Visual analytics for digital radiotherapy : towards a comprehensible pipeline

Document status and date:
Published: 23/03/2017

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

Visual Analytics for Digital Radiotherapy

Towards a Comprehensible Pipeline

Renata Georgia Raidou

Visual Analytics for Digital Radiotherapy

Towards a Comprehensible Pipeline

PROEFSCHRIFT

door

Renata Georgia Raidou

geboren te Vólos, Griekenland

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

A nonno Nicola

Contents

Contents

Summary

Visual Analytics for Digital Radiotherapy:
Towards a Comprehensible Pipeline

Prostate cancer is one of the most frequently occurring cancer types in males. It is often treated with radiation therapy, which aims at irradiating tumors with a high dose, while sparing the surrounding healthy tissues. In the course of the years, radiotherapy technology has undergone great advancements, but there is still room for further improvement.

Clinical researchers are now aware that tumors are not only different from each other. They are also highly heterogeneous within, consisting of regions with distinct tissue characteristics. According to these characteristics, clinical researchers should choose adequate radiation doses for each intra-tumor region. Tailoring radiotherapy planning to the specific needs and intra-tumor tissue characteristics of each patient is expected to lead to more effective treatment strategies. Currently, clinical research is moving towards this direction, but a deep understanding of the specific tumor characteristics of each patient, and the integration of all available knowledge into a new radiotherapy planning pipeline are required.

The patient- and tumor-specific radiotherapy planning pipeline involves complex multimodal and multi-valued data. Understanding and analyzing this data can be a demanding and time-consuming task, even for experienced clinical researchers. Additionally, all implicated data include sources of uncertainty, which can affect the accuracy and/or precision of the final planning outcome. Some of these uncertainties can be minimized. The rest, which cannot be avoided, need to be studied and their effect on radiation therapy planning needs to be predicted. As a consequence, there is an emerging need for solutions and tools, which can help clinical researchers explore, understand and analyze all the available patient- and tumor-specific information.

The goal of this dissertation is to investigate and provide solutions and tools, which empower clinical researchers to gain insight into the complex data and processes present at each step of the radiotherapy pipeline. This kind of solutions and tools can be sought within the domain of Visual Analytics, the scientific field that connects visualization techniques with other disciplines, such as data mining or statistics, in highly interactive environments. To this end, we designed comprehensible visualization strategies for the interactive exploration and analysis of the radiotherapy planning data and processes, enabling the integration of the knowledge and cognitive skills of the intended clinical users.

More in particular, we introduce a novel technique, the *Orientation-Enhanced Parallel Coordinates Plots*, for the representation and exploration of multi-variate, multi-dimensional data (*Chapter 4*). With our technique, which is an enhancement of Parallel Coordinate Plots, we improve the display of the data, addressing the challenging topic of reducing clutter due to overplotting of polylines.

Additionally, we propose a number of visualization designs, more specific to the radiotherapy planning pipeline. Initially, we provide a visualization design for the exploration and visual analysis of potential *variability in imaging-derived features*, caused by differences in the methods employed for their computation (*Chapter 5*). With this, we facilitate the – currently tedious – exploration of variations of the feature space, with respect to patient anatomy.

Furthermore, with a new visualization framework, we support the easy exploration and analysis of the feature space of imaging-derived *intra-tumor tissue characteristics* (*Chapter 6*). This component enables the identification of distinct intra-tumor regions, the exploration of tumor heterogeneity, new knowledge discovery within the feature space of tumor tissue characteristics, as well as hypothesis generation and confirmation, with respect to clinical reference data, such as histopathological data. We demonstrate the strengths of the proposed framework with an additional practical application, where we employ the proposed system to aid the design and to understand the behavior of *tissue characterization classifiers*.

Moreover, we propose a Visual Analytics tool for the exploration and assessment of errors occurring during the *segmentation* of the involved organs from the medical images (*Chapter 7*). Our proposed approach supplies new insight in the performance of the employed segmentation algorithms.

In addition, we demonstrate the use of a new visualization design to address the exploration and analysis of *variability in an ensemble of radiotherapy dose plans*, generated by adjustments in the previous steps of the pipeline (*Chapter 8*). With this work, we aim at providing new understanding into the impact of these adjustments on the final outcome.

Finally, we present a new Visual Analytics system for the exploration and analysis of *Tumor Control Probability modeling* to evaluate and predict the outcome of a radiotherapy treatment (*Chapter 9*). With our introduced approach, the – up to now disregarded – imaging-induced uncertainty and sensitivity analysis of the employed radiobiological parameters can be incorporated in the workflow of clinical researchers, providing new possibilities for the evaluation of the selected radiotherapy strategies.

All in all, this dissertation describes solutions from the field of Visual Analytics, aiming at incorporating information from the distinct steps of the radiotherapy planning pipeline, along with potential sources of uncertainty, into comprehensible visualizations. These approaches contribute towards the interactive exploration and visual analysis of the involved data and processes at each step of the radiotherapy planning pipeline, creating a fertile ground for future radiotherapy planning and visualization research.

Samenvatting

Visual Analytics voor Digitale Radiotherapie: Naar een Inzichtelijke Werkwijze

Prostaatkanker is een veelvoorkomende vorm van kanker, die vaak wordt behandeld door middel van radiotherapie. Hierbij wordt getracht de tumor met een hoge dosis ioniserende straling te bestralen en het omliggende gezonde weefsel te sparen. Radiotherapietechnieken hebben grote ontwikkelingen doorgemaakt in de afgelopen jaren, maar er is nog altijd ruimte voor verdere verbeteringen.

Tumoren kunnen niet alleen verschillend zijn ten opzichte van elkaar, maar een enkele tumor kan ook verschillende weefseltypen bevatten, die op een andere manier behandeld moeten worden. Het aanpassen van radiotherapiebehandelingen voor de specifieke weefselkarakteristieken binnen de tumor in iedere patiënt zou kunnen leiden tot effectievere behandelingen. Het huidige klinische onderzoek richt zich steeds meer op dergelijke behandelmethoden, maar een goed begrip van de specifieke tumorkarakteristieken en de integratie van deze kennis in een nieuwe patiënt-specifieke planningsmethode voor radiotherapiebehandelingen ontbreekt nog.

Patiënt- en tumorspecifieke planning van radiotherapiebehandelingen vergt analyse van complexe data uit verschillende modaliteiten. Het analyseren en begrijpen van dit soort data kan een moeilijke en tijdrovende taak zijn, zelfs voor ervaren klinische onderzoekers. Daarnaast is iedere soort data onderhevig aan bepaalde onzekerheden die de nauwkeurigheid en precisie van de behandeling kunnen beïnvloeden. Sommige onzekerheden kunnen geminimaliseerd worden, andere kunnen niet worden vermeden en dienen daarom te worden onderzocht en het effect ervan op de behandeling dient te worden voorspeld. Daarom is van belang dat er methoden worden ontwikkeld die klinische onderzoekers kunnen helpen alle beschikbare patiënt- en tumorspecifieke data te bekijken, begrijpen en analyseren.

Het doel van dit proefschrift is het onderzoeken en ontwikkelen van methoden die klinische onderzoekers kunnen helpen inzicht te krijgen in complexe data en processen in iedere stap in het radiotherapie proces. Voor dit soort methoden maken we gebruik van Visual Analytics, waarbij visualisatietechnieken worden gecombineerd met andere onderzoekgebieden zoals datamining en statistiek in interactieve computerprogramma's. Daarvoor hebben we begrijpelijke visualisatietechnieken ontwikkeld om interactief onderzoek en analyse te kunnen doen naar data en processen met betrekking tot radiotherapiebehandelingen, die het mogelijk maken de kennis en cognitieve vaardigheden van de klinische gebruikers te integreren in de planning.

In Hoofdstuk 4 introduceren we een nieuwe techniek genaamd Orientation-Enhanced Parallel Coordinate Plots voor de representatie en exploratie van multivariate en multidimensionale data. Met deze techniek, een uitbreiding van Parallel Coordinate Plots, verbeteren we de visualisatie van polylines.

Daarnaast beschrijven we een aantal specifieke visualisatiemethoden voor de planning van radiotherapiebehandelingen. In Hoofdstuk 5 beschrijven we een visualisatiemethode voor de exploratie en visuele analyse van potentiële variabiliteit in features (eigenschappen) die zijn berekend uit beelden, veroorzaakt door verschillen in de methoden waarmee ze berekend zijn. Hiermee maken we de exploratie van de feature space met betrekking tot variatie in de anatomie mogelijk.

In Hoofdstuk 6 beschrijven we een nieuw visualisatie framework voor de gemakkelijke exploratie van de feature space van de weefselkarakteristieken binnen de tumor berekend uit de beelden. Dit framework maakt mogelijk: identificatie van onderscheidende gebieden binnen de tumor, exploratie van de heterogeniteit van de tumor, nieuwe ontdekkingen binnen de feature space en het opstellen en testen van hypotheses met betrekking tot klinische referentiedata, zoals histopathologische data. Daarnaast laten we de voordelen van dit framework zien met een praktische toepassing waar we het systeem gebruiken om het ontwerp en het begrip van classifiers voor het herkennen van weefseltypen te ondersteunen.

In Hoofdstuk 7 introduceren we een methode voor de exploratie en beoordeling van fouten die zijn ontstaan bij segmentatie van organen in medische beelden. De methode zorgt voor nieuwe inzichten in de kwaliteit van de gebruikte segmentatiemethodes.

In Hoofdstuk 8 laten we het gebruik van een nieuwe visualisatiemethode zien voor de exploratie en analyse van variabiliteit in verschillende radiotherapiedosisplanningen resulterend uit aanpassingen in de voorgaande stappen in de radiotherapie pipeline. Met dit werk hopen de invloed van deze aanpassingen om het uiteindelijke resultaat van de behandeling beter te kunnen begrijpen.

In Hoofdstuk 9 presenteren we een nieuw Visual Analytics systeem voor de exploratie en analyse van Tumor Control Probability modeling voor de evaluatie en predictie van het resultaat van radiotherapiebehandelingen. Met deze methode kan de onzekerheid in de gebruikte radiobiologische parameters worden meegenomen in de workflow van klinische onderzoekers, wat zorgt voor nieuwe mogelijkheden met betrekking tot de evaluatie van radiotherapiebehandelingen.

Samengevat beschrijft dit proefschrift verschillende Visual Analytics methoden gericht op het gebruiken van informatie en potentiële bronnen van onzekerheid binnen de verschillende stappen van radiotherapiebehandelingen om deze te kunnen presenteren in begrijpbare visualisaties. Deze methodes dragen bij aan de interactieve exploratie en visuele analyse van data en processen in iedere stap van de pipeline van radiotherapiebehandelingen en kunnen worden gebruikt voor verder onderzoek naar radiotherapiebehandelingen en visualisatiemethoden.

Περίληψη

Η οπτική ανάλυση ως μια πιο κατανοητή προσέγγιση
της ψηφιακής ακτινοθεραπείας

Ο καρκίνος του προστάτη είναι ένα από τα συνηθέστερα είδη καρκίνου που συναντάται στον ανδρικό πληθυσμό. Για τη θεραπεία του συνιστάται η χρήση ακτινοθεραπείας, η οποία στοχεύει στο να παράσχει υψηλή δόση ακτινοβολίας που καταστρέφει τους παθολογικούς ιστούς, με όσο δυνατόν λιγότερες παρενέργειες για τους περιβάλλοντες φυσιολογικούς ιστούς. Τα τελευταία χρόνια, η ακτινοθεραπευτική τεχνολογία έχει σημειώσει ιδιαίτερη πρόοδο, αλλά επιδέχεται περαιτέρω βελτίωση.

Πλέον είναι γνωστό ότι οι καρκινικοί όγκοι δεν διαφέρουν μόνο μεταξύ τους, αλλά ταυτόχρονα αποτελούνται από ετερογενείς παθολογικούς ιστούς, οι οποίοι φέρουν διακριτά ιστολογικά χαρακτηριστικά που χρήζουν ξεχωριστής αντιμετώπισης. Προσαρμόζοντας κατάλληλα την ακτινοθεραπεία στις ιδιαίτερες ανάγκες και στα ενδότερα ογκολογικά χαρακτηριστικά του κάθε ασθενούς, μπορούμε να επιτύχουμε πιο τελεσφόρες στρατηγικές θεραπείας. Επί του παρόντος, κλινικές έρευνες οδεύουν προς αυτή την κατεύθυνση, απαιτώντας μια πιο ουσιαστική κατανόηση των ιδιαίτερων ογκολογικών χαρακτηριστικών του κάθε ασθενούς και την ενσωμάτωση όλης αυτής της γνώσης στην θεραπεία.

Αυτή η προσέγγιση, που λαμβάνει υπόψιν τις ιδιότητες των ασθενών και τα χαρακτηριστικά των όγκων τους, εμπλέκει πολυδιάστατα και πολυσύνθετα δεδομένα. Η κατανόηση και ανάλυση αυτών των δεδομένων είναι μια απαιτητική και χρονοβόρα διαδικασία, δύσκολη ακόμα και για τους πιο έμπειρους κλινικούς ερευνητές. Επιπροσθέτως, τα εμπλεκόμενα δεδομένα περιλαμβάνουν αβεβαιότητες, που μπορούν να επηρεάσουν την ακρίβεια και την ορθότητα του επιθυμητού αποτελέσματος. Μερικές από αυτές τις αβεβαιότητες μπορούν να καταπολεμηθούν. Όσες δεν μπορούν να αποφευχθούν, οφείλουν να μελετηθούν και η επίδρασή τους στο τελικό αποτέλεσμα πρέπει να προβλεφθεί. Ως εκ τούτου, έχει προκύψει τα τελευταία χρόνια η ανάγκη για νέες λύσεις και εργαλεία, που μπορούν να βοηθήσουν τους κλινικούς ερευνητές να διερευνήσουν, να κατανοήσουν και να αναλύσουν όλες τις διαθέσιμες ογκολογικές πληροφορίες των ασθενών.

Ο στόχος αυτής της διατριβής είναι η διερεύνηση και η παροχή λύσεων και εργαλείων, που ενισχύουν τους κλινικούς ερευνητές στην προσπάθειά τους να εμπλουτίσουν τις γνώσεις τους σχετικά με τα πολυσύνθετα και πολυδιάστατα δεδομένα τις διαδικασίας του προγραμματισμού της ακτινολογικής θεραπείας. Τέτοιου είδους λύσεις μπορούν να αναζητηθούν στον επιστημονικό τομέα της Οπτικής Ανάλυσης (*Visual Analytics*), ο οποίος συνδέει τεχνικές οπτικοποίησης (*Visualization*) με άλλους επιστημονικούς κλάδους, όπως τον κλάδο της εξόρυξης δεδομένων (*data mining*) και της στατιστικής, μέσω διαδραστικότητας. Μέσα στα πλαίσια αυτής της διατριβής, σχεδιάστηκαν νέες και εύληπτες στρατηγικές οπτικοποίησης για την διαδραστική διερεύνηση και ανάλυση των προαναφερθέντων δεδομένων και των διαδικασιών από τις οποίες προκύπτουν. Αυτές οι στρατηγικές επιτρέπουν την ενσωμάτωση

των γνώσεων και των νοητικών ικανοτήτων των εμπλεκόμενων κλινικών χρηστών, καθ᾽ όλη την αναλυτική προσέγγιση.

Πιο συγκεκριμένα, εισάγουμε μία πρότυπη τεχνική, ονομαζόμενη *Orientation-Enhanced Parallel Coordinates Plots*, για την απεικόνιση και διερεύνηση των εμπλεκόμενων πολυδιάστατων δεδομένων (Κεφάλαιο 4). Με την παρούσα τεχνική, η οποία αποτελεί ουσιαστική ενίσχυση των ήδη υπαρχόντων *Parallel Coordinates Plots*, βελτιώνουμε τον τρόπο παρουσίασης πολυδιάστατων δεδομένων.

Επιπλέον, παρατίθεται ένα σύστημα για τη διερεύνηση και την οπτική ανάλυση ενδεχόμενης μεταβλητότητας στα γνωρίσματα που προέρχονται από τις ιατρικές εικόνες του ασθενούς (*imaging-derived features*) (Κεφάλαιο 5). Αυτή η μεταβλητότητα μπορεί να προέρθει από τη χρήση διαφορετικών μεθόδων υπολογισμού των προαναφερθέντων γνωρισμάτων. Με τον σχεδιασμό που προτείνουμε, διευκολύνουμε την δύσκολη – μέχρι πρότινος – διερεύνηση των αποκλίσεων αυτών, σε σχέση με την ανατομία του ασθενούς.

Στη συνέχεια, με ένα νέο πλαίσιο οπτικοποίησης, υποστηρίζουμε την εξερεύνηση και ανάλυση του συνόλου των ενδότερων ογκολογικών χαρακτηριστικών, όπως αυτά έχουν εξαχθεί από τις διάφορες μεθόδους ιατρικής απεικόνισης του ασθενούς (*imaging-derived features*) (Κεφάλαιο 6). Αυτό το νέο εργαλείο επιτρέπει την ταυτοποίηση των διακριτών ενδοογκολογικών περιοχών, την διερεύνηση της ετερογένειας του όγκου, την απόκτηση νέων γνώσεων σχετικά με τα χαρακτηριστικά γνωρίσματα (*feature space*) των ετερογενών περιοχών του όγκου, καθώς επίσης και τη δημιουργία ή την επιβεβαίωση υποθέσεων, σε σχέση με άλλα κλινικά δεδομένα, όπως ιστοπαθολογικά δεδομένα. Παραθέτουμε μια επιπρόσθετη πρακτική εφαρμογή του στον εύκολο σχεδιασμό και στην κατανόηση μεθόδων ταξινόμησης και χαρακτηρισμού ιστών (*tissue characterization classifiers*).

Επιπροσθέτως, παρουσιάζουμε ένα νέο εργαλείο Οπτικής Ανάλυσης για την διερεύνηση και αξιολόγηση σφαλμάτων που προκύπτουν κατά τον καταμερισμό των οργάνων στις ιατρικές εικόνες (*segmentation*) (Κεφάλαιο 7). Η προτεινόμενη προσέγγιση προσφέρει μια καινούρια εικόνα σχετικά με την αποτελεσματικότητα των χρησιμοποιούμενων μεθόδων.

Ακόμη, προτείνουμε τη χρήση ενός νέου εργαλείου απεικόνισης για τη διερεύνηση και ανάλυση της ενδεχόμενης μεταβλητότητας μέσα σε ένα σύνολο προγραμματισμένων ακτινολογικών πλάνων, τα οποία έχουν προκύψει από διάφορες τροποποιήσεις καθ᾽ όλη τη διάρκεια της διαδικασίας προγραμματισμού της θεραπείας (Κεφάλαιο 8). Μέσω αυτού, στοχεύουμε στο να προσφέρουμε μια νέα αντίληψη για τον αντίκτυπο που μπορούν να έχουν ενδεχόμενες τροποποιήσεις πάνω στο τελικό αποτέλεσμα.

Τέλος, παρουσιάζουμε ένα νέο σύστημα Οπτικής Ανάλυσης για την εξερεύνηση και ανάλυση των μοντέλων που χρησιμοποιούνται κατά την πρόβλεψη του αποτελέσματος της ακτινοθεραπείας (*Tumor Control Probability modeling*) (Κεφάλαιο 9). Με αυτή μας την προσέγγιση, οι – παραβλεφθείσες, έως τώρα – αβεβαιότητες των μοντέλων μπορούν να ενσωματωθούν στη ροή εργασίας των κλινικών ερευνητών, προσφέροντας νέες δυνατότητες για την εκτίμηση του αποτελέσματος των επιλεγμένων ακτινοθεραπευτικών στρατηγικών.

Συνολικά, αυτή η διατριβή περιγράφει λύσεις από τον κλάδο της Οπτικής Ανάλυσης, που στοχεύουν στην ενσωμάτωση πληροφοριών από τα διάφορα βήματα της διαδικασίας του ακτινολογικού προγραμματισμού και ενδεχόμενων πηγών αβεβαιότητας, μέσα σε εύληπτες απεικονίσεις. Αυτές οι λύσεις συνεισφέρουν στην διαδραστική διερεύνηση και ανάλυση των εμπλεκόμενων δεδομένων και μεθόδων, δημιουργώντας κατάλληλο έδαφος για μελλοντική έρευνα τόσο στον κλινικό τομέα όσο και στον κλάδο της Οπτικής Ανάλυσης.

1

Introduction

Πάντες ἄνθρωποι τοῦ εἰδέναι ὀρέγονται φύσει.

Aristotle (384-322BC)

1.1. Motivation

Cancer involves a group of diseases, which are characterized by the uncontrollable and ab-normal division of cells. These cells have also the potential to invade or spread from one organ or body part to another [183, 309]. In 2012, 14.1 million people were globally diagnosed with cancer, while 8.2 million deaths have been attributed to malignancies [262]. Yet, the constantly rising life expectancy of the population, as well as the changes in the lifestyle in the developed world, are expected to induce an increase in the current rates. In the next two decades, cancer incidence is expected to rise by 60% among the population [262].

Prostate cancer is the specific type of cancer that develops in the prostate gland of the male reproductive system. Being first in incidence and second in mortality, it is one of the most common malignancies in males [6, 309]. However, this type of cancer can be success-fully treated, usually through radical prostatectomy, chemotherapy or radiotherapy. Among these, the latter is the most common, with 60% of all prostate cancer patients being referred to radiotherapy at some stage of their treatment [66].

In the past decade, radiotherapy technology has undergone a big revolution, offering exceptional flexibility in dose delivery. It has managed to improve treatment by irradiating tu-mors with a high dose, while minimizing the side effects of radiation on the adjacent healthy organ tissues [300]. Despite the significant achievements of radiotherapy, there is still room for further improvement. Each patient case should be handled individually, as tumors might differ in type, aggressiveness, location, or size. Additionally, tumors are heterogeneous tis-sues, consisting of distinct regions with different characteristics. Understanding better the specific anatomical and intra-tumor characteristics of each patient and incorporating these into treatment planning, by selecting the most adequate radiation strategy for each tumor region, can lead to the design of more effective treatments [266].

To this end, a standardized process, which will include all available patient- and tumor-specific information, needs to be designed and employed. Developing a reusable, but per-sonalizable radiotherapy pipeline, tailored to the specific anatomical and intra-tumor charac-teristics of each patient, is the main goal of the *FP7 European Project DR THERAPAT – Digital Radiation Therapy Patient* [89], part of which is this dissertation. The pipeline proposed by the DR THERAPAT Project consists of several steps, which are depicted in Figure 1.1 and will be discussed in detail in the upcoming chapters.

In this pipeline, imaging data of the prostate of the patient are initially acquired, using several modalities. From this, additional features indicative of tissue characteristics may be computed. Subsequently, tumor tissue characterization takes place to enable the iden-tification of intra-tumor regions. At this point, the specific characteristics of each region, such as aggressiveness or resistance to treatment, are derived. The structures surrounding the prostate, which need to be spared during treatment, must be identified as well. This is performed during the segmentation step. Based on all this tumor tissue and anatomical in-formation, radiation doses can be selected adequately, to more effectively treat the different tumor parts, without harming the adjacent healthy organs, during the planning phase. After the radiotherapy plan is designed, the eventual response of the tumor to the employed radio-therapy treatment strategy is modeled. From the latter, clinical researchers can predict the outcome of the treatment.

This patient- and tumor-specific radiotherapy planning pipeline involves data, which are
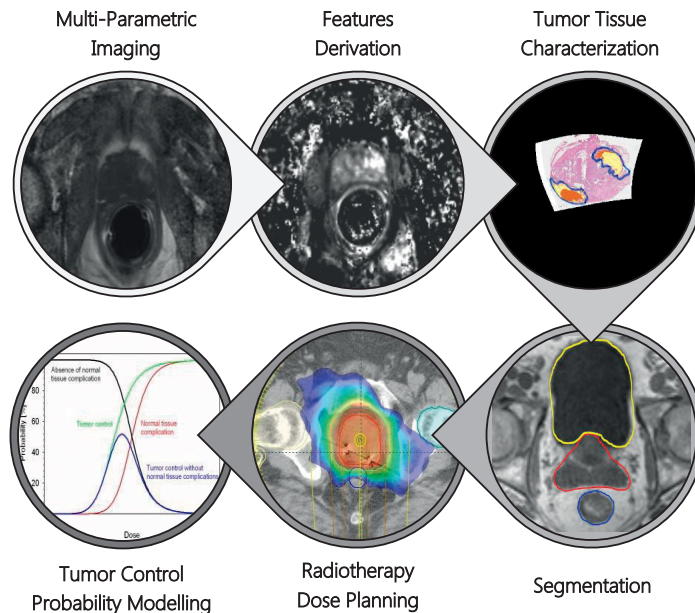
Figure 1.1: Schematic depiction of the steps of the patient- and tumor-specific radiotherapy planning pipeline, employed in clinical research to create personalized, tumor-tailored radiotherapy plans.

constantly growing in complexity. In addition, this data might include noise, inaccuracies, errors and uncertainties, which also need to be considered, studied, or predicted. Currently, the mental processing and cognitive analysis of this complex information is a tedious task, even for experienced clinical researchers. Hence, there is an emerging need for tools and solutions, to help clinical researchers explore, understand and analyze the complex information involved in the tumor-tailored radiotherapy planning pipeline.

Conventional computational solutions would not be sufficient for exploration and hypothesis generation purposes, as there is one essential aspect that cannot be easily automatized and incorporated: the prior knowledge and cognitive skills of the involved clinical experts. Yet, in visualization, human vision and perception can be exploited to amplify cognition and provide new insight for clinical researchers, helping them to achieve better understanding of the information from the radiotherapy planning pipeline. Of particular interest is the field of *Visual Analytics* [142]. This is, in essence, the integration of visualization techniques with other disciplines, such as data mining or statistics, further enhanced by the capabilities of interactive visual interfaces. Visual Analytics is particularly suitable for facilitating interactive data exploration and analysis, and for providing a deeper cognitive insight for the intended users [142, 269].

In the present context, Visual Analytics has the potential to provide a direct means of feedback on imaging data, tumor tissue characterization, segmentation and modeling of the tumor response to treatment for clinical researchers. This can be attained by design-

ing comprehensible visualizations for the interactive exploration and visual analysis of the involved data and processes. In this way, the intended clinical researchers are empowered to derive information from their data, to understand the newly discovered knowledge, and to formulate or confirm hypotheses with respect to their observations and findings. The main research goal of this dissertation can be summarized as:

> ***Investigating Visual Analytics strategies***
> ***to aid clinical researchers gain insight into***
> ***the complex processes and data,***
> ***present in the radiotherapy pipeline.***

To the best of our knowledge, involving clinical experts through visual analysis and interaction in the workflow of the entire personalized and tumor-specific radiotherapy pipeline has not been addressed before. In addition to that, despite the broadness of applications that the field of Visual Analytics has tackled, the exploration and analysis of the data involved in the radiotherapy planning pipeline has also not been addressed. Although there are numerous ways of visualizing multi-dimensional and complex data [143], the most relevant previous approaches are not fully compatible with our specific field of application, nor with the involved data and requirements of clinical researchers.

1.2. Contribution and Outline

The primary *contribution* of this work is the design and implementation of Visual Analytics solutions, which are able to increase the knowledge and understanding of clinical researchers into the data and processes of the radiotherapy planning pipeline. We propose new visualization strategies in the form of novel application prototypes, which advance the state-of-art in visualization. Our proposed solutions are structured following the radiotherapy pipeline, as depicted in Figure 1.2. In particular:

- We introduce a novel technique for the representation and exploration of *multi-variate, multi-dimensional data*, as an extension of Parallel Coordinate Plots (PCPs) [128]. Reducing clutter due to overplotting of polylines in this representation is a challenging topic, and the proposed technique addresses it, by improving the display of the data and by emphasizing their underlying structure. Pattern and outlier discernibility is enhanced, especially, in cases where multiple patterns are present or when the view on certain patterns is obstructed by noise. This technique, called Orientation-Enhanced Parallel Coordinate Plots, can be applicable to several steps of the pipeline, as well as to other broader applications.

- We provide a visualization design for the exploration and visual analysis of potential variability in *imaging-derived features*, caused by differences in the methods employed for their computation. Currently, this is a tedious and time-consuming task for clinical researchers, which does not provide all the necessary insight. The presented ap-
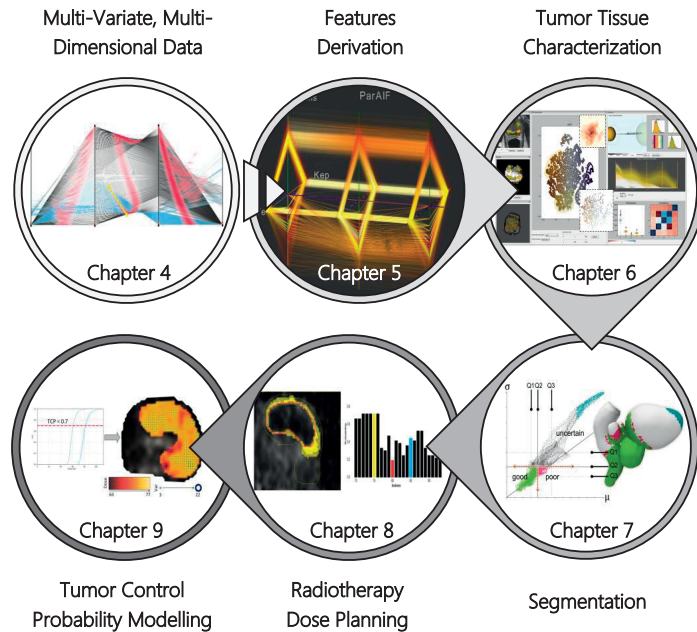
Figure 1.2: Overview of the contributions of the core chapters of the dissertation. Each chapter addresses one or more steps of the tumor-tailored radiotherapy pipeline, depicted in Figure 1.1.

proach provides a solution for the easy exploration and analysis of variations in a single combined view. Particular attention is given to the association of observations from the feature space to patient anatomy.

- We present a new Visual Analytics system that addresses the easy exploration and analysis of the feature space of imaging-derived *intra-tumor tissue characteristics*. This work employs dimensionality reduction, enabling the identification of distinct intra-tumor regions, the exploration of tumor heterogeneity, new knowledge discovery within the feature space, as well as hypothesis generation and confirmation, with respect to clinical reference data, such as histopathological data. An additional practical application of the proposed system, namely aiding the design of tissue characterization classifiers and understanding their behavior, is also introduced.

- We propose a new Visual Analytics tool for the exploration and assessment of errors occurring in the *segmentation* step of the pipeline. Here, the focus is on the visual analysis of errors in the segmentation of the involved pelvic structures surrounding the prostate and on how Visual Analytics can supply insight into the prediction of the performance of the employed segmentation algorithms. These two key-points of the segmentation step of the radiotherapy pipeline have not been addressed before.

- We demonstrate a new visualization design to address the exploration and analysis of *variability in an ensemble of radiotherapy dose plans*, generated by adjustments in

the previous steps of the pipeline. For example, parameter adaptations in one of the previous steps might result into differences in the final dose planning outcome. The goal of this design is to provide insight into the effect of these adjustments in the pipeline on the resulting radiation dose plan.

- We establish a new Visual Analytics system for the exploration and analysis of *tumor control probability modeling*. This modeling is often performed in clinical research to evaluate and predict the outcome of a radiotherapy treatment. With our introduced approach, the – up to now disregarded – imaging-induced uncertainty and sensitivity analysis of the radiobiological parameters employed in the model can be incorporated in the workflow of clinical researchers, providing new possibilities for the evaluation of the selected radiotherapy strategies.

All proposed solutions share three common main characteristics. First of all, they can be described as highly interactive exploratory environments that amplify the understanding of clinical researchers into data or processes of the pipeline. In all of them, insight is achieved through a high level of involvement of the intended users. Secondly, they are all designed taking into consideration the characteristics of the data, the users – clinical researchers – and the specific tasks that these users need to perform at each step of the pipeline, following the *Data-Users-Tasks Design Triangle* concept [178]. Thirdly, evaluation plays an important role in all of the proposed solutions. After the design and implementation, all of them are carefully evaluated with their intended users, following commonly employed evaluation guidelines [159].

The remainder of this dissertation is structured as follows. Chapters 2 and 3 are establishing the clinical and technical background of this work, respectively. Chapter 2 gives an overview of the clinical background, providing all the necessary information for understanding the radiotherapy pipeline, its components and the data or processes involved at each step. Chapter 3 provides a summary of the technical background of this work. Basic concepts of the visualization field are presented and previous related work is discussed, with particular focus on the existing limitations, open problems and challenges. Chapters 4 to 9 comprise the core of this dissertation. More in particular, Chapter 4 proposes a novel technique for the representation of multi-variate, multi-dimensional data, which is applicable to the whole radiotherapy pipeline. Chapters 5 to 9 are structured to address each of the steps of the radiotherapy pipeline, as shown in Figure 1.2. Finally, Chapter 10 concludes the dissertation, providing an overview and discussion of the obtained results. A reflection regarding lessons learned and directions for future work are also presented in the final chapter of this dissertation.

2

Clinical Background

Medicine is a science of uncertainty and an art of probability.

Sir William Osler (1849-1919)

In this chapter, a brief overview on the basics of cancer and, in particular, of prostate cancer is provided. Basic concepts of radiotherapy are introduced, with emphasis on the radiotherapy planning pipeline and the steps, which need to be performed prior to radiotherapy treatment. The multi-varied and multi-dimensional data involved at each step of the pipeline, along with potential sources of uncertainty, are also presented.

2.1. Cancer: An Overview

2

The oldest reference to cancer dates back to 1600 BC, in an ancient Egyptian textbook on trauma surgery, called the Edwin Smith Papyrus [300]. The papyrus did not refer to the disease as cancer, but it described several cases of breast tumors, and proposed surgical treatments. The term cancer (in greek, καρκίνος or καρκίνωμα) is credited to Hippocrates (c. 460 - 370 BC), due to the resemblance of the shape of tumors to a crab [200]. Hippocrates attributed cancer to an excess of black bile – a belief that defined cancer for many centuries, as a systemic disease or a natural process. According to this theory, any kind of local treatment, such as surgery, was considered unsuitable [200, 300].

The theory of Hippocrates continued to prevail, until the French physician Claude Gendron (1663-1750) described cancer as a locally hard-growing mass, which cannot be treated by drugs, but needs to be removed with all its *filaments* [200]. In the 18th century, Jean Astruc and Berhard Peyrilhe were the first to conduct experiments, seeking better diagnosis, new treatments and deeper understanding of the causes of the disease [200]. The development of microscopes, in the late 19th century, made possible the examination of cancer tissues and tumors [192]. Researchers discovered that cancer cells differ in appearance from normal cells, and started focusing more on investigating the behavior of cancerous tissues. In the following years, several different theories and studies about the origin, the causes and the treatment of cancer emerged, out of which, many were discarded [192].

Nowadays, researchers are able to diagnose tumors, even in early stages. Through the examination of the genetic information of cells, scientists are able to determine the processes that are responsible for the initiation of cancer in a patient. Current theories explain that cancer is the result of abnormal cellular growth, during which cells proliferate and divide in an unregulated way. This occurs due to genetic mutations in otherwise non-reproductive cells [300]. In many tumor types, mutations both in the genes that promote the proliferation of cells and in the genes that stop cell division, are encountered. These mutations may be triggered by exposure to carcinogens, such as radiation, sunlight or smoking [171, 300]. However, gene mutations can also be passed on through generations in a familial expression of the disease, while random mutations are also possible [300]. A schematic depiction of the abnormal processes taking place before and during the multiplication of cancer cells is presented in Figure 2.1.

Tumors can originate from any cell and are, in general, divided into two categories: *benign* and *malignant* [300]. The former are well differentiated and do not invade surrounding normal tissues, but can compress it causing damage. The latter have the ability to spread to other body parts though the bloodstream or lymphatic system, in a process called *metastasis* [300]. They can also divide and grow by creating new blood vessels to sustain themselves, in a process called *angiogenesis* [87, 300]. In addition to these two types, pre-malignant cases are also possible. These involve abnormal cells that may develop into malignancies, if left untreated.

Early detection and diagnosis is considered an important factor in tumor treatment, as chances of metastasis are lower [300]. However, not all cancer types are accompanied with clear symptoms, making detection and diagnosis difficult in the early stages. Advances in medical imaging have increased the capabilities of physicians to detect tumors, and – in some cases – screening tests are performed. Examples of screening tests include mam-
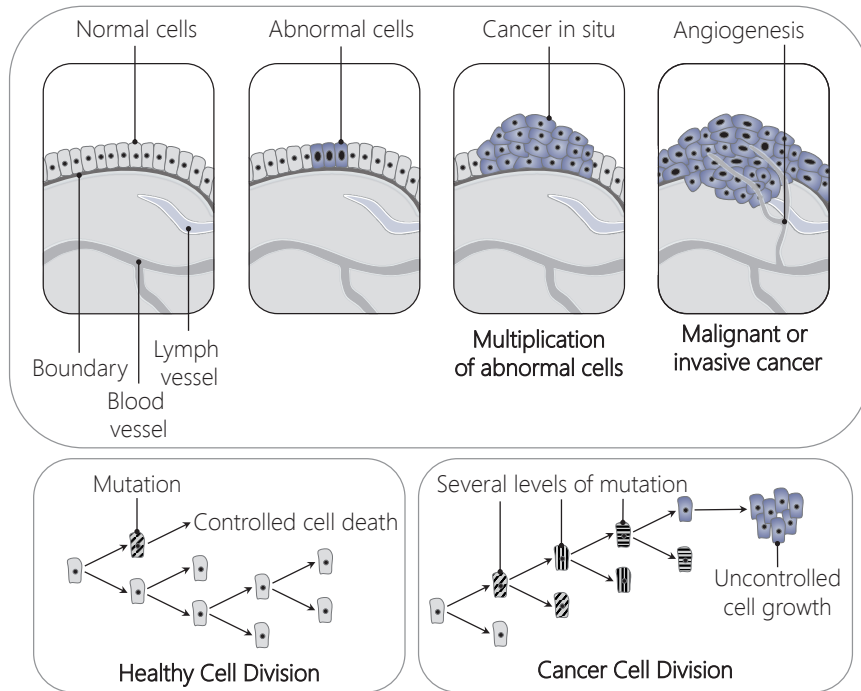
Figure 2.1: Schematic depiction of all processes occurring during the uncontrolled and abnormal division and growth of cancer cells, in a malignant tumor. *Figure created by the author, inspired by figures in the website of the National Institute of Health* [184].

mography for breast cancer, or colonoscopy for colorectal cancer [163, 264]. Medical imaging, such as Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), has allowed a view on the inside of their patients, aiding clinicians to determine more appropriate treatments. Currently, cancer can be successfully treated through a vast selection of techniques [300], among which surgery, chemotherapy and radiotherapy.

2.2. Prostate Cancer

Prostate cancer – or carcinoma of the prostate – refers to the development of cancer in the prostate, a gland in the male reproductive system [275]. The prostate gland is a walnut-shaped and sized organ that consists of fibrous, glandular and muscular tissues. It surrounds the urethra, and is located between the bladder and the urogenital diaphragm.

In pathology, the prostate is divided into four zones [191]. These include the *peripheral zone*, which is the biggest part of the gland and the location where 70-80% of cancer originates; the *central zone*, which surrounds the ejaculatory ducts and accounts for more aggressive types of cancer; the *transition zone*, which surrounds the urethra; and the *fibro-muscular*

2

Bladder

Vas Deferens

Seminal Vesicles

Central Zone
Prostate Capsule
Prostate
Transition Zone
Verumontanum
Peripheral Zone
Urethra
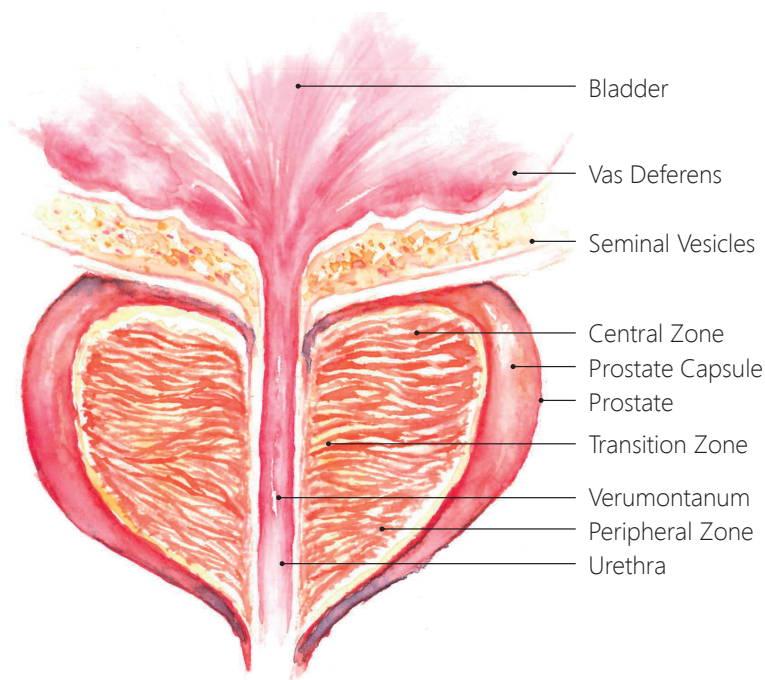
Figure 2.2: The anatomy of the prostate and its main surrounding structures. *Drawing created by Peter Mindek for this dissertation.*

zone or *stroma*. The anatomy of the prostate, as well as the main structures around and inside it, are presented in Figure 2.2.

Prostate cancer is the most common malignancy in males. Approximately 1 out of 6 men are estimated to develop the disease in their lifetime [300]. Several factors may increase the risk of prostate cancer, among which older age, hereditary predisposition and race [7, 300]. More than 60% of prostate cancer cases occur in men above 65 years, while having a first-degree relative with the disease doubles the risk [7]. Additionally, African Americans have a higher incidence of the disease in comparison to other races of similar age [300].

Often, prostate cancer screening and diagnosis are performed through digital rectal examination, prostate-specific antigen (PSA) testing, biopsy, and medical imaging [300]. Staging of the disease is necessary to determine the most suitable treatment, and is performed using the so-called TNM Classification of Malignant Tumors [77, 300]. This system takes into consideration the size of the primary tumor (T), the number of involved lymph nodes (N) and the presence of metastasis (M). Another standardized scoring method, which is often employed, is the Prostate Imaging Reporting and Data System (PI-RADS), which is assessed on imaging data of the prostate of the patient [8].

If a biopsy is performed, then the Gleason grading system [82, 125] can be used, as well. For this, the degree of differentiation and the morphology of the tumor are graded to deter-

Increasing tumor aggressiveness / Increasing Gleason grade →

1   2   3   4   5
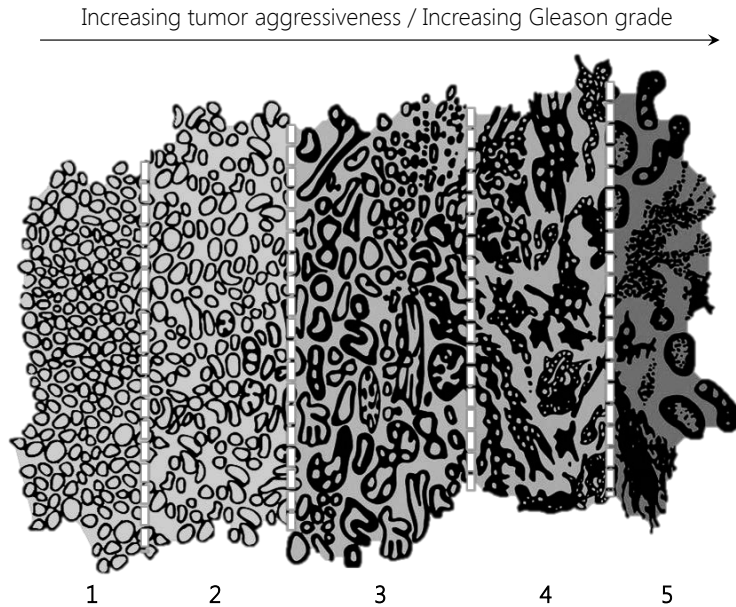
Figure 2.3: Increasing Gleason grades indicate increasing aggressiveness of the prostate tumor, which is accompanied by changes in the formation, differentiation and structure of the glands. *Figure adapted from* [300].

mine a Gleason score (GS). Lower grades denote slowly growing, non-aggressive tumors, while higher grades denote invasive and metastatic malignancies. A schematic example of the Gleason grading system is depicted in Figure 2.3.

Common treatments for prostate cancer include radical prostatectomy, hormonotherapy, chemotherapy and radiotherapy [300]. Yet, the most frequently adopted treatment is radiotherapy, addressing up to 60% of all prostate cancer patients at some stage during their treatment [66].

2.3. Radiotherapy

Radiotherapy, or radiation therapy (RT), is one of the most common approaches for cancer treatment, which uses ionizing radiation to treat malignant cells. It is used as *therapeutic treatment* to cure the disease, as *adjuvant therapy* to prevent tumor recurrence, or as *palliative treatment* to relieve patients from several symptoms [300]. Often, radiotherapy complements surgery, chemotherapy, hormonotherapy, immunotherapy or a combination of those [300].

