, pp. 887–97, 2003.

[19] M. Ding, B. Chiu, I. Gyacskov, X. Yuan, M. Drangova, D. B. Downey, and A. Fenster, "Fast prostate segmentation in 3D TRUS images based on continuity constraint using an autoregressive model," *Med. Phys.*, vol. 34, pp. 4109–25, 2007.

[20] C. Knoll, M. Alcañiz, V. Grau, C. Monserrat, and M. C. Juan, "Outlining of the prostate using snakes with shape restrictions based on the wavelet transform," *Pattern Recognit.*, vol. 32, no. 10, pp. 1767–1781, 1999.

[21] W. Qiu, J. Yuan, E. Ukwatta, D. Tessier, and A. Fenster, "Rotational-slice-based prostate segmentation using level set with shape constraint for 3D end-firing TRUS guided biopsy," *Med. Image Comput. Comput. Interv. MICCAI*, vol. 15, no. Pt 1, pp. 537–44, Jan. 2012.

[22] A. C. Hodge, A. Fenster, D. B. Downey, and H. M. Ladak, "Prostate boundary segmentation from ultrasound images using 2D active shape models: Optimisation and extension to 3D," *Comput. Methods Programs Biomed.*, vol. 84, no. 2, pp. 99–113, 2006.

[23] L. Gong, S. D. Pathak, D. R. Haynor, P. S. Cho, and Y. Kim, "Parametric shape modeling using deformable superellipses for prostate segmentation," *IEEE Trans. Med. Imaging*, vol. 23, no. 3, pp. 340–349, 2004.

[24] S. S. Mahdavi, I. Spadinger, N. Chng, S. E. Salcudean, and W. J. Morris, "Semiautomatic segmentation for prostate brachytherapy: dosimetric evaluation," *Brachytherapy*, vol. 12, no. 1, pp. 65–76, Jan. 2013.

[25] S. S. Mahdavi, M. Moradi, W. J. Morris, S. L. Goldenberg, and S. E. Salcudean, "Fusion of ultrasound B-mode and vibro-elastography images for automatic 3D segmentation of the prostate," *IEEE Trans. Med. Imaging*, vol. 31, no. 11, pp. 2073–2082, Jul. 2012.

[26] S. Ghose, A. Oliver, J. Mitra, R. Martí, X. Lladó, J. Freixenet, D. Sidibé, J. C. Vilanova, J. Comet, and F. Meriaudeau, "A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images," *Med. Image Anal.*, vol. 17, no. 6, pp. 587–600, Aug. 2013.

[27] Y. Zhan and D. Shen, "Deformable segmentation of 3D ultrasound prostate images using statistical texture matching method," *IEEE Trans. Med. Imaging*, vol. 25, no. 3, pp. 256–272, 2006.

[28] P. Yan, S. Xu, B. Turkbey, and J. Kruecker, "Discrete Deformable Model Guided by Partial Active Shape Model for TRUS Image Segmentation," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1158–66, May 2010.

[29] W. Qiu, J. Yuan, E. Ukwatta, and A. Fenster, "Rotationally resliced 3D prostate TRUS segmentation using convex optimization with shape priors," *Med. Phys.*, vol. 42, no. 2, pp. 877–91, Feb. 2015.

[30] P. Wu, Y. Liu, Y. Li, and B. Liu, "Robust Prostate Segmentation Using Intrinsic Properties of TRUS Images," *IEEE Trans. Med. Imaging*, vol. 34, no. 6, pp. 1321–1335, Jan. 2015.

[31] Q. Xie and D. Ruan, "Low-complexity atlas-based prostate segmentation by combining global, regional, and local metrics," *Med. Phys.*, vol. 41, no. 4, p. 041909, Apr. 2014.

[32] H. Akbari, X. Yang, L. V. Halig, and B. Fei, "3D segmentation of prostate ultrasound images using wavelet transform," in *Proc. SPIE*, vol. 7962, 2011, p. 79622K.

[33] X. Yang and B. Fei, "3D prostate segmentation of ultrasound images combining longitudinal image registration and machine learning," in *Proc. SPIE*, vol. 8316, 2012, p. 83162O.