Radiotherapy has been used as cancer treatment for more than a century [254, 267] – starting in 1896, when x-rays were first employed to treat breast tumors [108]. With the discovery of radioactive elements by Marie Curie, a new era in medical research and treatment

began [267]. Although the hazards of radiation exposure were still unknown, radiotherapy was already applied to many diseases. Prior to World War II, only radium was known and used as source of radiation for radiotherapy, but later additional artificial radioisotopes were employed [146, 254]. In the 50s, linear particle accelerators were developed, improving many aspects of radiotherapy [38, 254]. In the 70-80s, new imaging technologies allowed for 3D radiation delivery, Intensity-Modulated Radiation Therapy (IMRT) [162, 165] and Image-Guided Radiation Therapy (IGRT) [311]. All these advances enabled better targeting tumors, with better treatment outcomes and less side effects for healthy tissues [267].

Radiotherapy is based on the concept that ionizing radiation can damage the genetic information of cancerous tissues [300]. The administered radiation dose used in radiation therapy is measured in Grays (Gy), and varies depending on the type and stage of cancer being treated. In general, the response of tumor tissues to radiation is determined by their size, but especially by their radiosensitivity [267, 300]. More radiosensitive cancers, such as leukemia or epithelial tumors, can be treated by moderate doses, while more resistant ones, such as renal cancer or melanoma, require much higher doses. *Hypofractionation* [267, 300] of the dose is another common practice, where the total dose is spread out in small amounts over time, to allow the recovery of normal cells and to prevent the repair of tumor cells between fractions.

Currently, the two most common technologies of radiotherapy treatment are External Beam Radiation Therapy (EBRT) and Brachytherapy (BT) [267, 300]. In the former, the radiation source is located outside of the patient. In the latter, the radioactive sources are positioned precisely in the area to be treated, affecting only a very localized area [93]. An example of EBRT is Intensity-Modulated Radiation Therapy (IMRT), which can precisely address concave tumors, such as tumors enclosing the spinal cord or major blood vessels. This is achieved, by modulating the intensity of the radiation beam near the tumor volume, while decreasing or avoiding the radiation among the surrounding healthy tissues [301]. In order to reduce the toxicity to the normal tissues, radiation beams need to be shaped and aimed from several angles of exposure, to cumulatively target the tumor [300].

Still, radiotherapy may result in a number of side effects, most commonly fatigue or skin irritations [300]. In particular, prostate cancer radiotherapy may cause rectal bleeding, incontinence, impotence, as well as other urinary or bowel-related side effects.

2.4. The Radiotherapy Planning Pipeline

Prior to the administration of the radiation dose, the delivery strategy and prescribed dose need to be determined, in a process called *treatment planning* [300]. Conventional radiotherapy treatment planning consists of a number of steps, which are depicted in the upper part of Figure 2.4.

After the patient has been diagnosed and referred to radiotherapy by a physician, medical imaging is employed to obtain a view into the patient. This is done, typically, using Computed Tomography (CT). Subsequently, the target tumor and the surrounding organs at risks are defined. In order to account for patient setup errors [263] and for other sources of inaccuracy [131, 234], safety margins are added around the organ volumes [19, 131, 176]. These safety margins are shown schematically in Figure 2.5.
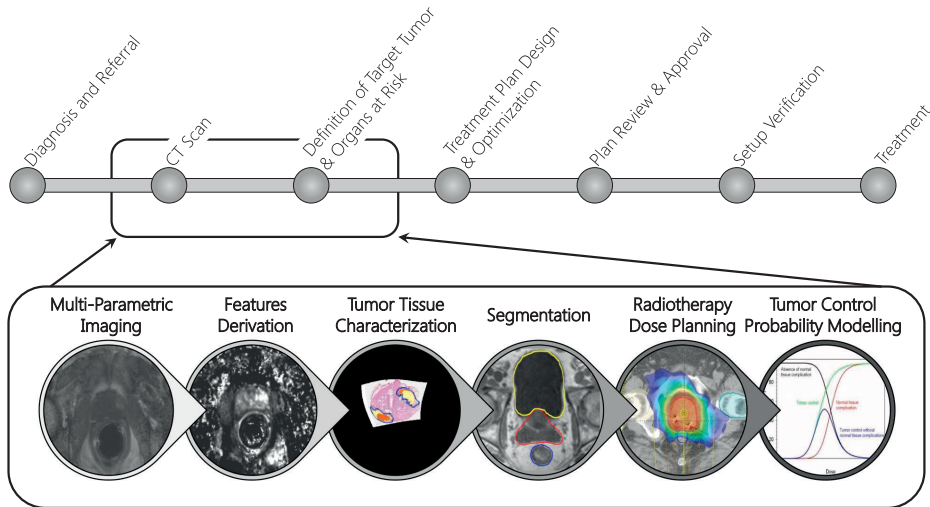
Figure 2.4: Schematic depiction of the conventional radiotherapy pipeline (top, *inspired from* [300]) and its modification (bottom), with the integration of patient- and tumor-specific knowledge, in order to achieve a treatment planning, tailored to the intra-tumor characteristics of each subject, as proposed by the DR THERAPAT project [89].

The volume that contains the palpable, or visible in imaging techniques tumor is called Gross Tumor Volume (GTV). The Clinical Target Volume (CTV) contains the GTV with an additional a safety margin around it, based on anatomical and biological guidelines. Finally, there is the Planning Target Volume (PTV), which accounts for the effect of the movement of the tissues of the CTV, the movements of the patient, variations in the size of the CTV tissues and variations in beam geometry characteristics, to ensure that the CTV actually receives the prescribed dose [19, 131, 176].

After the localization of the tumor and adjacent organs, an initial treatment plan is designed, using dedicated treatment planning software. This treatment planning software defines the geometric, radiobiological and dosimetric aspects of the treatment, optimizing for tumor treatment and for healthy tissue preservation [300]. Once the plan is performed, it is reviewed and approved. Still, before the treatment delivery, the location of all implicated structures, as well as the selected radiation strategy are verified, once more.

In the past decades, radiotherapy treatment has managed to improve tumor control, and to minimize the radiation-induced toxicity in healthy tissues around the tumor. However, a standardized process, which, as described in Chapter 1, will include all available patient- and tumor-specific information, still needs to be designed and employed [266]. To achieve this, a number of additional steps need to be conducted. These steps have already been briefly presented, in Figure 1.1 (Chapter 1).

In the lower part of Figure 2.4, we depict schematically how the basic radiotherapy planning pipeline is affected, when additional patient-specific and tumor tissue information are
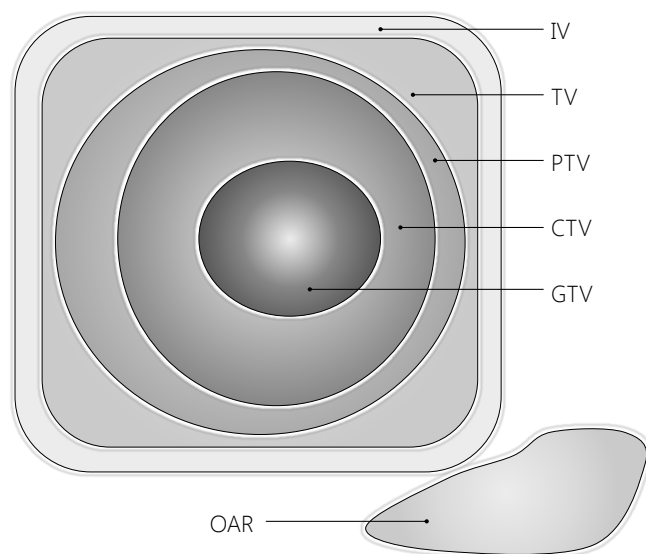
Figure 2.5: Safety margins to account for patient setup errors and other sources of inaccuracy, during the administration of the radiation dose. The Gross Tumor Volume (GTV), Clinical Target Volume (CTV) and Planning Target Volume (PTV), as well as the organs at risk (OAR) are denoted. Outside of the PTV, the Treated Volume (TV) and the Irradiated Volume (IV) are additionally defined. *Figure adapted from* [300].

included, in order to achieve a more personalized planning, tailored to the specific intra-tumor properties of each patient. In this new, modified pipeline, data from different acquisition modalities, their derived features or parameters, segmented data, tumor-specific tissue characteristics, as well as relationships between them, need to be integrated, as shown in Figure 2.4. The modified pipeline is the one proposed by the DR THERAPAT project [89].

Apart from the aforementioned multi-modal and multi-valued data, another important aspect is the uncertainty, which is present at all steps of the planning pipeline. In literature, there is no unanimous opinion on the definition of uncertainty. According to the National Institute of Standards and Technology (NIST), *data uncertainty includes concepts, such as statistical variation or spread, error or inaccuracy and minimum-maximum ranges* [167]. Another definition is provided by Griethe et al. [104], as *a composition of different concepts, such as error (outlier or deviation from a true value), imprecision (resolution of a value compared to the needed resolution), subjectivity (degree of subjective influence in the data) and non-specificity (lack of distinction for objects)*. In the current case of radiotherapy planning, we define uncertainty as any variation in the dose planning outcome, which is produced by an ad-hoc choice or a stochastic process, in one or more steps of the radiotherapy planning pipeline.

In all steps of the radiotherapy planning pipeline, there are several sources of uncertainty. Although some of these uncertainties can be minimized, there are others that cannot be avoided. The accumulation and propagation of uncertainties, throughout the entire pipeline,

2

may have an influence on the dosage planning, and the final outcome of the planning procedure. Therefore, the analysis and communication of uncertainty to the clinical user would also be essential for the interpretation of the outcome, for reducing the existing uncertainties and, potentially, for improving the final outcome.

Incorporating patient-specific tumor and anatomical characteristics, as well as the involved uncertainty, in the radiotherapy planning pipeline is currently addressed only in research – not in clinical practice. Researchers are interested in exploring and analyzing the complex, multi-valued data involved in the tumor-tailored planning pipeline, in order to understand the data and formulate hypotheses concerning their patients, the approaches that are being followed during the pipeline and the treatment strategies that need to be pursued. These researchers could be radiation oncologists, radiologists and clinical physicists, and also researchers that are working on algorithms and methods used through the pipeline.

In the following sections, we will separately address each step of the personalized radiotherapy planning pipeline, along with the data, uncertainty, and processes that are of particular interest for researchers and for the present dissertation.

2.4.1. Multi-Parametric Imaging

The purpose of the first part of the pipeline is to obtain the images needed for radiotherapy planning, from a multitude of acquired medical imaging sequences [124]. The multi-parametric imaging that is employed in prostate cancer research includes T2-Weighted Magnetic Resonance Imaging (MRI), Diffusion-Weighted Imaging (DWI), Dynamic-Contrast Enhanced MR Imaging (DCE-MRI) and, optionally, MR Spectroscopy [16, 25, 51, 124]. Additionally, Computed Tomography (CT) Imaging may be performed [124]. All aforementioned images need to be registered, in order to be transformed into the same coordinate system. An example of a multi-parametric imaging acquisition is shown in Figure 2.6.

Computed Tomography (CT) Imaging is the acquisition of a series of X-ray images at different angles, which are composed together into a volume [220]. Each single X-ray image exploits the different absorption properties of tissues with different densities. CT imaging is not suitable for soft tissue differentiation, but it provides a very good contrast between soft tissue and bones. Hence, it can be used to provide anatomical context, with respect to the pelvic bones, while it can also aid to identify whether the cancer has already spread into the lymph nodes, other organs or boney structures [124].

Magnetic resonance imaging (MRI) is a medical imaging technique, which exploits the properties of human tissues in magnetic fields [220]. The basic concept behind MRI is that, when the protons of nuclei of atoms – in particular hydrogen, which is in abundance in the water molecules of the human body – are placed in a magnetic field, they align themselves along this field. When a radio frequency pulse is emitted and tuned to a specific range of frequencies, at which the hydrogen protons precess, the magnetic spins of some hydrogen protons flip with a specific angle. When the radio pulse is stopped, these hydrogen protons align back with the magnetic field, losing energy and emitting an electromagnetic signal. This emitted signal is detected by receiver coils located around the body and the intensity of the received signal is used to build up the cross sectional images of the patient.

By using different pulse sequences, different tissue characteristics are brought forward, creating different kinds of images from inside the human body [220]. T2-Weighted imaging
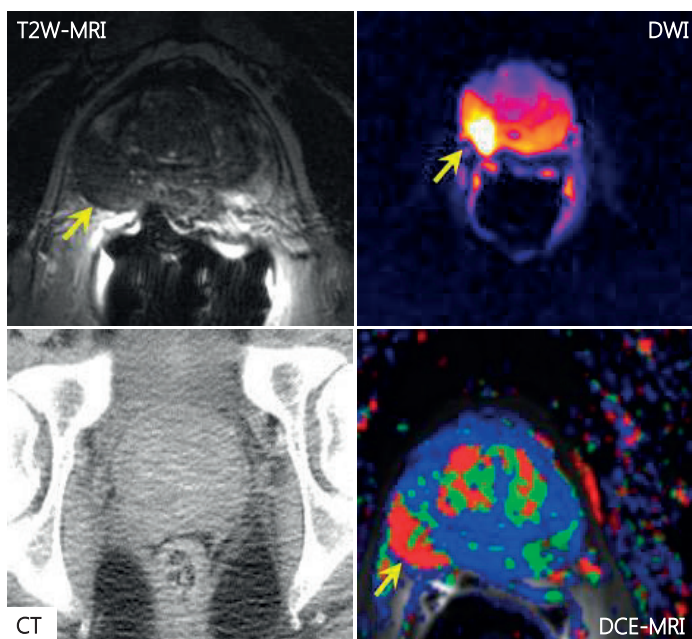
Figure 2.6: The four main medical imaging acquisitions that are used for prostate cancer diagnosis and detection: T2-Weighted Magnetic Resonance Imaging (MRI), Diffusion-Weighted Imaging (DWI), Dynamic-Contrast Enhanced (DCE) MR Imaging and Computed Tomography (CT). The latter (CT) is used to provide anatomical context.

is one of the basic sequences of MRI and is commonly employed in prostate imaging [16, 25, 51, 124], in order to highlight the peripheral zone, which is the biggest part of the gland and the location where 70-80% of cancer originates [191]. The central zone, the transitional zones and the stroma have a lower signal intensity in T2-Weighted images.

Diffusion-Weighted (DW or DWI) MRI is an imaging method that produces in-vivo images of biological tissues, exploiting the diffusion of water molecules [256]. It is based on the concept that the diffusion of water molecules inside a voxel of tissue, which is the random Brownian motion of these water molecules, is constrained by the boundaries of the cell membranes. Different tissues or pathologies can affect the properties of diffusion, which can be used to identify highly dense cellular tissue, like tumors [16, 25, 51, 124]. As a measure of the magnitude of diffusion and to quantitatively asses the cell membrane restriction of water, the apparent diffusion coefficient (ADC) maps are used [256]. In prostate cancer, there are indications that lower diffusion values relate to higher tissue cellularity, which is a sign of the existence of a tumor [16, 25, 51, 124].

Dynamic-Contrast Enhanced (DCE) MRI is based on the idea that different tissues have different uptake properties, when they are injected with a contrast agent [16, 25, 51, 124]. Tumors tend to develop new, disorganized and permeable vessels, which have thinner and
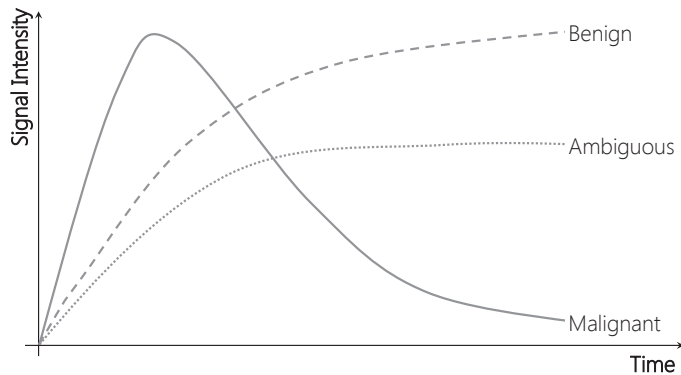
Figure 2.7: Signal Intensity enhancement curves for three different types of tissue (malignant, benign and ambiguous) across time. *Figure inspired from* [190].

weaker walls [283]. Thus, they show early and rapid enhancement and early wash-out of contrast agents, unlike healthy tissue, as shown in Figure 2.7. In order to study the vascularity, the vascular perfusion and vessel permeability characteristics of the tissues, pharmacokinetic (PK) modeling is performed. From that, different maps, indicative of tissue characteristics are calculated [257]. These characteristics will be discussed in the upcoming section.

All these imaging acquisitions are accompanied by uncertainty, mainly from noise or artifacts inherent in each scanning procedure. More specifically, T2-Weighted MR imaging has high sensitivity and poor specificity for tumor detection and characterization [156, 282], while it usually suffers from artifacts, due to patient motion.

DW-MRI has highly varying sensitivity and specificity for tumor detection [147, 282], depending on patient population characteristics, on the prostate tissue zone and on the scanning procedure itself. In DW-MRI, different degrees of diffusion weighting can be applied, in order to retrieve the ADC maps. This degree of diffusion is described by the so-called b-value and, traditionally, a value of 1000 sec/mm^2 is used, due to hardware restrictions for achieving acceptable echo times. However, there is no general consensus and several recommendations for b-values have been made in literature [193]. Still, the use of higher b-values can enhance the sensitivity and the specificity of the acquisition [25, 51], but they can also deteriorate the contrast resolution among healthy and tumorous tissues [51]. Among others, additional common problems in DW-MRI can be motion artifacts [25, 244, 248] even from small movements of the rectum or the bladder, poor spatial imaging resolution and image distortions, due to magnetic field inhomogeneities at the interfaces between different tissues [25, 51, 248].

DCE-MRI alone has reported highly varying sensitivity and specificity for the detection and characterization of tumors [156, 282]. This means that several benign conditions, such as prostatitis or infections, may have the same appearance as tumors in DCE-MRI [26, 119, 282, 288]. In addition to this, the employed pharmacokinetic modeling may also be a source of uncertainty, which will be further discussed in the upcoming section. Another major source of uncertainty in DCE-MRI is related to motion artifacts, such as patient motion dur-

2

ing acquisition, involuntary rectal motion or bladder filling [39, 283, 288], which can affect the quality of the images.

Several sources of uncertainty, such as the ones related to the specificity and sensitivity of each one of the employed acquisitions, may be minimized by combining the different imaging modalities. Many studies have demonstrated that the combination of the different acquisitions can improve detection, diagnosis and staging [16, 25, 51, 124]. Nevertheless, there are others sources of uncertainties that cannot be avoided, and their effect needs to be analyzed and explored.

In this first step of the radiotherapy planning pipeline, clinical researchers are interested in exploring the involved multi-varied, multi-dimensional data effectively. Patterns and trends within these data need to be identified and relationships among data dimensions require to be discovered, especially when noisy, complex data are present. The current method of exploration of this data is a rudimentary slice-based technique, where all slices of the different imaging volumes are visually inspected and the relationships between different acquisitions are mentally performed. This exploratory method provides limited insight and is time-consuming – hence, new techniques need to be investigated and designed.

2.4.2. Features Derivation from Medical Images

As mentioned in the previous section, it is common practice to derive several features from medical images, in order to obtain information about patient-specific tissue characteristics. In the present case, we focus on the derivation of features from DCE-MRI data.

A way of measuring tissue properties from DCE-MRI data is to extract several measures from the enhancement curves, presented in Figure 2.7. In this case, several measures can be considered: the *wash-in velocity of enhancement*, which is the flow of the contrast agent, entering the blood plasma; the *wash-out velocity of enhancement*, which the flow of the contrast agent, being diffused out of the blood plasma into the extracellular extravascular space; the *peak or maximum enhancement intensity*; the *time* between the beginning of the acquisition and the peak of contrast; and the *area under the curve (AUC)*, to relatively quantify the enhancement of contrast agent over time. Additional measures can also be calculated.

Another way of measuring tissue properties from DCE-MRI data is to employ one of the established pharmacokinetic (PK) models [153, 257, 271]. These models are employed to derive per voxel a set of PK parameters, which describe the distribution of the contrast agent inside the tissue and are indicative of tissue characteristics [257].

One of the most common and simple PK models is the Extended Tofts model (ETM) [271], which is a mathematical model that quantifies per voxel the exchange of contrast agent between the vessel and the extracellular extravascular space, as shown in Figure 2.8. In this case, four parameters are derived: K^{trans} (min^{-1}), which denotes the concentration of contrast agent, transferring from the blood plasma into the extracellular extravascular space; k_{ep} (min^{-1}), which denotes the concentration of contrast agent, transferring from the extracellular extravascular space to the blood plasma; V_e (mL/100mL of tissue; %), which denotes the extravascular extracellular space volume per unit of volume of tissue; and V_p (mL/100mL of tissue; %), which is the blood plasma volume per unit of volume of tissue.

Apart from the Extended Tofts Models, there are more complex mathematical mod-
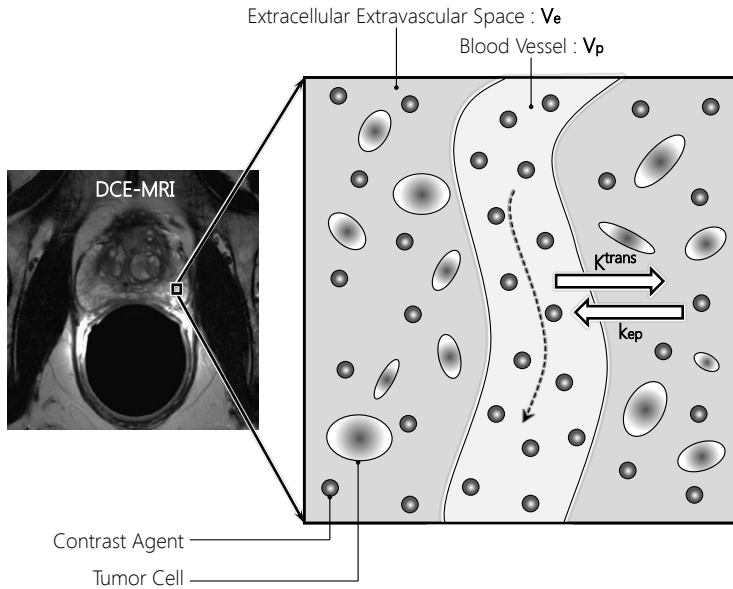
Figure 2.8: The Extended Tofts Model, as an example of pharmacokinetic (PK) modeling applied on DCE-MRI data. *Figure inspired from* [271].

els [61, 152, 197, 283]. Each of these models considers a number of assumptions or crisp choices, depending on which, the number or values of the obtained parameters, or even the parameters themselves may differ.

In general, most uncertainties and errors in DCE-MRI data can be associated to the general lack of standardized acquisition and modeling protocols, which results into problems in the repeatability and the reproducibility of the computed parameters [244, 288]. To name a few, a high temporal resolution is required, so that the enhancement curves can be described completely and significant errors in the estimation of the pharmacokinetic parameters during the fitting procedure can be avoided. However, high temporal resolution is conflicting with high spatial resolution and different trade-off strategies in spatiotemporal resolution can be followed [39, 113, 151, 288]. Another source of uncertainty can originate from the choice of the arterial input function (AIF), such as reference-based AIFs, population-based AIFs and AIFs measured from the DCE-MRI magnitude or phase signal [61, 152, 197, 283, 314]. Quantification of the AIF is usually demanding, because it depends on the saturation, the blood flow and the eventual inhomogeneity of the magnetic field [39]. The choice can be either based on literature knowledge or on calculations, which might both introduce uncertainty [197]. Other assumptions in the pharmacokinetic modeling approach include assumptions on blood hematocrit values, contrast agent relaxivity, models that lead to the oversimplification of the contrast agent injection and the determination of blood plasma flow rates [197], which can also induce uncertainty to the calculation of the DCE-MRI derived parameter maps [39, 63, 68, 197].

2
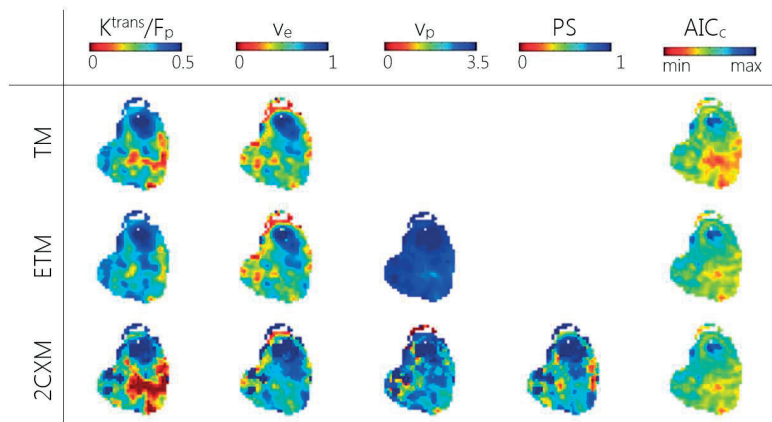
Figure 2.9: Standard layout for the inspection of pharmacokinetic (PK) parameter maps (columns) of a tumor, modeled using three different PK models (rows). Some parameters of the two-Compartmental Exchange model (2CXM) are not involved in the Tofts model (TM) and Extended Tofts model (ETM), and are missing [257]. *Figure courtesy of J. F. Kallehauge, Aarhus University Hospital.*

The impact of the modeling choice, as well as the suitability of each model for every patient is not fully comprehended [139]. For clinical researchers working on PK modeling, it is valuable to investigate how the derived PK parameters behave with different modeling choices. To this end, they currently employ a slice-based technique in their exploration and analysis, similar to the one described in the previous section, which is sub-optimal in providing insight and tedious. An example of this method is presented in Figure 2.9.

2.4.3. Tumor Tissue Characterization

For many years, it has been common practice to consider tumors homogeneous masses. In reality, tumors are heterogeneous tissues, enclosing multiple regions with distinct characteristics. Incorporating patient-specific intra-tumor tissue information into radiotherapy planning can play an important role in tumor diagnosis and in designing more effective treatment strategies, where distinct intra-tumor tissues are irradiated with adequately selected radiation doses [106, 107].

Currently, the only way to investigate intra-tumor tissue heterogeneity is to study data acquired from invasive procedures, such as from biopsies or from the inspection of histopathological slices. To perform a non-invasive in-vivo identification and exploration of intra-tumor tissues, clinical researchers need to associate histopathological findings, such as Gleason scores [82] with features derived from co-registered imaging data, such as perfusion from DCE-MRI, or diffusion from DW-MRI data.

The exploration and analysis of the characteristics of distinct intra-tumor regions is not a trivial subject, but it is an essential component of the radiotherapy planning pipeline. The intra-tumor tissue classification exploration can be additionally used by clinical researchers to evaluate and assess supplementary clinical data that are often used as reference or as a

means of diagnosis and outcome prediction [106, 107]. These clinical reference data may be, for instance, data from risk prediction models [106]. Furthermore, the exploration of tumor tissue characterization could be of particular interest for researchers developing classification algorithms, to aid the design of classifiers that can differentiate between distinct tissue types, as well as to understand the behavior of such classifiers.

2.4.4. Segmentation

The segmentation step of the radiotherapy planning pipeline aims at constructing models of the prostate and the organs at risk in the area involved, such as the rectum and the bladder. Segmentation of the bones and the body surface is also important, as these structures provide information for anatomical reference.

The results of this step are highly dependent on the selected segmentation method and the eventual parameter settings of the chosen algorithms. Segmentation can be either performed manually, semi-automatically, or automatically. In manual segmentation, medical experts inspect the imaging slices one-by-one and delineate the structures of interest. This procedure can be time consuming, but it can also create inter-observer variability, which may pose critical questions concerning the reproducibility and accuracy of the procedure [24, 59]. Therefore, automatic methods are often preferred, with a lot of effort being invested in the development of robust algorithms [43, 76, 148, 239]. However, opting for an automatic algorithm may also come with several limitations and challenges.

The segmentation of structures in the pelvic area involves organs with a large variability in shape, size and imaging intensity [59]. The noise, inappropriate resolution or geometrical distortions induced by the MR scanner can be detrimental when we need to segment smaller prostate volumes [148]. In addition to that, it has been documented that prostate volumes can vary by an average of ±10%, the bladder and the rectal volumes can vary by ±30%, while the seminal vesicles can vary up to 100% [232]. The variation of the prostate center of mass has been found to vary less than 1 mm in the left-right direction, but up to 1 cm in the anterior-posterior and the superior-inferior direction [232].

Organ motion is another important factor that needs to be taken into account in the radiation therapy planning pipeline [145, 148, 157, 199, 239, 270]. The prostate and the seminal vesicles move not only relatively to the bony structures, but also within the pelvic region, due to their position close to the bladder and the rectum. For example, the distension of the adjacent organs, due to filling, may result into the displacement of the prostate [62, 232].

In this step of the radiotherapy planning pipeline, the focus of clinical researchers is on creating robust segmentation algorithms. Still, their implemented methods might not be able to account for all cases and may perform sub-optimally. In such cases, it is required to predict anatomic regions and circumstances under which these methods are more prone to inaccuracies. The ultimate goal, in this case, would be to determine how to improve the segmentation process, namely the selected methods and their settings.

2.4.5. Radiotherapy Dose Planning

After segmentation, the actual radiotherapy plans is designed. A simulation of the treatment planning is performed in dedicated software, as described in the previous sections. This software takes care of maximizing the effect of the dose in the tumor areas, while minimiz-

ing the toxicity in the adjacent healthy organs. Our focus in this work is not on the actual dose planning procedure, but on the incorporation of the variability in planning, which can be induced as a result of adjustments or choices in the previous steps of the pipeline.

All previously mentioned steps introduce their own specific uncertainties. In particular, the results of each step also highly depend on the employed methods and their assumptions, or chosen algorithms and their respective parameter settings. It is valuable for clinical researchers to understand and evaluate the sensitivity of the treatment plan to different assumptions and parameterizations, from the previous steps of the pipeline. In that way, they can assess whether different choices in the planning pipeline can have an impact on the final treatment planning and be aware of this, when designing their treatment plans.

2.4.6. Tumor Control Probability Modeling

Clinical practice aims at choosing the most effective radiotherapy strategy, based on clinical knowledge and guidelines. However, clinical research aims at thoroughly evaluating all possible treatment alternatives. These take into account several points, such as dose escalation, uniform or non-uniform tumor irradiation, the amount of the received dose and eventual fractionation of the treatment. To simulate and evaluate the effects of these different treatment strategies, clinical researchers need radiobiological modeling. This involves two aspects: Tumor Control Probability (TCP) modeling [302] and Normal Tissue Complication Probability (NTCP) modeling [172]. TCP models are statistical models that quantify the probability that a tumor is effectively controlled, i.e. treated, given a specific radiation dose. NTCP models are statistical models that quantify the probability that normal tissue around the tumor is harmed, given a specific radiation dose. Figure 2.10 presents an example of TCP/NTCP modeling prediction. In the present work, only the TCP modeling part will be addressed.

Conventional TCP models are linear regression models, based only on statistical and literature knowledge. Recently, novel TCP models [48] started incorporating additional information from imaging modalities, such as DW-MRI [48]. In this way, patient-specific properties of tumor tissues are included, improving the radiobiological accuracy of TCP modeling. As a consequence, these image-based TCP models are subject to uncertainties, inherent in their employed imaging modalities, with significant impact on the outcome. In addition to this, the modeling step includes, amongst all, parameter assumptions, which are not always crisp choices [97, 258, 302]. The parameter sensitivity of the model is also an aspect, which needs to be taken into consideration when predicting the outcome of a specific radiotherapy strategy. Currently, these two aspects are not incorporated into clinical research, neglecting significant information for the outcome prediction.

2.4.7. Supplementary Steps of the Radiotherapy Planning Pipeline

In the presented radiotherapy planning pipeline, an additional step has been implicitly considered – registration. MRI and CT data need to be registered in order to combine functional and anatomical information, while multi-parametric MR images also need to be registered to ensure the same coordinate space for each image acquisition. In addition to that, the histopathological data that are used in clinical research, also need to be registered to all other available data. Although we will not address registration in the present work, we need to underline the importance of this process. The literature on registration methods is vast [321]
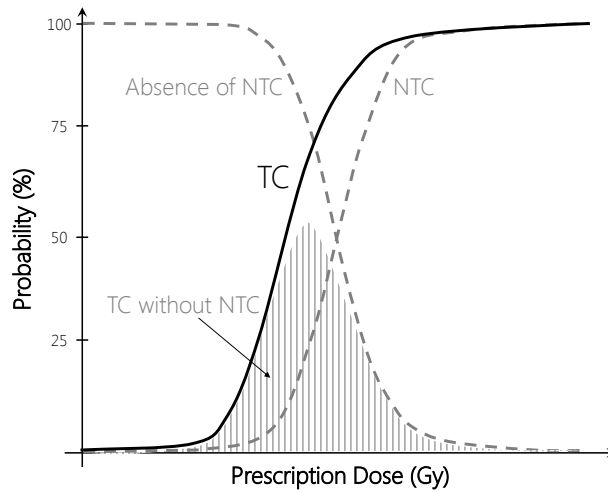
Figure 2.10: The concept of Tumor Control (TC) Probability [302] and Normal Tissue Complication (NTC) Probability [172].

and different algorithms can be employed, each with different strengths and implications.

Registration may be accompanied by uncertainty, which may have an impact on the remainder of the radiotherapy planning pipeline, as well. Uncertainty in image registration is primarily related to the inherent characteristics of the different imaging modalities that are co-registered. In addition to this, different registration algorithms may bring different types of uncertainty, related to the localization accuracy or robustness [141]. In particular, the use of non-rigid registration requires the selection of parameters, which can yield results with large variability [231]. In other cases, the lack of objective ground truth in the validation of registration creates the need for manual registrations by experts, which introduces uncertainty that is related to inter-observer variability.

3

Visualization: State of the Art

In this chapter, we identify and present the research field, to which the current work belongs, providing the necessary technical background and setting the context of this dissertation. For this, we introduce the field of Visualization and, in particular, the field of Visual Analytics to the reader. Basic terminology or techniques, which are vastly used throughout the entire dissertation, are also presented. Additionally, we provide a general overview of state-of-the-art visualization systems and solutions dealing with multi-dimensional or multi-varied data, and with uncertainty. Other related work – more specific to methods and techniques presented in the core chapters of the dissertation – will be discussed in detail within each one of the following chapters.

3.1. Visualization

As already mentioned in Chapter 1, the field of visualization can provide solutions for researchers, investigating the different steps of the radiotherapy planning pipeline. These solutions aim at providing better understanding and deeper insight into the complex data and processes, involved in radiotherapy planning.

According to Cliff Pickover, Visualization is *the art and science of making the unseen workings of nature visible* [204]. In essence, Visualization is the scientific field that employs elements from the field of computer graphics to create meaningful visual representations of the data [46], which can facilitate interpretation and can provide access to – or deepen – the understanding of users, taking advantage of the human vision and cognition.

Visualization can be described as a tool to enhance human cognitive capabilities – and not a set of automatic computational methods, which aims at replacing humans [182]. Therefore, Visualization can be particularly suitable and powerful, in cases where users know little about their data and processes [182], but need to explore and discover new knowledge within these, generate or confirm hypotheses, and make decisions.

Card et al. [46] describe the Visualization process, depicted in Figure 3.1, as a number of steps which enable visual sense making. Initially, raw data is collected and transformed into data tables, which are derived data that are easier to manipulate and understand. Then, these data are turn into visual structures, with the aid of visual mappings. Subsequently, a transformation is employed to provide perspective on the data, in a view that is presented to the user. The user, finally, interprets the view and gains a deeper insight into the data. Perception, cognition and interaction play a very important factor in the explorative process, as users navigate through the data and interact with them, in order to understand the involved information.
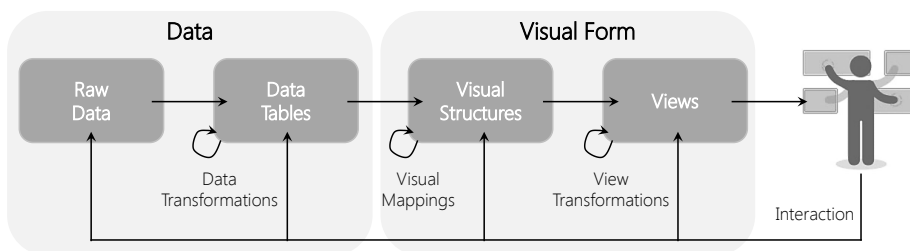
Figure 3.1: Overview of the visualization process, described by Card et al. [46] as the mapping of data to visual forms that support human interaction for visual sense making.

Visualization can be applied to many different application fields, among which to processes and data from the medical field. In this case, we refer to it as Medical Visualization [217]. A common taxonomy of Visualization is done, with respect to the characteristics of the data. The term *Scientific Visualization* refers to the discipline that engages data with a geometric structure or inherent spatial information, which are typically related to scientific applications, such as medical data [277]. *Information Visualization* refers to the discipline

that deals with abstract, non-physical data [277, 299]. In addition to these two fields, there is the distinct discipline of *Visual Analytics* [58], which is discussed in the following section.

3.2. Visual Analytics

Visual Analytics is one of the three main flavors of Visualization. It is the result of integrating concepts from Scientific and Information Visualization, with other disciplines and analytical processes, such as data mining or statistics. This integration usually happens in a highly interactive environment, where analytical sense making and reasoning are supported. Thomas and Cook [58] describe Visual Analytics as *the science of analytical reasoning facilitated by interactive visual interfaces*, while Keim et al. [144] state that:

> *Visual Analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets.*

The human factor, or human-in-the-loop concept, is an essential component of Visual Analytics solutions, which focus on integrating and combining the strengths of human abilities for sense and decision making, with semi-automated methods for data analysis [182, 276]. The human is deeply involved in the process, while steering the exploration through visualization and interaction. The process of Visual Analytics is depicted in Figure 3.2.

In the present dissertation, Visual Analytics will be employed to provide insight into the data, information and knowledge at each distinct step of the radiotherapy planning pipeline, through the interactive exploration and analysis of the involved data and processes.

3.3. Fundamental Techniques and Methods

In the coming chapters, we will present Visualization solutions for the exploration and analysis of each step of the radiotherapy planning pipeline. All of them incorporate several design methods and interaction techniques, which are briefly described below:

- The *Information Seeking Mantra* [249] can be summarized as: *Overview first – Zoom and filter – Details-on-demand*. According to this, users should first have an overview on the data, which can give a global impression about the information within the data. Then, users should be able to zoom in and filter the data, in order to obtain more detailed information. At the end, all detailed information should be shown on demand. This process should be enabled through interaction.

- The *Visual Analytics Seeking Mantra* [142] is an adaptation of the Information Seeking Mantra, adapted to fit within the field of Visual Analytics: *Analyze first – Show the*
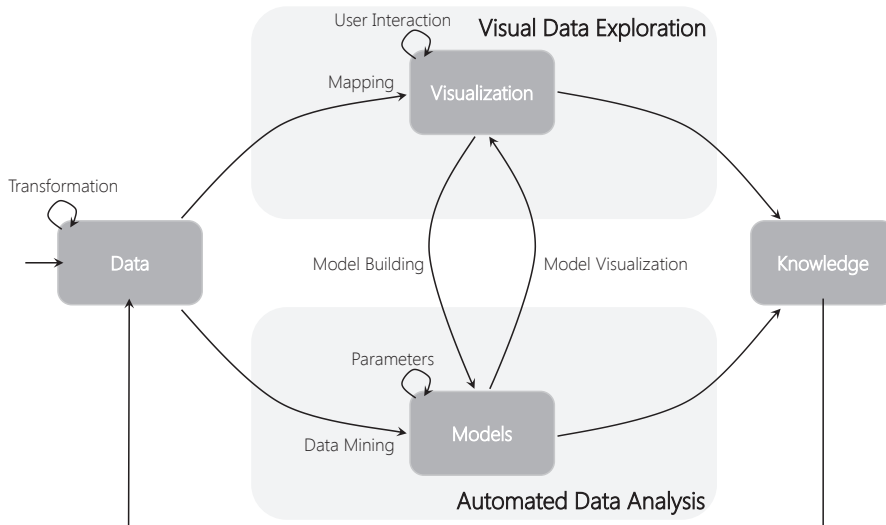
Figure 3.2: Overview of the Visual Analytics process, described by Keim et al. [142]. In Visual Analytics, the strengths of visual data exploration and automated data analysis are combined with the data, visualization and models, to obtain knowledge within a highly interactive environment.

important – Zoom, filter and analyze further – Details-on-demand. In contrast to the original Mantra, the data need to be initially visually analyzed to show the most interesting or relevant aspects.

- The *Data-Users-Tasks Design Triangle* [178] is a concept that dictates that, during the design of a Visual Analytics system, three aspects should be taken into consideration: the data involved in the process, the users for which the system is being designed and their respective needs, and the tasks that the system is required to fulfill.

- *Multiple (Coordinated) Views* [294] is a widely used design method in Visual Analytics. This refers to employing multiple views that provide different viewpoints on the data, in order to observe the data and their in-between relations though different perspectives. Multiple Views are usually combined with Brushing and Linking, described below, to facilitate the identification of relationships between data.

- *Brushing and Linking (B/L)* [18, 40, 143] is a concept that involves selecting one or several interesting items in one view, and highlighting corresponding items in another. As stated by Keim, this method is meant to *overcome the shortcomings of single techniques*, and provides more information than the exploration of individual views.

- The idea behind *Focus + Context (F+C)* [46] is to present items at different levels of detail. More interesting or relevant items are presented with more detail, while less important items are presented with less detail, but are retained in the view, in order to provide context for a better understanding and insight.

- The notion of the *Overview + Detail* technique [54] is related to the combined use of Multiple Views and Focus + Context. Here, at least two views are presented to the users: one with a rough overview on the entire visualization space, and one with a detailed view of a smaller portion of the space.

3.4. Related Work

The field of Visual Analytics has addressed numerous fields of application in the past. The exploration and visual analysis of the data involved in the radiotherapy pipeline, however, has not been explicitly tackled before. Each step of this pipeline requires different handling, depending on the involved data and processes to be explored, on the tasks that have to be fulfilled, and on the intended users and their needs.

During all steps of the pipeline, the topics revolve around the exploration and visual analysis of multi-dimensional, multi-varied data with complex relationships, or the incorporation of uncertainty into the designed visualizations. Therefore, in this section, we will address previous related work, with respect to these two categories. In the other chapters, we will deepen into literature that is more specific to each step of the pipeline.

3.4.1. Visualizing Multi-Dimensional, Multi-Varied Data

The literature that addresses the visualization of multi-dimensional and complex data is vast [143]. In this section, we will focus on previous related work that is relevant to ours. Initially, we will present traditional, widely used techniques from the field of Information Visualization and, subsequently, a number of Visual Analytics system designs and methods that support the exploration and analysis of multi-variate, multi-dimensional complex data. At the end of this section, we go one step beyond the exploration and understanding of data, by discussing the use of Visual Analytics to understand dimensionality reduction and clustering processes.

Multi-Dimensional Information Visualization Representations

Among all distinct techniques that have been designed in the field of Information Visualization [143], two have been extensively employed: the Scatterplot Matrix and the Parallel Coordinates Plot.

The Scatterplot Matrix, or SPLOM, is a matrix-like configuration of all pairwise scatterplot views of the dimensions of the data [50, 110, 280]. It is based on the concept of small multiples by Tufte [279]. Being easy to comprehend and interpret, it has become a powerful representation of multi-dimensional data and it can be used as a good starting point for the exploration of relationships within the data. However, with an increasing number of dimensions, the number of scatterplots – and subsequently, the demand for screen space – increases rapidly. In addition to that, it may cause significant cognitive load, as the part of the SPLOM below the diagonal is redundantly encoding the same information, as the part above the diagonal. Without interaction, SPLOM may also require from the user to use their memory, in order to identify and analyze relationships within the data. Furthermore, each scatterplot of the SPLOM is able to show only pairwise relations within the data. An example of the use of the this representation is depicted in Figure 3.3 - a.

Parallel Coordinates Plot, or PCPs, is another widely known and used representation for high-dimensional data [128]. In this case, multiple data dimensions are mapped one-by-one to a number of parallel vertical axes. Each multi-dimensional data object is mapped to a polyline that intersects the axes, connecting the scalar values of every dimension. PCPs efficiently display in a single view all 2D projections of adjacent data dimensions [126, 129, 303], enabling the identification of relations and the detection of data patterns or trends, especially with the help of interaction [112, 251] such as brushing [111] or reordering [13, 201, 285]. A limitation of PCPs is that they might suffer from clutter due to overplotting [112]. This causes problems in the exploration and interpretation, especially in high density data. Still, scalability is not only an issue with respect to the number of data points; it is also a matter of an acceptable number of axes. The order of the latter is also important as it implicitly determines the relationships that can be identified and explored. An example of PCPs is presented in Figure 3.3 - b.

(a)   (b)
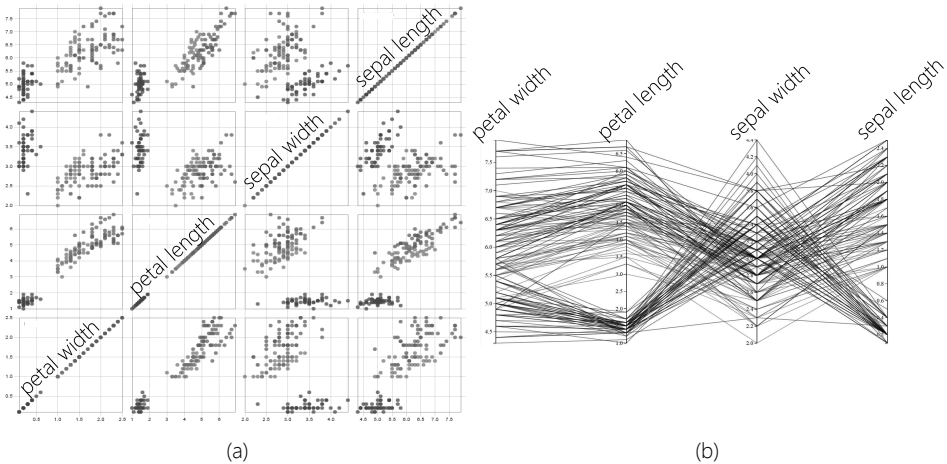
Figure 3.3: An example of visualizing the four dimensions of the *iris dataset* [86], using (a) a Scatterplot Matrix (SPLOM) and (b) the Parallel Coordinates Plot (PCPs). The two representations were created, using D3.js [29].

Visual Analytics for the Exploration of Multi-Dimensional Data

In the field of Visual Analytics, there are many solutions and systems, which deal with the exploration of multi-dimensional data. In this section, we select the most relevant to the content of this dissertation and we present them, dividing them them into four main categories. We present solutions where *only two data dimensions* are visualized after being selected by the user; solutions where *linear projections* of the data are visualized; solutions where *non-linear projections* of the data are visualized; and solutions visualizing *all data dimensions* through projections and *providing understanding* in the employed projections. At the end of this section, we provide a summarized view and connection to our own work.

■ *Selection and visualization of two data dimensions.* The exploration of multi-dimensional data has been tackled multiple times before, by allowing the user to select two data dimensions of particular interest and visualizing only these. Early examples include the Grand Tour [15] and the XmdvTool [295].

More recently, WEAVE [103] was introduced as an environment for the interactive visualization of multi-dimensional data. It employed linking and brushing between custom three-dimensional visualizations and multi-dimensional statistical representations. The WEAVE system was an improvement to mere visualizations of scientific data and to existing statistical visualization packages, by allowing the user to quickly compare and to correlate variables, while also visualizing the spatial data. However, the system did not support the simultaneous incorporation of data from different modalities.

Coto et al. proposed the MammoExplorer system [60], to enable the exploration and analysis of breast DCE-MRI data. In this approach, segmentation techniques were linked to visualization, in an interactive environment. Scatterplots were used to show the enhancement of the contrast agent in DCE-MRI data. Combined with two-dimensional and three-dimensional anatomic representations of the data through brushing and linking, they enabled the identification and characterization of breast lesions. For conveying temporal information, the authors proposed using multiple scatterplots, with all the limitations that such an approach could entail. Color encoding was employed to highlight brushed areas, while volume rendering was performed using Maximum Intensity Projection (MIP) or Composite Ray Casting (CRC). Focus+Context was employed to selectively visualize the brushed data, while keeping the rest of the context visible with less details.

The work of Hennemuth et al. [116] presented a method for the exploration and visualization of the perfusion and late enhancement in myocardial tissue MRI data, for the detection of distortions due to motion. The authors implemented an approach to inspect the enhancement curves and their derived parameter distributions, in different areas of the myocardium. For this, they enabled users to segment regions by thresholding the parameter space. They also provided a functionality to compare regions segmented with different settings. Brushing and linking was a necessary interactive component, also in this work.

Another well-known system, which was employed for the interactive exploration and visual analysis of various multi-dimensional and time-varying data, was SimVis [72, 73, 74, 158, 179]. SimVis was using multiple linked views, interactive feature derivation and selection through smooth brushing, and Focus+ Context visualizations, to visualize, explore and analyze data from a multitude of domains, from engines to hurricane data, and from air flow to medical datasets, such as brain perfusion data.

■ *Visualization of linear data projections.* Up to now, all presented systems faced issues, either with respect to the acceptable number of display views or with respect to the displayed dimensions of the data, which were limited to two, by user selection. The latter started being addressed more suitably with the incorporation of linear projection methods [154] in Visual Analytics approaches.

The work of Oeltze et al. [189] enabled the exploration of the correlations and relations between several features and parameters of perfusion data. In their approach, they initially extracted time-intensity curves (TICs) that characterize the amount of contrast-agent enhancement at each voxel in the perfusion imaging data. From these curves, they derived

parameters that could be used as indicators in the diagnosis of tumors. To check the eventual correlation between these parameters, they employed Principle Component Analysis (PCA) [137] to reduce the dimensionality of their parameter space. After that, they used multiple linked views to enable the user to explore and analyze complex or multi-dimensional features within their data. This work allowed the localization of specific characteristics of the parameter space in the anatomic and temporal domain, it enabled a multi-variate analysis of the parameter space and it facilitated the local exploration of the data.

In a comparable approach, Fang et al. [84] proposed a method for the visualization and exploration of time-varying medical image datasets. In their work, a time-activity curve (TAC) was extracted from each voxel of the time-varying dataset, similarly to the TICs of Oeltze et al. [189]. Using linear projections techniques, all the voxels with similar TACs were retrieved and meaningful features that form these similarities were brought forward. With this approach, they managed to distinguish between tissues with different features, such as heart, liver and lung tissues.

■ *Visualization of non-linear data projections.* The frameworks, which were discussed up to this point, were able to visualize either user-selected pairs of dimensions or two-dimensional linear projections of the multi-dimensional data. This entails the limitation of providing a restricted view on the entire data, as well as assuming that the structure of the involved feature space is characterized by linearity.

To address these points, Blaas et al. presented a new approach to handle large multi-field data, showing both the anatomical domain and the high-dimensional feature space of the data in an interactive environment, with multiple linked views [23]. They employed this approach to investigate its utility in data segmentation. To this end, they demonstrated its classification functionality by finding clusters and their relations within the data, as well as its capability for selective data visualization based on feature values. Interaction, brushing and linking, feature selection and pattern analysis were used again to enhance the data exploration process.

To extend this work, Steenwijk et al. [261] proposed a Visual Analytics system for cohort studies. This approach enabled inter-patient studies, where users could easily easy explore multi-modal and multi-timepoint parameters across patients by extracting and visualizing parameters of interest. In their approach, the investigation of a full medical cohort was made possible in a highly interactive framework for the visual and statistical analysis of the involved data.

Recently, the Cytosplore application of Höllt et al. [122] was introduced to provide understanding into the cellular composition of the immune system and the properties of the cells, as derived from mass cytometry data. It consisted of an interactive environment, which was built on the basis of dimensionality reduction to represent a high-dimensional feature space of cells. Multiple linked views on the data, at different level of details, facilitated the exploration of the exploding feature space of mass cytometry data. The interface of Cytosplore is presented in Figure 3.4.

■ *Visualization of data projections, providing insight into the employed projection techniques.* The frameworks, which were discussed up to now, aimed only at the exploration and understanding of the involved data. There is a number of approaches that aims – additionally, or
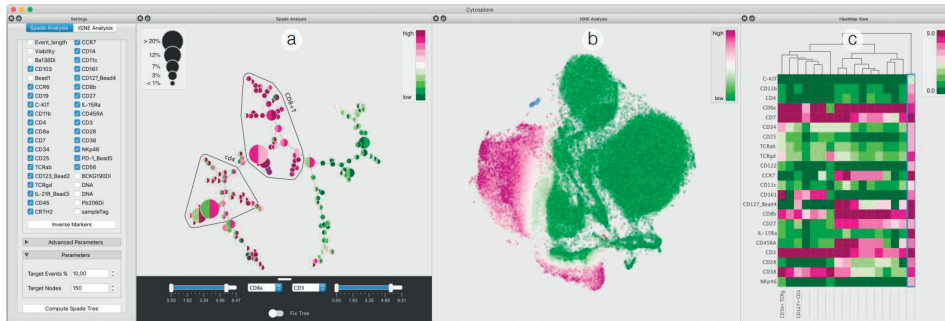
Figure 3.4: An example of Cytosplore, as proposed by Höllt et al. [122], with its four main components: (a) overview, (b) embedding and (c) heatmap. *Figure courtesy of Thomas Höllt, Delft University of Technology.*

instead – at providing insight and means for direct feedback on an employed projection or clustering techniques.

Jeong et al. [133] proposed a system for the evaluation and understanding of the results of Principal Component Analysis (PCA) [137]. Although PCA is widely used, many times it becomes difficult to understand why a specific result is achieved. In order to aid the user to understand and use PCA appropriately, the iPCA system was developed. It consisted of multiple interactive coordinated views for the visualization of multi-variate data and their correlations in three spaces, i.e., the original data space, the eigenspace and the projected space.

The iVisClassifier [52] was another Visual Analytics system for obtaining insight into clustered data and the classification process performed with linear discriminant analysis (LDA). It enabled users to explore high-dimensional data, using a combination of different views. These included parallel coordinate plots, scatterplots and heat maps for an overview on the cluster relationships in both high and low dimensional domains. In this way, a new interactive interpretation of LDA and its outcomes was facilitated.

DimStiller [127] was another tool for the visual analysis and exploration of dimensionality reduction. It used a set of abstractions to structure and navigate through the complex processes of dimensionality reduction, by providing guidance in the feature space, through interaction.

Poco et al. [207] proposed a system where the user could modify, control and improve 2D or 3D projections of complex diffusion tensor imaging feature spaces. The system was targeting the exploration of large collections of fiber tracts for diagnosis and for understanding brain functions. To do so, users could interact with or modify and improve the generated Local Affine Multidimensional projections (LAMP). Views on the fiber space and the projections of the feature space were presented to the user and linked bi-directionally, to enhance the exploration of the data.

Seo et al. [247] designed a system, where multiple linked views were employed to analyze clustering results in genome data. This work focused on the analysis of hierarchical clustering, which may be complex and difficult to understand and analyze. With the proposed

system, the user was presented with several views and features to obtain an overview of the data and the features that contribute to the clustering outcome. The user could also control the clustering process and interact with the generated outcomes.

Finally, Turkay et al. [281] presented a visual analysis method to analyze how clustering results change in time. The structural and quality changes of the clusters were encoded in a Parallel Coordinates view, while another view visually summarized properties of the clusters over time. These two components were linked through brushing and enabled users to enhance the understanding of the temporal evolution of clusters.

■ *Connection to our work.* All previously presented systems and solutions aim at dealing with the representation and exploration of multi-dimensional data. In the first category, the user needed to select two data dimensions and visualize them. However, this kind of approaches would not be sufficient for the data at the different steps of the radiotherapy planning pipeline, because of the implicated dimensionality and complexity. Furthermore, such approaches would require prior knowledge about which dimensions are more significant, which is not feasible in our case. The second category included solutions employing linear projections to represent and visualize the data. As already explained, these solutions make specific assumptions about linearity in the structure of the data, which is not guaranteed in the case of the complex data of intra-tumor tissue characteristics. Therefore, these two categories are not suitable for the exploration and analysis of the data involved in radiotherapy planning.

Closer to the approaches that we will introduce in the upcoming chapters are the solutions employing non-linear projection techniques, as well as the solutions that aim at understanding the underlying structure of the data and at providing means for feedback on the result of the employed techniques. Yet, all previous related work is not fully applicable to our application domain and to the specific radiotherapy pipeline steps, as we will describe in detail in the following chapters.

3.4.2. Uncertainty Visualization

Uncertainty visualization is a relatively new and popular domain [31, 136]. Although the importance of raising awareness on uncertainty information and its influence on the data has been stressed multiple times [30, 105], in many cases this concept is still overlooked with serious implications [30, 167]. For example, in the present case of the radiotherapy planning pipeline, the precision and accuracy of the outcome of the radiotherapy procedure can have consequences on the treatment outcome. Hence, it is crucial to incorporate information about eventual sources of uncertainty into the radiotherapy pipeline.

Uncertainty visualization is a difficult and demanding task. Often, uncertainty comes as an additional channel of information, which needs to be visualized on top of other underlying data. This can increase the complexity of the view and the visual overload, decreasing the understanding of the user about the original data and the implicated uncertainty. When approaching an uncertainty visualization problem, the choice of the design methods depends on the nature of the uncertainty data itself, on the uncertainty data type and on the already employed visualizations of the remainder of the data [104]. This design choice is often not easy, as uncertainty tends to dominate over certainty in the data [36], which results into visualizations where the underlying data are distorted or obscured, while uncertainty is em-

phasized [49, 104].

Some of the most common design approaches for the visualization of uncertainty, which have been vastly used in the past [36, 27, 195], are summarized in the scheme of Figure 3.5. All of them entail a number of limitations. Using free graphical variables is suitable in cases, where a representation within the already existing geometry of a structure is required. The choice of which variable should be used to encode uncertainty is not easy, and the combination of these attributes with the visualization for the other information within the data must be done in moderation and with consideration. Using additional graphical objects needs sensible and careful management, as the underlying data should not be overshadowed by the uncertainty visualization. Opting for this approach might also mean that the user faces the chance of a cluttered visualization, where visual overload cannot be avoided [185]. Employing interactive representations for uncertainty visualization requires a lot of interaction with the interface, which may be distractive for the user, if not properly designed. Animations can also be distractive, as they are known for causing visual fatigue to the user [168].

In the following two sections, we will present previously designed methods for visualizing uncertainty. We separate them into uncertainty incorporated in Scientific Visualization solutions, and uncertainty in Information Visualization or Visual Analytics systems.

Uncertainty in Scientific Visualization Systems

The literature on solutions for the Scientific Visualization of uncertainty is vast. In the current section, we introduce the most relevant solutions, using the categorization presented in Figure 3.5. In the upcoming chapters, several of the described methods and visual cues will be used as inspiration for the visualization of the uncertainty, implicated in the different steps of the radiotherapy pipeline.

■ *Using free graphical variables.* Free graphical objects, such as color, brightness, fuzziness,

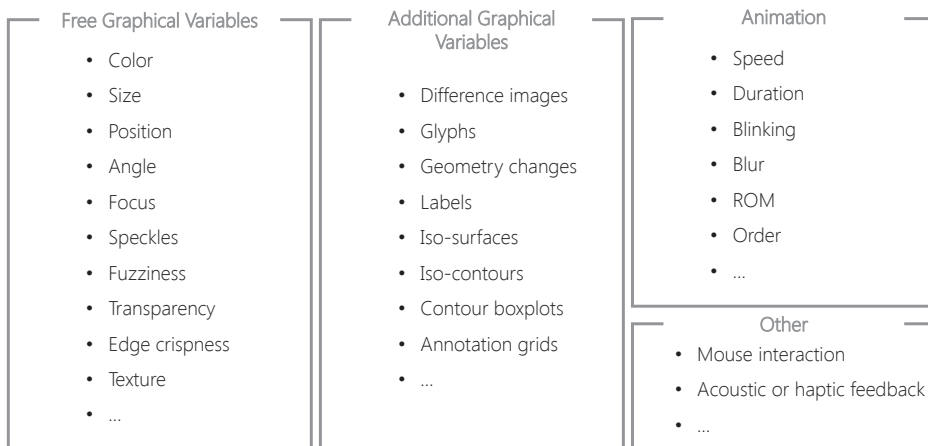| Free Graphical Variables	Additional Graphical Variables	Animation
• Color	• Difference images	• Speed
• Size	• Glyphs	• Duration
• Position	• Geometry changes	• Blinking
• Angle	• Labels	• Blur
• Focus	• Iso-surfaces	• ROM
• Speckles	• Iso-contours	• Order
• Fuzziness	• Contour boxplots	• ...
• Transparency	• Annotation grids	**Other**
• Edge crispness	• ...	• Mouse interaction
• Texture		• Acoustic or haptic feedback
• ...		• ...

Figure 3.5: A categorization of different methods, which have been employed in literature for the visualization of uncertainty.

texture, or combinations of the previous, have been employed for the visualization of uncertainty in a multitude of application fields. However, the available non-conflicting free graphical objects are often limited, limiting also the potential to encode simultaneously several sources of uncertainty.

Color and texture are the most frequently encountered techniques. The concept behind the use of these methods is that color is an easy and effective attribute, which can be used as an additional parameter, without adding much distraction, if adequately selected. On the other hand, when texture is employed to encode uncertainty, the surface color is available for the visualization of an additional variable. For example, the paper of Botchen et al. [30] presented a texture-based technique for visualizing uncertainties in real-world measured data or inaccuracies in simulated flow data. This was performed either with a generic texture filtering process to improve the perception of uncertainty affected regions or with a user-adjusted color coding of uncertainty. Color and texture was also employed by Rhodes et al. [230], in a surface visualization of the structure under investigation.

In the same category, Twiddy et al. [284] proposed to visualize missing data using some neutral and non-distracting shades of grey, Davis et al. [65] employed an approach where different shades or different levels of hue represented different levels of uncertainty, and Roth et al. [233] proposed to use a two-dimensional discrete scheme that varied in color and the alpha channel value. An extension to a continuous scheme was introduced by Hengl [115] for visualizing uncertainty, using the HSI color space. Finally, Coninx et al. [57] visualized uncertain scalar data fields by combining color encoded with animated, perceptually adapted Perlin noise, while Drapikowski et al. [75] presented a method for depicting uncertainty in surface-based models from fragments of CT or MRI data, with the combined use of a color scale to show the quality of uncertainty present at various locations of the data.

Apart from color and texture, other free graphical objects were used by Djurcilov et al. [71]. In this work, the uncertainty was incorporated directly into volume rendering using grids for one-dimensional data and transfer functions for two-dimensional data. Additionally, post-processing by adding speckles, holes, noise and texture to locations of uncertainty was also proposed. In this case, the simultaneous use of color, transparency, noise, speckles could be conflicting and, hence, difficult for the user to interpret.

A different approach is encountered in the paper of Grigoryan et al. [105], who adopted point-based surfaces that showed the uncertainty of a tumor surface. The surfaces were rendered as a collection of points, where each point was displaced from its original location along the surface normal, proportionally to the uncertainty. Combined with pseudo-coloring and transparency, this method could handle up to six dimensional uncertainties. However, the employed deformation of the shape of the rendered structures may not always be an adequate choice. This is the case for medical applications, where the anatomical shape of structures should be preserved.

■ *Using additional graphical variables.* The literature presented in this section includes methods where additional graphical variables, meaning additional objects, were used. This is a good solution to the limited amount of free graphical variables, but it can entail clutter in the visualization, encumbering the exploration and interpretation of the underlying data.

Pang et al. [195] and Johnson [136] gave an interesting overview of uncertainty visualization techniques that use additional graphical objects. In the paper of Pang et al. the authors

discuss uncertainty visualization methods, such as adding glyphs or geometry, and modifying attributes or geometry. There are more papers proposing the use of additional graphical objects, such as uncertainty glyphs [306] or distorted annotation grids [49] to show the uncertainty along with the data.

Glyphs have also been used in the paper of Lodha et al. [167]. In this work, six methods for visualizing uncertainty in flow visualization with the use of glyphs, envelopes, animations, priority sequencing, twirling baton displays of trace viewpoints and rakes were presented. All those methods were tested on artificial datasets and demonstrated to be valuable for people interested in decision-making, especially when missing or imperfect flow data were involved. In a medical visualization application, Schultz et al. [243] propose a method to model and visualize the probability distribution of fiber directions from diffusion MRI data. These directions were integrated into a new glyph representation, which significantly enhanced the insight into the direction of fibers and their uncertainty in comparison to previously employed glyph designs for the same purpose [138].

Other methods, which involved the calculation and visualization of iso-surfaces in an uncertain field, were discussed by Pothkow et al. [209, 210, 211, 212], Pfaffelmoser et al. [203] and Ferstl et al. [85]. A particularly interesting approach was described in the paper of Whitaker et al. [305], where the authors introduced an abstraction of the boxplot metaphor to encode the variability and quantify the uncertainty in ensembles of contours. The proposed method was called *contour boxplots* and was based on the concept of data depth to generate an ordering of the data, based on the location of each sample within the density function.

■ *Using animations.* Animation has been used for the visualization of fuzzy data, taking advantage of the sensitivity of the human visual system to motion and dynamic changes in a display [95]. A medical application of uncertainty visualization using animations can be found in the paper of Lundström et al. [168]. Here, uncertainty in tissue classification was addressed with the use of animation methods. Uncertainty was animated in the volume rendering of anatomical structures, with the use of transfer functions. Regions that were certain remained static, while uncertain structures changed with time in an animation cycle. This method was, though, evaluated and documented to cause visual fatigue [168].

Uncertainty in Information Visualization and Visual Analytics Systems

Uncertainty has been tackled less often in Information Visualization and Visual Analytics systems. Below, we present the most relevant approaches, diving them with respect to the aforementioned two fields of Visualization.

■ *Uncertainty in Information Visualization.* A traditional way of displaying uncertainty in Information Visualization was to use error bars to convey accuracy in measurements or boxplots to convey information on the value ranges and outliers [213]. Modifications of these techniques have allowed the incorporation of additional statistical information. An example of that was the modified boxplot or violin plot. In this representation, the size or the skew of the boxes have been used to encode additional statistical information [27]. Also, the 2D boxplot has been used for the same purpose [27, 213].

However, the dimensions of the conveyed information were still limited to a small number. In order to solve this issue, the summary boxplot was proposed in the paper of Potter

et al. [214]. In this approach, the summary boxplot incorporated a collection of descriptive statistics that were able to highlight features of the data, related to uncertainty due to errors. The features of summary boxplots were easy to identify and their two-dimensional extension facilitated the comparison of correlations and the ability to highlight variations.

Although the research on uncertainty visualization in univariate data is rich, an extension to multivariate data is encountered less frequently. A method for the visualization of uncertainty in high-dimensional data involved the incorporation of this information into Parallel Coordinate Plots. Barlowe et al. [17] proposed a novel visualization pipeline for exploring interactively multi-variate data and their relationships. Having mentioned the inherent limitations of parallel coordinate plots, it is not difficult to imagine that incorporating uncertainty in this representation can encumber the visualization and exploration of the data and their correlations, and increase the visual overload for the user.

Two more examples of methods for exploring multi-valued data, along with their uncertainty, were proposed by Elmqvist et al. [81] and by Xie et al. [310]. Elmqvist et al. took advantage of the simplicity, familiarity and clarity of scatterplots, expanding their use into a matrix of scatterplots to interactively visualize multi-dimensional data. They called their approach *rolling the dice*, because the transition between different scatterplots in the matrix was performed as animated rotations in 3D space. The users could build queries to refine the visualization and could re-order the animation space to highlight correlations, inaccuracies and differences among them. In the approach by Xie et al., the authors investigated different approaches for including uncertainty in visualizations using different viewing methods, such as Scatterplot Matrices, Parallel Coordinates and glyphs.

■ *Uncertainty in Visual Analytics.* Berger et al. [20] implemented an interactive system for the continuous analysis of a sampled parameter space with respect to multiple target values, using multiple linked views of 2D scatterplots and Parallel Coordinate Plots. Their approach could guide the users to explore the data and to find interesting patterns in them, as well as to detect inaccuracies and uncertainties. A similar approach was presented in the paper of Matkovic et al. [174], where they used multiple linked views, including 2D and 3D scatterplots, histograms, Parallel Coordinates Plots and pie charts. The goal was to visualize and explore data from simulations, which can usually be complicated and their correlations or comparisons can be difficult to identify.

In forecast and meteorology visualization, uncertainty has also been often addressed. Potter et al. [216] first created an interactive and dynamic framework for the visualization of uncertainty in the field of climate modeling and meteorology, by using multiple linked displays. Noodles, a system for the exploration and visual analysis of forecast uncertainties, was later proposed by Sanyal et al. [238]. In this approach, Multiple Coordinated Views were employed to provide different views on uncertainty, such as with ribbons, glyphs, spaghetti plots and colormaps. This visualization was implemented to be used interactively for the detection of the effect of a meteorological event on weather prediction.

In the medical field of application, Saad et al. [236] proposed an interactive tool for the exploration and analysis of probabilistic segmentation results. This approach was meant for the analysis of regions with segmentation uncertainty, using a number of widgets that were integrating the analysis of multivariate probabilistic field data with direct volume rendering. Another paper proposing an interactive visualization of uncertainty in the medical field was published by Brecheisen et al. [33]. In this paper, the authors proposed a visualization tool

that allows the visual exploration of the impact of small parameter variations on the result of fiber tracking in Diffusion Tensor Imaging (DTI) data. The purpose of this paper was the assessment of sensitivity of chosen parameters and the evaluation of intra-patient results.

Additional or more specific systems and methods for uncertainty visualization are going to be addressed in the upcoming chapters. These systems will be more relevant to the topic of each chapter and will be discussed there in detail.

3.5. Evaluation of Visualization Solutions

Evaluation is an important aspect of the design and implementation process of visualization systems. In the field of visualization, many discussions have been conducted on this topic. Recently, many papers providing guidelines and address methods for performing an evaluation have been published [12, 79, 130, 159, 181, 205, 245], also in the field of medical visualization [100, 255].

However, in this work, we have based the evaluation of all our visualization designs and implementations on the paper of Lam et al. [159], which is more general and includes also notions, concepts and methodologies from the other evaluations schemes. Lam et al. categorize evaluation approaches into seven scenarios for Information Visualization systems, which can also be applicable in Scientific Visualization or Visual Analytics systems. These categories include the following seven scenarios:

- Understanding Environments and Work Practices (*EWP*), where feedback is requested from a group of evaluation participants, with or without using the visualization systems, through understanding the work, analysis or information processing practices. This can be performed with field observations or interviews.

- Evaluating Visual Data Analysis and Reasoning (*VDAR*), where a visualization system is assessed on the basis of whether it supports analysis and reasoning and helps to derive knowledge in a domain. This is usually performed with case studies or controlled experiments.

- Evaluating Communication Through Visualization (*CTV*), where a visualization system is evaluated based on whether it can communicate information for teaching/learning or it can help in presenting an idea. This can be performed with controlled experiments or field observations and interviews.

- Evaluating Collaborative Data Analysis (*CDA*), where the value of a visualization system to aid the collaboration of people in a team for data analysis is assessed.

- Evaluating User Performance (*UP*), which aims at measuring how specific features of the visualization system affect the performance of the user. This can be performed with controlled experiments.

- Evaluating User Experience (*UE*), which aims at obtaining feedback and the opinion of the users on the visualization system. Informal evaluations, but also usability tests are employed to measure UP.

- Automated Evaluation of Visualization (*AEV*), which is related to the measurement of the performance of the algorithm.

This categorization has been extended by Isenberg et al. [130] with one additional category: Qualitative Result Inspection (*QRI*), where evaluations are conducted by means of qualitative discussions and assessments of visualization results, not by end users of the visualizations, but by general viewers. The categories that are more relevant for our work are the *EWP*, *VDAR*, *UP* and *UE*. In the upcoming chapters, we will discuss in detail, for each step of the pipeline, how we built and performed our evaluations to assess the value of our designed visualization solutions.

4

Visualization of Multi-Dimensional Data

Above all else, show the data.

Edward R. Tufte (1942 -)

In this chapter, we present a novel technique for the extension of Parallel Coordinate Plots (PCPs), a common choice for the representation and exploration of multi-variate, multi-dimensional data. Among others, this kind of data are encountered throughout the entire radiotherapy planning pipeline. The proposed enhancement aims at improving the display of the data by emphasizing their underlying structure and patterns. Hence, it could be employed at all steps of the pipeline, but also in other applications outside this domain. Chapter 4 is based on the paper:

■ Raidou, R.G., Eisemann, M., Breeuwer, M., Eisemann, E. and Vilanova, A., 2016. Orientation-Enhanced Parallel Coordinate Plots. IEEE Transactions on Visualization and Computer Graphics (Proceedings of the Information Visualization 2016), 22(1), pp.589-598 [222].

4.1. Abstract

Parallel Coordinate Plots (PCPs) is one of the most powerful techniques for the visualization of multi-dimensional, multi-variate data. However, for large datasets the representation suffers from clutter due to overplotting. In this case, discerning the underlying data information and selecting specific interesting patterns can become difficult. We propose a new and simple technique to improve the display of PCPs by emphasizing the underlying data structure. We call this technique Orientation-Enhanced Parallel Coordinate Plots (OPCPs).

Our proposed OPCPs improve pattern and outlier discernibility by visually enhancing parts of each PCP polyline with respect to its slope. This enhancement also allows us to introduce a novel and efficient selection method, the Orientation-Enhanced Brushing (O-Brushing). Our solution is particularly useful when multiple patterns are present or when the view on certain patterns is obstructed by noise.

We present the results of our approach with several synthetic and real-world datasets. Finally, we present the results of a user evaluation, which verifies the advantages of the OPCPs in terms of discernibility of information in complex data. The results also confirm that O-Brushing eases the selection of data patterns in PCPs and reduces the amount of necessary user interactions compared with state-of-the-art brushing techniques.

4.2. Introduction

Parallel Coordinate Plots (PCPs) [128] are used for the visualization of multi-dimensional, multi-variate data. With the use of PCPs, multiple data dimensions are mapped one-by-one to a number of parallel vertical axes, as described in Chapter 3. Each multi-dimensional data object is mapped to a polyline that intersects the axes, connecting the scalar values of every dimension [128], as depicted in Figure 3.3. PCPs are able to efficiently display in a single view all 2D projections of adjacent data dimensions [126, 128, 303]. In this way, they enable the identification of relations and the detection of data patterns or trends – especially with the help of interaction [112, 251] such as brushing [111] or reordering [13, 201, 285].

A limitation of PCPs is that they might suffer from clutter due to overplotting [112]. This causes problems in data exploration and interpretation, especially in high density data. Reducing visual clutter in PCPs is an important topic [14, 91, 175, 188, 320]. However, most of the previous solutions are complex and focus mainly on aiding the detection of clusters in the data [14, 320], not in revealing the overall data structure. In other cases, the proposed visualizations may even unintentionally lead to concealing patterns and outliers [14, 188]. Finally, other solutions require interaction to achieve clutter reduction [91, 175], which is not always possible.

We propose a simple technique to improve the representation of datasets in PCPs: the *Orientation-Enhanced Parallel Coordinates* (OPCPs). Our technique visually enhances specific parts of each PCP line, depending on its slope. It enables discerning individual trends and patterns, while it may even reveal patterns that are potentially obscured in traditional PCPs. This enhancement also allows us to introduce a new brushing technique to facilitate pattern selection in complex data, the *Orientation-Enhanced Brushing* (O-Brushing).

Our work consists of the following two major *contributions*:

· The concept of Orientation-Enhanced Parallel Coordinates (OPCPs) to improve the view and discernibility of patterns in otherwise cluttered PCPs, without loss of low density data information or outliers.

· A versatile brushing technique based on the OPCPs: the Orientation-Enhanced Brushing (O-Brushing). It enables efficient selection of individual data structures, with reduced user interaction.

The remainder of this chapter is organized as follows: Section 4.3 provides an overview on previous work related to the topic of clutter reduction and discernibility improvement in PCPs. Section 4.4 is the core of this chapter, where the OPCPs and the O-Brushing techniques are presented and discussed. Section 4.5 presents the results of our proposed approach, while Section 4.6 presents the results of our user evaluation. Finally, Sections 4.7 and 4.8 conclude the chapter with a discussion on several points with respect to OPCPs and propositions for future work.

4.3. Related Work

Many different techniques have been proposed for enhancing the display of multi-variate data and for reducing clutter in Information Visualization representations [80], including PCPs. Some approaches require the manipulation of the axes of the representation, using *reordering* [13, 201, 285, 313]. These approaches are able to reveal hidden patterns and facilitate data interpretation. However, in data with a large number of points reordering is insufficient. Other approaches involve *visual enhancement* of PCPs by rendering curves or splines instead of lines [11, 102, 268, 320]. Such approaches are especially effective in reducing clutter at the crossings of PCP lines, but they might suppress data patterns, such as outliers.

Another commonly encountered group of techniques requires *clustering*, combined with different kinds of visual enhancements. Among these, we often encounter:

· manipulating PCPs by averaging polylines and visualizing correlation coefficients between polyline subsets [250],

· filtering PCPs based on frequency or density of the data [14],

· combining polyline splatting for cluster detection and segment splatting for clutter reduction [319],

· using cluster-based hierarchical enhancements and proximity-based coloring schemes to provide a multi-resolution view to the data [91],

· enabling context visualization at several levels of abstraction, both for the representation of outliers and trends [188],

· using several transfer functions to reveal specific clusters and patterns in the data [134].

All previously mentioned cases involve clustering methods and focus on detecting and differentiating specific clusters or trends in the data – not data patterns or underlying structures. In certain cases, clustering solutions inevitably lose information in low density areas,

when reducing overplotting in high density areas. The approach of Zhou et al. [319] even requires animation, which is not always feasible. Finally, a more artistic approach was proposed by McDonnell et al. [175]. It incorporates a variety of techniques when rendering PCPs, such as edge-bundling, visualization of the distribution and density of the data via opacity and shading or silhouettes for easy distinction of overlapping clusters. However, not all techniques can be used in a single view, as some of them do not work well, if combined.

PCPs have also been used in *combination* with other representations, such as Star Glyphs [83, 227], radviz [21], or scatterplots [315]. As recognized by Holten et al. [123], combining scatterplots with PCPs outperforms many other PCP variants, such as combining with colors, opacity, curved polylines or animations. PCPs have also been combined with histograms [92] to simultaneously show the density and slopes of polylines. This combination enables the exploration of clusters, linear correlations and outliers in large datasets, with more emphasis on data-driven and not pattern-revealing exploration.

Interaction makes local and dynamic data enhancements possible. The use of lenses [78, 307] or brushing are typical examples. As part of the XmdvTool [295], a number of different brushes have been proposed by Martin et al. [173] and Ward [296]. Depending on the information that needs to be shown in the data, different brushes are used for highlighting, linking or masking the underlying data. Additionally, wavelet approximations are used to enhance brushing [308], by showing different parts of the polylines at different resolutions. However, brushing two variables in a non-separable way has only been enabled by the angular brushing proposed by Hauser et al. [111]. The most important state-of-the-art brushing approaches related to our approach are presented in Figure 4.1.

To sum up, there are different approaches for data enhancement and readability improvement in PCPs. However, most of the solutions aim at reducing clutter in PCPs by clustering the data, without giving a better understanding of the overall underlying structure. Data details such as outliers are often unintentionally hidden. Additionally, some solutions work better, or only on a screen, either because they are animated or because they require interaction. Finally, most of the approaches require complex steps, which means that they cannot always be easily reproduced or used. In the following sections, we present our approach to handle all aforementioned challenges.

4.4. Orientation-Enhanced Approach for PCPs

Our solution consists of two main components: the Orientation-Enhanced Parallel Coordinate Plots (OPCPs) for the visual enhancement of PCPs, which will be described in section 4.4.2, and the Orientation-Enhanced Brushing (O-Brushing) for the interactive selection and analysis in OPCPs, which will be presented in section 4.4.3.

4.4.1. Background: Parallel Coordinate Plots

In a simple two-dimensional dataset with dimensions d_1 and d_2, a data point $D = (y_1, y_2)$ is plotted as a PCP line. This line is intersecting the two vertical axes d_1 and d_2 at the positions y_1 and y_2, respectively, as shown in Figure 4.2. When plotting the PCP lines, it is common to employ opacity as a simple way of representing the density of the lines [112]. From now on we will refer to this enhancement as *density PCPs*.

Probing   Area Brush   Lasso Brush

Composite Slider   Composite Slider
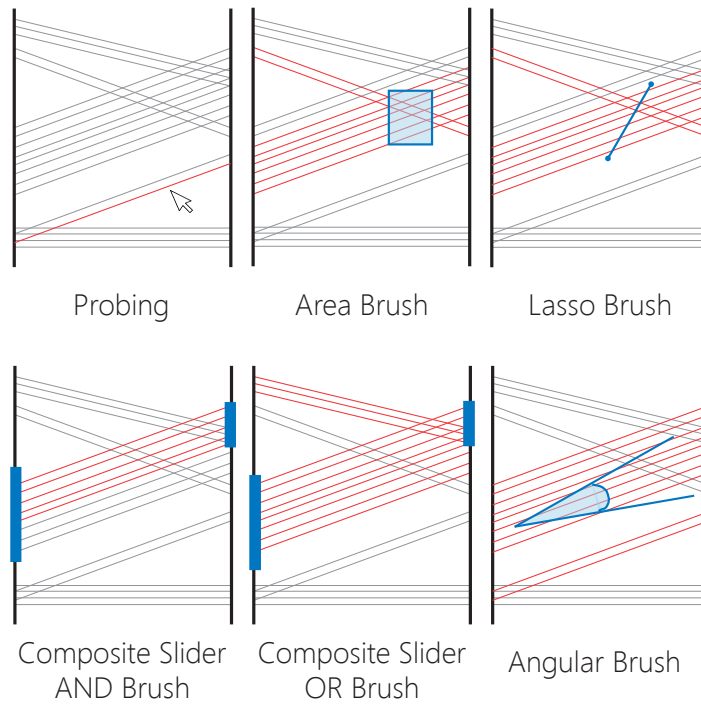AND Brush   OR Brush   Angular Brush

Figure 4.1: Overview of different state-of-the-art brushing approaches for PCPs. With the red lines, we denote the PCP lines in each case, which were selected with the brushing operations illustrated in blue.
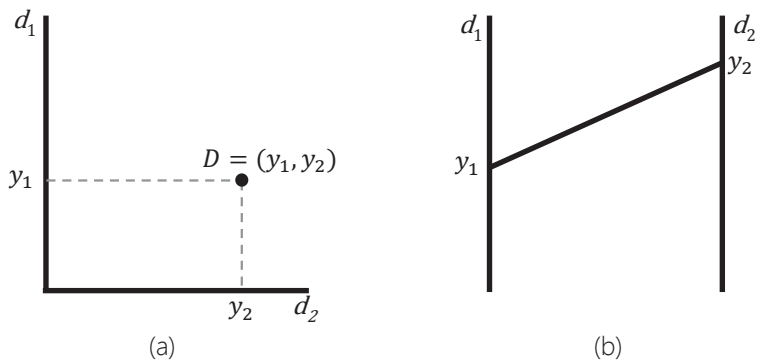

$D = (y_1, y_2)$

(a)   (b)

Figure 4.2: Schematic representation of the concept of PCPs for the simple case of a two-dimensional point D with dimensions d_1 and d_2 and dimension values (y_1, y_2): (a) In a scatterplot. (b) In a PCP.

4.4.2.   Orientation-Enhanced Parallel Coordinate Plots

Holten et al. [123] conducted an evaluation of PCP variants, where they demonstrated that no other enhancement from the examined alternatives improves PCPs significantly apart from combining scatterplots with PCPs. Inspired by this paper, we investigated a simple way to combine effectively the two representations to enhance the display of PCPs. In many papers, the combination of PCPs with scatterplots has been limited to having multiple interactive linked views. This might entail memory limitations for the user, caused by switching the view between the two separate representations.

The goal of the proposed visual enhancement of PCPs is to provide a better understanding in the visualized data, by integrating PCPs and their corresponding scatterplots in one view. A similar approach was followed by Yuan et al. [315]. However, this technique is complex and requires bending the polylines to fit to the points of the scatterplots. In contrast, we are looking for a simple approach that keeps the original appearance of PCPs intact. In the proposed OPCPs, the basic principle is to enhance the PCP lines with respect to their slope. This solution links PCPs and the corresponding scatterplot of the neighboring two axes, in a natural way. In the remainder of the subsection, we describe the steps that we followed for the design of our OPCPs.

■ *Mapping.* For illustration purposes, we demonstrate our concept using a two-dimensional case. For simplicity, we also assume that the data values for each dimension have been normalized to the range $[0, 1]$. In Figure 4.3 - a, we show a PCP line which is defined by its dimension values (y_1, y_2) and a slope α:

$$\alpha = \frac{y_2 - y_1}{d_x}, \tag{4.1}$$

where d_x is the distance between the two vertical PCP axes. In essence, we map the PCP line to a unique reference point $P = (x_p, y_p)$ in the space between the two PCP axes, with $x_p \in [0, d_x]$ and $y_p \in [0, d_y]$, where d_y is the length of the vertical axis, as shown in Figure 4.3 - a. The slope in equation (4.1) is linearly mapped to x_p, while y_p is chosen to make P lie on its corresponding PCP line:

$$x_p \quad = \quad \frac{d_x^2}{2d_y} \cdot \alpha + \frac{d_x}{2} \tag{4.2}$$

$$y_p \quad = \quad y_1 + x_p \cdot \alpha \tag{4.3}$$

In Figure 4.3 - b, we show an example with multiple PCP lines and their respective reference points. Equations (4.2, 4.3) result into a point-to-point transformation – or warping – of a 2D scatterplot space to the OPCP space, as shown in Figure 4.4. This illustration shows the link between the scatterplot points and the reference point positions on the PCP lines.

■ *Representation.* To visually enhance each reference point and to preserve the orientation and context of its PCP line, we create a small line segment, which we call Orientation-Enhanced PCP (OPCP). It is a small segment that shares the original PCP line orientation and is centered at the reference point P.
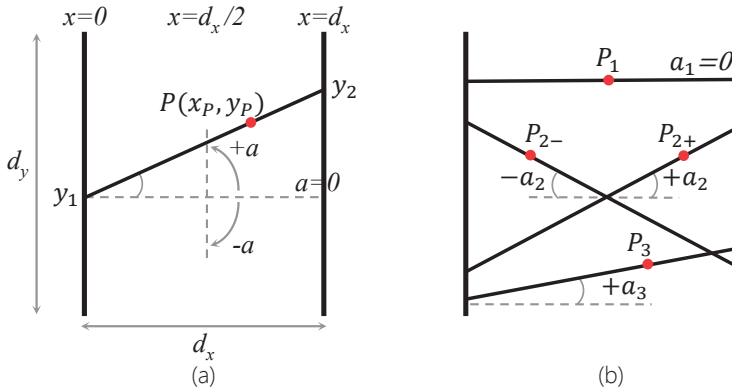
Figure 4.3: Schematic representation of the concept behind OPCPs: (a) Mapping of the slope α of the PCP line (y_1, y_2) to the reference point $P = (x_p, y_p)$ between the two PCP axes. (b) Mapping of the slopes of multiple PCP lines to their corresponding reference points in the PCP space.
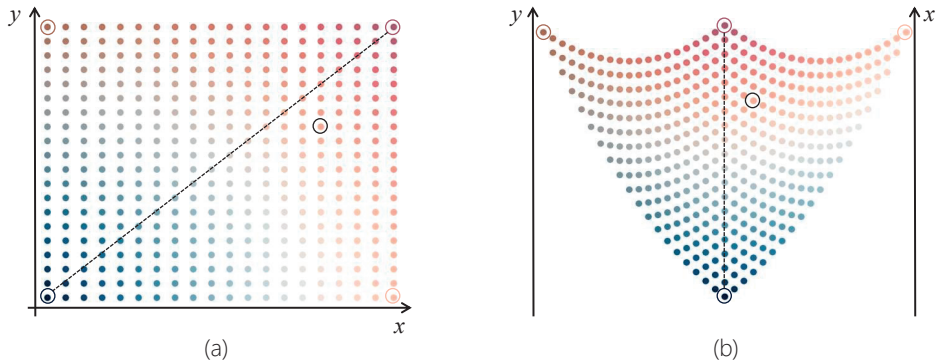

Figure 4.4: Schematic representation of the transformation from (a) the scatterplot space to (b) the OPCP space. Here, we use a 2D colormap [229] and annotations to show the point-to-point correspondence from one space to the other.

Assigning a constant intensity and a given length to each segment would result into OPCP segments that would not be visually separated, if they would be very close to each other. Therefore, we vary the intensity of the segments using a kernel smoother. We smooth the edges of the segments and assign higher intensities in the middle, which is the location of the reference point P, as shown in Figure 4.5.