[34] T. Heimann, M. Baumhauer, T. Simpfendörfer, H. P. Meinzer, and I. Wolf, "Prostate segmentation from 3D transrectal ultrasound using statistical shape models and various appearance models," in *Proc. SPIE*, vol. 6914, 2008, p. 69141P.

[35] R. Toth, J. Ribault, J. Gentile, D. Sperling, and A. Madabhushi, "Simultaneous Segmentation of Prostatic Zones Using Active Appearance Models With Multiple Coupled Levelsets," *Comput. Vis. Image Underst.*, vol. 117, no. 9, pp. 1051–1060, Sep. 2013.

[36] J. Yang and J. S. Duncan, "3D image segmentation of deformable objects with joint shape-intensity prior models using level sets," *Med. Image Anal.*, vol. 8, no. 3, pp. 285–94, Sep. 2004.

[37] S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. W. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Med. Phys.*, vol. 35, p. 1407, 2008.

[38] J. A. Dowling, J. Fripp, S. Chandra, J. P. W. Pluim, J. Lambert, J. Parker, J. Denham, P. B. Greer, and O. Salvado, "Fast automatic multi-atlas segmentation of the prostate from 3D MR images," *Prostate Cancer Imaging Image Anal. Image-Guided Interv.*, pp. 10–21, 2011.

[39] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *Neuroimage*, vol. 21, no. 4, pp. 1428–1442, 2004.

[40] C. Lu, S. Chelikani, X. Papademetris, J. P. Knisely, M. F. Milosevic, Z. Chen, D. A. Jaffray, L. H. Staib, and J. S. Duncan, "An integrated approach to segmentation and nonrigid registration for application in image-guided pelvic radiotherapy," *Med. Image Anal.*, vol. 15, no. 5, p. 772, 2011.

[41] O. Acosta, A. Simon, F. Monge, F. Commandeur, C. Bassirou, G. Cazoulat, R. de Crevoisier, and P. Haigron, "Evaluation of multi-atlas-based segmentation of CT scans in prostate cancer radiotherapy," in *Biomed. Imaging (ISBI), IEEE Int. Symp.* IEEE, 2011, pp. 1966–1969.

[42] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solorzano, "Combination strategies in multi-atlas image segmentation: application to brain MR data," *IEEE Trans. Med. Imaging*, vol. 28, no. 8, pp. 1266–77, Aug. 2009.

[43] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label

propagation and decision fusion," *Neuroimage*, vol. 33, no. 1, pp. 115–26, Oct. 2006.

[44] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy," *Neuroimage*, vol. 46, no. 3, pp. 726–38, Jul. 2009.

[45] A. Redpath, "Automatic determination of needle and source positions for brachytherapy of the prostate using 125Iodine Rapid Strand," *Radiother. Oncol.*, vol. 64, no. 2, pp. 215–27, Aug. 2002.

[46] G. Ferrari, Y. Kazareski, F. Laca, and C. E. Testuri, "A model for prostate brachytherapy planning with sources and needles position optimization," *Oper. Res. Heal. Care*, vol. 3, no. 1, pp. 31–39, Mar. 2014.

[47] J. Pouliot, D. Tremblay, J. Roy, and S. Filice, "Optimization of permanent 125I prostate implants using fast simulated annealing," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 36, no. 3, pp. 711–20, Oct. 1996.

[48] R. R. Meyer, W. D. D'Souza, M. C. Ferris, and B. R. Thomadsen, "MIP Models and BB Strategies in Brachytherapy Treatment Optimization," *J. Glob. Optim.*, vol. 25, no. 1, pp. 23–42, Jan. 2003.

[49] E. Lessard, S. L. S. Kwa, B. Pickett, M. Roach, and J. Pouliot, "Class solution for inversely planned permanent prostate implants to mimic an experienced dosimetrist," *Med. Phys.*, vol. 33, no. 8, p. 2773, Jul. 2006.

[50] Y. Yu, J. B. Zhang, R. A. Brasacchio, P. G. Okunieff, D. J. Rubens, J. G. Strang, A. Soni, and E. M. Messing, "Automated treatment planning engine for prostate seed implant brachytherapy," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 43, no. 3, pp. 647–52, Feb. 1999.