This desired intensity profile can be achieved with peak-shape kernels, such as a Gaussian kernel [101]. The intensity I of the OPCP segment, resulting from a reference point $P = (x_p, y_p)$ after applying the Gaussian kernel, will be described as:
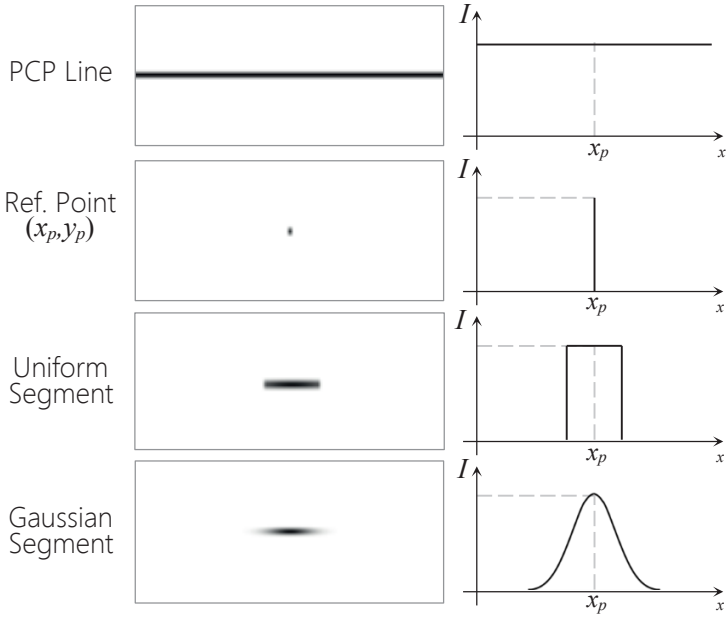
Figure 4.5: Alternatives considered for the intensity encoding of the OPCP segments. Next to each case, we show also the intensity profile.

$$I(x_s) = k \cdot \exp\left(-\frac{(x_s - x_p)^2}{2 \cdot \sigma^2}\right), \tag{4.4}$$

where σ is the bandwidth of the kernel, which is user-defined, and k is a scale factor, which relates to the height of the peak and is given by: $\frac{1}{\sigma\sqrt{2\pi}}$. The bandwidth σ has an impact on the length of the OPCP segment: larger σ values result in smoother and wider-spread segments. Figure 4.6 shows an example of OPCPs applied to three simple synthetic cases and the effect of σ on their appearance.

■ *Visual Enhancement.* In the paper of Harrison et al. [109], it is stated that PCPs can emphasize specific correlations more than others. Depending on the data aspects that need to be emphasized, we propose to employ three enhancement methods: gamma correction, transfer functions, and histogram equalization.

Gamma correction [101] allows the user to remap the levels of the intensity range, in order to discern more details in the darker parts of the OPCP segments. This can be accomplished with low values of gamma, while increasing values of gamma sharpen the OPCPs. Gamma correction is applied per pixel, transforming the intensity I to $I_{gcorr} = I^\gamma$. The effect of the parameter γ is depicted in Figure 4.7.

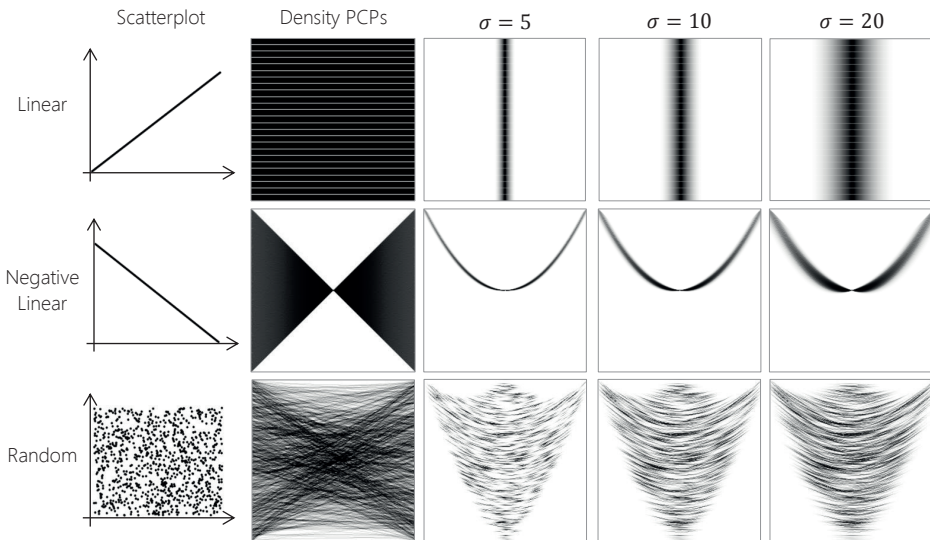The effect of gamma correction can be generalized by applying a transfer function (TF),

Figure 4.6: Effect of the σ value of the Gaussian kernel on OPCPs, for three simple synthetic cases. To increase the visibility of the segments, we have linearly scaled the image intensities to the range $[0, 1]$.

aiming at controlling the contrast in the representation. As introduced in the work of Johansson et al. [134], different TFs affect the appearance of different data aspects. A linear TF gives an overview on the data, a logarithmic TF enhances low density areas, a square root TF emphasizes outliers in the data, and a quadratic TF enhances the high density areas. The effect of the four previously mentioned TFs is shown in Figure 4.8.

Optionally, histogram equalization [101] reassigns the intensity values of an image, such that the output will exhibit a uniform distribution of intensities. Histogram equalization can create a background-foreground effect and enable better discernibility of different patterns in the data, especially in the presence of noise or of strong patterns. The impact of histogram equalization in OPCPs is depicted in Figure 4.9.

■ *Overlay.* We enhance PCPs by overlaying the OPCP segments on top of the traditional PCP polylines – for example, on density PCPs. To this end, we employ alpha blending [208], as shown in Figure 4.10 - a. Overlaying OPCPs on top of PCPs helps in preserving the main benefit of the latter, which is the connectivity across data dimensions. In this way, PCP polyline bundles can still be traced.

Additional color encoding of OPCPs can enhance and visually separate them from the underlying PCPs, as shown in Figure 4.10 - b. To reduce as much as possible the chances of distracting the user by overlaying OPCPs on the PCP polylines and interfering with PCP bundle tracking, the appearance of OPCPs can be further adjusted. The user can modify the color and opacity of the OPCP segments, but also to fine-tune the σ and γ values to make the OPCPs more or less prominent. For the purpose of this work, we encode the OPCP intensities as black color values in the explanatory examples and red in the overlay examples.
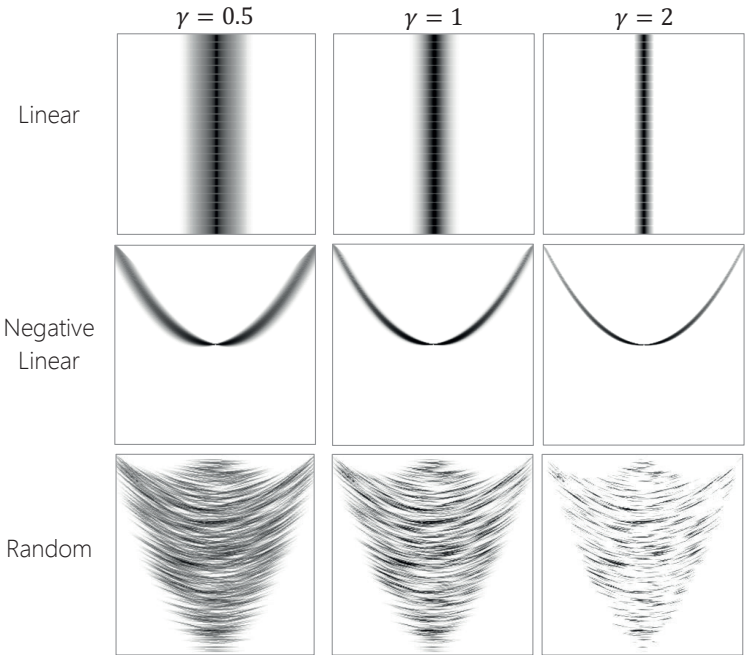
Figure 4.7: Effect of the gamma correction on the appearance of the OPCPs. Here, the σ value was set to 10 and the image intensities were scaled to the range $[0, 1]$.
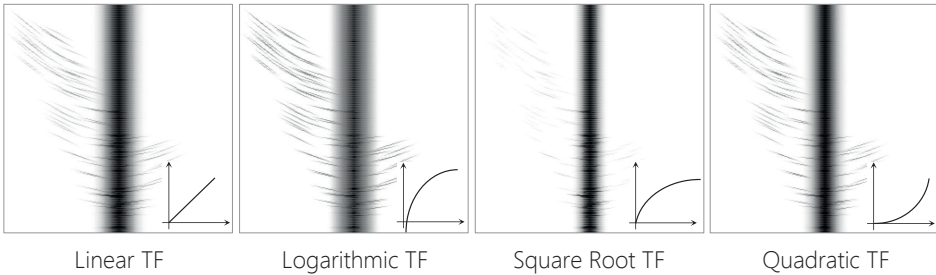

Figure 4.8: Enhancement of different data aspects, using transfer functions (TFs). Here, we use a synthetic dataset with a dominant linear relation among the dimensions and a few outliers. The bandwidth σ was set to 10 and the image intensities were scaled to the range $[0, 1]$.

■ *Parameter values.* The parameters involved in the visual enhancement of the OPCPs, such as the bandwidth σ and the gamma correction value γ, should depend on the specific aspects of the data that need to be brought forward. Therefore, we do not assign a specific set of values, but allow them to be user-controllable. In our interactive tool, we initially assign a set of values (σ=10 and γ=1), which give already a good result, but can be changed adequately

| No Histogram Equalization | With Histogram Equalization |
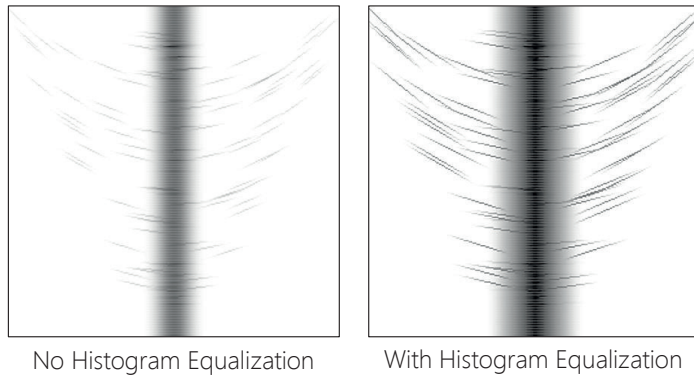
Figure 4.9: Effect of histogram equalization on the appearance of OPCPs. Here, the σ was set to 10 and γ to 1.
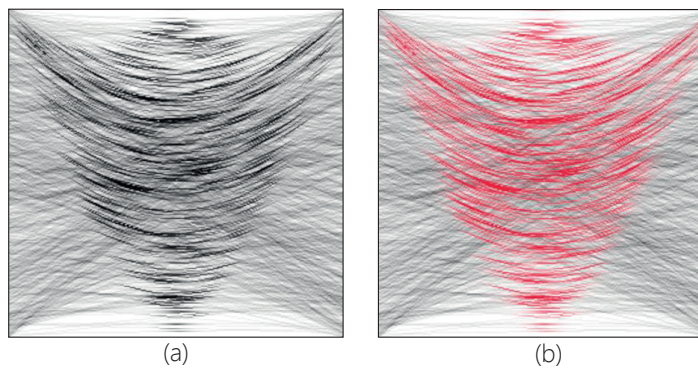

(a)　　　　　　　　　　(b)

Figure 4.10: Example of alpha blending in random noisy data: (a) Alpha blending of histogram equalized OPCPs with density PCPs. (b) Color encoding of the foreground blended OPCPs (red) and background density PCPs (black).

by the user.

4.4.3. Orientation-Enhanced Brush (O-Brushing)

Brushing [112, 251] is a common strategy for the selection of polylines of interest in PCPs. However, when the amount of plotted lines increases, this selection becomes difficult. OPCPs have an important property: they establish for each 2D data point a unique position in the space between each pair of the PCP axes. This allows us to introduce a new brushing approach that is applied in the OPCP space, for easier and more efficient selection of specific data patterns, which we name Orientation-Enhanced Brushing (O-Brushing).

O-Brushing is performed on OPCP segments in two ways: either with a traditional brush

metaphor (O-Brush) or with a prober (O-Prober). These two methods are shown in Figure 4.11. The O-Brush acts as a lasso brush [112], applied only in the OPCP space. It requires two user interactions, namely two clicks. The O-Brush is shown in Figure 4.11 - a. The O-Prober is an interactive rectangle that can be resized and moved around the representation, as depicted in Figure 4.11 - b. It works similarly to an area brush applied only in the OPCP space. It requires maximally three interactions, namely, 2D resizing and translation.

Compared to traditional brushing methods, the O-Brush and O-Prober act only in the OPCP space, hereby allowing for a more precise and local selection and resulting in a reduced amount of required user interaction. The O-Prober and the O-Brush can produce the same result, but their main difference is that the former can be used to probe through the dataset for multiple similar patterns – for example, for lines with a given slope or a range of slopes – employing minimal user interaction. In our current implementation, the O-Prober is a simple movable rectangle of user-defined size, but it could be easily extended to arbitrary shapes.

Figure 4.12 shows a comparison of the O-Brushing methods with their alternatives from existing literature. It depicts for each brushing technique the best achieved result, and the user interactions required to select one specific pattern of interest. In this example, the specific selection is only possible with the composite slider brush and the two proposed O-Brushing methods. However, the two O-Brushing methods require fewer user interactions than the composite sliders. In our interactive implementation, we included also the state-of-the-art brushes, to enable users to perform selections both in the traditional PCP space and the OPCP space.

4.5. Results

In this section, we present the results obtained by the application of OPCPs and the O-Brushing to different datasets, intending to provide a deeper understanding into the OPCPs space and its characteristics. To this end, the visualization of the OPCPs was implemented
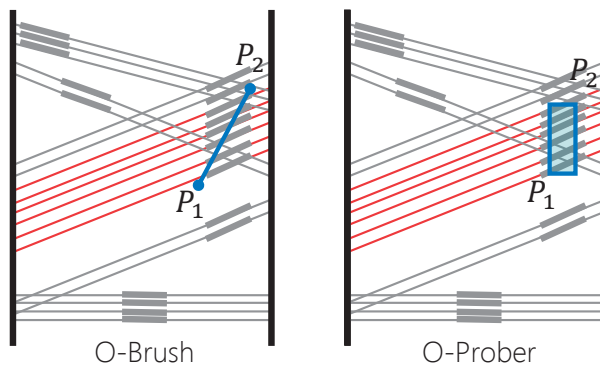

Figure 4.11: Schematic representation of the concept behind O-Brushing. The thick gray segments represent an OPCP for each underlying PCP line. With red we denote the selections in each case, while with blue we depict the brushing operation.
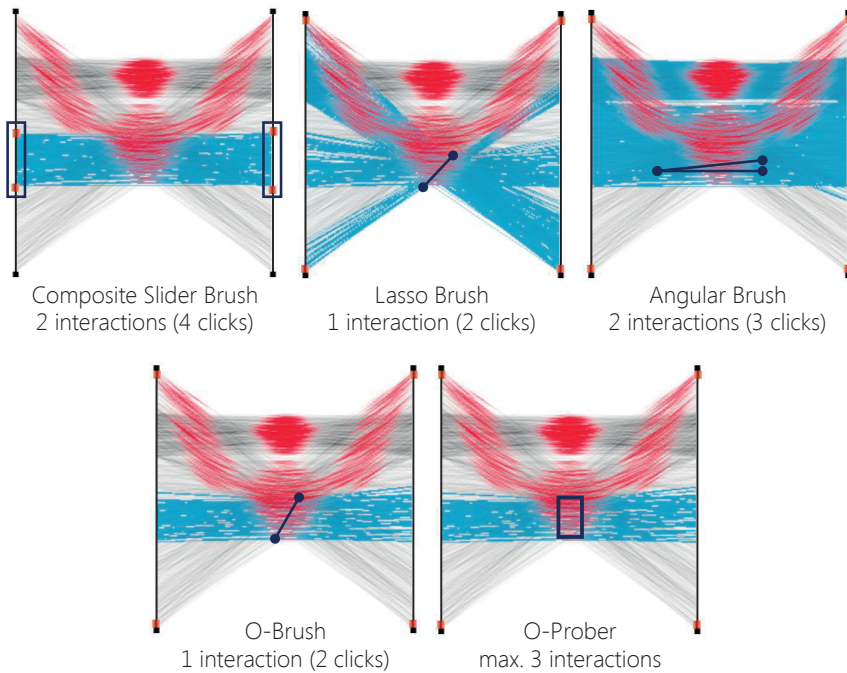
Figure 4.12: Example for the comparison of O-Brush and O-Prober against traditional brushing methods, when attempting to select the same part, i.e., data points with middle values of both dimensions. All brushes have been applied individually to the data. In this example, the lasso and area brush do not succeed in selecting the specific data region. We show also the number of user interactions, i.e., the number of clicks, required for each of the brushes. The composite slider brush requires maximally four clicks, the O-Brush requires two clicks and the O-Prober requires maximally 3 user interactions (resize in both dimensions and translate). Here, red is used to encode the OPCPs and blue to denote the selected PCP lines, using each of the brushes.

in Python on the GPU, using OpenCL. The interaction for the brushing was realized using the Visualization ToolKit (VTK) [2].

We tested the OPCPs on two different types of data. First, inspired by previous work [134, 315], we tested the behavior of OPCPs on a number of synthetic cases with two-dimensional data with 1000 data points, containing predefined patterns and structures. PCPs and OPCPs are meant for multi-dimensional data, but we employ these two-dimensional examples for illustration purposes. Secondly, to demonstrate a real usage scenario of OPCPs, we used multi-variate data obtained from various databases [1, 4, 135, 164].

4.5.1.   Results with Two-Dimensional Synthetic Stimuli

In Figure 4.13, we show our approach as applied to the synthetic stimuli, together with their corresponding scatterplots and density PCPs. From this figure, the warping transformation

of the scatterplot space on the PCPs, which was described in section 4.4.2, becomes obvious. From these examples, we confirmed that the OPCPs, in comparison to PCPs, facilitate the discernibility of multiple data patterns, data outliers and also data structures obstructed by noise. In the remainder of the section, we illustrate this with additional examples.

■ *Discernibility of data patterns.* We assume two main subcategories of relationships in the data: either there is a single relationship in the data, which is not immediately recognizable; or there are multiple and more complex relationships. Examples of both results were shown in Figure 4.13. In the first category, we include four different stimuli. For the cubic and square root stimuli, the OPCPs facilitate the identification of the different patterns compared to PCPs, because of the visible correspondence to the scatterplot space. Additionally, the double spread and sinusoidal stimuli, have a similar appearance when shown in the density PCPs. However, the OPCPs allow to see that these are different patterns. Finally, for the second category, it is also easier to identify the multiple relations between the two data dimensions - or data clusters - when employing the OPCPs, as depicted in Figure 4.13.

■ *Discernibility of data outliers.* We use two synthetic stimuli, for which the dimensions are linearly correlated, as shown in Figure 4.14. In addition to this, the second stimulus contains some outliers. By overlaying the OPCPs on the density PCPs, we enhance the main pattern in the data, which is a linear relationship, without obscuring the outliers.

■ *Discernibility of noise-obstructed data structures.* We created two stimuli with initially no correlation between the two dimensions, as shown in Figure 4.15. In addition to this, a structure with a linear relationship between the dimensions was then added to the second stimu-
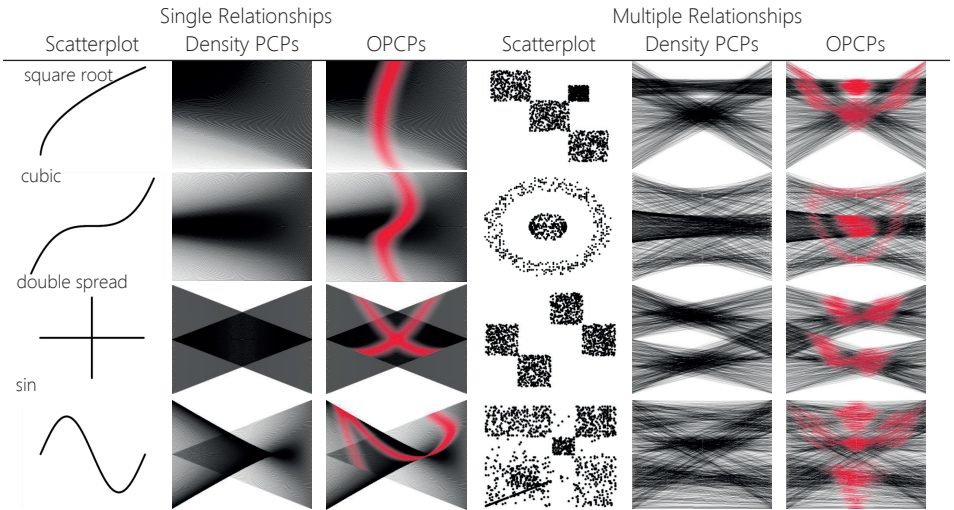

Figure 4.13: Examples showing that the OPCPs allow the discernibility of (multiple) patterns or clusters in the synthetic stimuli.
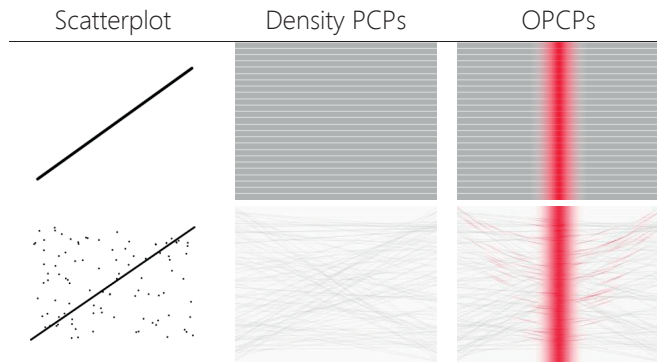
Figure 4.14: Examples showing that the OPCPs enable the discernibility of outliers in the synthetic stimuli.
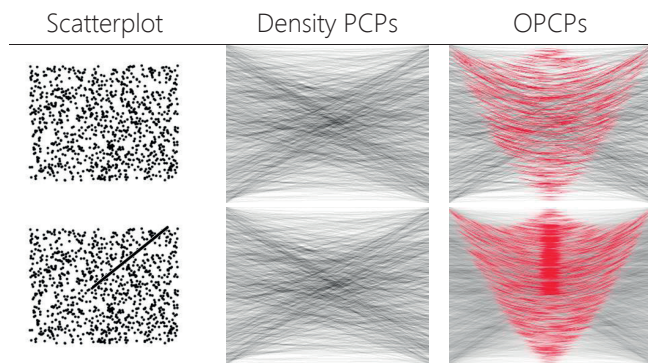

Figure 4.15: Examples showing that the OPCPs enable the discernibility of noise-obstructed structures in the synthetic stimuli.

lus. In density PCPs, this structure is hidden. By overlaying the OPCPs on the density PCPs, we can visually enhance the obstructed data structure and recover the underlying relation.

4.5.2. Results with Multi-Variate Synthetic and Real Data

In a real-world analysis, PCPs are used to visualize multi-variate data. To additionally assess our approach, we employ more complex data with more realistic data patterns across their dimensions. Among these datasets, we included four well-known datasets from various databases. The employed datasets are the `apartments` dataset from the database of TU Braunschweig [1] with 2,290 data points, the `Venus` dataset from the database of the XmdvTool [4] with 8,784 data points, the `Out5d` dataset from the database of the Xmdv-Tool [4] with 16,384 data points and the `household` dataset from UCI repository [164] with
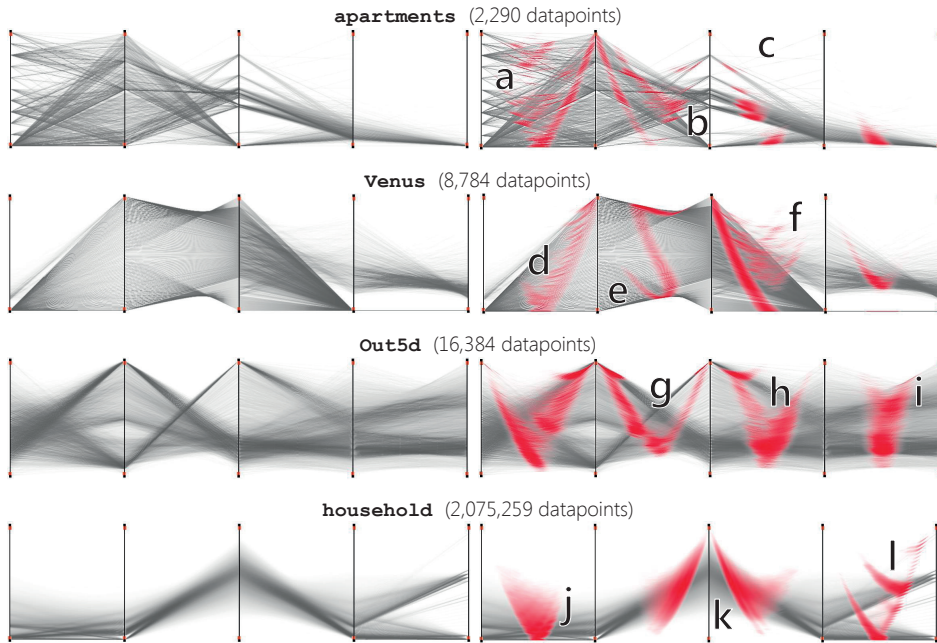
Figure 4.16: Application of OPCPs to multi-variate synthetic or real data obtained from various databases [1, 4, 164].

2,075,259 data points. Figure 4.16 shows the results of using OPCPs to represent the above mentioned datasets. The OPCP advantages discussed in section 4.5.1 are again apparent in these cases.

In the `apartments` dataset, especially between the first two data dimensions, multiple data patterns are emphasized. They are also more discernible when using OPCPs, as shown in Figure 4.16 - a. This also occurs between the second and third dimension of the same dataset, as depicted by Figure 4.16 - b. In this dataset, OPCPs are also able to bring forward outliers – for example, between the second-third and third-fourth data dimensions, which were not easily discernible in the density PCPs, as can be seen in Figure 4.16 - c.

In the `Venus` dataset, the OPCPs facilitate the identification of distinct patterns – for example, three patterns between the second and third data dimension, as depicted in Figure 4.16 - e. In addition to this, OPCPs allow to visually enhance the multiple small clusters between the first two dimensions, which is shown in Figure 4.16 - d, as well as outliers between the third and fourth data dimension that are not visible in the respective density PCPs, presented in Figure 4.16 - f.

In the `Out5d` dataset, pattern identification becomes easier throughout all dimensions, especially in parts of the representation, where the patterns are obstructed by noise. An example of this is visible in Figure 4.16 - g,h,i, in the last three dimensions of the data.

Finally, in the `household` dataset, between the first two dimensions, but also between the third and fourth dimensions of the data, patterns that were not visible in the traditional density PCPs are brought forward with the use of OPCPs. This is demonstrated in Figure 4.16 - j,k. Especially, in Figure 4.16 - k, there are two main patterns in the data, which appear as a single pattern in the density PCPs. Also, between the last two data dimensions there are several outliers, which are significantly enhanced with the OPCPs, as can be seen in Figure 4.16 - l.

■ *Advantages of the O-Brushing.* Based on the added benefits of OPCPs, it is expected that O-Brushing will facilitate the selection of the respective data structures in the OPCP space, in comparison to state-of-the-art brushing methods, which act in the PCP space. In this work, we compare our proposed O-Brushing to methods, such as the lasso brush, the angular brush and the composite slider brush. Examples of the use of O-Brushing are shown in Figure 4.12 and Figure 4.17.
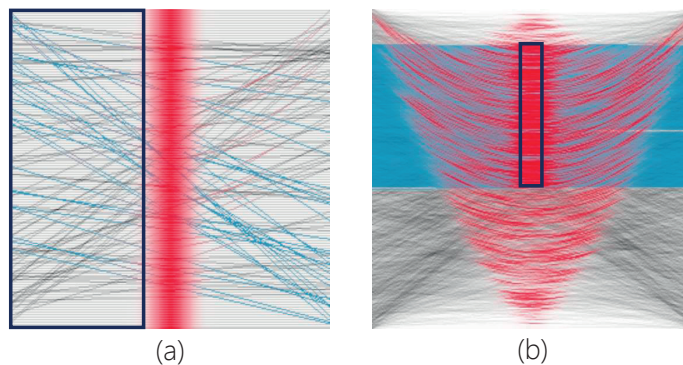

(a)   (b)

Figure 4.17:   Examples showing that the OPCPs allow the selection of (a) outliers and (b) noise-obstructed structures in the data.

4.5.3.   Performance

The performance of our approach was tested on several datasets from various databases [1, 4, 135, 164]. The datasets vary between 1,000 and 2 million data points. The test was conducted on an Alienware Aurora R4 with an Intel Core i7-4820K @ CPU 3.70GHz Processor, 16GB RAM and NVIDIA GeForce GTX 780. The performance results are depicted in Figure 4.18. The system is implemented on the GPU and enables interactive brushing. The O-Brush and O-Prober can be employed for almost real-time data-driven selection.

4.6.   Evaluation

To test our approach with respect to the state-of-the-art, we conducted a user evaluation. We used the implemented interactive prototype, which enables visualizing data with density PCPs and OPCPs, as well as data selection with all five brushing techniques: compos-
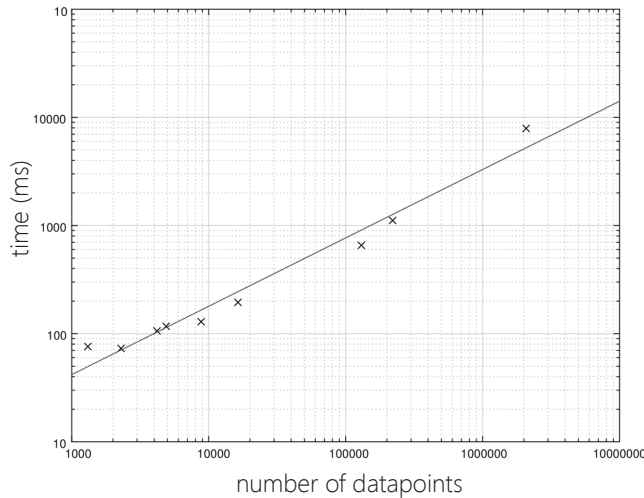
Figure 4.18: Performance times of OPCPs for multi-variate synthetic or real data from various databases [1, 4, 135, 164].

ite slider brushes, classical lasso brush [112], angular brushing [111], as well as our two O-Brushing approaches.

The evaluation was designed based on the paper of Lam et al. [159] and consisted of two main parts. The first part was a controlled user study to measure User Performance (UP) [159] with the OPCPs and O-Brushing, against density PCPs and traditional brushing. For this part, we performed three experiments, which are described in detail in the following subsections. The second part consisted of answering a questionnaire to measure User Experience (UE) [159], using Likert scales, ranking, and open questions.

We employed 16 participants, with various backgrounds: Computer Science (11, out of which 5 from Computer Graphics and 4 from Visualization), Electrical Engineering (3), Physics (1) and Biomedical Engineering (1). Most of them (9) had preliminary knowledge of PCPs, although only one participant had worked with PCPs before. Before the evaluation, we gave a short introduction, we demonstrated the functionality of the prototype – for example, how to perform data selections with each method –, and we allowed participants to use it, until they felt confident with it. In average, people spent around 5 minutes on the prototype before the experiment.

4.6.1. First Part: User Performance

We performed three experiments. The *first experiment* aimed at measuring the performance of users in discerning patterns, outliers, and data structures obstructed by noise using density PCPs or OPCPs. We created two comparable, two-dimensional synthetic datasets per case, and we visualized them with both representations. Then, we showed static images of the representations to the users in a randomized order, and we asked them to perform tasks,

such as identifying and pointing out patterns in the data, outliers and noise-obstructed data structures. For each one of the static images, we measured the time that users needed to give a conclusive answer and accuracy of their answers. Since the data were synthetic, we already knew the exact number of the required structures in the data. Thus, every wrong or unidentified pattern was penalized in the accuracy measurement.

The *second experiment* aimed at measuring user performance in selecting specific patterns, outliers, and noise-obstructed data structures, with state-of-the-art brushing or O-Brushing. We created a two-dimensional synthetic dataset per case. We showed static images to the users, explaining which part of the data needed to be selected. Then, we asked them to perform a selection of the previously specified data part, using only one of the brushing techniques at a time, in a random order. All tasks were possible with all methods. Again, we measured time, accuracy, and number of interactions, or number of clicks, required for task completion.

The *third experiment* aimed at measuring performance with multi-variate, complex data and tasks. We created two comparable five-dimensional synthetic datasets, using the PCDC tool [34]. Then, we designed a set of questions, which were related to identifying and/or selecting data patterns, outliers and noise-obstructed structures. The users were asked to apply traditional brushing to one of the datasets with PCPs, and O-Brushing to the other dataset with OPCPs to perform the given tasks. The order of the dataset and approach, as well as their combination, was alternated randomly, to reduce bias from learning. We measured completion time, accuracy and number of interactions, or clicks, required from the user for the task completion.

The outcomes of the statistical analysis of the experiments are summarized in Figures 4.19 - 4.21, at the end of this chapter. The first and third experiment were analyzed with paired t-tests, while the second was analyzed with ANOVA and Tukey's HSD test.

The results of the *first experiment* (Figure 4.19) indicate that identification of patterns, outliers and noise-obstructed structures is more accurate with OPCPs than with PCPs ($\rho < 0.01$). Especially, in case where a structure is obstructed by noise in the data, the OPCPs were much more accurate ($\mu = 1, \sigma = 0$) than PCPs ($\mu = 0.06, \sigma = 0.25$). The distinction of noise-obstructed structures is also faster ($\rho < 0.05$) in OPCPs: users required half the time to recognize these kinds of structures in the data with OPCPs than with PCPs. For pattern and outlier detection, there is no conclusive result for the time performance, but the accuracy is significantly improved with OPCPs.

The outcome of the *second experiment* (Figure 4.20) shows that O-Brushing is faster and more accurate ($\rho < 0.01$), in all cases. For pattern and outlier selection, O-Brushing also requires significantly less interactions ($\rho < 0.05$). From Tukey's HSD test, it results that there is no statistically significant difference between the performance of users when using the O-Brush or the O-Prober. Based on this test, the overall ranking of the different brushing techniques for the three investigated tasks results as: the two variants of O-Brushing, angular brushing, composite brushing using sliders and lasso brushing.

The results of the *third experiment* (Figure 4.21) demonstrate that our approach is more accurate than the state-of-the-art approach, for the four given tasks. The combined use of OPCs with O-Brushing had an average accuracy of **0.96** for all tasks, while traditional PCPs with standard brushing only **0.75**. In this experiment, there were no indications that pattern

discernibility requires less time with OPCPs. However, for the other three tasks the use of OPCPs and the proposed O-Brushing makes a big difference in performance times. For example, for pattern selection our approach requires half the time of the traditional approach. Overall, there is an indication that when selection is involved, our approach is also significantly faster and requires less interaction ($\rho < 0.05$).

4.6.2. Second Part: User Experience

The second part of the evaluation consisted of conducting a survey. First, we asked users to grade PCPs and OPCPs and, also, the five previously used brushing methods, using *Likert scales*. The outcome of the statistical analysis of the experiments is summarized in Figure 4.22. From the statistical analysis, it resulted that the PCPs were easier to understand, but the OPCPs were considered significantly easier to use (**4.13**), more useful (**4.44**) and also more suitable for the identification of patterns (**4.44**), outliers (**4.38**) and noise-obstructed data structures (**4.31**), compared to traditional PCPs. Moreover, the composite sliders and O-Brushing were considered easier to use and useful, while the easiest to understand were composite sliders and the O-Prober. The sliders and O-Brushing were considered most suitable for pattern selection, while for outlier and obstructed structure selection only O-Brushing was preferred.

The next part of the questionnaire consisted on ranking the two representations using the same scale and the five brushing methods. The OPCPs were ranked significantly higher (**8.31**) than the PCPs (**5.81**) ($\rho < 0.05$), while the O-Brush and the O-Prober were ranked significantly higher (**8.25** and **8.29**, respectively) than the sliders (**6.50**), the lasso (**4.81**) and the angular brushing (**5.25**) ($\rho << 0.01$).

The questionnaire was concluded with open questions. The participants replied that the OPCPs *can be very strong in structure detection in the data*, especially *when there is a lot of overlap in the data*. However, *the OPCPs take more time to get used to* and *might require some training for naive users*. Also, finding *simple correlations across dimensions can be easier sometimes with PCPs only*. O-Brushing makes it *easier to select patterns locally*, but *O-Prober could be improved by using also different shapes, other than the rectangle*. Most users commented that our approach supported them more in the identification and selection of patterns and outliers, in particular. For simple cases, due to the fact that OPCPs require prior familiarization and training, they might be less suitable. However, for cluttered data, the advantages are straightforward.

4.7. Discussion

The results of the application of OPCPs on synthetic and real datasets presented in section 4.5, as well as the evaluation results of section 4.6, brought forward a number of limitations and raised several points for discussion.

Firstly, the proposed OPCPs are a visual enhancement of PCPs that enables the discernibility of patterns, outliers and noise-obstructed structures in the data. In the paper of Holten et al. [123], it is stated that combining scatterplots with PCPs can result in significant performance gains for the users. In many papers, the combination of PCPs with scatterplots is limited to having multiple linked views, where interactive linking and brushing can reflect selections from one representation to the other. However, in this case, users need to switch

between windows and use their mental memory for data exploration and analysis – for example, when the user performs an operation and sees the result in another window.

Our OPCPs, instead, are not aiming at substituting scatterplots or at using linked scatterplot views. They focus on giving a better understanding of the data represented by PCPs, by integrating in a seamless way the two representations in one, combining their benefits and reducing the memory limitations that result from switching between the two separate representations. We consider that a comparison between scatterplots linked to PCPs against OPCPs is out of the scope of this work, as the latter is a visual enhancement of PCPs and not a new representation on its own. Still, if a user would consider it necessary, OPCPs could be linked to additional scatterplots. In this case, it would make sense to investigate a comparison between scatterplots linked to PCPs, and scatterplots linked to OPCPs.

Additionally, from our evaluation it resulted that the interpretability of the patterns might not be straightforward. Our approach requires a certain level of familiarization with the enhancement. However, during the evaluation, the users were able to identify patterns more accurately than with traditional PCPs. Also, the cognitive load induced by the use of OPCPs is not so significant, to slow down the analysis of the data. As it can be seen in the evaluation results, in the vast majority of the tasks, the time needed to perform an operation using OPCPs and the related O-Brushing is significantly less than the time needed to perform the same operation with state-of-the-art techniques. This is a first indication that the interpretability of patterns in OPCPs is not compromised. In a future additional evaluation, it would be interesting to research this further.

Moreover, there was no evidence so far that the use of OPCPs might be distractive for the user or interfering with bundle tracking. OPCPs are a new visual enhancement that requires some training, as pointed out by users. However, in our interactive tool the appearance of OPCPs can be adjusted by fine-tuning the σ and γ values to make the enhancement as prominent as the user would like. Also, there is always the option to adjust the color and opacity of the OPCP segments, to interfere less with the underlying PCPs and the polyline bundles. In the user evaluation, we included tasks where bundle tracking was necessary. In these cases, the users could perform the tasks without problems. However, for a more conclusive answer to this point, a more extensive study would be required.

For the brushing functionality, in the interactive version of our tool, the users can select in which of the two spaces – PCP space or OPCP space – they would like to brush. In the OPCP space, the two proposed O-Brushing methods can be employed, while in the PCP space, state-of-the-art brushing, such as angular or lasso or composite brushing, can be used. As some users stated during the evaluation, having the possibility to choose the space to perform selections on the data is useful in different occasions: for example, if the user needs to perform selections based on the range values of some dimensions, state-of-the-art brushing methods are more appropriate and more straightforward to use. However, if specific patterns, or outliers or structures in the data need to be selected, then O-Brushing is more efficient.

To conclude with, we foresee some limitations of our approach. First, OPCPs require some familiarization, as they are not immediately intuitive. Additionally, they require a wider spacing between the dimension axes as compared to traditional PCPs in order to be effective. Moreover, the OPCPs should be accompanied by PCPs, to preserve context and connectivity across dimensions. Finally, the O-Prober could improve by using free-hand shapes

4

or a scribbling interface instead of the predefined rectangle. This would enable easier, faster and more accurate selection of specific patterns with OPCPs, similarly to the shape-based method proposed by Muigg et al. [179].

4.8. Conclusions and Future Work

Parallel Coordinate Plots exhibit overplotting, which results in a cluttered view on the data. Therefore, discerning the underlying data information and selecting interesting patterns can become difficult. We proposed a new technique, the Orientation-Enhanced Parallel Coordinate Plots, to improve the view and discernibility of patterns in otherwise cluttered PCPs. We achieved our goal by visually enhancing parts of each PCP line with respect to its slope, hereby incorporating information from scatterplots in the representation [123].

Compared to the state-of-the-art, our approach is simple and provides better discernibility of data patterns, especially when there are multiple overlapping patterns or when there are outliers and structures, obstructed by noise. We evaluated our approach with several synthetic and real-world datasets. One of the main advantages of OPCPs is that they allow a new and versatile selection method, the Orientation-Enhanced Brushing. Brushing in the OPCPs space enables an efficient selection of individual data structures involving a reduced user interaction when compared to the state-of-the-art selection tools in PCPs. On the other hand, OPCPs require more training, compared to PCPs.

A direction for future work includes employing color transfer functions in the OPCPs for better discrimination of the different data patterns, or even clustering. Moreover, it would be interesting to extend the evaluation of our proposed visual enhancement, but also of the related brushing method, to cover the points discussed in section 4.7. Finally, the extension of the O-Prober to other shapes should allow easier, faster, and more interactive selections of data patterns.
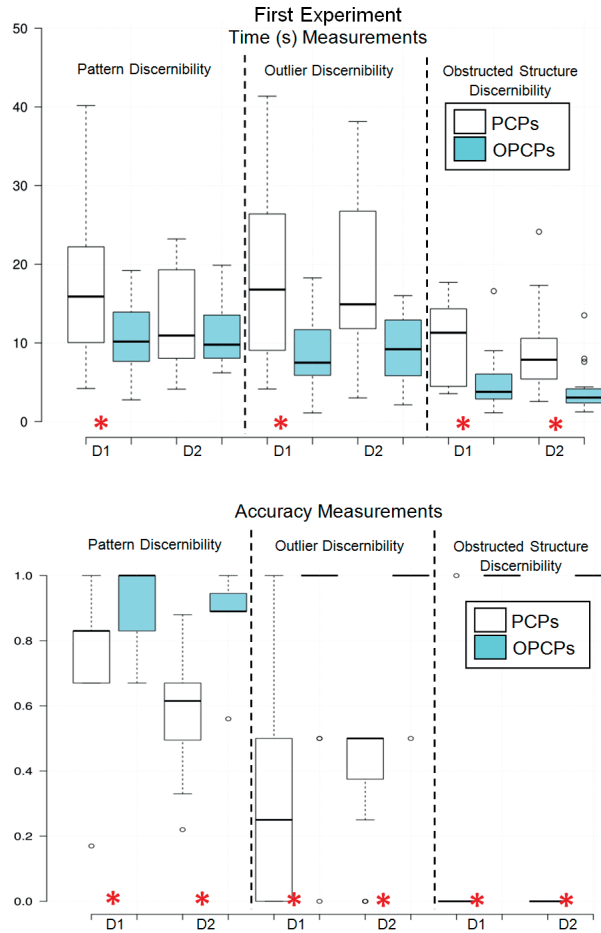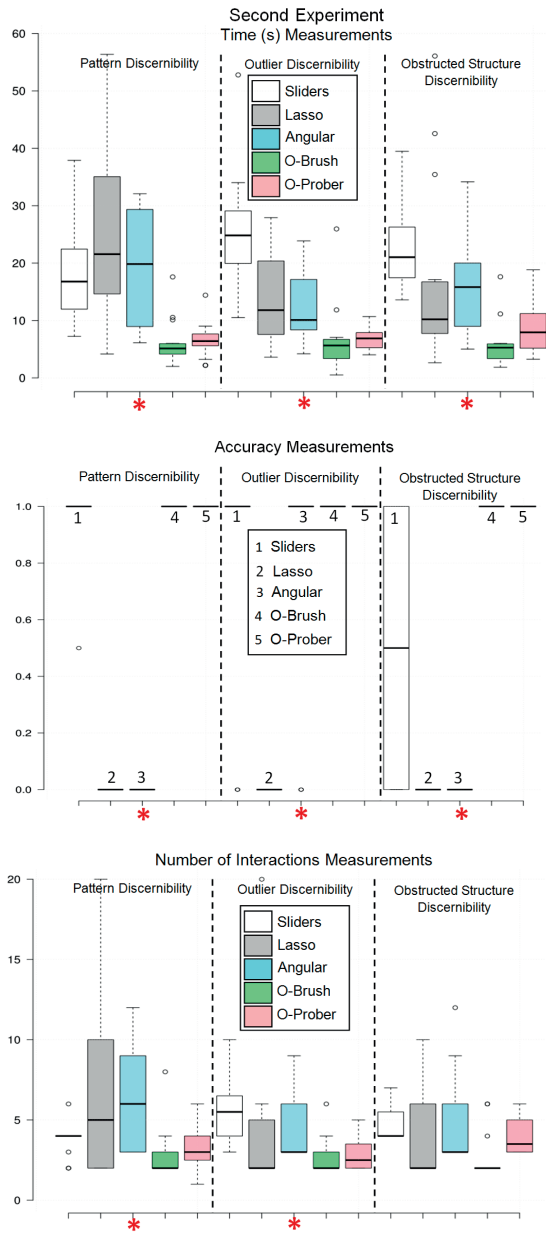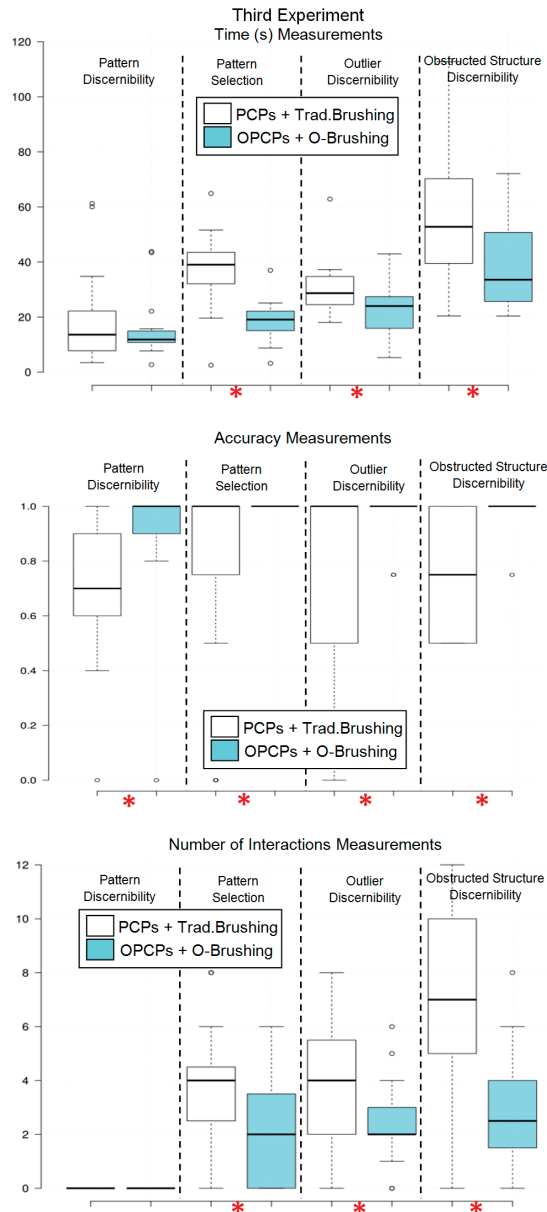
Figure 4.19: Results for the experiments conducted as part of the evaluation, for the first experiment of the User Performance (UP) part (performance in discerning patterns, outliers, and data structures obstructed by noise using density PCPs or OPCPs) with two datasets (D1, D2). The small white circles denote outliers in the measurements. The asterisks denote a statistically significant difference ($p < 0.05$) between the measurements, as it resulted from our statistical analysis using a paired t-test.
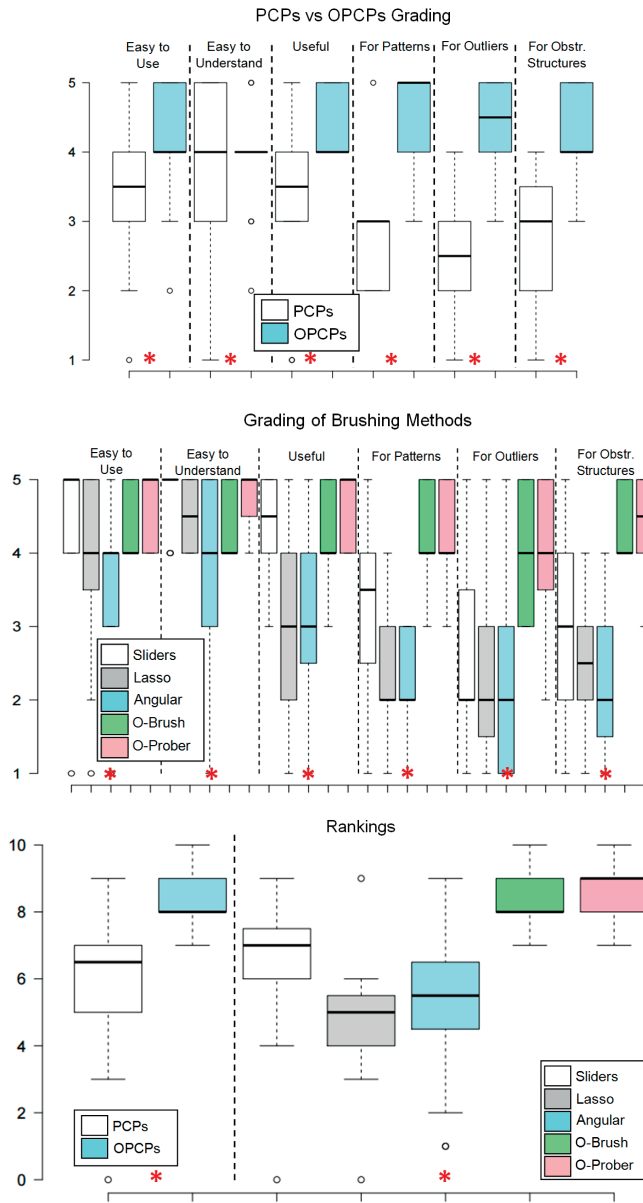
Figure 4.20: Results for the experiments conducted as part of the evaluation, for the second experiment of the User Performance (UP) part (performance in selecting specific patterns, outliers, and noise-obstructed data structures, with state-of-the-art brushing or O-Brushing). The small white circles denote outliers in the measurements. The asterisks denote a statistically significant difference (p <0.05) between the measurements, as it resulted from our statistical analysis using ANOVA and Tukey's HSD test.

Figure 4.21: Results for the experiments conducted as part of the evaluation, for the third experiment of the User Performance (UP) part (performance with multi-variate, complex data and tasks). The small white circles denote outliers in the measurements. The asterisks denote a statistically significant difference (p <0.05) between the measurements, as it resulted from our statistical analysis using a paired t-test.

Figure 4.22: Results for the experiments conducted as part of the evaluation, for the User Experience (UE) part. The asterisks denote a statistically significant difference ($p < 0.05$) between the measurements, as it resulted from our statistical analysis using ANOVA and Tukey's HSD test.

5

Visual Analytics for the Exploration of Imaging Modeling

Well, I must endure the presence of a few caterpillars,
if I wish to become acquainted with butterflies.

Antoine de Saint-Exupéry (1900 - 1944), The Little Prince

In this chapter, we present a novel technique for the representation, exploration and visual analysis of variations in imaging features, derived by means of pharmacokinetic modeling. Our approach supports the exploration of variations in the derived features, across modeling choices, in a single combined view. Particular emphasis is given on the association of observations from the feature space to patient anatomy. Chapter 5 is based on the paper:

■ Raidou, R.G., van der Heide, U.A., van Houdt, P.J., Breeuwer, M. and Vilanova, A., 2014. The iCoCooN: Integration of Cobweb Charts with Parallel Coordinates for Visual Analysis of DCE-MRI Modeling Variations. Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM 2014), pp. 11-20 [227].

5.1. Abstract

Efficacy of radiotherapy treatment depends on the specific characteristics of tumor tissues. To determine these characteristics, clinical practice uses Dynamic Contrast Enhanced (DCE) Magnetic Resonance Imaging (MRI), as already described in Chapter 2. DCE-MRI data is acquired and modeled using pharmacokinetic modeling, to derive per voxel a set of parameters that are indicative of tissue properties. However, different pharmacokinetic modeling approaches can be employed – thus, making different assumptions and resulting in parameters with different distributions or even in different parameters. A priori, it is not known whether there are significant differences under different modeling assumptions, and which assumption is best to apply. Therefore, clinical researchers need to know at least how different choices in modeling affect the resulting pharmacokinetic parameters and the locations, where parameter variations are occurring more frequently .

In this paper, we introduce the iCoCooN: a visualization application for the exploration and analysis of model-induced variations in pharmacokinetic parameters. We designed a visual representation, the Cocoon, by integrating in a perpendicular setting, Parallel Coordinate Plots (PCPs) with Cobweb Charts (CCs). PCPs display the variations in each parameter between modeling choices, while CCs present the relations in a whole parameter set, for each modeling choice. The Cocoon is equipped with interactive features to support the exploration of all data aspects, in a single combined view. Additionally, interactive brushing allows to link the observations from the Cocoon to the anatomy of the patient.

As part of this work, we conducted evaluations with experts and also with general users. The clinical experts judged that the Cocoon, in combination with its features, facilitates the exploration of all significant information and especially enables them to find anatomical correspondences for their observations from the feature space. The results of the evaluation with general users indicate that the Cocoon produces more accurate results, compared to independent small multiple views.

5.2. Introduction

As mentioned in the early chapters of this dissertation, the efficacy of radiotherapy is hypothesized to depend on the specific characteristics of tumor tissues. Tumor characterization requires the use of different imaging modalities, among which Dynamic Contrast Enhanced (DCE) Magnetic Resonance Imaging (MRI). These are time series of three-dimensional image volumes, which reflect the absorption of a contrast agent by tissues as a function of time after injection. After DCE-MRI data acquisition, clinical practice uses one of the established models to derive a set of output parameters per voxel, which is indicative of tissue characteristics. However, different modeling approaches require different assumptions or choices. Depending on these choices, the resulting parameters might present different values or might be fundamentally different.

It is difficult to decide beforehand whether different assumptions or choices lead to significant parameter differences and which assumption leads to a better description of the tumor tissues. Therefore, it is valuable for clinical researchers to explore the variability in the parameter values, as given by the different alternatives. In this way, they can identify which anatomical regions are affected more by the modeling choices and whether this has

an impact on the final clinical decision and treatment. Apart from exploring variations in the parameters independently, it is also important to know how the relationships between parameters change. Still, these relationships are complex and the effect of different assumptions on DCE-MRI modeling is difficult to predict. To the best of our knowledge, there is no tool that allows this kind of inspection.

In this paper, we introduce a new application: the iCoCooN (Integrated Cobweb Charts and Parallel Coordinate Plots for Visual ANalysis of DCE-MRI Modeling Variations). It is a visualization tool for the exploration and analysis of variability in the parameter values that result from different choices during DCE-MRI modeling. As part of the iCoCooN, we designed a new representation: the Cocoon. For this, we integrated Parallel Coordinate Plots (PCPs) with Cobweb Charts (CCs), which are also called Star or Spider Plots, Radar Charts or Kiviats in literature. In this way, we manage to simultaneously show different significant aspects of the data and to provide a more effective exploration. The interactive features of iCoCooN facilitate data exploration and improve anatomical interpretation.

After the design, we conducted an evaluation of iCoCooN with clinical experts, where we performed two case studies, with patient data. Due to the limited amount of field experts, we conducted an additional usability study with non-experts, to increase the statistical power of our evaluation. Our work presents the following *contributions*:

- The iCoCooN is a new interactive application that enables the visual analysis of DCE-MRI modeling variations and, especially, the association of the observations from the parameter space to the patient anatomy.

- The Cocoon is a visual representation within the iCoCooN, which results from the integration of PCPs with CCs. It enables the exploration and analysis of the DCE-MRI modeling information in a single combined view.

- The evaluations demonstrate the potential of the iCoCooN for the analysis of DCE-MRI modeling variations.

The remainder of this chapter is organized, as follows: Section 5.3 summarizes the necessary clinical background information. Section 5.4 provides an overview on previous work. Section 5.5 is the core of this chapter, where the design of our proposed approach is presented and discussed. Section 5.6 presents the results of our evaluations. Finally, Section 5.7 concludes the chapter with a discussion and propositions for future work.

5.3. Clinical Background

Dynamic-Contrast Enhanced (DCE) Imaging is a commonly used MRI technique in cancer diagnosis. The basic concepts of DCE-MRI data have already been described in Chapter 2. As a short reminder, these are 4D data, which depict the absorption and washout of a contrast agent (CA) in tissue over time, based on the idea that this process happens differently, in different tissues. Usually, pharmacokinetic (PK) models [153, 257, 272] are employed to derive per voxel a set of PK parameters, indicative of tissue characteristics, as shown in Figure 2.8 (Chapter 2).

Each PK model contains a number of assumptions or crisp choices [61, 152, 197, 283], which may affect the values of the resulting parameters, as shown in Figure 2.9 (Chapter 2).

Clinical researchers do not know a-priori whether there are significant differences among models or their assumptions, and which is the optimal to apply. Therefore, it is valuable for them to investigate how the derived PK parameters behave with different modeling choices. In this way, they can explore the impact of these choices on the precision of the treatment outcome. The current practice for this exploration has been shown in Figure 2.9 (Chapter 2). Here, the so-called PK parameter maps are explored visually side-by-side, making the whole procedure time consuming and limited in insight. More specifically, in a real-world analysis, clinical researchers are interested in performing the following tasks, at a voxel level:

- *Identification of Variability* (`T1`) – How does each one of the PK parameters vary across different modeling choices?

- *Identification of Relations* (`T2`) – What kind of relations develop within each of the PK parameter sets?

- *Comparison of Behaviors* (`T3`) – How modeling choices affect the relations within each of the PK parameter sets?

- *Detection of Patterns* (`T4`) – How modeling choices affect patterns within the data?

- *Association to Anatomical Reference* (`T5`) – What is the anatomical location of specific interesting behaviors or features?

5.4. Related Work

In this section, we present the most relevant work related to the iCoCooN. The exploration and analysis of perfusion parameters that characterize the shape of the DCE-derived enhancement curves and their correlation with the data has been presented in a survey of existing applications, by Preim et al. [218]. Yet, model-induced variability of PK parameters has not been addressed by any of these applications. The minimization of uncertainty in kinetic PET modeling parameters has been addressed in the work of Nguyen et al. [186], which allows the exploration of variabilities in the parameters. However, the capabilities to show relations between parameters and the effect of variability on these are limited.

Combining PCPs with CCs has also been tackled in other applications. In Temporal Stars [187], discrete multiples of radial graphs, with each axis representing a variable, are set along a central time axis to describe variation with time. This representation is useful for comparing different star glyphs, or for monitoring time evolution of the variables. As an extension, the 3D Kiviat [273] combines variable axes circularly arranged to a central time axis and a surface rendering around the 2D Kiviats for each time step to show the evolution of the 2D Kiviats in time. In this way, 3D Kiviats show correlations between attributes, while preserving the focus on time evolution.

The visualization proposed by Fanea et al. [83] also employs a combination of PCPs and Star Glyphs in a single configuration, to address clutter in the former. Each Star results from PCP polylines unfolding around the central axis and depicts a data item or a data dimension, thus maintaining the total number of dimensions. The PCPs of this visualization do not provide additional dimensions with additional information or insight in data relations or

patterns. Finally, in VisLink [56], inter-relationships between multiple visualizations can be interactively explored, with the form of multiple 2D layouts positioned in space and linked together to show data associations.

Our approach combines specific attributes of the previous work with new ones, to accommodate all the requirements of our application, as mentioned in Section 5.3. We keep the Star Glyphs for the exploration of relations and behaviors in each parameter set [83, 187, 273], but we provide additional dimensions with PCPs for the exploration of trends and relations in each one of the parameters across these different choices. In this way, the PCPs provide additional information and are not restricted to time representation [187, 273] or to linking multiple representations [56]. Finally, we incorporate functionality to link observations from the parameter space to the anatomical space, which is not provided in any previously published work.

5.5. The Design of the iCoCooN

After DCE-MRI acquisition, different output parameters are derived per voxel, through PK modeling. At this point, different established models and/or several assumption alternatives can be considered. In order to visualize all data aspects, as described in Section 5.3, we employ the workflow proposed in Figure 5.1.

For the visualization of each one of the PK parameters across the different choices, we decided to employ *PCPs* [128]. This representation allows the user to visualize multiple data dimensions in limited space and to detect trends and patterns. In our case, each line in the PCPs corresponds to a location in the medical volume and each dimension to a PK parameter derived from different modeling choices, as depicted in Figure 5.2. This is a suitable choice for the identification of model-induced variability in each of the PK parameters.

For the visualization of each modeling choice, we decided to employ *CCs* [50]. This compact iconography representation combines the advantages of PCPs with glyphs, for the easy detection of patterns among different plots [143]. This is based on the human perceptual ability to easily discern shape differences. In the present application, each line in the CCs corresponds to a location in the medical volume and each dimension to a PK parameter de-
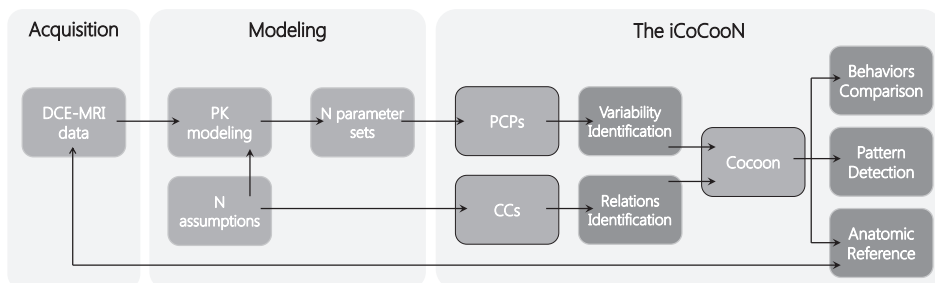
Figure 5.1: Workflow considered for the construction of the iCoCooN. With dark grey, we denote the five requirements discussed in Section 5.3, for the visualization and exploration of the required data aspects.
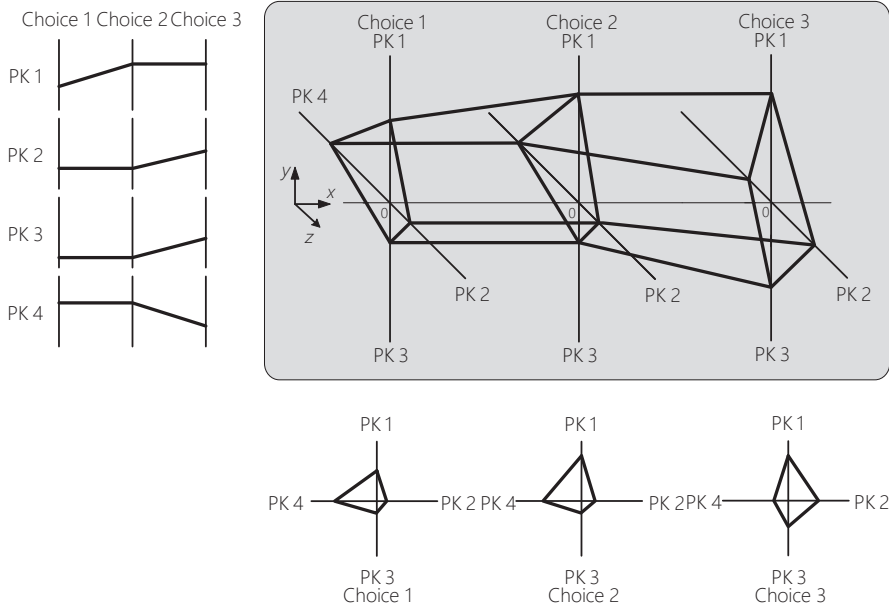
Figure 5.2: Configuration of Cocoon from perpendicular integration of PCPs and CCs. Here, we consider four PK parameter sets (PK 1-4), from three different modeling assumptions (Choice 1-3).

rived from one modeling choice, as depicted in Figure 5.2. It is a proper choice for comparing qualitatively a whole PK parameter set against another, or for locating relations, similar and dissimilar behaviors or outliers in a parameter set.

Using a single PCP to encode the PK parameter values, and an additional discrete axis for the different modeling choices, would not be an appropriate option. This solution would add clutter in the representation, due to an increasing number of overlapping lines from the different modeling choices. Although this limitation could be solved with the technique presented in Chapter 4, an additional issue with this solution is that it would only provide an overview on global shifts in the parameters, disregarding spatial relationships between parameter sets. Therefore, for visualizing and linking both intra- and inter-model information, both PCPs and CCs are needed. However, if many variables are involved in the exploration, independent multiple PCPs and CCs might be visually cumbersome, due to the amount of windows. For these reasons, we decided to create a new representation, the *Cocoon*, from the perpendicular integration of PCPs and CCs in a single 3D view with orthographic projection [180].

With the 3D Cocoon, we facilitate the exploration, by conveying different views of the data in an equal context. This is also one of our main differences with respect to previous work [56, 83, 186, 187, 218, 273]: the PCPs of the Cocoon are not linking multiple CC representations, but they show additional dimensions of the data. In this way, we facilitate linking between parameters and modeling choices in a compact view, without forcing the user to use his memory during tasks, as in the case of the independent small multiples.

Figure 5.2 depicts the concept behind the configuration of the Cocoon, when we have four parameters to compare across three assumptions. For our application, we need to show a limited number of different modeling assumptions and a maximum of six independent parameters, which is the maximum number of parameters encountered in the two-Compartment Exchange Model [257], previously described in Chapter 2. The optimal number of parameters is four, but for more than four parameters in the CCs, the user can selectively switch on and off axes in the representation. Therefore, the scalability of the Cocoon is adequate for this concrete application, and we consider out of the scope of this paper to study higher dimensionality. Figure 5.3 shows an implementation of the Cocoon for the concept of Figure 5.2.

At this point, we needed to address three main issues with respect to the Cocoon: the complexity of interaction, eventual perception constraints due to limited short term memory [206] and clutter [64, 112, 180]. For the first issue, we facilitate interaction by reducing the degrees of freedom. Not all orientations of the Cocoon are sensible, so the user can rotate the Cocoon only around the latitudinal and longitudinal axes, represented in Figure 5.2 by x and y, to adapt the view. For the second issue, the specific tasks in our application do not require to rely on short term memory. However, the user can still selectively show the independent multiple PCPs and CCs, to clarify information from the Cocoon. The small multiple PCPs and CCs are linked to the Cocoon and viceversa [180, 298].

For the last issue, we increase the visibility of the polylines of the Cocoon using low alpha values, but also a colormap based on the lines density, as shown in Figure 5.3. The density colormap enables better discernibility of overlapping bundles, giving an impression of texture and a first idea about the variability of the data. For example, high density lines that highly vary throughout the assumptions may indicate that the assumptions have a strong impact

5
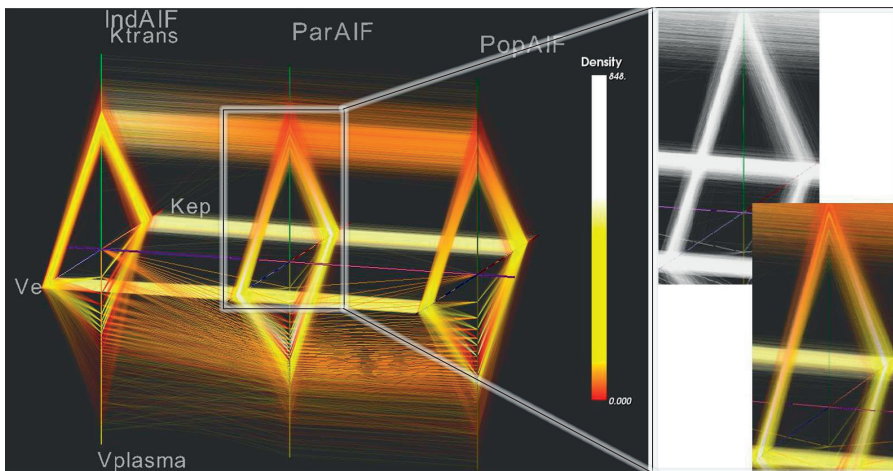
Figure 5.3: An implementation of the Cocoon for four PK parameters (K^{trans}, v_e, v_{plasma} and k_{ep} [271]) and three assumptions (IndAIF, PopAIF and ParAIF). In the zoomed view, a density colormap is employed for better line visibility. Here, red denotes very low density or outliers, yellow low density and white high density.

on parameters. On the other side, very low density lines that highly vary throughout the assumptions may indicate outliers. Other methods to improve visibility and reduce clutter, such as the ones described in Chapter 4, could have also been employed, but adjustments to suit the 3D design of the Cocoon should be first considered.

Clinical researchers usually look at specific combinations of PK parameters, which we offer as the default view in the axes configuration. To improve the view and to increase visibility further, we use a number of interaction features. *Axes reordering* often affects positively the view and the ability to see relations and trends in the data [293], while *scaling* allows to pull apart dense parts of the representation for better discernibility. Although automatic algorithms for axes reordering can be helpful, given the dimensionality of our application, they were not considered further.

Additionally, *band coloring* allows the user to extend a primitive clustering, based on the values of one dimension of the Cocoon, to all. We employ a divergent colorblind safe colormap from ColorBrewer [44], in order to differentiate between low, medium and high parameter values. Figure 5.4 - a shows how band coloring aids the rough detection of trends in the data, visualized as color bands. Moreover, *brushing* allows the users to explore, analyze and detect interesting trends and relations by selecting areas or values of specific interest in the Cocoon [143, 295], as presented in Figure 5.4 - b. Finally, *linking* the brushed observations to the medical data, by highlighting the corresponding regions in a 2D slice viewer and a 3D model, establishes correspondence to the patient anatomy, as shown in Figure 5.4 - c, d.

■ We implemented the iCoCooN in Python as a DeVIDE module [32], employing the Visualization Toolkit (VTK) [2].

5.6. Evaluation

In order to assess the value of the iCoCooN, we conducted an evaluation, inspired by the article of Lam et al. [159]. It consists of two parts. First, we performed an evaluation of the iCoCooN with clinical researchers. Then, we conducted a general evaluation to test the usability and effectiveness of the Cocoon and to increase the statistical power of the evaluation, given the limited availability of field experts. For the second part, we abstracted the tasks from the clinical field to a more general domain, so that they could also be performed by non-experts.

5.6.1. Evaluation with Clinical Researchers

In order to evaluate whether the designed visualization meets its requirements, we conducted individual evaluations with intended users. We included four field experts from two different institutions (Netherlands Cancer Institute and Aarhus University Hospital), representing two types of clinical researchers: two clinical physicists and two biomedical engineers. As a proof of concept, we used two cases. First, we employ a prostate tumor case, where clinical researchers want to explore and analyze the effect of different clinically established choices of arterial input function (AIF) [61, 152, 198] within the same model on the PK parameters and their in-between relationships. Second, we use a cervical tumor case, where clinical researchers want to explore and analyze how the parameters and their in-between relationships vary when derived using three different established models [153, 257, 272]. These

datasets were provided by the clinical researchers and are described in Table 5.1.

The evaluation consisted of four phases. The third phase was task-specific, while the others were general and aimed at clarifying three broader aspects:

- Does the proposed Cocoon offer new understanding in the data, in comparison to current practice and in comparison to the independent multiple PCPs and CCs? If yes, how? If no, why?

- Do the features (axes reordering, scaling, band coloring, brushing and linking with respect to the anatomy) of the Cocoon contribute to the visualization and facilitate cognition? If yes, how? If no, why?

- Does brushing and linking contribute to the exploration and interpretation of the data? If yes, how? If no, why?

In the *first phase* of the evaluation, we simulated the visual environment for the exploration of the PK parameter space in prostate and cervical data. In this phase, we were interested in obtaining an initial opinion on the individual features of the iCoCooN. This part
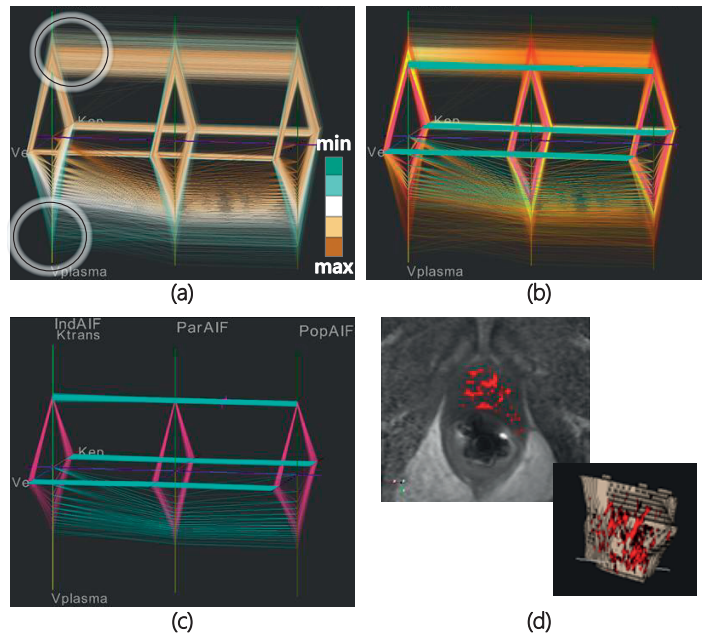
Figure 5.4: The interactive features of the iCoCooN. (a) Band coloring on the v_{plasma} axis of the first AIF for the identification of trends in the data. The zero values of v_{plasma} for the first AIF (brown lines), correspond to the middle values of K^{trans}, in the circle. The legend shows the employed colormap. (b) Brushing the Cocoon. Cyan brush in PCPs and magenta in CCs are employed to reduce occlusion due to overlapping polylines and glyphs. (c) Visualization of the brushed bundles from (b) only. (d) Linking to 2D anatomy and to 3D model.

Table 5.1: Description of the datasets used in the evaluation with clinical researchers. We work on the whole prostate region, but only on the specific tumor region in the cervix, due to the high anatomical variability of the organ between individuals.

	DCE-MRI data (4D)				Modeled data (3D)			
	Resolution (voxels)	Voxel size (mm³)	Time resolution (s)	ROI size (voxels)	Modelling choices	PK maps	PK maps resolution (voxels)	PK maps voxel size (mm³)
Prostate dataset	256×256× 20 for 120 timepoints	1 × 1 × 3	2.5	80×80×20 (usual prostate size)	3 AIF alternatives	4 3D maps per choice	256×256× 20 (as in DCE-MRI)	1×1×3 (as in DCE-MRI)
Cervix dataset	176×176× 20 for 120 timepoints	2.273 × 2.273 × 3	2.5	33×33×18 (specific tumor size)	3 different models	3, 4 and 6 3D maps per model	176×176× 20 (as in DCE-MRI)	2.273× 2.273 × 3 (as in DCE-MRI)

was also used as training, so the tool was initially operated by the first author, while the test subject first observed a demonstration and then explored the functionality. We asked the test subjects with a questionnaire to comment on the clarity and potential usefulness of each one of these features, but also to quantify their value using a grading scale $(1-5)$. The *second phase* required from the test subjects to grade $(1-5)$ the ability to identify variations and relationships in the data, but also to relate them to the anatomy and pathology of the patient. The *third phase* required a more detailed hands-on exploration of the data, aiming at evaluating the insight provided by the iCoCooN. In this phase, the test subjects operated the tool themselves and they were asked to explore the data in the iCoCooN and to analyze their observations, as they would do in a real case, performing the tasks of Section 5.3. In the *last phase*, they evaluated the tool as a whole, based on their experience, commenting also on the strengths, limitations and missing features of the iCoCooN.

First Phase: the iCoCooN Features

In the first phase, the test subjects evaluated the features of the iCoCooN individually, following a questionnaire. The quantitative results from this evaluation are summarized in Table 5.2, with a convention of 1 for negative to 5 for positive. All features of iCoCooN received scores above 4, apart from two cases with a neutral grade (3), where the subjects commented that they needed to form a more concrete opinion by performing actual tasks.

Using a questionnaire, we also asked for general comments on each one of the features. First of all, the *Cocoon* was considered understandable and relatively easy to use. The test subjects confirmed that it enables the identification of variations, relations and trends in the multidimensional data in a combined view, even for parameters that do not seem to have an obvious association in the independent multiple views. The *restricted manipulation of the Cocoon* in space is according to the test subjects appropriate. Yet, the Cocoon requires training for learning how to obtain the most adequate view and to interpret the conveyed information. They were inclined, though, to say that there is no need to additionally inspect the *independent multiples*. The independent multiples might be used selectively, only for details-on-demand or for easier tasks that involve single parameters. However, they stated that they needed to confirm this impression with a more exploratory task.

Table 5.2: First phase of the evaluation: Evaluation of the features of iCoCooN by the four clinical researchers, using a grading scale (1=negative, 5=positive).

	#1	#2	#3	#4
Cocoon Concept	3	4	4	4
Cocoon Manipulation	5	4	5	4
Cocoon instead of Small Multiples	5	3	4	4
Scaling/Reordering	4	5	5	4
Band/Density Coloring	4	4	5	4
Brushing/Linking	4	5	5	5

The advantages of *reordering* and *scaling* the Cocoon axes are also straightforward for data interpretation, while *band coloring* provides a visual context of how the values of one parameter behave with respect to the rest and aids the detection of basic patterns in the data. The application of the *density* colormap was considered useful for distinguishing overlapping lines and for deciding on the importance of the variations, relations or trends. This feature requires training, but was regarded as easy to learn. Finally, according to the test subjects, currently there is no tool with *brushing and linking* functionality for their purposes. They described it as potentially easy to learn and use – an appropriate and appealing feature for data exploration, especially for relating observations from the Cocoon to the patient anatomy.

Second Phase: Information Identifiability

In the second phase, the test subjects had to grade (1-5) the ability to identify specific information using the iCoCooN, namely the variability and relations or trends, and the ability to relate findings from the iCoCooN to anatomy. The quantitative results from this evaluation are summarized in Table 5.3, with a grading convention of 1 for easy to identify to 5 for difficult. All but one gradings were below 2, indicating that the test subjects considered the three tasks easy. In only one case, the relation to the anatomy received a neutral grade (3), as the test subject explained that he would need to interact with the Cocoon more, to form a more concrete opinion.

Table 5.3: Second phase of the evaluation: Evaluation of the ability to identify information in iCoCooN by clinical researchers, using a grading scale (1=easy, 5=difficult).

	#1	#2	#3	#4
Ability to identify variability	2	2	1	2
Ability to identify relations in the features	2	1	2	2
Ability to identify relations to anatomy	2	1	2	3

Third Phase: Case Studies

In this phase, the test subjects used iCoCooN for the exploration and analysis of a prostate dataset and a cervical dataset from the DR THERAPAT project, as they would do in a real case, executing the tasks of Section 5.3. The patient-specific findings presented below are observations of the clinical researchers.

■ *Case study – Prostate.* The test subjects were interested in exploring how the four PK parameters of the extended Tofts model (K^{trans}, v_e, v_{plasma} and k_{ep}) [272] vary across three different AIF choices (Individual, Population-based and Parker-based AIF) [61, 152, 198]. The AIF is used as an input assumption to the model and there are different options for its shape [152]. The behavior of the parameters for the different AIF choices, presented these major patient-specific findings:

- *Identification of Variability* (`T1`) – The K^{trans} parameter, which relates to the tissue permeability, remains highly unaltered throughout the different AIF choices, as shown in Figure 5.5 - a:1. Minor changes in the distributions are reflected by changes in the density colors. The v_{plasma}, which relates to the blood plasma volume, presents significant variations throughout the AIF choices, depicted in Figure 5.5 - a:2.

- *Identification of Relations* (`T2`) – Lower and slightly decreasing v_e, which relates to the volume of the extracellular extravascular space, is associated to slightly increasing k_{ep} and lower values of K^{trans}, as presented in Figure 5.5 - a:3.

- *Comparison of Behaviors* (`T3`) – Although the general behavior of the parameter sets seems stable across the AIF choices, the differences in the density colors of the CCs reflect slight changes in K^{trans}, v_{plasma} and k_{ep}, which can be seen in Figure 5.5 - a.

- *Detection of Patterns* (`T4`) – The highest values of K^{trans} are related to high values of k_{ep} and v_e in all three AIF choices, as shown in Figure 5.5 - b:1. This is also related to values of v_{plasma} that are highly variable between the AIF choices, visible in Figure 5.5 - b:2.

- *Exploration and Anatomical Reference* (`T5`) – Low v_e and low K^{trans} regions reflect the necrotic core of the tumor and some outliers at the border of the prostate, which is depicted in Figure 5.5 - a:3,4.

■ *Case study – Cervix.* The test subjects explored how the application of different models (Tofts model: TM, Extended Tofts model: ETM and Two-Compartmental Exchange model: 2CXM) [257] affects the behavior of the respectively derived PK parameters in the tumor. The first model results in two parameters (K^{trans} and v_e). The second has three parameters (K^{trans}, v_e and v_p). The third gives five parameters (K^{trans}, v_e, v_p, F_p and PS) [257]. An opposite Akaike information criterion (AIC_c) is additionally included to each PK set. The AIC_c relates to the relative quality of fit of each model. High values of AIC_c mean high relative quality. The test subjects identified the following major patient-specific findings:

- *Identification of Variability* (`T1`) – The K^{trans} parameter does not present significant changes despite the application of different models, apart from slight decreases between the ETM and the 2CXM, as can be seen in Figure 5.6 - a:1. This, together with

(a)

5
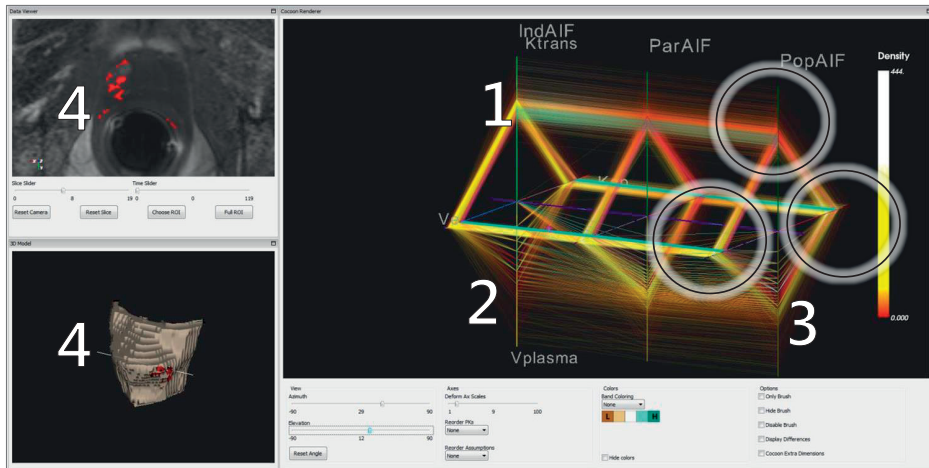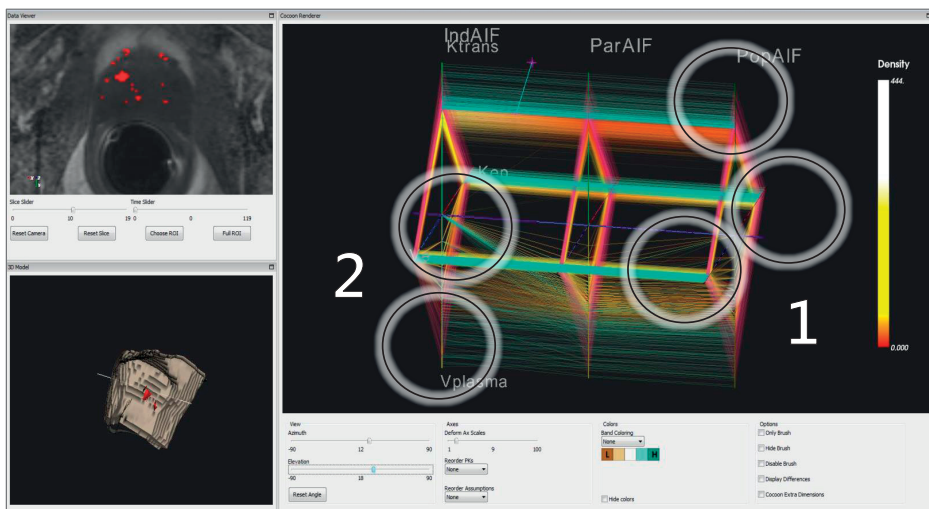
(b)

Figure 5.5: Exploration and visual analysis of a prostate tumor case with iCoCooN during the third phase of the evaluation with clinical researchers.

high and stable AIC_C in the TM and ETM models, is an indication that these models had a reasonable fit. This finding can be seen in Figure 5.6 - a:2.

- *Identification of Relations* (`T2`) – In the 2CXM, high AIC_C is highly correlated with high flow F_p and permeability-surface area product PS, as depicted in Figure 5.6 - b.

- *Comparison of Behaviors* (`T3`) – Regions with lower K^{trans} and lower v_e present lower v_p in the ETM and 2CXM, as shown in Figure 5.6 - a:3. In the ETM model, v_p presents a big spread. This is bigger than in the 2CXM, as presented in Figure 5.6 - a:4,5. This variability in the v_p values of the 2CXM is related to the slight variability in K^{trans} values, which is visible in Figure 5.6 - a:5.

- *Detection of Patterns* (`T4`) – In regions, where the K^{trans} and the v_e values are low, the AIC_C values do not present significant changes across the three models. This is shown in Figure 5.6 - a:2,3. This finding means that clinical researchers expect that they will all give similar classifications.

- *Exploration and Anatomical Reference* (`T5`) – A usual indication of malignancy is the combination of low v_e and low K^{trans}. This is the necrotic core of the tumor, shown in Figure 5.6 - a:6, which has the worst responses to treatment. In these parts, the model that fits better is TM, which is confirmed by slightly higher AIC_C values in the Cocoon.

Fourth phase: the iCoCooN Overall

In the last phase of the evaluation, the test subjects commented on the iCoCooN overall and, mainly, on its strengths, limitations and missing features. According to their experience, they judged the application as useful: the Cocoon in combination with its features provides the necessary information and the user requirements are met. They were also asked to compare the tool to current clinical practice, namely, the slice-based inspection of all alternative parameter maps. To that, they commented that current practice is mainly manual and mental work, which hampers the identification of variabilities and relations in the data, since it requires from them to go through the multiple slices of all the parameter maps. Instead, the Cocoon is versatile in showing multiple dimensions in one view and in aiding their investigation. Thus, it enables them to see the consequences of each modeling choice, to perform and analyze selections and see their exact relation to the anatomy.

Although the first results of the evaluation are promising, they also exposed some limitations. The iCoCooN is a tool that needs training and time to learn. This is also supported by the statements of the test subjects in the previous phases of the evaluation. They agreed that once the user becomes familiar with the visual mapping of the dimensions, the exploration and analysis is faster and easier than with current practice. In order to improve this further, semi-automation of the data exploration and analysis would be required. Moreover, the iCoCooN misses the functionality for the inspection of the related enhancement curves, in order to quantitatively see the absorption of the contrast agent in the tissue. This would give additional information on the physical meaning of the visualized data and their patterns.
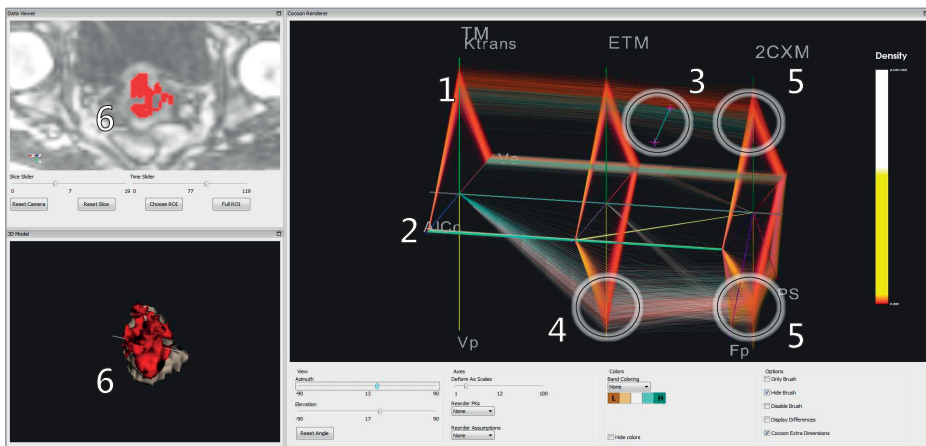
(a)
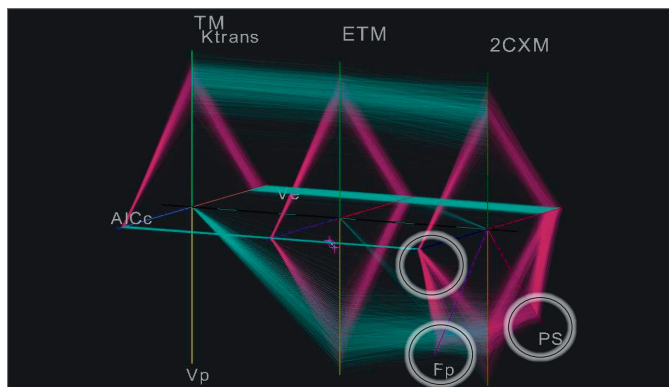
(b)

Figure 5.6: Exploration and visual analysis of a cervical tumor case with iCoCooN during the third phase of the evaluation with clinical researchers.

5.6.2. Usability and Effectiveness Evaluation

Our application is built for a specific target group and, so far, we have only conducted user studies with a limited amount of experts. In order to strengthen our evaluation, we decided to perform a controlled study with non-experts, particularly for the usability and effectiveness of the Cocoon, compared to independent multiples. For this, we translated the clinical tasks of Section 5.3 to two domains more accessible to general users. The generalization of the tasks is described in Table 5.4. As we are outside the clinical domain, linking to the anatomy, which was already positively judged by clinical researchers, is not tested here.