[51] E. K. Lee, R. J. Gallagher, D. Silvern, C. S. Wuu, and M. Zaider, "Treatment planning for brachytherapy: an integer programming model, two computational approaches and experiments with permanent prostate implant planning," *Phys. Med. Biol.*, vol. 44, no. 1, pp. 145–65, Jan. 1999.

[52] G. Yang, L. E. Reinstein, S. Pai, Z. Xu, and D. L. Carroll, "A new genetic algorithm technique in optimization of permanent 125I prostate implants," *Med. Phys.*, vol. 25, no. 12, pp. 2308–15, Dec. 1998.

[53] R. J. Gallagher and E. K. Lee, "Mixed integer programming optimization models for brachytherapy treatment planning," *Proc. AMIA Annu. Fall Symp.*, pp. 278–82, Jan. 1997.

[54] S. Nouranian, S. S. Mahdavi, I. Spadinger, W. Morris, S. Salcudean, and P. Abolmaesumi, "A Multi-Atlas-Based Segmentation Framework for Prostate Brachytherapy," *IEEE Trans. Med. Imaging*, vol. 34, no. 4, pp. 950–961, Dec. 2015.

[55] S. Liao, Y. Gao, J. Lian, and D. Shen, "Sparse patch-based label propagation for accurate prostate localization in CT images," *IEEE Trans. Med. Imaging*, vol. 32, no. 2, pp. 419–34, Feb. 2013.

[56] M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *IEEE Trans. Med. Imaging*, vol. 29, no. 10, pp. 1714–29, Oct. 2010.

[57] O. Commowick, A. Akhondi-Asl, and S. K. Warfield, "Estimating a reference standard segmentation with spatially varying performance parameters: local MAP STAPLE," *IEEE Trans. Med. Imaging*, vol. 31, no. 8, pp. 1593–606, Aug. 2012.

[58] A. Akhondi-Asl, L. Hoyte, M. Lockhart, and S. Warfield, "A Logarithmic Opinion Pool Based STAPLE Algorithm for the Fusion of Segmentations With Associated Reliability Weights," *IEEE Trans. Med. Imaging*, Jun. 2014.

[59] T. R. Langerak, F. F. Berendsen, U. A. Van der Heide, A. N. T. J. Kotte, and J. P. W. Pluim, "Multiatlas-based segmentation with preregistration atlas selection," *Med. Phys.*, vol. 40, no. 9, p. 091701, Sep. 2013.

[60] S. Nouranian, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "An Automatic Multi-atlas Segmentation of the Prostate in Transrectal Ultrasound Images Using Pairwise Atlas Shape Similarity," in *Med. Image Comput. Comput. Interv. MICCAI*, 2013, vol. 8150, pp. 173–180.

[61] ——, "Multi-atlas-based automatic 3D segmentation for prostate brachytherapy in transrectal ultrasound images," *Proc. SPIE Med. Imaging*, vol. 8671, pp. 86 710O–86 710O–7, 2013.

[62] J. P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.

[63] X. Pennec, P. Cachier, and N. Ayache, "Understanding the demons algorithm: 3D non-rigid registration by gradient descent," in *Med. Image Comput. Comput. Interv. MICCAI.* Springer, 1999, pp. 597–605.

[64] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Non-parametric Diffeomorphic Image Registration with Demons Algorithm," in *Med. Image Comput. Comput. Interv.*, ser. Lecture Notes in Computer Science, N. Ayache, S. Ourselin, and A. J. Maeder, Eds., vol. 10, no. Pt 2, Asclepios Research Group, INRIA Sophia-Antipolis, France. Springer, 2007, pp. 319–326.

[65] S. Khallaghi, C. G. M. Leung, K. Hastrudi-Zaad, P. Foroughi, C. Nguan, and P. Abolmaesumi, "Experimental validation of an intra-subject elastic registration algorithm for dynamic-3D ultrasound images," *Med. Phys.*, vol. 39, no. 9, pp. 5488–97, Sep. 2012.