Table 5.4: Translation of clinical tasks to tasks with the synthetic datasets, for the controlled study.

	Clinical Study (with experts)	Controlled Study (with general users)
Data Dimensions	A number of PK parameters for different modeling assumptions.	A number of parameters from different measurements (e.g. in weather forecast from different weather stations).
Identification of Variability (T1)	Identify which of the PK parameters remains unaltered/presents variability throughout the different modeling assumptions.	Identify which of the parameters remains unaltered/presents variability throughout the different measurements.
Identification of Relations (T2)	Find a relationship or trend between two or more parameters throughout the different modeling assumptions.	Find a relationship between two or more specific parameters throughout the measurements.
Comparison of Behaviors (T3)	In which of the assumptions do we encounter a specific behavior of two or more parameters?	In which of the measurements do we encounter a specific behavior of two or more parameters?
Detection of Patterns (T4)	Find a specific pattern in the parameter behaviors. How does it change throughout the assumptions?	Find a given pattern in the parameter behaviors. How does it change throughout the measurements?
Exploration	e.g. Which assumption do you trust more for each zone of the tumor based on all the parameters?	e.g. Which measurement do you choose if you demand specific characteristics?

For the remaining tasks, we created two synthetic, but realistically sized datasets using the PCDC tool [34]. Since we focus on evaluating the Cocoon for our concrete application, we use similar dimensions to the ones of our case. Exploring the limits of scalability of the representation is not expected to be of interest in our application. Therefore, it was considered out of the scope of this paper and was not tested further.

We had 15 test subjects: 8 females and 7 males, between 23 and 44 years old. All of them had normal vision, with or without glasses and none of them was colorblind. Their background included electrical engineering (4), computer sciences (3), biomedical engineering (2), mechanical engineering (2), chemistry (2), hydraulics engineering (1), and mathematics (1). They described their computer expertise as medium (9) to high (6). Only two test

subjects had previous familiarity with PCPs and CCs.

First, we gave them a small *introduction* to explain the reasoning behind the study and to present notions such as PCPs, CCs and the Cocoon. Secondly, we demonstrated the basic *functionality* of the Cocoon. Thirdly, we did some *first exercises*, until the test subjects were confident with the visualizations. Fourthly, we conducted two *experiments* with the synthetic datasets. For each one of the experiments, the test subjects needed to perform the five tasks given in Table 5.4 as fast and accurate as possible, once using only the independent multiples and once only the Cocoon. We exchanged the order of the two representations both between the experiments and across users, to avoid bias. For all tasks, we measured completion times and correctness of answers, which were known to us. Finally, we asked the test subjects to complete a small *survey*, consisting of a grading scale $(1-5)$, an overall ranking $(1-10)$ and open questions.

Results of the Evaluation with General Users

The completion times and correctness for the five tasks of the two experiments, as well as the gradings of the test subjects are summarized in Tables 5.5 and 5.6, respectively. In these tables, we color-encoded their statistically significant difference, as it resulted from t-tests analysis. For the more demanding tasks, i.e. comparison of behaviors, detection of patterns and exploration, the observations from the Cocoon were more *correct* than the independent multiples in both experiments $(p < 0.05)$. The *time* difference was in favor of the independent multiples, but not statistically significant between the two representations $(p>0.05)$, except for the variability task in the second experiment. Additionally, there is a statistically significant difference between the *gradings* in the two representations in favor of the Cocoon, which was judged more suitable for the more demanding tasks and for the identification of relations task $(p < 0.05)$. Also, it was considered less overloading than the independent multiples $(p \ll 0.01)$. Both representations, were considered comparably useful, easy to use and understand $(p > 0.05)$. In the *overall ranking*, the Cocoon comes first with a difference of 1.67 points $(p = 0.0012)$, while only two test subjects ranked the independent multiples higher. Nine test subjects commented that they would not need the independent multiples at all, on top or instead of the Cocoon. The rest commented that the independent multiples could be useful and faster in certain instances, such as simple tasks, for comparison of few parameters or for beginners.

In their *general comments*, the test subjects stated that when higher dimensionality is involved in the tasks, they preferred the Cocoon, because they could check everything in a single, compact view (*The Cocoon offers compact information giving a more intuitive understanding of complex relations. With the independent multiples, I had to compile all the information from multiple windows in my mind first*). They also stated that Cocoon made them more attentive, efficient and eventually more accurate (*I felt that I could be faster with PCPs. However, I realized that I was jumping easily to conclusions and making more mistakes, because I was not paying attention to the multiple relations that affected my observations*). Yet, choosing the most effective view is time demanding and requires training (*I needed to think and learn which was the most effective view. Getting the correct view of the Cocoon takes time, The Cocoon has a higher learning curve, but can provide more information at a glance with adequate training*).

5.7. Conclusions and Future Work

In this chapter, we presented the iCoCooN: a visualization tool that aids clinical researchers to explore and analyze how different choices in modeling affect the parameter space derived from modeling DCE-MRI data. The contribution of our work lies within the design of the Cocoon that allows users to explore the required DCE-MRI data aspects in a single combined view. The interactive features of the tool facilitate the exploration and interpretation of the data and, especially, the correspondence to anatomy.

The value of the iCoCooN for our application was confirmed by an evaluation with clinical experts. An additional evaluation with general users indicated that the Cocoon produces more accurate results compared to the independent PCPs and CCs, especially for more complex tasks. At any case, adequate training of the users is essential. The evaluations also provided feedback towards future work. The tool can still improve by reducing interaction workload and time for adjusting the view. In addition to this, reducing clutter in the Cocoon with lines illumination or bundling [123] or the incorporation of OPCPs (Chapter 4) could further be considered. Finally, although the tool can be extended to other similar applications, scalability needs to be examined by additional evaluations. First indications show that iCoCooN has good potential of use for the easier exploration and analysis of model-induced variations in DCE-MRI data.

5

Table 5.5: Results from the evaluation with general users. We present the mean (μ) and standard deviation (σ) of completion times (t in s) and task correctness ($corr$, 0=wrong to 1=correct) for the two experiments. Bold indicates lower time and higher correctness, while color relates to statistically significant difference (dark green: $p < 0.01$, green: $p < 0.05$, orange: $p > 0.05$).

First Experiment	Cocoon				Independent Multiples			
	μ_t	σ_t	μ_{corr}	σ_{corr}	μ_t	σ_t	μ_{corr}	σ_{corr}
Variability (T1)	55.67	31.35	**1.00**	0.00	**31.47**	26.26	0.97	0.13
Relations (T2)	**46.80**	30.90	**1.00**	0.00	46.87	30.01	**1.00**	0.00
Behaviors (T3)	61.60	25.04	**0.87**	0.35	**56.00**	35.99	0.40	0.57
Patterns (T4)	93.67	39.52	**0.73**	0.46	**82.73**	45.14	0.33	0.49
Exploration	**66.47**	26.95	**0.97**	0.13	79.27	36.59	0.73	0.32
Second Experiment	μ_t	σ_t	μ_{corr}	σ_{corr}	μ_t	σ_t	μ_{corr}	σ_{corr}
Variability (T1)	24.60	9.92	**1.00**	0.00	**23.60**	9.68	0.80	0.17
Relations (T2)	46.87	27.22	**0.93**	0.26	**38.60**	30.67	**0.93**	0.26
Behaviors (T3)	**41.60**	14.21	**1.00**	0.00	45.20	25.51	0.73	0.46
Patterns (T4)	77.73	34.29	**0.93**	0.27	**77.00**	33.47	0.67	0.49
Exploration	**38.53**	17.18	**1.00**	0.00	44.40	27.49	0.87	0.23

Table 5.6: Results from the evaluation with general users. We present the mean (μ) and standard deviation (σ) of gradings $(1-5)$ and the overall rankings $(1-10)$ for the two representations. Bold indicates higher grading, while color relates to statistically significant difference (dark green: $p < 0.01$, green: $p < 0.05$, orange: $p > 0.05$).

Characteristics	Cocoon		Independent Multiples	
	μ	σ	μ	σ
Easy to Use	3.87	0.74	**4.33**	0.62
Easy to Understand	3.73	0.80	**4.13**	0.74
Useful	**4.13**	0.74	3.87	0.74
Suitable for (T1)	4.13	0.83	**4.20**	0.68
Suitable for (T2)	**4.20**	0.68	3.73	0.70
Suitable for (T3)	**4.47**	0.74	3.87	0.99
Suitable for (T4)	**4.27**	0.70	3.53	0.83
Suitable for Exploration	**4.73**	0.59	2.73	1.03
Not Overloading	**4.27**	0.70	2.27	1.03
Overall Ranking	**8.13**	1.36	6.47	0.99

6

Visual Analytics for the Exploration of Tissue Characterization

Chapter 6 is divided into two parts. In the first part (Part I), we propose a visual analysis tool that enables clinical researchers to identify, explore and analyze the tissue characteristics of distinct intra-tumor regions. The second part (Part II) is an extension of the application of this visual analysis tool, with the purpose of enabling a more insightful design of tissue classifiers. The latter was performed outside of the field of prostate cancer to demonstrate the breadth of applications that can be addressed with the proposed tool. Chapter 6 is based on the following publications:

Part I:

■ Raidou, R.G., Van Der Heide, U.A., Dinh, C.V., Ghobadi, G., Kallehauge, J.F., Breeuwer, M. and Vilanova, A., 2015. Visual analytics for the exploration of tumor tissue characterization. In Computer Graphics Forum (CGF), Vol. 34, No. 3, pp. 11-20 [226].

■ Vilanova, A., Raidou, R.G., and Pezzotti, N., 2015. Visual Analysis for Hypothesis Generation in Medical Imaging Research. Proceedings of the 9th MedViz Conference (MedViz 2015), pp.11-12 [289].

■ Raidou, R.G., Moreira, M.P., van Elmpt, W., Breeuwer, M. and Vilanova, A., 2014. Visual analytics for the exploration of multiparametric cancer imaging. In IEEE Conference on Visual Analytics Science and Technology (VAST) 2014, pp. 263-264 [225].

Part II:

■ Raidou, R.G., Kuijf, H.J., Sepasian, N., Pezzotti, N., Bouvy, W.H., Breeuwer, M., and Vilanova, A., 2016. Employing Visual Analytics to Aid the Design of White Matter Hyperintensity Classifiers. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Proceedings, Part II, Springer International Publishing, pp. 97-105 [223].

6

Part I – Exploration and Analysis of Tumor Tissue Characterization

6.1. Abstract

Tumors are heterogeneous tissues consisting of multiple regions with distinct characteristics. Characterization of these intra-tumor regions can improve patient diagnosis and enable a better targeted treatment, where radiation doses are chosen adequately for each region. Ideally, tissue characterization should be performed non-invasively, using medical imaging data to derive per voxel a number of features, indicative of tissue properties. However, the high dimensionality and complexity of this imaging-derived feature space prohibits easy exploration and analysis – especially when clinical researchers require to associate observations from the feature space to other reference data, such as features derived from histopathological data. Currently, the exploratory approach used in clinical research consists of juxtaposing these data, visually comparing them and mentally reconstructing their relationships. This is a time consuming and tedious process, from which it is often difficult to obtain the required insight.

We propose a visual tool for the easy exploration and visual analysis of the feature space of imaging-derived tissue characteristics. In addition to this, the tool enables knowledge discovery and hypothesis generation and confirmation, with respect to reference data used in clinical research. We employ, as central view, a reduced 2D embedding of the imaging-derived feature space. Multiple linked interactive views provide functionality for the exploration and analysis of the local structure of the feature space, enabling linking to patient anatomy and other clinical reference data. We performed an initial evaluation with ten clinical researchers. All participants agreed that, unlike current practice, the proposed visual tool enables them to identify, explore and analyze heterogeneous intra-tumor regions and to generate and confirm hypotheses, with respect to clinical reference data.

6.2. Introduction

As described in Chapter 2, tumors are heterogeneous tissues, enclosing multiple regions with distinct characteristics. Incorporating these distinct tissue characteristics information into radiotherapy planning can play an important role in tumor diagnosis and in designing more effective treatment strategies. For example, better targeted radiotherapy planning, where distinct intra-tumor tissues are irradiated with tailored doses, can be achieved.

Currently, intra-tumor tissue heterogeneity is investigated by studying histopathological data acquired from biopsies. To substitute histopathology, clinical researchers need to associate histopathological findings, such as aggressiveness or resistance of a part of the tumor, to *features derived from imaging data*. For example, Dynamic Contrast Enhanced (DCE) and Diffusion Weighted (DW) Magnetic Resonance Imaging (MRI) are employed in tumor analysis to derive per voxel, using mathematical models, features indicative of tissue characteristics. In addition to histopathology, clinical researchers often employ in their analysis *supplementary clinical data* that they use as reference, such as maps that predict high-risk tumor zones or maps depicting tumor control probability. These data also need to be explored in association with the imaging-derived features.

Moreover, the derivation of tumor tissue characteristics from imaging data is based on complex mathematical models [257]. Different modeling approaches make different assumptions, resulting in features with different distributions, as presented in Chapter 5. A

priori, it is not known whether different alternatives in modeling present significant differences or which option leads to better results. Additionally, all modeling approaches introduce model-fitting inaccuracy in the derived features, as described in Chapter 5. Exploring *variability* and incorporating *inaccuracy* induced by different modeling approaches, with respect to anatomical or clinical reference data, can have significant impact on the final clinical decision and treatment.

However, the high dimensionality and complexity of the imaging-derived feature space, along with the model-induced variability and inaccuracy, makes exploration a difficult task. The structure of this feature space is also not fully comprehended and, in state-of-the-art clinical research workflow, there is no easy and insightful way to link observations from histopathology or clinical reference to features derived from imaging.

In this chapter, we introduce a visual analysis tool for tumor tissue characterization. We mainly focus on knowledge discovery and hypothesis generation and confirmation, by linking imaging-derived tissue characteristics to anatomical and clinical reference data. In our approach, we employ as central view, a reduced embedding, which maps the feature space of imaging-derived tissue characteristics into a 2D information space. Multiple linked interactive views provide functionality for the exploration and analysis of the local structure of the feature space, enabling the user to retrieve information on distinct intra-tumor regions and to associate observations between the feature space and clinical reference data.

The *contribution* of this work is the design and implementation of a visual tool for the exploration of tumor tissue characterization. To the best of our knowledge, there is no other tool to serve this purpose. The proposed visual tool incorporates the following components:

- It supports the identification and exploration of intra-tumor regions, which have distinct imaging-derived tissue characteristics.

- It facilitates the exploration and analysis of the feature space structure in relation to patient anatomy.

- It enables the association of observations from clinical reference information to the feature space and vice versa.

- It allows the exploration and analysis of model-induced variability and inaccuracy of the feature space.

The remainder of the first part of this chapter is organized as follows: Section 6.3 includes all the necessary background information. Section 6.4 provides an overview on previous work, and Section 6.5 is the core of this chapter, where our proposed approach is presented and discussed. Section 6.6 presents the results of our user evaluation and the conducted case studies. Finally, Section 6.7 concludes the chapter with a discussion and propositions for future work.

6.3. Clinical Background

Several imaging modalities are used in tumor analysis and diagnosis. In this dissertation, we focus on prostate and cervical tumors, which are investigated using MRI data, such as DCE and DW images, which have been already presented in Chapters 2 and 5. However, our

approach can be easily generalized to include other modalities, such as Computed Tomography (CT) or Positron Emission Tomography (PET). In this work, we will employ the tissue properties that can be derived from DCE-MRI data using one of the established pharmacokinetic (PK) models [257], as well as from DWI to derive the apparent diffusion coefficient (ADC) [256]. These tissue properties need to be explored and analyzed, and the distinct intra-tumor regions that these features form, need to be identified.

Additionally, clinical researchers require to associate findings in the imaging-derived feature space to observations from reference data and vice versa. This data might be anatomical images, such as T2-weighted MRI scans, or derived clinical data. The latter can be, for example, derived from linear regression models predicting high-risk tumor zones [106], or histopathological data that are obtained invasively from patient biopsies and can reveal information on tumor aggressiveness or resistance [82]. As already described in Chapter 2, Gleason Scores (GS) are assigned to delineated tumor tissue foci based on their microscopic appearance. For example, high GS tumors are more aggressive and have a worse prognosis [82]. To substitute invasive histopathology, information such as the GS need to be associated with non-invasive, imaging-derived features.

In current clinical research, the exploration and analysis of the feature space of imaging-derived tissue characteristics is conducted using a simplistic slice-based technique, shown in Figure 2.9 (Chapter 2). The values of each imaging-derived feature are encoded with a colormap [139]. Then, these so-called parameter maps are juxtaposed, manually inspected slice-by-slice and mentally compared and analyzed. The cross-sectional analysis of the imaging-derived features with clinical reference data is also done in a similar way. This approach has high memory demands, as it requires from the user to mentally reconstruct relationships and patterns in the data and it is sub-optimal in insight.

Clinical researchers require an exploratory tool that allows them to perform the following tasks:

- To identify and explore intra-tumor regions with distinct imaging-derived tumor tissue characteristics. (T1)

- To analyze and understand the structure of each distinct intra-tumor region and, then, to compare these intra-tumor regions. (T2)

- To discover relations between the feature space of imaging-derived tumor tissue characteristics and clinical reference data. Also, to generate and confirm hypotheses that can associate these two. (T3)

- To identify, explore and analyze the impact of variability or inaccuracy, in the derivation of tumor tissue characteristics from imaging data. (T4)

6.4. Related Work

Several systems have been proposed for the exploration and visual analysis of feature spaces derived from medical imaging. The most similar to our field of application were proposed by Preim et al. [218], Glasser et al. [99] and Fang et al. [84], as well as our own previous work [227] that was presented in Chapter 5. However, these papers are mainly focusing on the exploration and analysis of perfusion parameters or on the distinction of tumors from

healthy tissues. None of these applications addresses intra-tumor tissue characterization based on imaging-derived features or cross-analysis with clinical reference and histopathological data.

Our approach is centered around a dimensionality reduced view, namely a 2D embedding that preserves local structure in the feature space and that allows visual analysis of clusters and their intrinsic feature characteristics. For this reason, we reviewed also the literature related to visualizations employing dimensionality reduction techniques. Some of these previous approaches have also been discussed in detail in Chapter 3.

■ *Dimensionality reduction.* Traditionally employed approaches for 2D projection of feature spaces include the Grand Tour [15], XmdvTool [295], WEAVE [103] and SimVis [72, 189], as described in Chapter 3. These frameworks visualize high-dimensional datasets through projections combined with a number of linked techniques for the visualization of the underlying feature space. However, they all support the visualization of 2D projections of the multidimensional data after the selection of two specific variables from the entire feature space. As explained also in Chapter 3, user selection of two specific variables is not an appropriate choice for our case as it requires prior knowledge about the data and their eventual relationships. In addition, none of these systems supports linked brushing or selection from a spatial view – the anatomical or clinical reference space, in our case – to an abstract view. Cluster analysis visualization is also not possible.

Frameworks to overcome some of these limitations were proposed by Blaas et al. [23], Steenwijk et al. [261], Jeong et al. [133], Choo et al. [52], Ingram et al. [127] and Poco et al. [207]. Most of these techniques were presented in detail in Chapter 3. They all integrate projection techniques with multiple information visualization views and bi-directionally linked scientific visualization views, to enhance data exploration. Nevertheless, the previously mentioned work is not fully applicable to our case. Even if all applications enable to identify clusters and to explore the respective feature space of the data, none of them supports incorporation of variability and inaccuracy in their analysis. Especially, they do not enable to cross-analyze the feature space with clinical reference data and histopathology, for knowledge discovery and hypothesis generation and confirmation.

■ *Visual analysis of clusters.* We are mainly interested in visualizing clusters that are present in the multidimensional feature space of the data, as identified by the user, but also in visualizing the feature characteristics of each cluster and their inbetween relationships. There are various methods for visual analysis of clusters, of which the most relevant to our work are H-BLOB [259], Narcissus [114], SmallWorlds [287], but also the frameworks proposed by Seo and Shneiderman [247], Linsen et al. [166], Glasser et al. [98] and Turkay et al. [281]. However, in some of these frameworks, the visual complexity of the visualizations makes them unsuitable for cluster visualization of large data sets. Also, in all previously mentioned work no information is provided about the structure of a cluster and other metrics, such as the quality of separation of two clusters.

To sum up, there are several systems that are able to manage feature spaces similar to ours. There are others that employ projections and multiple views in their approaches and several that facilitate visual analysis of clusters. However, to the best of our knowledge, there is no visual tool incorporating all the required functionality for our field of application.

6.5. Visual Analysis of Tumor Tissue Characterization

After the acquisition of the imaging data, tumor tissue characteristics are derived from medical images. However, the dimensionality and complexity of this imaging-derived feature space does not allow the easy visual exploration and data-driven analysis. The proposed visual tool aims at satisfying the specific exploratory needs of clinical researchers, working on tumor tissue characterization. Therefore, our design choices are oriented to fulfill their requirements, as described in Section 6.3.

Our visual tool consists of three mutually linked components, as illustrated in Figure 6.1. A 2D embedding of the feature space forms the central view in our visual tool. This view is complemented by an anatomical view of a clinical reference space and multiple linked interactive views for the visual analysis of clusters.

The first requirement is the *identification and exploration of distinct intra-tumor regions with similar tissue characteristics* (T1). An option for the completion of this task could be to employ a scatterplot matrix to visualize all imaging-derived features. As already mentioned in Chapter 3, this would allow only pairwise comparison of the features. However, relations between features in different intra-tumor regions are expected to be more complex. To incorporate information from the feature space of tumor tissue characteristics in one simple view that enables visual exploration of all features, dimensionality reduction is required. In this way, we map the high dimensional imaging-derived feature space (N-D) to a lower dimensional space (2D). This is performed as part of a pre-processing step and several dimensionality reduction techniques could, therefore, be used.

In our approach, we choose to employ t-Distributed Stochastic Neighbor Embedding (t-SNE) [286]. t-SNE has the ability to map data from a high dimensional feature space into a 2D embedding space, preserving the local structure of the initial N-D space. In this way, the *embedding* can be represented in a simple 2D scatterplot, shown in Figure 6.2 - a, the dimensions of which do not have a direct association to the initial features. In this embed-
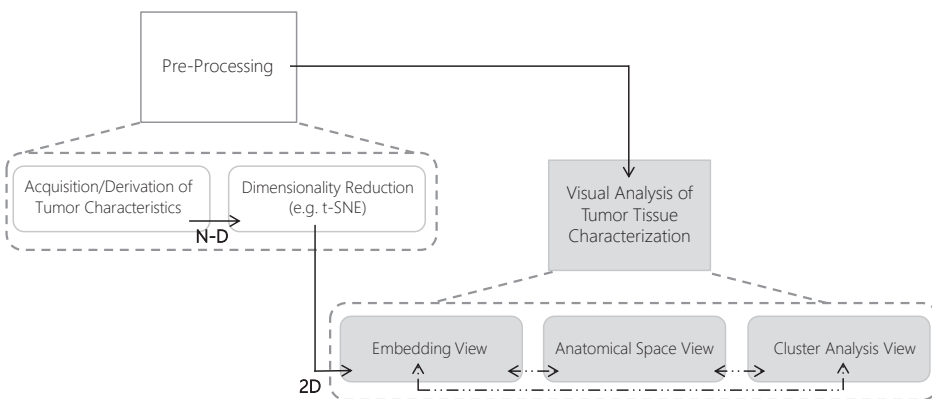
Figure 6.1: General profile of the proposed visual tool. The core of the tool consists of three components (all denoted in grey), linked to each other (links denoted with dotted lines).
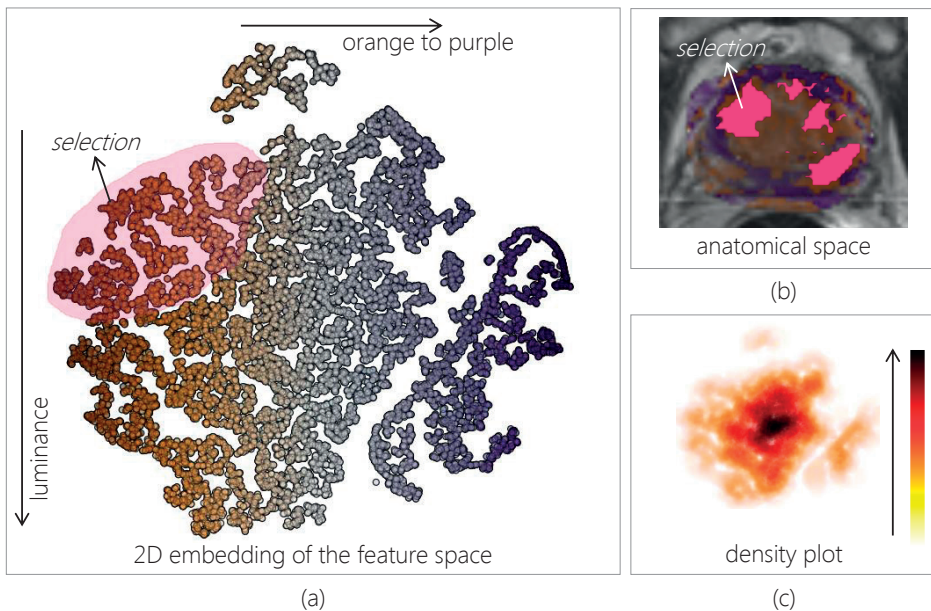
Figure 6.2: Visualizations used for the identification and exploration of distinct intra-tumor regions with similar tissue characteristics (T1). (a) The 2D embedding of the feature space. (b) The representation of the anatomical space. (c) The density plot of the embedding to aid the visibility of visual clusters.

ding, each high dimensional data point is embedded in an abstract 2D space, in such way that the resulting 2D data points plotted nearby represent N-D data points with similar values in the high dimensional feature space. In our case, nearby 2D data points in the embedding represent volumetric positions with similar imaging-derived tissue characteristics, while voxels with dissimilar tissue characteristics are plotted further apart. We remark that the voxel location is not used in the embedding – only the imaging-derived features.

To provide context from the embedding map to the anatomical space, we apply a *2D colormap* [35], to the points of the embedding, based on the position that they have in the 2D scatterplot. Then, this 2D colormap is propagated in the form of an overlay on the anatomical images, as shown in Figure 6.2 - b. The 2D embedding map also supports interactive selection of regions of points, which we call *visual clusters*. An example of this kind of selection is shown in Figure 6.2 - a, in purple. When one visual cluster is selected, a color is assigned to it and this color is used coherently in all views to represent it. To visualize the selections, we did not employ transparency in the points for *Focus+Context (F+C)*, because the visualization of the whole space is equally important, throughout the whole exploratory process. We use color, instead, to enable multiple selections. The association between the two spaces, performed through *brushing and linking*, provides also an evaluation of the localization of the visual clusters selected in the embedding space.

In some cases, visual clusters in the 2D embedding are not well-defined, or overlapping points may give a misleading representation of the space. For these reasons, we introduce

a dual-view in the scatterplot, to depict also the point *density*, as illustrated in Figure 6.2 - c. Hereby, the exploration can be guided by the density peaks, which might reflect the location of densely populated visual clusters. To intuitively depict high density regions versus low density, we employ the heated body colormap [35].

The embedding provides a view on a reduced abstract 2D space, as the dimensions of the embedding do not have a direct relation to the imaging-derived features. Yet, clinical researchers need to be able to link back to this feature space and *analyze the underlying high dimensional feature structure of the identified intra-tumor regions* (T2). This part is achieved with several linked interactive views, which hold complementary information and are interactively updated when the user selects one or more visual clusters in the embedding.

Two initial interesting aspects of the underlying data are the *distributions* of the features of the selected regions and the *pairwise correlations* among these features, which we visualize respectively with boxplots and a simplified color-coded scatterplot matrix (SPLOM) [47], as shown in Figure 6.3 - a,b. These representations were chosen, as they are intuitive and well-known to the intended clinical users. For a simplified correlation view, we abstract the SPLOM by calculating and visualizing directly Pearson's correlation ρ, instead of all the points. We color-code the calculated correlation value to the divergent red-to-blue colormap [35] to show the range from $\rho = -1$ (red) to $\rho = +1$ (blue), as presented in Figure 6.3 - b.

To identify relations that go *beyond two dimensions* of the selected visual clusters, we additionally use parallel coordinate plots (PCPs) [128], as depicted in Figure 6.3 - c. PCPs are versatile in representing multiple dimensions in a single view and are commonly used to easily detect patterns, trends and outliers. In our design, PCPs can be either straight or curved polylines, rendered with low opacity for clutter reduction and improved readability. They also support bi-directional brushing and linking to the embedding, to establish connection with the high dimensional space. As a common way to link observations from different windows, colors are used coherently to denote in all views the same visual cluster.

For the easy comparison and assessment of the feature characteristics of distinct intratumor regions, we employ an additional view on the selected visual clusters. In this view, we provide information on the validity of each visual cluster, visualizing three commonly used internal validity measures: cohesion, separation and the average silhouette coefficient [265]. *Cohesion* (WSS) is a measure of how closely objects are related within a visual cluster and is measured by the within sum of square distances to the mean feature vector:

$$WSS = \sum_{\mathbf{x} \in C} (\mathbf{x} - \mathbf{m})^2 \tag{6.1}$$

where C is the selected visual cluster, $\mathbf{m}$ the mean feature vector of the visual cluster and $\mathbf{x}$ a feature vector element of the visual cluster [265]. *Separation* (BSS) reflects how distinct or well-separated a visual cluster is from another, using the between visual clusters sum of square distances:

$$BSS = \sum_{i} |C_i|(\mathbf{m} - \mathbf{m}_i)^2 \tag{6.2}$$

where C_i are the selected visual clusters, $\mathbf{m}_i$ their respective mean feature vectors, $|C_i|$ the
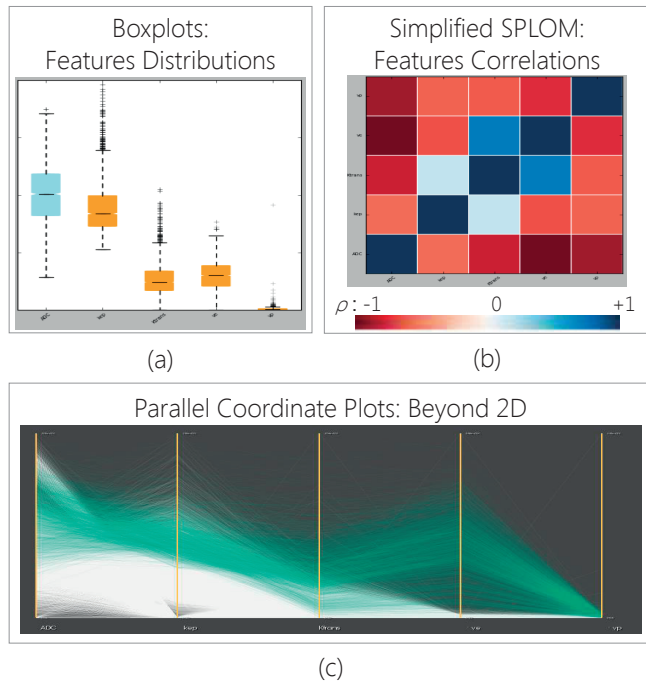
Figure 6.3: Visualizations used to analyze the underlying structure of the features of the each intra-tumor region (`T2`). (a) Boxplots are employed to show the distribution of features, (b) SPLOMs depict the correlations between features and (c) PCPs illustrate relations beyond two dimensions of the data.

size of visual clusters and $\mathbf{m}$ is the overall mean feature vector [265]. The *average silhouette coefficient* (s) combines the notion of cohesion and separation:

$$s = \frac{BSS - WSS}{\max(BSS, WSS)} \tag{6.3}$$

The average silhouette coefficient ranges between 0 and 1, but it is usually interpreted in an ordinal way. Values between $0 - 0.25$ denote bad-defined visual clusters, $0.26 - 0.5$ weak clusters, $0.51 - 0.75$ reasonable clusters and $0.76 - 1$ stand for excellent clusters [265]. An initial option for the visualization of these indices would be a table, which would, however, need sequential analysis and be time-consuming, especially for the comparison of more than two clusters. For this reason, we abstract these numbers to glyph attributes in a 2D view, as shown on Figure 6.4 - a. First, we abstract each visual cluster to a sphere, which is an intuitive encoding for a group notion. The color of each sphere is used for the easy distinction of multiple visual clusters and is the same as the color used in all other views, for each visual cluster.

Cohesion is mapped to the area and opacity of each sphere. As presented in the legend of Figure 6.4 - b, small and opaque spheres depict firm and coherent visual clusters, while
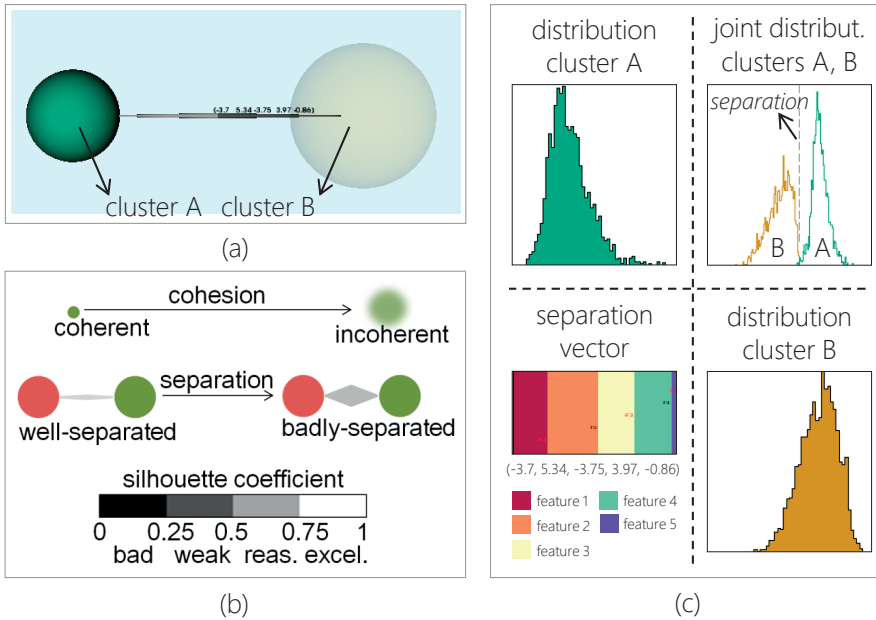
Figure 6.4: Cluster analysis view for comparison and assessment of two distinct intra-tumor regions (`T2`). On the left side (a,b), internal validity measures of each cluster are encoded and presented to the user. On the right side, (c) LDA allows the user to identify which features contribute most in the separation of two visual clusters.

6

large and transparent spheres depict incoherent, cloudy-like visual clusters. Transparency is needed to avoid occlusion and to emphasize the coherent visual clusters. Hereby, to preserve context despite the combined use of size and transparency, we force a minimum limit in both visual encodings.

For the *separation* between two visual clusters, we considered two alternatives: encoding its value to the distance of the spheres or using an additional glyph in-between the spheres. The first option results in a cluttered view, where small spheres were often included or even hidden by the larger ones. Thus, we position among each pair of spheres a double-ended arrow glyph, whose thickness encodes the separation of the two visual clusters. Thin arrows depict well-separated visual clusters giving the illusion of distance, while thick arrows depict badly-separated visual clusters, as shown in the legend of Figure 6.4 - b. We also force a minimum limit in the arrow thickness.

The choice of the double-ended arrow allows the incorporation of the *silhouette coefficient* in the visualization. In this glyph arrow, each end belongs to one visual cluster of the pair and the color encodes the value of the coefficient, using a luminance color scale. This is presented also in the legend of Figure 6.4 - b.

For a more detailed comparison of the visual clusters in the feature space, it is also necessary to show the most prominent features that differentiate them. At this point, we perform

a *Linear Discriminant Analysis (LDA)* [22] between each pair of visual clusters. This is used to calculate the vector that maximizes the linear separation between the means of these clusters, while it minimizes the variance within each cluster. For each pair of clusters, we also obtain the separation histograms resulting from the projection of the high dimensional feature space of the visual clusters on the separation vector. We initially overlay the separation vector as text over the separation arrows of the visual cluster validity representation, as shown in Figure 6.4 - a.

A more intuitive and faster-to-perceive choice is to visualize the separation vector in a separate view as a stacked bar, to show which feature or combination of features contributes more to the separation. As we also want to show how good this separation is for each visual cluster pair, we use a matrix-like configuration to additionally show the distribution and the pairwise cluster projected histograms. This view can help to identify which multiple features contribute to the separation of clusters, as depicted in the bottom right side of Figure 6.4 - c. The colors in the histograms are used for visual cluster distinction and are consistent to the colors employed in all views, while the colors in the stacked bar are used for the distinction of the different features.

The *association of the feature space with the anatomical or clinical reference space and vice versa* (T3) is one of the most important aspects for clinical users. Without the clinical reference context, users would not able to generate and confirm hypotheses. For anatomical inspection, the visual tool provides functionality to slice through any kind of volumetric imaging data and a linked 3D view of the organ where the tumor is located. To enable simultaneous inspection of clinical reference data, we need another 2D slice-based view. The anatomical and clinical reference views are bi-directionally linked to the embedding space. Selections in one space are reflected in the other.

Linking the feature space to the anatomical or reference space is performed as described in (T1). Linking the clinical reference space to the feature space is possible in two ways: by *color-encoding* or *brushing*. In the first case, the entire discretized regions of the reference data can be reflected on the embedding, with the use of a qualitative colormap [35] to match each distinct region of the reference to the respective embedding points, as depicted in Figure 6.5 - a. In the second case, specific regions of interest can be interactively brushed in the reference space and linked to the other views, as illustrated in Figure 6.5 - b.

Finally, the *easy exploration of the effect of modeling-induced variability and incorporation of measurements of model-induced inaccuracy* (T4) are also required. For *variability*, the visual tool enables side-by-side visualization of multiple linked 2D embedding maps, namely one per feature space. Different models result in different feature spaces and embeddings. Therefore, direct comparison of embedding spaces based on positions is not meaningful. To preserve context across multiple embedding maps and to link the points of the embedding map to their volumetric position in the anatomical space, we decided to use a simple visualization. For example, visualizations which would require to trace lines across maps, would be too complex and clutter the view. In our design, one map is used as reference. Then, the position-based 2D colormap discussed in (T1) is used to retain the information of the 2D position of each point across the multiple embedding spaces, as shown in Figure 6.6 - a. All interactions in one map or in the anatomy are reflected on all maps, to strengthen the link between the different spaces.

For *inaccuracy*, clinical researchers are interested either in exploring regions with low

(a)

(b)
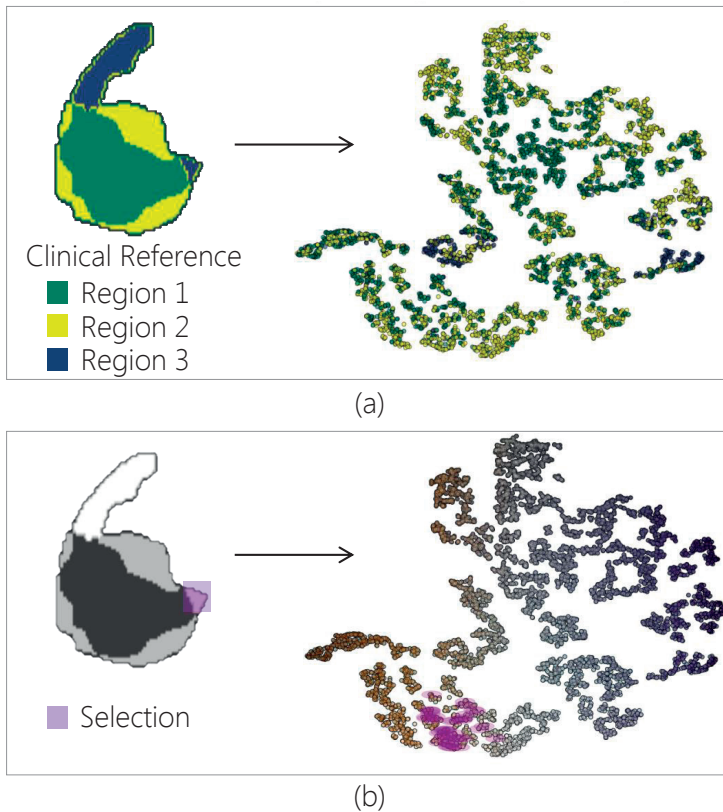
Figure 6.5: Linking the anatomical/clinical reference to the feature space (`T3`) using (a) color-encoding and (b) selection. The reverse linking has been already presented in Figure 6.2.

accuracy to explain why inaccuracy occurs or in restricting their analysis to regions with high accuracy. To simplify this, we employ a double slider that allows to selectively visualize accuracy ranges in the embedding. To enhance the visualization, points of the embedding map within the selected accuracy range are visualized as firm, opaque and with well-defined edges, if they have high accuracy, while lower accuracy points are transparent and with blurry edges, as illustrated in Figure 6.6 - b.

■ *Implementation.* The entire user interface of the proposed visual tool is depicted in Figure 6.7. We implemented the visualization tool in Python as a DeVIDE module [32], employing the Visualization Toolkit (VTK) [2], matplotlib and scikit-learn. The design of the visual tool at this point requires a large number of windows, each one of which is needed to show a specific complementary data aspect. The interface cannot be simplified drastically, given all the tasks that we want to accommodate. Yet, we enable users to selectively manipulate the profile of the tool and to selectively show the most useful representations for their specific exploratory tasks.
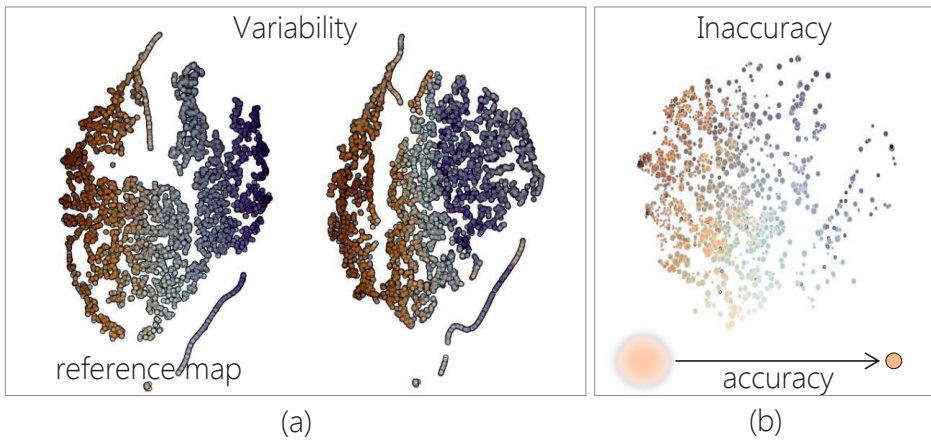
Figure 6.6: Visualizations for the easy exploration of the effect of (a) variability and (b) inaccuracy (`T4`).
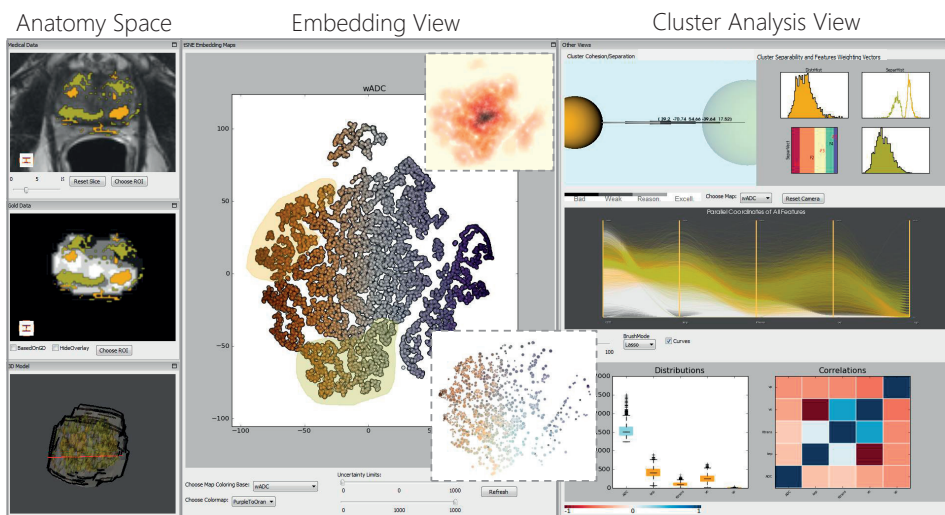
Figure 6.7: The entire interface of our proposed visual tool, for the exploration and analysis of tumor tissue characteristics.

6.6. Evaluation

In order to assess the value of our visual tool, we performed an evaluation, inspired by the paper of Lam et al. [159]. The evaluation was performed with ten domain experts from three different institutions (Netherlands Cancer Institute, Aarhus University Hospital and MAAS-TRO Clinic): three women and seven men. The group of participants included two research physicists, three medical physicists, four biomedical engineers and one computer scientist,

who works on the automatic classification of tumor tissue. All participants have normal vision, five with and five without glasses, and nobody is colorblind. They ranked their computer expertise as intermediate to high and all have a high expertise in tumor tissue characterization. Four of the participants were already familiar with the visual tool, as they were actively involved in its design. In our results, we will treat these four participants separately from the rest, to remove eventual bias.

Before the evaluation, we gave an introduction to the visual tool, where we explained basic notions and main components. We simulated the visual environment for the exploration and analysis of the high dimensional feature space of tumor tissue characteristics, as the clinical researchers would do under real-life circumstances. At this point, the visual tool was initially operated by the first author, while the clinical researchers observed the demonstration of the individual components. Nevertheless, they could also operate the visual tool themselves anytime to better understand the functionality.

The first part of the evaluation included two case studies, which required hands-on exploration of the data, aiming at analyzing the actual relevance of the insight provided by the visual tool, as it would occur in real clinical research settings. Each of the four tasks of Section 6.3 was performed with the *thinking-out-loud* method, as the clinical researchers explained and reasoned on the findings in the data. In the second part, the participants answered a questionnaire.

6.6.1. Case Studies

In this section, we present the analysis performed by the evaluation participants, during the case studies.

■ *Case study: Prostate tumor.* In this case, data from a patient with advanced prostate tumor were employed. DCE-MRI and DW-MRI data were acquired. From these, four pharmacokinetic parameters maps (K^{trans}, k_{ep}, v_e, v_p) [257] and the additional ADC map with a b-value of 1000, were respectively derived. Afterwards, t-SNE was applied to obtain a 2D embedding of the six-dimensional feature space. This case consists of two subcases, where two different clinical reference data were used:

- A linear regression model [106] predicting four prostate risk zones (high to low risk zone).

- Gleason Scores (GS) retrieved from histopathological data that reflect the aggressiveness of the tumor foci based on their microscopic appearance. To obtain the latter, the patient was scanned and thereafter, the prostate was resected and histopathological slices were prepared and registered to the imaging data.

The results of the first case are summarized in Figure 6.8. Here, two distant, well separated regions of points were initially identified in the embedding map for further exploration, as shown in Figure 6.8 - a. Linking to the clinical reference data shows that *region 1* corresponds to part of the high risk zone of the prostate, while *region 2* corresponds to part of the low risk zone, as depicted in Figure 6.8 - b. Using the cluster analysis view of Figure 6.8 - c, the abstraction of *region 1* shows that it is more coherent than *region 2*, but also that the two regions are well separated from each other. The joint histogram of the two regions, in the upper right view of Figure 6.8 - d, also shows this separation, which is mainly due to a combined

effect of three parameters (k_{ep}, K^{trans} and v_e), as it results from the vector of linear separation. This is also confirmed by the different patterns and relationships between the imaging-derived features of the two regions, as shown by the parallel coordinates of Figure 6.8 - e and the correlation matrix view of Figure 6.8 - f. The evaluation participants commented that the findings of this part of the analysis correspond to their theoretical knowledge [257].

When the analysis is done by color-coding the points in the projection maps based on the four discrete risk zones, as illustrated in Figure 6.8 - g, the *low risk cluster 1* is partially separated from the rest, but it is incoherent. Also, the two *highest risk clusters 3* and *4* are reasonably coherent, despite some dispersion in the map, but not well-separated from each other. The employed risk prediction statistical model is built based on the same feature space as the one employed in the current analysis. Thus, the results from the analysis should be matching the model, but they do not. This can be an indication that the model is not able to detect potential sub-clusters in the high risk zones. To cross-check that this is not caused by inaccuracies during the derivation procedure of the features from the imaging data, high values of the goodness of fit, relating to model accuracy, were selected.

In the second subcase, presented in Figure 6.9, the analysis was done using the GS
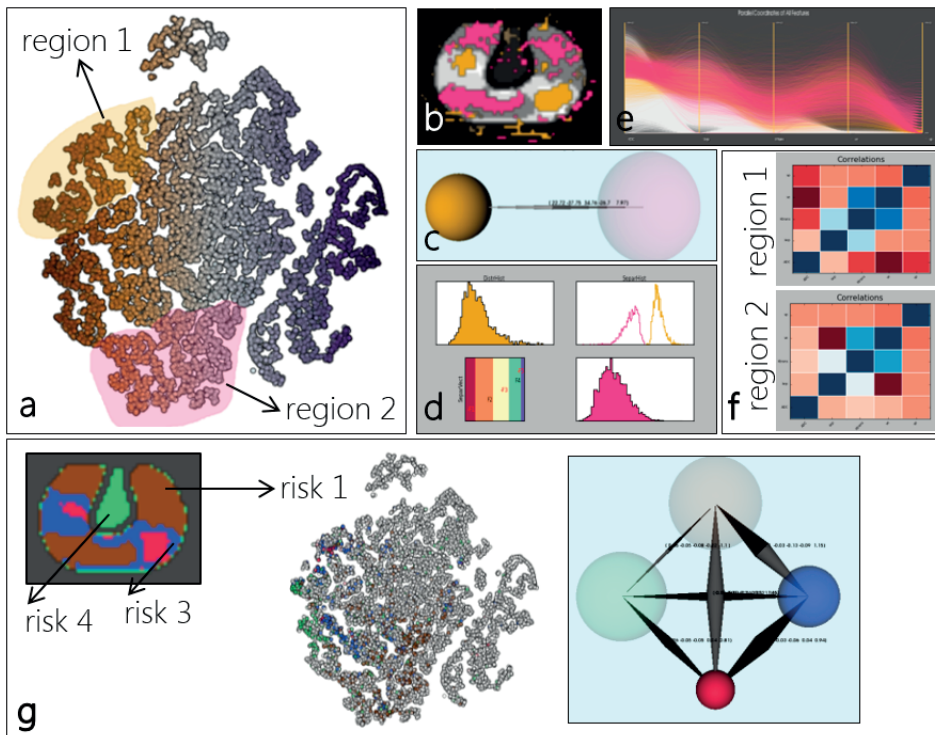
Figure 6.8: Case study of a prostate tumor patient, using as a clinical reference a linear regression model [106] predicting the prostate risk zones.

retrieved microscopically from the histopathology [82]. In this case, the analysis was performed by going from the reference data to the embedding space. There are two regions in the histological data: *GS3* and *GS4*, where higher GS corresponds to more aggressive tumor. We color-code the points of the embedding map with respect to the GS region that they correspond. The points of the two GS regions did not correspond to well-defined, separated visual clusters in the embedding space, which is an indication that the specific feature space is not able to reflect the GS system. However, the visual tool could be used to detect additional imaging-derived features that reflect GS in a better way.
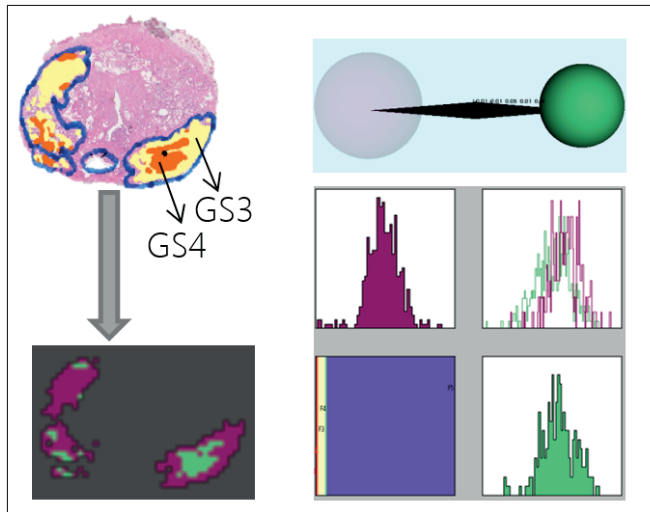
Figure 6.9: Case study of a prostate tumor patient, using as a clinical reference Gleason Scores (GS) retrieved from histopathological data that reflect the aggressiveness of the tumor foci based on their microscopic appearance [82].

■ *Case study: Cervical tumor.* In this case, data from a patient with an advanced cervix tumor were employed. Three different modeling approaches were used to derive pharmacokinetic parameters from DCE-MRI data. The employed pharmacokinetic models were the Tofts model (TM), the Extended Tofts model (ETM) and the Two-Compartment Exchange Model (2CXM)[257].

The first model results in two features (K^{trans} and v_e), the second, in three (K^{trans}, v_e and v_p), and the third in five (K^{trans}, v_e, v_p, F_p and PS) [257]. The Akaike information criterion (AIC_c) that relates to the quality of fit of each model is included as an extra dimension in each feature space.

For each one of the derived feature spaces, we computed a 2D embedding map. As clinical reference, we used segmented data that depict the three regions of the tumor: the periphery, the center and the in-between region. In literature [139], it has been hypothesized that different models perform better in different tumor regions and our visual tool was used for validation.

Initially, the three embedding maps need to be compared. By colorcoding the points based on the TM, it appears that the 2CXM behaves differently than the other two models, as shown in Figure 6.10 - a. For example, the region denoted in the red box is consistent in the ETM and TM, despite the fact that the cluster position is different across maps, as depicted in Figure 6.10 - a. However, in the 2CXM map, this visual cluster does not exist, as the points are spread. This is a confirmation to the theory that the 2CXM in some cases is unnecessarily complex and thus may produce inaccurate results.

Then, two regions were selected, as illustrated in Figure 6.10 - b. From the anatomy, based on the theoretical hypothesis [139], *region 1* represents the areas where the simplest model, the TM, should fit better. On the other hand, *region 2* represents areas where the ETM model fits better. For *region 1*, this is supported by the ETM embedding map: the corresponding visual cluster presents a split, as shown in Figure 6.10 - b.

Similar observations can be drawn for the points of *region 2* that also form a well-defined region in the ETM map, but present a bigger spread in the TM map. These observations are also validated by the parallel coordinates, visualizing the values of the features in Figure 6.10 - c. To compare the underlying structure of these two regions, we use the cluster analysis view of Figure 6.10 - d, from which we identify that the visual clusters are well separated from each other in both ETM and TM maps. The coherence, though, is better in the map of better fit: *region 1* in TM and *region 2* in ETM.

6.6.2. Interviews with Clinical Researchers

After the case studies, the participants were asked to complete a questionnaire. First, we asked four questions, related to the four main tasks of Section 6.3. Each question required an open answer, but also grading using Likert scales $(1 - 5)$ for the perceived effectiveness, perceived efficiency and perceived satisfaction. In order not to compromise the results, we separate in our statistical analysis the four people involved in the design and the other six who were not involved.

It was found that the two groups had in most cases comparable results. The measured variables received high scores, as summarized in Figure 6.11. The scores of the variability and inaccuracy incorporation tasks were slightly lower than the other three, as some of the participants noted that they would like to use the inaccuracy component more, before giving a conclusive answer. One participant graded the efficiency of the second task lower, because he stated that the interpretation of the cluster analysis view takes time when the user is not familiar with the employed representations. The others considered that the representations adopted in the multiple views of the visual tool were intuitive.

The participants were also asked to compare the visual tool to what they are currently using and to evaluate the visual tool, as a whole, by discussing its suitability, strengths, limitations and missing features. In comparison to the currently employed approach, they commented: *it is a much more elegant approach than what I am currently using. It is very intuitive and versatile.*, continuing: *I can learn more about the data and discover more about it, than with the current approach.*. They also said: *I think the tool can help us to explore the feature space better.*, and concluded that: *The information provided by this tool is very interesting and once collected for a wider population, it can be used to train a model based on more relevant features that provide a better separation of tissues.*

6
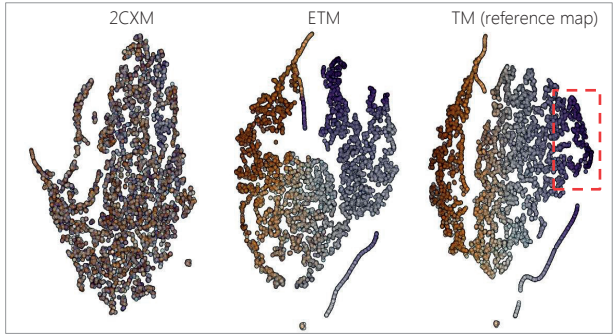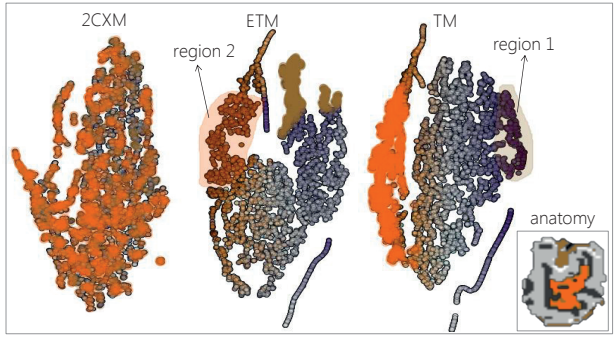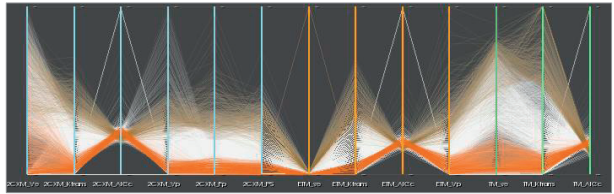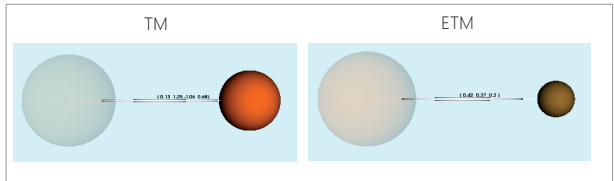
Figure 6.10: Case study of a cervical tumor patient, explored for three different modeling approaches [257].

All participants agreed that the visual tool is overall understandable and easy to learn. They also agreed that the suitability of the visual tool was mainly for data exploration, knowledge discovery and hypothesis validation or generation. Suitability for decision making might come as a result of the previous, after better familiarization with the visual tool. The strong features of the visual tool – and also what the evaluators liked more about it – are the multiple interactive linked views on the data, the link between anatomy and feature space and the incorporation of inaccuracy in the analysis, even if the latter received a lower mean grade.

According to the evaluation participants, a missing feature or limitation of the visual tool is that it currently does not provide functionality for viewing simultaneously the PK parameter maps in a supplementary slice-based view, which is still important for clinical users, due to familiarity. Finally, some of the users commented that the areas of interest do not always correspond with descriptive density variation areas in the 2D embedding. Therefore, it is not always obvious how to define the selection boundaries in the embedding.

6.7. Conclusions and Future Work

The current exploratory approach for tumor tissue characterization based on imaging is time consuming, making it difficult to obtain the required insight. In this chapter, we proposed a visual tool that enables clinical researchers to perform easy exploration and visual analysis of the feature space of imaging-derived tissue characteristics, and to discover new knowledge, with respect to reference data used in clinical research. We employ, as central view, a 2D embedding of the feature space, linked to multiple interactive views. These views provide information concerning the structure of the feature space and relations to anatomical and clinical reference information. We performed an initial evaluation with ten clinical researchers, who confirmed the usefulness of the visual tool in their analysis, as it opens new possibilities in the exploration of the feature space and provides access to new insight in the data. We illustrated this with two case studies performed during the evaluation.

A direction for future work includes the extension of the application to allow meaningful follow-up or inter-patient analysis. It would also be interesting to extend the visual tool or to generalize its functionality for other applications, where also a higher number of features is involved. An example of this is demonstrated in Part II of this chapter. Finally, the incorporation of embedding precision information in the analysis would also lead to more reliable observations. The proposed visual tool is a promising basis for clinical researchers to identify, explore and analyze heterogeneous intra-tumor regions and particularly, to generate and confirm hypotheses, with respect to clinical reference.

6
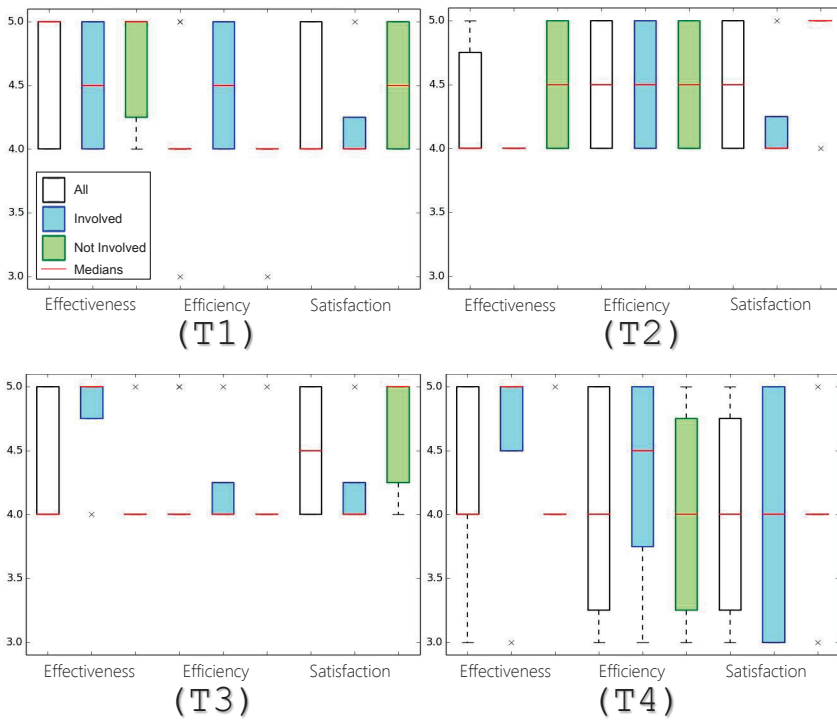
Figure 6.11: Schematic representation of the evaluation results, for each one of the tasks of Section 6.3.

Part II – Aiding the Design and Understanding the Behavior of Tissue Classifiers

6.8. Abstract

The previously proposed Visual Analytics tool can be employed in many different applications. Among these, it can be used to aid the design of tissue classifiers and to increase the understanding of their behaviors. We demonstrate this outside the field of prostate tumor tissue characterization.

Accurate segmentation of brain white matter hyperintensities (WMHs) is important for prognosis and disease monitoring. To this end, classifiers are often trained – usually, using T1 and FLAIR weighted MR images. Incorporating additional features, derived from diffusion weighted MRI, could improve classification. However, the multitude of diffusion-derived features requires selecting the most adequate. For this, automated feature selection is commonly employed, which can often be sub-optimal.

In this work, we propose a different approach, introducing a semi-automated pipeline to select interactively features for WMH classification. The advantage of this solution is the integration of the knowledge and skills of experts in the process. In our pipeline, the Visual Analytics system proposed in Part I [226] is employed, to enable a user-driven feature selection. The resulting selected features of our approach are the intensities from T1 and FLAIR images, Mean Diffusivity (MD), and Radial Diffusivity (RD) – and secondarily, C_S and Fractional Anisotropy (FA). The next step in the pipeline is to train a classifier with these features and compare its results to a similar classifier, used in previous work with automated feature selection. Finally, Visual Analytics is employed again to analyze and understand the classifier performance and results.

6.9. Introduction

White matter hyperintensities of presumed vascular origin (WMHs) are a common finding in MR images of elderly subjects. They are a manifestation of cerebral small vessel disease (SVD) and are associated with cognitive decline and dementia [196]. Accurate segmentation of WMHs is important for prognosis and disease monitoring. To this end, automated WMH classification techniques have been developed [10]. Conventional approaches include raw image intensities from T1 and FLAIR weighted MR images, but recently, it has been suggested that diffusion MRI can improve the segmentation [155, 169]. Multiple features can be derived from this imaging modality. Thus, careful feature selection is required.

In this work, we propose a semi-automated approach, to aid the design of WMH classifiers. Our novelty is the introduction of a user-driven, interactive pipeline that provides new insight into the entire classification procedure, especially in the identification of an adequate feature list and the outcome analysis. Up to now, the knowledge and cognitive skills of experts have not been intensively involved in the process. In the *first step* of our pipeline, we employ the Visual Analytics system [226], which was presented in Part I of this chapter. Using this system, expert users select interactively the most important features. In the *second step*, the resulting feature list is used to train a classifier for WMH segmentation. The performance and results of this classifier can be analyzed and interpreted in the *final step* of the pipeline, again using Visual Analytics. To the best of our knowledge, involving users through visual analysis and interaction in an entire pipeline for feature selection, classification and outcome evaluation for WMH structures, has not been addressed before.

The remainder of the second part of this chapter is organized, as follows: Section 6.10 is the core, where the novel pipeline for feature selection, classification and outcome evaluation for WMH structures is presented. Section 6.11 presents the results of using our proposed approach. Section 6.12 concludes with a discussion on several points with respect to our proposed pipeline, along with propositions for future work.

6.10. Materials and Method

6.10.1. Subjects and MRI Data

We used the subjects of the MRBrainS13 challenge [177], with additional manual WMH delineations. Subjects included patients with diabetes and matched controls. The cohort included 10 men and 7 women, with an age of 71±4 years. All subjects underwent a standardized 3 T MR exam, including a 3D T1-weighted, a multi-slice FLAIR, a multi-slice IR, and a single-shot EPI DTI sequence with 45 directions. All sequences were aligned with the FLAIR sequence [148]. The diffusion images were corrected for subject motion, eddy current induced geometric distortions and EPI distortions, including the required B-matrix adjustments, using ExploreDTI [161].

The dataset includes T1, FLAIR and IR weighted images, as well as the following diffusion features: Fractional Anisotropy (FA), Mean Diffusivity (MD), Axial Diffusivity (AD), Radial Diffusivity (RD), the Westin measures C_L, C_P, C_S [304], and MNI152-normalized spatial coordinates [148, 88]. This exact dataset has been previously reported in a study of Kuijf et al. [155] for the investigation of the added value of diffusion features in a WMH classifier. Since we could have access to the exact same data and we share the same goal, we will use this previous work of Kuijf et al. as a baseline for the evaluation of our results.

6.10.2. Method

In this section, we describe our new pipeline for the user-driven, interactive selection of features that can differentiate WMHs from healthy brain tissue. Our pipeline consists of three steps, depicted in Figure 6.12. First, the data are interactively explored and analyzed by expert users, using our tissue characterization system [226], which was presented in the previous part of this chapter. From this step, we obtain through interaction and visual analysis a list of features, adequate for WMH detection. These features are subsequently used to train a classifier. After classification, the system is used again to evaluate and better understand the classification process and outcome.

■ *Feature Selection using Visual Analytics.* Our previously described Visual Analytics systems [226] is employed to interactively explore the data of each one of the available subjects, as shown in Figure 6.13. Initially, t-Distributed Stochastic Neighbor Embedding (t-SNE) [286] is used to map the high-dimensional feature space of each subject, which was described in Section 6.10.1, into a reduced 2D abstract embedding view, preserving the local structure of the feature space. Spatial coordinates are excluded, as we are interested in preserving similarities in the feature space, whereas the voxel positions could introduce bias.

In the resulting embedding view, which is presented in Figure 6.13 - b, close-by 2D data points reflect voxels with similar behavior in the high-dimensional feature space. Therefore,

6
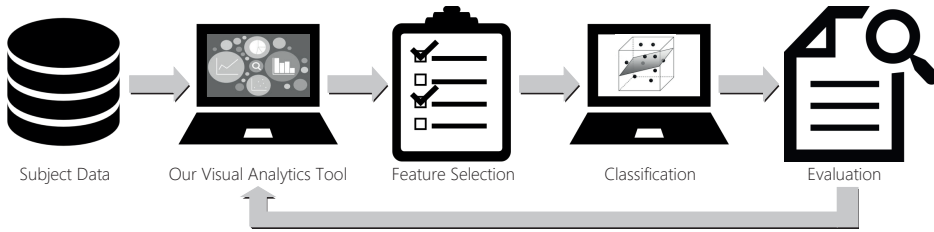
Figure 6.12: The pipeline proposed for the user-driven feature selection, classification and outcome evaluation for the segmentation of White Matter Hyperintensities (WMHs).

| (a) Anatomy Views | (b) t-SNE Embedding View | (c) Multiple Interactive Views |
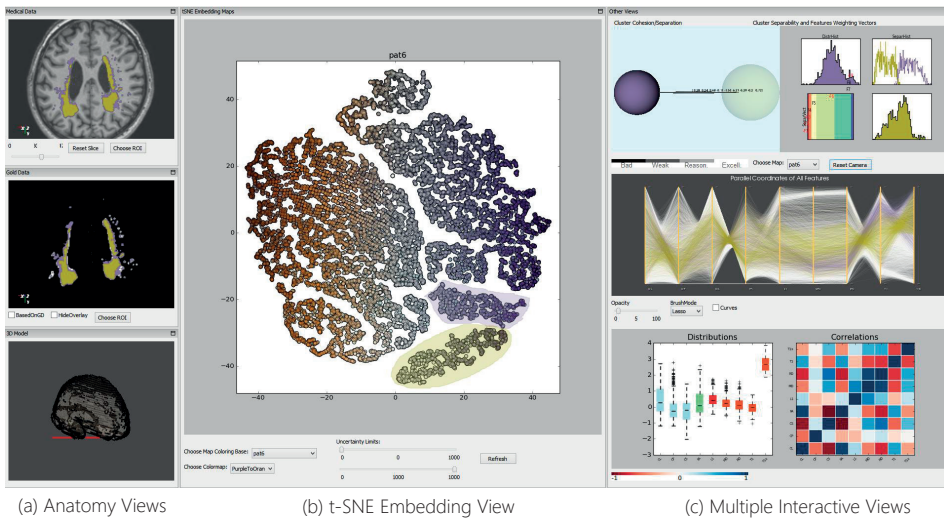
Figure 6.13: Our proposed Visual Analytics system [226] during the exploration of the data of a subject from the MRBrainS13 challenge [177]. The three main components of the system are denoted.

voxels from structures with similar imaging characteristics are expected to be grouped together in the embedding, in so-called visual clusters. The availability of manual delineations of the WMHs as ground truth data allows to associate visual clusters from the feature space to anatomy, and vice versa, as depicted in Figure 6.13 - a.

When a WMH-containing visual cluster is interactively selected, its intrinsic feature characteristics can be explored. This can be conducted, for example, against other structures of the brain, or against WMHs voxels that are not within the selected visual cluster. Then, several linked views, such as the ones depicted in Figure 6.13 - c, are interactively updated with complementary data information. This includes feature distributions and correlations, multidimensional data patterns, cluster validity analysis and information on features that

help separating visual clusters from each other, as given by the weights of the separation vector of Linear Discriminant Analysis (LDA). In this way, features suitable for the detection of WMHs are interactively identified.

For example, for the subject of Figure 6.13, two visual clusters have been selected in the t-SNE of the middle view. As depicted in the anatomical views, one corresponds to the WMH core, highlighted in green, and the other to the periphery, highlighted in purple. Together, they represent the biggest part of the structure. Still, several small parts are missed. The separation vector, resulting from LDA between the two visual clusters containing the WMHs against the rest of the brain, is extracted. From the weights of this vector, features adequate for differentiating the detected WMHs from the rest of the brain are identified. This analysis is subject-specific and has to be performed on a single-subject basis. When all subjects have been explored, the user decides on the most overall suitable feature list.

■ *Classification.* In this step, many different classification approaches could be followed, but comparing all would be out of scope, for this work. Recently, Kuijf et al. [155] presented an approach for WMH classification, using the same set of diffusion features. To evaluate whether our user-driven feature selection outperforms automated feature selection, we adopt a similar classification approach, as in the previous work of Kuijf et al. The list of features resulting from the system is used to train a k-nearest-neighbor classifier for WMH segmentation. For different feature combinations, several classifiers are trained with $k = 50$, 75, or 100, and the neighbor-weighted is either uniform or distance-based [155].

■ *Evaluation of Classification.* In many cases, classifiers are treated as black boxes, and users do not have actual insight into the achieved result. With this step, we want to provide a way for evaluating and understanding both the results of the classifier and the classifier itself. To this end, we import the binary masks resulting from the classification, which contain the detected and the missed WMHs, into our Visual Analytics system [226]. The user can interactively explore the high-dimensional feature space of the two regions of the mask, and generate hypotheses about why the classifier failed to detect parts of the WMHs, with respect to the imaging features.

6.11. Results

In this section, we present the results from each step of our proposed pipeline, for feature selection, classification and outcome evaluation for WMH structures in the subjects of the MRBrainS13 challenge [177].

■ *Feature Selection using Visual Analytics.* In most of t-SNE embeddings of the subjects, the majority of voxels of the WMHs are grouped together in one or two visual clusters, similar to the case depicted in Figure 6.13. From selecting these visual clusters, we could identify that for subjects with two visual clusters, these either correspond to the core and the periphery, or to anterior and posterior WMHs. For large WMHs (top 50%), the visual clusters of the embedding identify 84-98% of the structures. For the rest, the visual clusters can at least detect the core, with a minimum detection percentage of 54%.

The multiple interactive linked views of the system show that there are comparable behaviors, within all cases of visual clusters, especially for larger WMH structures. As mentioned before, the cluster analysis view of the system provides the separation vector, result-

ing from LDA between the visual cluster containing most of WMHs and the visual cluster of the rest of the brain. Table 6.1 depicts, for all investigated subjects, the weights of separation for these two visual clusters.

In all, but three cases, T1, FLAIR, RD and MD are more important and have a considerable weight. For bigger WMHs, C_S and FA also become important. The contribution of other features such as AD, C_L, C_P and IR seems not significant. Considering also the (cor-)relations between diffusion features, we decide on the overall set of features for the classifier: MD, RD,

Table 6.1: The most important features for each subject, as resulting from the weights of the LDA separation vector, performed for the detected visual clusters of WMH voxels against the rest of the brain. The second column denotes the size of WMHs in voxels. The third column shows the percentage of WMHs detected by visual clusters in our Visual Analytics tool. The other columns represent features, and their weights are color encoded per row. The resulting feature list is the set MD, RD, T1 and FLAIR (then, C_S and FA).

Subject Data			Features									
ID	Size	% Found	CL	CP	CS	FA	AD	MD	RD	IR	T1	FLAIR
1	1664	98	0.83	0.15	-1.65	-1.46	-0.06	-1.81	3.22	0.3	0.03	1.73
2	1370	93	0.95	0.08	-2.29	-1.99	-0.82	-4.26	5.93	0.16	0.56	2.43
3	1167	91	1.18	0.1	-3.39	-2.94	0	-6.51	7.86	0.1	1.72	2.97
4	819	89	0.75	0.06	-4.75	-4.27	-1.7	-23.9	19.35	0.12	9.48	5.14
5	613	87	1	-0.06	-4.75	-3.36	0.18	-16.9	17.29	-0.12	3.98	3.61
6	571	89	-2.22	-1.48	-8.66	-1.8	0.92	-15.7	21.23	0.46	2.52	6.7
7	494	87	-0.3	-0.17	-2.02	-0.84	0.48	-3.57	4.54	-0.34	1.3	1.83
8	461	84	2.39	1.22	6.61	0.35	-0.34	-8.63	13.37	-0.17	-0.47	-4.19
9	417	79	-3.96	-1.47	8.28	12.06	-0.22	-3.04	8.29	-0.11	-3.54	-7.51
10	353	60	2.6	1.68	-5.58	-9.09	-2.62	-25	5.9	-0.41	24.56	7.27
11	328	62	1.75	0.43	-2.07	-2.87	0.28	-8.88	8.56	0.43	2.34	1.6
12	273	65	0.1	-0.2	-2.17	-1.75	0	-10	8.06	-0.4	3.52	1.45
13	265	54	-0.54	-0.19	-4.28	-1.74	-0.84	-16	14.54	-0.19	4.87	2.73
14	239	68	0.36	-0.01	-1.79	-0.14	-0.62	-8.58	7.67	-0.01	2.47	1.51
15	228	63	2.82	1.64	7.67	-0.74	-2.06	-10.5	-2.19	1.28	7.31	-4.94
16	221	76	0.81	0.01	-5.09	-3.42	0.64	-8.19	10.32	0.02	2.01	3.35
17	220	84	-0.78	0.22	-5.53	-2.59	-1.74	-15.5	14.91	0.44	5.51	4.72

Negative High Weight   Low Weight   Positive High Weight
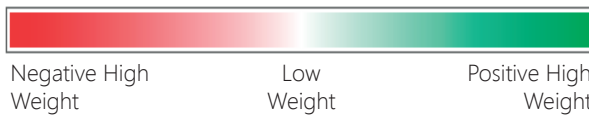
T1 and FLAIR – and secondarily, C_S and FA. Here, we add the MNI152-normalized coordinates (x, y, z) to better represent the brain volume and to suppress non-WMH structures.

■ *Classification.* Based on the results of the system, the following four combinations of feature sets $s_i \in S$ are chosen for our k-NN classifier:

- s_1: MD, RD, T1, FLAIR
- s_2: $s_1 + (x, y, z)$
- s_3: $s_1 + C_S$, FA
- s_4: $s_3 + (x, y, z)$

For each classifier trained on a feature set $s_i \in S$, we measure the sensitivity and Dice similarity coefficient (mean ± standard deviation), as shown in Table 6.2. These measurements are performed with respect to the available manual delineations of the WMH structures. Furthermore, our results are compared to the feature sets $f_i \in F$, previously used by Kuijf et al. [155]:

- f_1: T1, IR, FLAIR
- f_2: $f_1 + (x, y, z)$
- f_3: $f_1 +$ FA, MD
- f_4: $f_2 +$ FA, MD
- f_5: $f_4 + C_L$, C_P, C_S, AD, RD

The results of Table 6.2 demonstrate that our proposed Visual Analytics-guided feature selection can achieve similar or slightly better performance than the automated feature selection presented by Kuijf et al. [155]. The two best performing feature sets of Kuijf et al. used 8 (f_4) and 13 (f_5) features, while our current two best methods use less features, with comparable results. Our approach allows to discard C_L, C_P, AD and IR, which do not contribute in the classification, hereby saving scanning and computational time.

■ *Evaluation of Classification.* To evaluate the classification outcome, we introduce the results of the two best performing classifiers, s_2 and s_4, into the Visual Analytics system. One of the goals is to explore and analyze the parts of the WMHs that are missed, but also to understand better how these classifiers work and how they can be improved. From an initial inspection, it results that classifier s_2 is restricted to the core of the WMHs, while s_4 detects an extension of it. The WMH core is always detected by both classifiers, as it has consistent imaging characteristics and is well-clustered in the t-SNE embeddings. In subjects with bigger WMHs, s_4 misses only small or thin structures and part of the periphery. In subjects with smaller WMHs, there is a tendency to miss periphery parts and posterior structures more often than the anterior. For bigger WMHs, the core differs in T1, MD, RD with the missed structures. Also, the latter are not as good clustered in the t-SNE embeddings as the core, indicating that they are not coherent in their imaging characteristics. As WMHs become smaller, the influence of T1 becomes less strong, while MD and RD seem to become more important.

6

Table 6.2: Sensitivity, Dice similarity coefficient (SI, higher is better) and number of features for the classifiers, trained on combinations of features $s_i \in S$ (top, from our Visual Analytics-driven approach) and $f_i \in F$ (bottom, from the work of Kuijf et al. [155]), with respect to the available manual delineations.

Our Visual Analytics-driven approach		
S	Sensitivity (%)	Dice SI	Features
s_1	64.8 ± 0.2	0.460 ± 0.003	4
s_2	76.2 ± 0.4	0.560 ± 0.005	7
s_3	66.3 ± 0.2	0.471 ± 0.004	5
s_4	76.6 ± 0.5	0.576 ± 0.004	8

Automated approach of Kuijf et al. [155]		
F	Sensitivity (%)	Dice SI	Features
f_1	59.7 ± 0.2	0.349 ± 0.001	3
f_2	73.4 ± 0.4	0.536 ± 0.005	6
f_3	67.8 ± 0.3	0.411 ± 0.003	5
f_4	77.2 ± 0.4	0.565 ± 0.004	8
f_5	75.2 ± 0.6	0.561 ± 0.003	13

6.12. Discussion and Conclusions

We proposed a user-driven pipeline for aiding the design of classifiers, focusing on WMH segmentation. Using Visual Analytics and employing the cognitive skills of an expert user, we initially identified the list of features (MD, RD, T1, FLAIR, and secondarily, FA and C_S) that are suitable for the separation of WMHs. Then, this list was used for WMH classification. In respect of previous work [155], our results are comparable. Yet, our results are not achieved through a trial-and-error approach, but after a justifiable and understandable, interactive feature selection.

Additionally, our approach requires less features, which allows to skip several imaging sequences, making the feature calculation less computationally intensive and time consuming. For example, we concluded that C_L, C_P, AD and IR can be omitted, which saves valuable scanning time (IR: 3:49 min).

After classification, we evaluated the classifier outcome in our proposed Visual Analytics system. The periphery is constantly missed. Thin and small structures can be missed due to partial volume effect, while the MNI152-normalized spatial coordinates can influence the separation of posterior or anterior WMHs. For certain subjects, the missed structures have intrinsically different imaging characteristics.

In this case, more features, such as texture or tensor information, should be further investigated. The performance of the classifier could be further improved by adding additional post-processing, to remove false positive detection, which was not performed here, to be comparable to Kuijf et al. [155]. Also, it would be interesting to investigate what happens when our Visual Analytics-selected features are used with more sophisticated classification algorithms.

In the entire pipeline, the user interacts and guides the analysis. This has the advantage that the cognitive capabilities of the user, which are not easily automatized, can be included in feature selection. However, the results are user-dependent and it remains important to analyze the bias introduced by the user. Although t-SNE is widely used [9] for understanding and exploring high dimensional data, errors can also be introduced due to its use.

Adding more features for exploration in the Visual Analytics system, such as textural features or information from tensors, could give interesting results. However, certain visualizations of our system do not scale well to a high number of features. Therefore, new visualizations would be needed to tackle hundreds of features. Finally, evaluating the use of the pipeline with a user study, to define its general usefulness, is another point for future work.

Nevertheless, employing Visual Analytics in the design of classifiers provides potential for better understanding the data under exploration, and for obtaining more insight into classifiers and the frequently exploding set of imaging features.

6

7

Visual Analytics for the Exploration and Assessment of Segmentation Errors

Errors are portals to discovery.

James Joyce (1882 - 1941), Ulysses

In this chapter, we present a Visual Analytics solution for the exploration and assessment of errors occurring during the model-based segmentation of pelvic structures. Our focus is on how Visual Analytics can contribute towards the prediction of the performance of segmentation algorithms. Chapter 7 is based on the paper:

■ Raidou, R.G., Marcelis, F.J.J., Breeuwer, M., Gröller, E., Vilanova, A., and van de Wetering, H.M.M., 2016. Visual Analytics for the Exploration and Assessment of Segmentation Errors. Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM 2016), pp. 193-202 [224].

7.1. Abstract

Several diagnostic and treatment procedures require the segmentation of anatomical structures from medical images. However, the automatic model-based methods that are often employed may produce inaccurate segmentations. These, if used as input for diagnosis or treatment, can lead to sub-optimal results for the patients. Currently, experts working on the development of these segmentation algorithms cannot easily perform an analysis to predict which anatomic regions are more prone to inaccuracies, and to determine how to improve their algorithms.

In this chapter, we propose a visual tool to enable experts, working on model-based segmentation algorithms, to explore and analyze the outcomes and errors of their methods. Our approach supports the exploration of errors in a cohort of pelvic organ segmentations, where the performance of an algorithm can be assessed. Also, it enables the detailed exploration and assessment of segmentation errors in individual subjects.

A usage scenario is employed to explore and illustrate the capabilities of our visual tool. To further assess the value of the proposed tool, we performed an evaluation with five segmentation experts. The evaluation participants confirmed the potential of the tool in providing new insight into their data and employed algorithms. They also gave feedback for future improvements.

7.2. Introduction

Several diagnostic and treatment procedures require the segmentation of anatomical structures from medical images. This can be either performed manually, semi-automatically, or automatically. In manual segmentation, medical experts inspect 2D imaging slices one-by-one, and delineate structures. As this procedure can be time consuming, automatic methods are preferred and, therefore, a lot of effort is being invested in algorithm development. Still, automatic algorithms cannot account for all cases, and may perform sub-optimally.

Experts working on automatic segmentation algorithms can relatively easily detect the errors. However, even for them, it is not trivial to understand why or how inaccurate outcomes are produced. Exploring and assessing segmentation errors can provide experts with new knowledge about the performance of their algorithms. For example, it can help them to predict the anatomic locations and circumstances under which errors occur. Moreover, it can aid them in confirming or generating hypotheses about their methods and, on the long term, it can allow them to improve their segmentation results. Even if segmentations cannot be improved, it still remains important to be aware of potential inaccuracies. Disregarding this information might affect diagnosis or treatment, if the latter are based on erroneous segmentation outcomes.

As proof-of-concept, we consider the automatic model-based segmentation of pelvic structures [239], used as input to radiotherapy planning for prostate tumors. Planning such a radiotherapy procedure requires the accurate segmentation of the prostate and the organs at risk that need to be spared during irradiation. Also in this case, segmentation errors often occur, and need to be explored and assessed. In current practice, this is not feasible on a patient basis, as it can only be done if golden standard segmentations are available. Therefore, new means for the exploration and analysis of segmentation errors need to be devised.

Our *contribution* is a visual tool that allows experts working on algorithms for model-based segmentation of pelvic structures, to explore and assess the outcomes and errors of their methods. Our approach incorporates the following two capabilities:

- It supports the exploration and assessment of errors in a cohort of pelvic organ segmentations. These segmentations result from applying the same algorithm to several subjects. With this, experts inspect the general performance of the algorithm.

- It facilitates the detailed exploration and assessment of segmentation errors in the pelvic organs of individual subjects. With this, experts can identify the specific details about the performance of the algorithm, concerning each subject of the cohort.

To the best of our knowledge, there is no other tool with the comprehensive functionality that our work offers. Although we demonstrate our visual tool on a specific case, our methods could be generalized to other applications, and fit to other segmentation algorithms.

The remainder of this chapter is organized, as follows: Section 7.3 includes all the necessary background information. Section 7.4 provides an overview on previous work, and Section 7.5 is the core of this chapter, where our proposed approach is presented and discussed. Section 7.6 presents the results of a usage scenario, while Section 7.7 presents the results of our user evaluation. Finally, Section 7.8 concludes the chapter with a discussion and propositions for future work.

7.3. Model-Based Segmentation of Pelvic Organs

As described in Chapter 2, radiotherapy planning requires the accurate segmentation of the prostate and all the surrounding organs at risk, such as the bladder, rectum and seminal vesicles. For the segmentation of the involved pelvic organs, shown in Figure 7.1, automatic model-based methods are often employed [76, 239].

In the present work, we consider the algorithm of Schadewaldt et al. [239], for the segmentation of pelvic structures in CT images. In this method, structures are considered to have a known general shape. Training data are used to build probabilistic models that explain the shape variation of each structure. These models are used as prior information, and are positioned in the volume. Then, they are iteratively adapted to the boundaries of the structure of interest [76], using a combination of rules. These rules describe how to use features, such as gradient magnitude, which have been learned from training data. Different features might be employed for different organs, or parts of these. More details about the algorithm can be found in the papers of Schadewaldt et al. [239] and Ecabert et al. [76].

Although the selected segmentation method is robust, it is not always accurate. Yet, the resulting inaccuracies might be propagated to the radiotherapy dose administration to healthy organs, with unwanted side effects [300]. Our collaborating experts from Philips Healthcare in Hamburg, working on the segmentation of these pelvic structures, need to explore, understand, and assess the segmentation results, as well as their respective inaccuracies. To this end, they generate, using their in-house algorithm [239], segmentations of four organs – prostate, bladder, rectum, seminal vesicles – and their interfaces, in the form of triangulated meshes. Meshes from different subjects have a triangle-to-triangle correspondence.

7
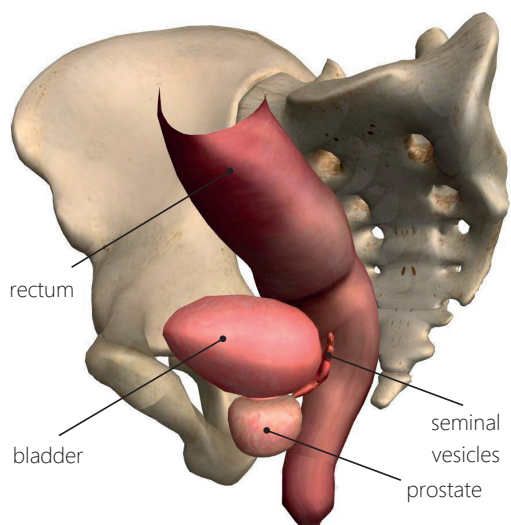
Figure 7.1: The anatomy of the pelvic structures involved in this work. *Image generated using ZygoteBody^{TM}.*

Additionally, ground truth for each subject is available from delineations of pathologists. Correspondence between the ground truth and the segmentation outcomes has been established, as described in the paper of Schadewaldt et al [239]. From the ground truth, our segmentation experts computed four local quality measures per triangle [239]. These are:

- The *target error*, which is the point-to-point distance from a triangle in the resulting mesh, to the target location in the ground truth data, measures in mm.

- The *features response*, which indicates the strength of a number of algorithm features at the target location.

- The *weighted features response*, which is the feature response, inversely weighted by the distance to the target.

- The *triangle area*, measured in mm^2.