[66] T. De Silva, A. Fenster, J. Bax, C. Romagnoli, J. Izawa, J. Samarabandu, and A. D. Ward, "Quantification of prostate deformation due to needle insertion during TRUS-guided biopsy: comparison of hand-held and mechanically stabilized systems," *Med. Phys.*, vol. 38, no. 3, pp. 1718–31, Mar. 2011.

[67] I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. van Ginneken, "Multi-atlas-based segmentation with local decision fusion–application to cardiac and aortic segmentation in CT scans," *IEEE Trans. Med. Imaging*, vol. 28, no. 7, pp. 1000–10, Jul. 2009.

[68] T. R. Langerak, U. A. van der Heide, A. N. T. J. Kotte, M. A. Viergever, M. van Vulpen, and J. P. W. Pluim, "Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE)," *IEEE Trans. Med. Imaging*, vol. 29, no. 12, pp. 2000–8, Dec. 2010.

[69] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans. Med. Imaging*, vol. 23, no. 7, pp. 903–921, 2004.

[70] T. Rohlfing, D. B. Russakoff, and C. R. Maurer, "Performance-based classifier combination in atlas-based image segmentation using

expectation-maximization parameter estimation," *IEEE Trans. Med. Imaging*, vol. 23, no. 8, pp. 983–94, Aug. 2004.

[71] A. J. Asman and B. A. Landman, "Formulating spatially varying performance in the statistical fusion framework," *IEEE Trans. Med. Imaging*, vol. 31, no. 6, pp. 1326–36, Jun. 2012.

[72] J. Nocedal, "Updating Quasi-Newton Matrices with Limited Storage," *Math. Comput.*, vol. 35, no. 151, pp. 773–782, 1980.

[73] T. Rohlfing, "Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable," *IEEE Trans. Med. Imaging*, vol. 31, no. 2, pp. 153–63, Feb. 2012.

[74] A. J. Asman, F. W. Bryan, S. A. Smith, D. S. Reich, and B. A. Landman, "Groupwise multi-atlas segmentation of the spinal cord's internal structure," *Med. Image Anal.*, vol. 18, no. 3, pp. 460–71, Apr. 2014.

[75] I. B. Tutar, S. D. Pathak, L. Gong, P. S. Cho, K. Wallner, and Y. Kim, "Semiautomatic 3D prostate segmentation from TRUS images using spherical harmonics," *IEEE Trans. Med. Imaging*, vol. 25, no. 12, pp. 1645–1654, 2006.

[76] M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin, "Fast free-form deformation using graphics processing units," *Comput. Methods Programs Biomed.*, vol. 98, no. 3, pp. 278–284, 2010.

[77] S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "Fast Prostate Segmentation for Brachytherapy based on Joint Fusion of Images and Labels," *Proc. SPIE Med. Imaging*, vol. 9036, pp. 90 361A–90 361A–7, 2014.

[78] V. D. Calhoun, T. Adali, and J. Liu, "A feature-based approach to combine functional MRI, structural MRI and EEG brain imaging data," in *Eng. Med. Biol. Soc. 28th Annu. Int. Conf. IEEE*, vol. 1, Jan. 2006, pp. 3672–5.

[79] V. D. Calhoun, T. Adali, N. R. Giuliani, J. J. Pekar, K. A. Kiehl, and G. D. Pearlson, "Method for multimodal analysis of independent source differences in schizophrenia: combining gray matter structural and auditory oddball functional data," *Hum. Brain Mapp.*, vol. 27, no. 1, pp. 47–62, Jan. 2006.

[80] K. Specht, R. Zahn, K. Willmes, S. Weis, C. Holtel, B. J. Krause, H. Herzog, and W. Huber, "Joint independent component analysis of structural and functional images reveals complex patterns of functional reorganisation in stroke aphasia," *Neuroimage*, vol. 47, no. 4, pp. 2057–63, Oct. 2009.

[81] J. Liu, G. Pearlson, A. Windemuth, G. Ruano, N. I. Perrone-Bizzozero, and V. D. Calhoun, "Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA," *Hum. Brain Mapp.*, vol. 30, no. 1, pp. 241–55, Jan. 2009.

[82] S. Nouranian, M. Ramezani, S. S. Mahdavi, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "Data fusion for planning target volume and isodose prediction in prostate brachytherapy," pp. 94 151I–94 151I–7, 2015.

[83] W. J. Morris, I. Spadinger, M. Keyes, J. Hamm, M. McKenzie, and T. Pickles, "Whole prostate D90 and V100: a dose-response analysis of 2000 consecutive (125)I monotherapy patients," *Brachytherapy*, vol. 13, no. 1, pp. 32–41, Jan. 2014.

[84] M. J. Rivard, B. M. Coursey, L. A. DeWerd, W. F. Hanson, M. S. Huq, G. S. Ibbott, M. G. Mitch, R. Nath, and J. F. Williamson, "Update of AAPM Task Group No. 43 Report: A revised AAPM protocol for brachytherapy dose calculations," *Med. Phys.*, vol. 31, no. 3, pp. 633–74, Mar. 2004.

[85] T. W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural Comput.*, vol. 11, no. 2, pp. 417–41, Feb. 1999.

[86] J. Himberg, A. Hyvärinen, and F. Esposito, "Validating the independent components of neuroimaging time series via clustering and visualization," *Neuroimage*, vol. 22, no. 3, pp. 1214–22, Jul. 2004.

[87] N. N. Stone, L. Potters, B. J. Davis, J. P. Ciezki, M. J. Zelefsky, M. Roach, P. A. Fearn, M. W. Kattan, and R. G. Stock, "Customized dose prescription for permanent prostate brachytherapy: insights from a multicenter analysis of dosimetry outcomes," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 69, no. 5, pp. 1472–7, Dec. 2007.

[88] M. Keyes, D. Schellenberg, V. Moravan, M. McKenzie, A. Agranovich, T. Pickles, J. Wu, M. Liu, J. Bucci, and W. J. Morris, "Decline in urinary retention incidence in 805 patients after prostate brachytherapy: the effect of learning curve?" *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 64, no. 3, pp. 825–34, Mar. 2006.

[89] S. Nouranian, M. Ramezani, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "Learning-based Multi-label Segmentation of Transrectal Ultrasound Images for Prostate Brachytherapy," *IEEE Trans. Med. Imaging*, vol. 35, no. 3, pp. 921–932, March 2016.

[90] I. Tosic and P. Frossard, "Dictionary Learning," *IEEE Signal Process. Mag.*, vol. 28, no. 2, pp. 27–38, Mar. 2011.

[91] G. Sanroma, G. Wu, K. Thung, Y. Guo, and D. Shen, "Novel Multi-Atlas Segmentation by Matrix Completion," in *Int. Work. Mach. Learn. Med. Imaging*, Sep. 2014.

[92] M. Ramezani, K. Marble, H. Trang, I. S. Johnsrude, and P. Abolmaesumi, "Joint sparse representation of brain activity patterns in multi-task fMRI data," *IEEE Trans. Med. Imaging*, vol. 34, no. 1, pp. 2–12, Jan. 2015.

[93] M. Aharon, M. Elad, and A. Bruckstein, "K -SVD : An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[94] J. Tropp, "Greed is Good: Algorithmic Results for Sparse Approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[95] F. Sardanelli and G. D. Leo, *Biostatistics for Radiologists: Planning, Performing, and Writing a Radiologic Study.* Springer Science & Business Media, 2009.

[96] M. Haber, H. X. Barnhart, J. Song, and J. Gruden, "Observer variability: a new approach in evaluating interobserver agreement," *J. Data Sci.*, vol. 3, pp. 69–83, Jan. 2005.

[97] S. Nouranian, M. Ramezani, I. Spadinger, W. J. Morris, S. E. Salcudean, and P. Abolmaesumi, "Automatic Prostate Brachytherapy Preplanning Using Joint Sparse Analysis," in *Med. Image Comput. Comput. Interv. MICCAI*, vol. 9350. Springer International Publishing, 2015, pp. 415–423.

[98] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets." *Neural Comput.*, vol. 18, no. 7, pp. 1527–54, Jul. 2006.

[99] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.