The above mentioned measures are indicative of segmentation accuracy and are extensively used by our intended users. For example, a triangle with high target error is expected to have low feature response, meaning that the selected features are not strong enough to attract the triangle towards the correct target position. Dramatic changes in the triangle areas can be another sign of erroneous segmentations.

Moreover, *feature response profiles* are computed by our collaborating segmentation experts per triangle, after the adaptation of each mesh. As shown in Figure 7.2, the provided data of the profile of each triangle result into a number of discrete point values. These values are located along a ray parallel to the normal of the triangle, centered to the adaptation
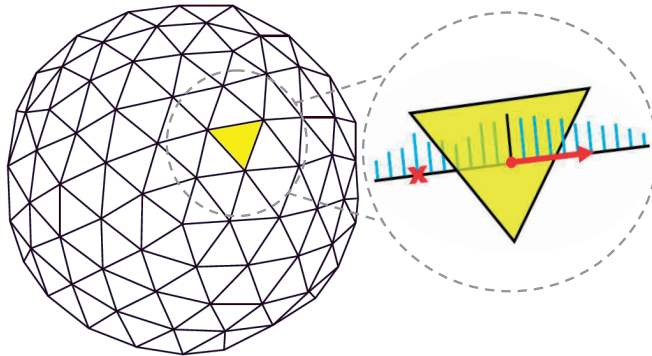
Figure 7.2: Response profile (cyan) of a triangle, after mesh adaptation, centered at the adaptation location (black) and parallel to the normal (red). There is also a latent peak, denoted with the cross.

location of each triangle. Each of these values indicate the strength of the features-rules that were used for the adaptation at each index position of this ray, and they relate to the above mentioned feature responses and target errors. During adaptation and profile computation, neighboring triangles are influencing each other, as well. For this, profile inspection in triangle neighborhoods, or in groups with similar response, can give a better idea of the reliability than individual triangles. In such an inspection, the number and locations of peaks – in essence, the local maxima – are important. Multiple peaks could indicate locations with high feature responses that are competing during the adaptation. Non-centered peaks could also be an indication of inaccuracy.

After a discussion with our collaborators, it resulted that they currently do not have an intuitive and easy-to-use way to obtain new insight into their segmentation outcomes with respect to the computed local quality measures, and the response profiles. They pointed out a number of *tasks* that they are interested in performing:

- For the *full cohort* of subjects:

    - Explore the distribution of local segmentation errors and response profiles `(T1-a)`.

    - Identify anatomical locations (organs or part of these) where the algorithm performs consistently `(T1-b)`.

    - Identify subjects that are special cases `(T1-c)`.

- For an *individual subject*:

    - Explore the distribution and anatomical location of the different local quality measures `(T2-a)`.

    - Discover relations between local quality measures `(T2-b)`.

    - Identify response patterns, for reliability evaluation `(T2-c)`.

7.4.   Related Work

Visual analytics for the exploration of segmentation outcomes has been addressed in several recent papers. Among them, there are frameworks for the *analysis of the impact of parameters* on segmentation algorithm outcomes, such as in the work of Torsney et al. [274] and Fröhler et al. [90]. Also, there is recent work on *shape variability analysis* [42, 149, 117]. However, the focus of these two categories is not on evaluating the employed segmentation algorithms and their results.

Another category comprises *comparative visualization*, which rather deals with the qualitative or visual evaluation of two segmentation outcomes. Busking et al. [41] proposed visualizations for the comparison of two surfaces, using different kinds of visual or graphical variables. In other papers, simple overlays [94] or extensions of checkerboard visualizations on 2D imaging slices [170, 240], but also side-by-side comparisons of 3D volumes have been used [5]. Visual variables, deformations, glyphs [317] and combinations of these have also been employed [194]. Specifically for mesh comparison, MeshLab [53] and PolyMeCo [253] have been proposed. Most of these papers refer to comparing two subjects, or one subject with a reference.

Comparison of multiple subjects was only recently tackled by Schmidt et al. [241]. In this work, a visual tool for the comparison of meshes is proposed, enabling the interactive exploration of their differences. This tool is meant for evaluating meshes generated by different algorithms with respect to a reference mesh and it is not fully applicable to the data that we are dealing with. It does not allow to explore and compare any local quality measures along with response profiles, which are necessary for our application. Additionally, it is limited to evaluating the visual quality of the resulting shapes. This is predominantly done in user-selected regions, which need to be interactively inspected.

For the *evaluation of the segmentation process and outcome* von Landesberger et al. [290] visualize the progress of quality during the segmentation of one organ. This approach enables the analysis of the segmentation process, but it is limited to one subject. Later, they improve this by proposing a method to show the distribution of quality values globally and to select cases with high or low quality values for a detailed inspection [292]. This strategy still does not allow the comparison of local quality measures across all subjects. In a more recent paper [291], von Landesberger et al. present a system for assessing and comparing segmentation quality across multiple datasets. A drill-down approach from an overview of a group of subjects to a detailed view of user-selected cases is employed.

As follow-up, Geurts et al. [96] propose a method for the visual comparison and evaluation of 3D segmentation algorithms. The goal is to determine the best segmentation algorithm, among different alternatives. To this end, they investigated both global and local approaches. Both previous works [291, 96] are similar to ours, but they are not fully applicable to our available data and tasks. This especially holds for the tasks related to the exploration of the segmentation response and the relations between local quality attributes. Table 7.1 shows schematically which requirements are (not) fulfilled by the most relevant previous related work.

Table 7.1: Requirement analysis concerning our application, for the tasks defined by the intended users and described in section 7.3 (✓: fulfilled; ✗: not fulfilled; o: partially fulfilled, or profile response not fulfilled; –: non-applicable).

	Compatibility with Data	Multiple Subjects	Compatibility with Tasks					
			(T1-a)	(T1-b)	(T1-c)	(T2-a)	(T2-b)	(T2-c)
Parameter Space Visualizations	✓	✗	–	–	–	–	–	–
Shape Variability Visualizations	✗	✓	–	–	–	–	–	–
Earlier Comparative Visualizations	✗	✗	–	–	–	–	–	–
[SPA*14]	✗	✓	✗	o	o	✗	✗	✗
[vLAA*13]	✗	✗	✗	✗	✗	✗	✗	✗
[VLBK*13]	✗	✓	✗	✗	✗	✗	✗	✗
[vLBB15]	✗	✓	o	✓	✓	✓	✗	✗
[GSK*13]	✗	✓	o	✓	o	✓	o	✗
Our proposed approach	✓	✓	✓	✓	✓	✓	✓	✓

7.5. Visual Analytics for the Exploration and Assessment of Segmentation Errors

The segmentation algorithm [239] is applied on imaging data from a cohort of subjects. Then, the respective triangulated meshes are generated, along with the measures described in section 7.3. Our approach enables the exploration and analysis of these measures using the components shown in Figure 7.3:

- The exploration of the *full cohort* of subjects.

- The exploration of an *error hierarchy* to detect special subject cases.

- The exploration of an *individual subject*.
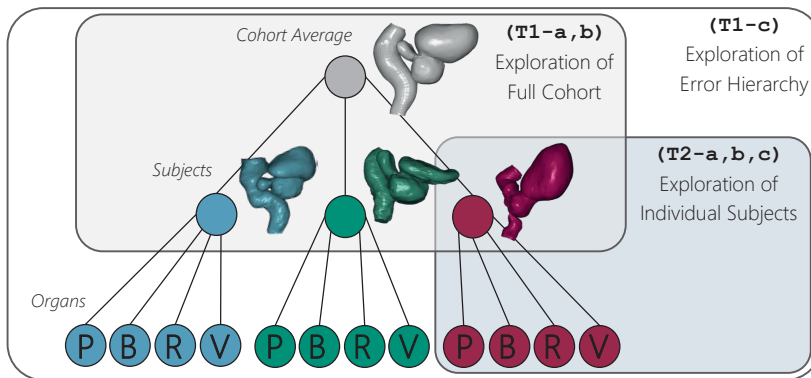
Figure 7.3: The three main components of our approach, together with the tasks from section 7.3 that they address. The abbreviations denote the different organs (P: prostate, B: bladder, R: rectum, V: seminal vesicles).

7.5.1.   Exploration of the Full Cohort

When exploring the full cohort of segmentation outcomes, segmentation experts initially need to explore the distribution of local segmentation errors and the respective response profile values `(T1-a)`. Visual comparison of individual outcomes, though, may be time con-suming, but also limited due to perception and screen space constraints [96]. For this reason, we decided to provide an overview at the triangle level. As mentioned in section 7.3, the in-dividual subjects of the cohort have a triangle-to-triangle correspondence. Thus, at each triangle position, we compute the mean and the standard deviation across all subjects, of both the target error and the response profiles.

For the *target errors*, mean and standard deviation are plotted in a scatterplot, where each data point represents one triangle location, as depicted in Figure 7.4. Data points in the scatterplot are rendered with a lowered opacity, to reduce clutter from overlapping points and as a density indication. We call this representation a *confidence scatterplot*, as it can provide information about three main regions of confidence based on the values of the mean and the standard deviation of the target error, as shown in Figure 7.4.

To convey additional information about the distribution of the mean and the standard deviation of the target error across the triangles of the mesh, we denote the first, second and third quartile of the respective distributions. In a confidence scatterplot, points with low mean and low standard deviation represent triangles where the algorithm performs system-atically well. Points with high mean and low standard deviation represent triangles where the algorithm performs systematically poorly. Finally, points with high standard deviation correspond to uncertain areas. This is related also to task `(T1-b)`.
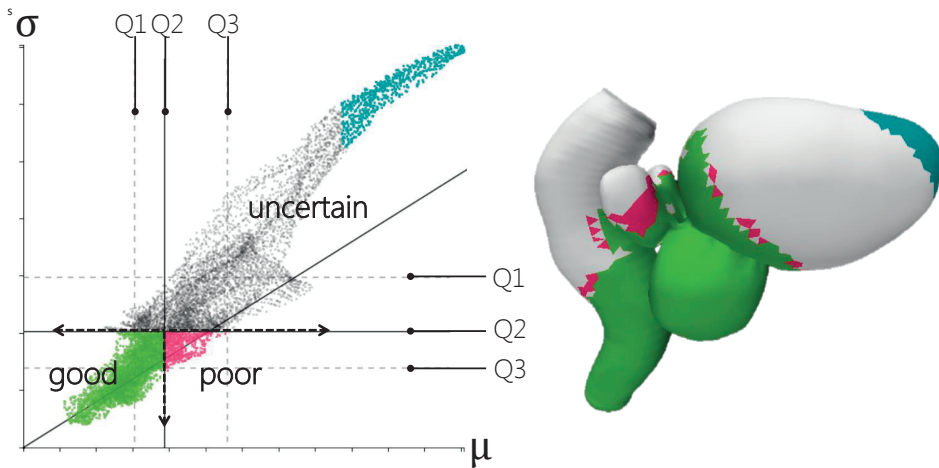
Figure 7.4: Confidence scatterplot of the mean error μ against the standard deviation σ, of the target error of all subjects. We denote the three areas of performance (good performance: $\downarrow\mu\downarrow\sigma$, poor perfor-mance: $\uparrow\mu\downarrow\sigma$, uncertain performance: $\uparrow\sigma$) `(T1-a)`. Three selections are made for good (green), poor (magenta) and uncertain (cyan) performance, and links to the anatomy are shown `(T1-b)`. $Q1 - Q3$ indicate the three quartiles.

For the *response profiles* a different approach is followed. As already mentioned in section 7.3, the reliability of the algorithm can be assessed from the inspection of profiles in triangle neighborhoods, and especially for triangles with similar response profiles. To this end, the peaks, namely the local maxima, are considered. A region of triangles with single-peaked response profile is more likely to be accurate than a region with multiple peaks. One option to illustrate this would be to reduce the mean profile information to a single scalar, representing the number of peaks. However, this would not convey the entire information about the mean profiles. For this, we retrieve clusters of mean profiles with a similar shape. These clusters can then be represented and visualized by an average profile.

Several *clustering* approaches can be employed [132]. However, determining a-priori the optimal value of clusters can be difficult and time consuming. For this reason, approaches such as *k*-means were discarded. For our application, we consider it more suitable to use a hierarchical clustering method. The computation of clusters with similar mean profiles is done, using an agglomerative hierarchical clustering method [297]. Initially, the number of clusters is equal to the number of triangles. This is followed by a phase, where iteratively the two most similar clusters are merged. Once a cluster is created, a representative, average profile is used in the next iteration. Clustering is performed with the similarity S between two normalized profiles p and q being:

$$S(p,q) = \sum_i 1 - |p[i] - q[i]| \qquad (7.1)$$

where i is the index location of each triangle profile.

In this way, two mean profiles with close-by peaks are assigned a higher similarity score, than two mean profiles with peaks further apart. After all iterations are finished, this algorithm results in a dendrogram, which can be interactively browsed. For visualization purposes we employ a collapsible *profile tree* metaphor, with the root being the average representative profile of all triangles. This can be expanded, revealing all underlying depth levels of clusters. The user can inspect the contents of the clusters interactively, without requiring to define a-priori the preferred number of clusters.

Each representative profile from a cluster is depicted in a one-dimensional visualization, also shown in Figure 7.5. In this visualization, the values at each index position of the representative profile are normalized to the range [0..1]. Each value corresponds to one square and is mapped linearly to a single hue, sequential color scale. In Figure 7.5, peaks are depicted in bright orange, while black denotes local minima. The size of the squares is inversely related to the standard deviation of a representative profile at each of the index positions. In essence, smaller squares indicate larger standard deviations, while bigger squares indicate smaller standard deviation. This encoding was inspired by the work of Höllt et al. [122].

After the exploration of errors and profile responses, segmentation experts need to identify whether their algorithm presents coherent behavior. For example, they need to identify the anatomic locations where their algorithm systematically fails or succeeds at a voxel level (T1-b). To this end, we enable *brushing and linking* both in the confidence scatterplot, and directly on the average reference mesh of the cohort. In this way, we establish a link between the anatomy and the computed target errors, as depicted in Figure 7.4. Brushing and linking is also applied from the profile tree to the scatterplots, which can be seen in Figure 7.5. Also,

7
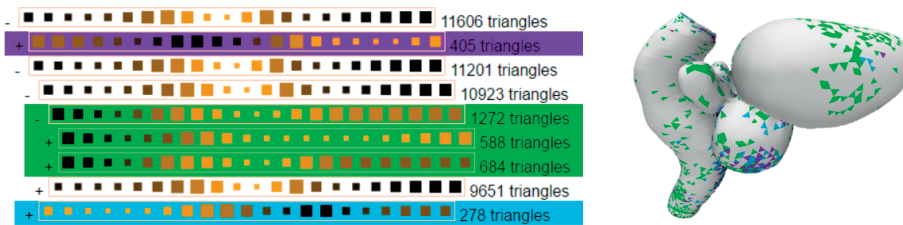
Figure 7.5: Profile tree visualization for the exploration of the clustering of profile responses (`T1-a`). The mean (encoded with the color of the squares) and the standard deviation (encoded with the size of the squares) are depicted for the 21 values (squares) of each representative profile (row). Three selections (purple, green, blue) are made to show the link to the anatomy (`T1-b`).

selections in the confidence scatterplot are followed by visualizing the respective average profile. In this way, all components are linked.

7.5.2. Exploration of the Error Hierarchy

The next required step is to provide an overview on the hierarchy of errors in the full cohort, and allow segmentation experts to identify subjects that are special, interesting cases (`T1-c`). For this, we employ a collapsible, undirected *tree graph*, to show an overview of the average target error and standard deviation in the full cohort. An example of this graph is illustrated in Figure 7.6.

The root of the tree represents the full cohort, which can be expanded to display the different subjects. These can be further expanded to depict the different organs. The size of the node encodes the magnitude of the average target error, while the magnitude of the standard deviation of the target error is encoded in the opacity of the node, but also in a halo around the circumference, as shown in Figure 7.6. To increase legibility, the nodes are sorted based on the average target error at each depth level. Also, to save screen space nodes of the tree that are not interesting for the analysis can be interactively collapsed. This representation summarizes the distribution of target errors in the cohort, across all patients and their respective organs. From this, users can be guided to select individual subjects that need further exploration in the next stage.

7.5.3. Exploration of Individual Subjects

Our tool fulfills also the requirement for a detailed exploration of segmentation errors in individual subjects. The first step involves a qualitative exploration of the resulting segmentation with respect to the *imaging slice data*. This qualitative exploration is shown in Figure 7.7. This exploration can give an initial indication of the outcome of the segmentation, as it shows the intersection of the resulting mesh with the imaging data. For the exploration of the distribution and anatomical locations of the different local quality measures (`T2-a`), *histograms* are employed, as depicted in Figure 7.8. Here, interactive selections provide a link to anatomy.

Discovering relations between local quality measures is also necessary (`T2-b`). For this, we initially enable the pairwise inspection of two measures, directly on the *mesh surface*. This is done by color encoding one local quality measure with a reduced heated body col-
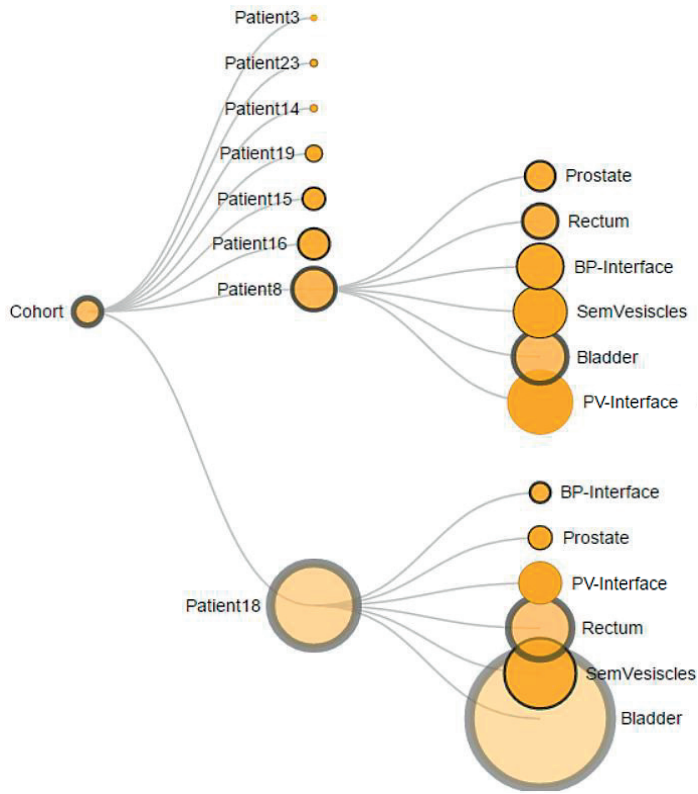
Figure 7.6: Exploration of the error hierarchy in the cohort using a collapsible tree graph representation (`T1-c`).

sagittal plane   coronal plane   axial plane
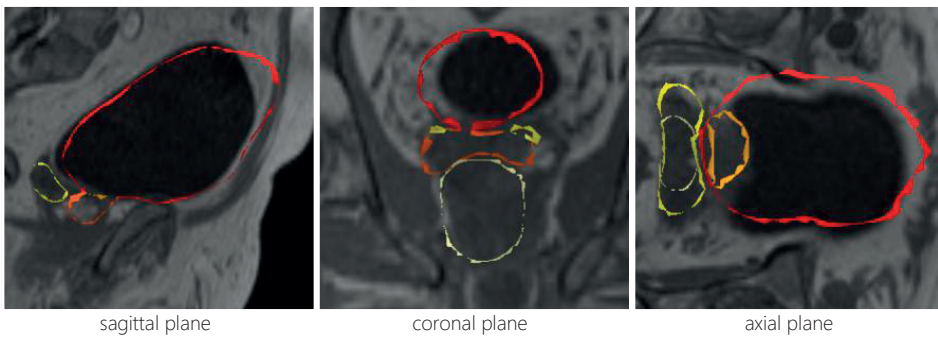
Figure 7.7: Qualitative exploration of the intersection of the segmented mesh with the imaging slice data (`T2-a`) (red: bladder, orange: prostate, yellow: seminal vesicles, white: rectum).
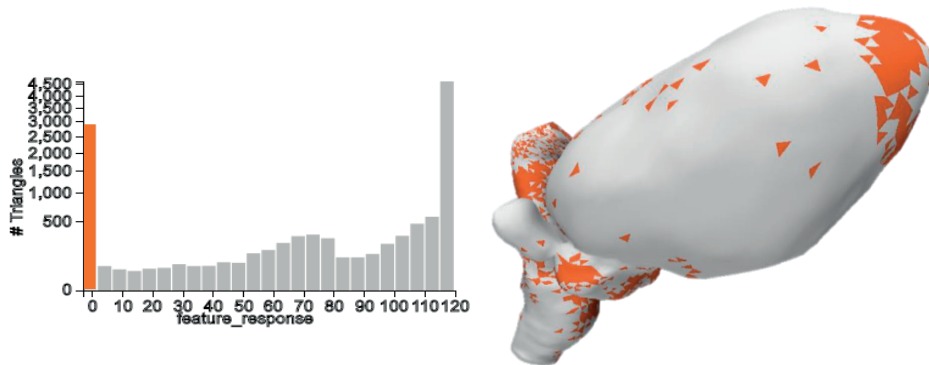
Figure 7.8: Interactive exploration of the distribution of local quality measures (`T2-a`). Selections in the histograms provide a link to the anatomy (orange).

(a)   (b)
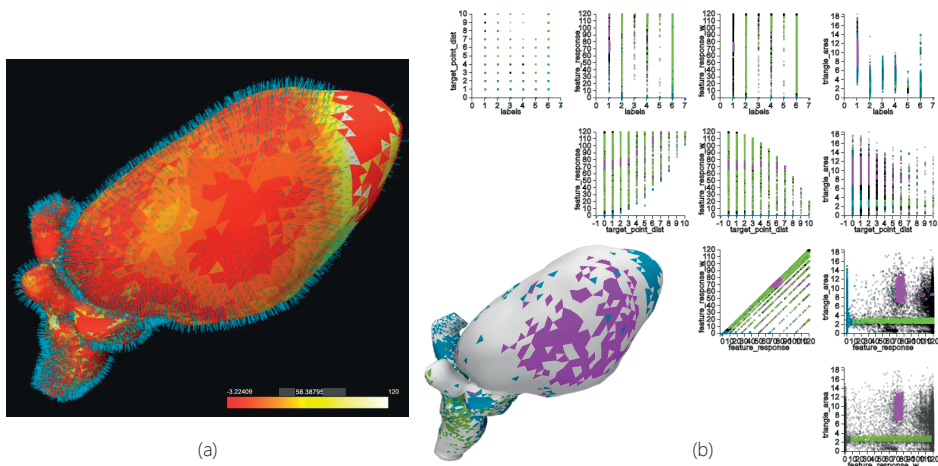
Figure 7.9: Discovering the relations between local quality measures. (a) Comparison of two measures, with color encoding and glyphs, directly on the mesh surface. (b) Multi-dimensional visualization of local quality measures in a scatterplot matrix, where selections (blue, purple and green) provide a link between different scatterplots, and also to the anatomy (`T2-b`).

ormap and a second measure with line glyphs, along the normal of each triangle of the mesh, as presented in Figure 7.9 - a. The size of the glyphs encodes the magnitude of the measure, at each triangle position. Still, this representation limits the exploration of relations to only two dimensions, while glyphs may introduce occlusion. To overcome this, a *scatterplot matrix (SPLOM)* is employed, as depicted in Figure 7.9 - b. The SPLOM was preferred over other multi-dimensional representations, e.g., parallel coordinate plots, due to the previous familiarity of the intended users. Brushing and linking in the scatterplot matrix facilitates finding and analyzing relations and patterns across multiple quality measures.
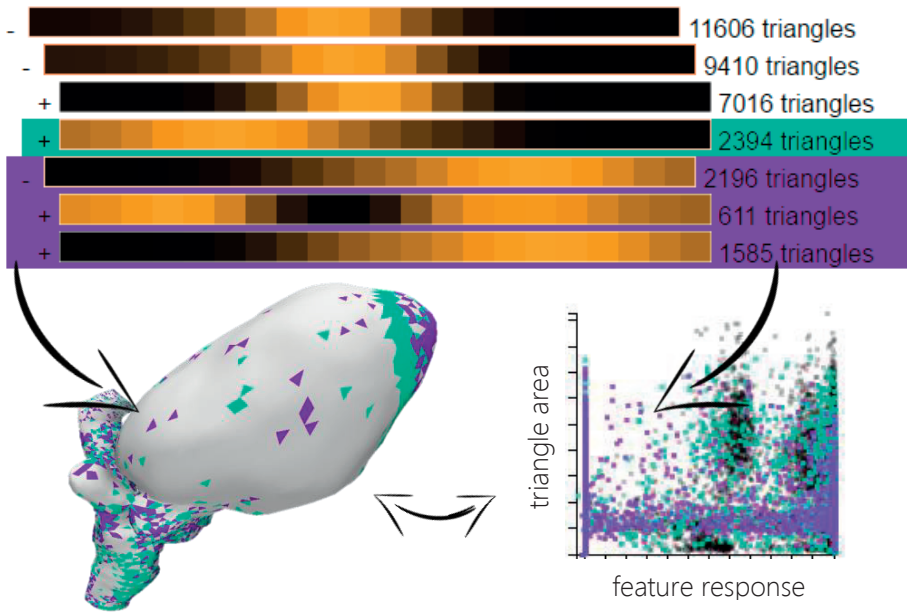
Figure 7.10: Profile tree visualization for the exploration of the clustering of profile responses of an individual subject (`T2-c`). Selections (cyan, purple) enable a link to the anatomy and the local quality measures.

Finally, the identification of patterns in the algorithm response enables segmentation experts to evaluate the algorithm reliability for each individual subject (`T2-c`). To this end, we use the same approach as the one proposed for the cohort exploration. Initially, we retrieve clusters of profiles with similar behavior, using the same hierarchical clustering method, as in task (`T1-c`). Then, a similar *profile tree* metaphor is employed. Here, each representative profile is depicted in a one-dimensional visualization that highlights the peaks of the profile clusters, as shown in Figure 7.10. As we have only one subject, the standard deviation encoding is not necessary. Interaction is employed to enable browsing the clustering hierarchy. Also, if a cluster is selected in the profile tree, the respective quality measures and the anatomical location are highlighted in the SPLOM and the mesh, respectively, as shown in Figure 7.10.

■ The application was developed in WebGL, using Three.js [3] and D3.js [29]. It is compatible with all browsers and platforms.

7.6. Usage Scenario Results

In this section, we elaborate on a usage scenario. Our purpose is to illustrate the functionality and some initial results that were achieved with our proposed visual tool. This usage sce-

nario has been guided by our collaborating segmentation experts, based on their previous knowledge and expectations. It was used to *explore* their data and to *confirm hypotheses* about their algorithm.

7.6.1. Dataset

The dataset employed for this usage scenario consists of a cohort of eight subjects. The explored data consisted of:

- CT volumetric data for all eight subjects, with dimensions $320 \times 320 \times 120$ and a spatial resolution of $1.563 \times 1.563 \times 1.000 mm$.

- A reference (average) mesh and the meshes of the eight subjects, each containing $11,606$ triangles with a triangle-to-triangle correspondence and organ labels.

- The respective local quality measures per triangle, as described in section 7.3.

- The profile response data per triangle. In this case, the profile responses consist of 21 scalar values at each triangle location.

7.6.2. Exploration of the Full Cohort

For the exploration of the full cohort, the average reference mesh is employed, together with the mean and standard deviation of the target error, and the mean and the standard deviation of the profile responses per triangle. Initially, the distribution of the local segmentation errors and profiles (`T1-a`), and also their anatomical correspondence (`T1-b`) are explored. In Figure 7.4, we illustrate in the confidence scatterplot the mean target error against the standard deviation at each triangle position, for the full cohort. From this, we can divide the algorithm performance into three categories: good, poor and uncertain. The good (green) and poor (magenta) categories are much less dispersed than the uncertain one.

Through brushing and linking, we can identify the anatomic regions of good performance, which correspond to the prostate and its very adjacent surfaces. In Figure 7.4, we denote some selections in green, which are the parts where the algorithm achieves high precision and high accuracy. Poor performance can be seen mainly in the seminal vesicles, shown in Figure 7.4 with magenta, which can be explained by the fact that seminal vesicles are small structures that may be hard to discern, and also are highly variable in shape. These are the parts, where the algorithm achieves low accuracy, but high precision. The rest, which is the biggest part of the bladder and also the top half of the rectum belong to the uncertain performance category. In particular, triangles of the bladder or the rectum that are further away from the prostate, denoted in cyan, are more uncertain. They have a low accuracy and low precision. This might be related to the high variability in the shape of these two organs. An additional reason for this might be that the employed algorithm produces segmentations used for radiotherapy planning in patients with prostate tumors. This is expected to affect structures closer to the prostate. Thus, the segmentation algorithm might be promoting better results for parts closer to the prostate.

The profile exploration in Figure 7.5 shows interesting results as well. Despite the fact that some triangles have an unusual profile response, where the peak was not centered.

These triangles are highlighted by the purple, green and blue selections. Still, these triangles manage to achieve a low target error, as they are influenced by neighboring triangles.

7.6.3. Exploration of Error Hierarchy

By exploring the error hierarchy in the tree graph (`T1-c`), we can identify the subjects and organs, where larger errors appear. From the representation illustrated in Figure 7.6, we identify *Patient18* as the subject with larger errors, and *Patient3* as the subject with smaller errors. For *Patient18*, the segmentation of the bladder has the largest error, while the interface between the bladder and the prostate seems to be well-segmented. For *Patient3*, the segmentation of all organs and their interfaces has small errors. *Patient8* is also another interesting case, where most of the organs have a high error. Here, the bladder and the interface between the prostate and the seminal vesicles segmentations have the highest errors, as depicted in Figure 7.6. From this exploration, we can select which subjects need to be explored individually, in more detail.

7.6.4. Exploration of Individual Subjects

In this section, we provide two usage scenarios, where we explore individually two cases – *Patient18* and *Patient3* – identified previously as the worst and best results, respectively.

■ *First case* – Patient18. For this subject, the segmentation outcome had the largest error. An initial qualitative exploration of the intersection of the segmented mesh, with respect to the imaging slice data (`T2-a`), as illustrated in Figure 7.7, indicates the locations where the algorithm had a bad performance. The prostate contour, denoted with white in the coronal slice in the middle view, seems to be well aligned with the borders of the organ on the CT slice. However, the bladder, shown in red, is not. The tip of the bladder has been missed and also a distal shift is visible.

The histograms in Figure 7.8 (`T2-a`) show a large peak in the distribution of *feature response* measure at the zero value, but also at the maximum value. The first peak indicates that for many triangles no suitable feature could be discovered, and that there may have been a problem with the feature selection. The respective triangles are located at the top of the rectum and on the tip of the bladder, denoted in Figure 7.8 with orange. The second peak corresponded to well-segmented triangles.

Upon inspection of the *feature response* against the *triangle area* measure (`T2-b`) in Figure 7.9 - a, we see that the tip of the bladder corresponds to a low feature response, denoted with the red color, and to low triangle area values, denoted with smaller-sized glyphs. We confirm this also in the SPLOM, shown in Figure 7.9 - b. Several clusters are easily identified when plotting the two measures, in the scatterplot in the fourth column and third row. These clusters include one with low response values (blue), one with low triangle area (green) and one cluster in the middle (purple). Selections provide insight into the physical location of these clusters, as can be seen in Figure 7.9 - b, revealing interesting information. The purple cluster corresponds to the areas at the sides of the bladder that presented the distal shift, in the previous qualitative exploration. The blue cluster corresponds to the wrongly segmented tip of the bladder and top of the rectum, and the green cluster corresponds to the well segmented regions of the bottom part of the rectum and the prostate. This exploration suggests a lack of strong features in the bladder and top of the rectum.

The profile information of this subject is also investigated (`T2-c`). Several triangles highlighted in Figure 7.10 are triangles where the feature response profile did not contain any peak close to the middle, denoted by cyan and purple. In practice, it is acceptable if a triangle has a feature response profile without peaks, as long as most of the neighbors do not present this same behavior. The position of neighboring triangles can positively influence the position of a triangle. In our case, these cyan and purple triangles add up to a total of $4,590$, which is almost $\frac{1}{3}$ of the mesh and they seem to form in their majority coherent regions. Therefore, an absence of peaks in these profiles indicates that no information was available on how to modify the triangles location and that the current location of the triangle is not supported by any of the selected features.

■ *Second case* – Patient3. For this subject, the segmentation outcome had the smallest error. An initial qualitative exploration of the intersection of the segmented mesh with respect to the imaging data (`T2-a`) showed that the segmentation outcome matches well the borders of the organs in the slices, as presented in Figure 7.11 - a. The histograms (`T2-a`) show a large peak in the distribution of the *feature response* measure at the zero value, but also at the maximum value (second row, left). The first peak indicates that for some triangles no suitable feature could be discovered, while the second peak corresponded to well-segmented triangles with a high feature response, as shown in Figure 7.11 - b with orange.

In the SPLOM (`T2-b`), the majority of points has a low triangle area, but there are several data points - triangles, with an outlier behavior, as denoted in Figure 7.11 - c, with green. Upon selecting those data points for further exploration, they correspond to the triangles on the top and the bottom of the rectum. We are also interested in seeing which parts of the meshes correspond to a low feature response, as shown in Figure 7.11 - c, with blue. These parts are few and scattered around the mesh. They have mostly a target error smaller than $4mm$ (second row, fourth column).

Another approach to investigate the segmentation quality is to select triangles on the mesh and inspect the attribute value distributions in the scatterplots. Figure 7.11 - d shows a selection containing the prostate and a small part of the adjacent organs and, also, a selection far away from the prostate, on the rectum (green). The scatterplots show the distribution of the selected triangles mapping the feature response against the triangle area and the target point distance. The majority of triangles far from the prostate (blue) have both high and low triangle area. The selection on the rectum (green) has low triangle area. As mentioned before, dramatic changes in these values indicate segmentation errors. The profile information of this subject is also investigated (`T2-c`). Only few triangles ($1,231$) have a profile without a peak in the middle. These are almost $\frac{1}{10}$ of all triangles, and they are evenly spread through the whole mesh. This is presented in Figure 7.11 - e, with the cyan and magenta selections.

7.7. Evaluation

To assess the value of our visual tool, we designed an evaluation, inspired by the paper of Lam et al. [159]. The evaluation was performed with five experts, working on developing segmentation algorithms. The group of participants included one professor in the field of Medical Image Analysis, three research scientists in the field of Image Processing and one
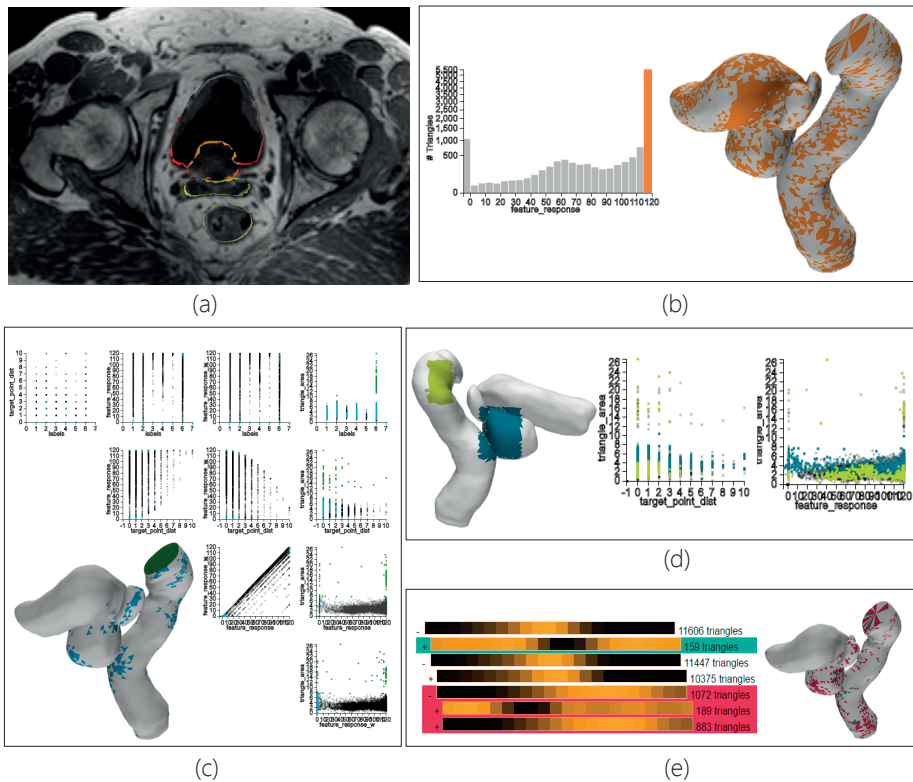
Figure 7.11: Usage scenario for a subject, where the algorithm has a good performance (Patient3). (a) Exploration of the intersection of the mesh with imaging slice data. (b) Exploration of the distributions of the quality measures and link to the anatomy (orange). (c) Exploration of relations between quality measures and link to the anatomy. (d) Selections directly on the mesh (green and blue), and exploration of corresponding quality measures. (e) Profile response exploration, and link to the anatomy.

scientist in the field of Computer Science. We did not include clinical experts, as they are not the intended users of our tool. Their experience with segmentation algorithms varies from seven years (for two people) to more than twenty years (for one person). All of them have also a radiological background. Four evaluators are male and one is female. They all have normal vision, two wear glasses and nobody is colorblind.

The evaluation had to be conducted remotely, and the participants were not able to interact with the tool. During the session, we demonstrated step-by-step the visual tool, using data provided by the experts and well-known to them. We demonstrated the main components of the tool, simulating the visual environment for the exploration and analysis of segmentation errors. The evaluation participants followed the demonstration. They could interrupt at any moment to make exploratory requests, such as selections and interactions that could help them analyze and understand their observations. We allowed them to discuss with each other these observations, but not their opinion on the tool.

To document their opinion on the demonstrated visual tool, they completed a questionnaire. This consisted of two parts. The questions of the first part were related to the tasks, presented in section 7.3. We divided task `(T1-a)`, to evaluate separately the visualizations employed for the error distributions and for the profile responses. Each question required an open answer, and also rating in a seven-point scale $[-3..+3]$. We evaluated four aspects:

- Utility (*Does it do what it is meant to do?*)

- Perceived Usability (*Would I be able to learn and use it?*)

- Appeal (*Do I like it?*)

- Overall Feeling (*How do I feel about it, in general?*)

The second part of the questionnaire included several questions regarding strengths, limitations, missing features of the tool and proposals for improvement.

■ *Ratings.* Figure 7.12 summarizes the results of the first part of our evaluation. There was no correlation with respect to the experience level of the evaluation participants. Most of the evaluated aspects ranked on the positive side of the scale (≥ 1), while only two received a neutral grade (0). All aspects have a median value of at least two, apart from one that has a median of one. The lower values were all documented for the profile response part of task `(T1-a, profiles)` and were all given by the same person. The error distribution part of task `(T1-a, errors)` and the error hierarchy exploration `(T1-c)` were, in general, rated higher than the rest.

■ *Open Answers.* The above mentioned ratings are in agreement with the open answers of the first part, and also with the second part of the questionnaire. The evaluators considered the tool to be overall intuitive and potentially easy to use. One evaluator commented that *it is a light-weight web-based tool, which makes it highly optimized for model-based segmentation analysis*, due to the involved large data. The feature that received most positive comments was the dynamic selection of triangles on the meshes, on the scatterplot and also their in-between link, i.e., `(T1-a, errors)`, `(T1-b)`, `(T2-a)` and `(T2-b)`. Yet, for the selection on the mesh, an evaluator commented that he would like to have visual feedback for the selections. Another appreciated feature was the tree graph, for exploring the error hierarchy in the cohort `(T1-c)`. One evaluator commented that he would actually like to use it, to explore a much larger cohort of segmentation outcomes.

■ *Feedback for Improvement.* Most of the participants gave feedback about improving the cohort profile response part `(T1-a, profiles)`. First, they commented that the visualization of the cohort profile responses takes some time to understand. It does not allow to change the similarity measure, apart from the mean values during clustering. One participant commented that the representation for the profiles in the cohort can be even reworked to be presented as an average curve, with a confidence band that denotes variability. This is in contrast to the positive opinion that the evaluators expressed for the individual profile response part `(T2-c)`, which was considered more intuitive and rated much higher. Another participant disliked, in particular, the glyphs used in task `(T2-b)`. These limitations were proposed as points for future work, along with a simultaneous visualization of multiple profile
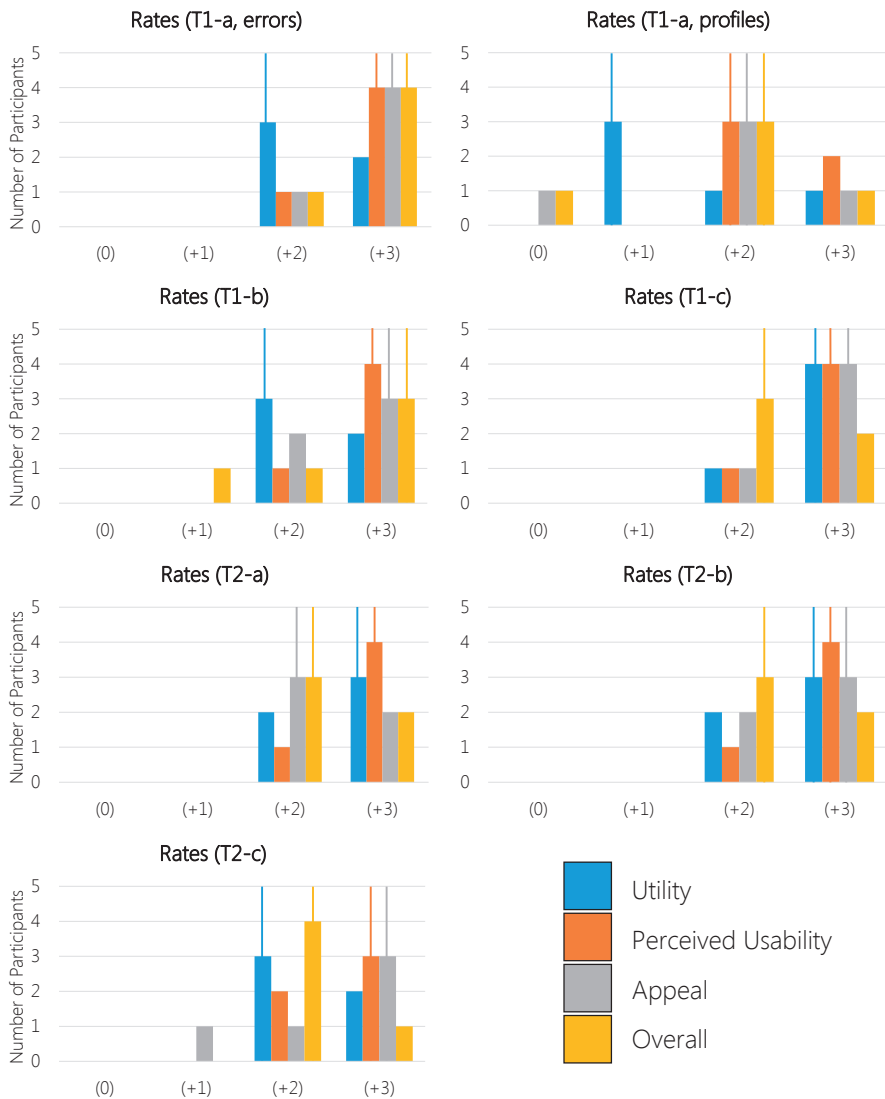
Figure 7.12: Rating results for the first part of our evaluation, for each of the tasks of our proposed tool. The scale range is $[-3..+3]$, but we only received answers higher than 0. With the additional vertical lines, we denote the median of each rating.

data. Also, functionality for annotating observations and for making a report from these, along with captured screenshots was proposed as future work.

7.8. Conclusions and Future Work

We introduced a visual tool to enable experts, working on algorithms for the model-based segmentation of pelvic organs, to explore and analyze the outcomes and errors of their methods. Our approach supports the global exploration of errors in a cohort of pelvic organ segmentations, where the performance of the algorithm can be assessed. Also, it enables the exploration and assessment of segmentation errors for individual subjects. We demonstrated the functionality of our tool with a usage scenario. Also, we performed an initial evaluation with five segmentation algorithm researchers, who confirmed the exploratory value of the tool, and gave feedback for future improvements.

A direction for future work includes improving the functionality for the exploration of the profile responses in the cohort. Adding functionality for the comparison of different aspects of the data, such as the local quality errors and profiles of different subjects, is also important. Exploring the impact of parameters used in the segmentation, and also the relation of the shapes of the various organs to the algorithm performance would be another interesting enrichment. An additional evaluation to quantify the experience of the user is needed, and should be conducted in the future. The proposed visual tool is a promising basis for segmentation experts. It allows them to gain more knowledge on the performance of their segmentation algorithms, and to determine strategies to improve their segmentation results.

7

8

Visual Analytics for the Exploration of Variability in Dose Planning Alternatives

Mr. Scott, there are always alternatives.

Mr. Spock, Star Trek

In this chapter, we present a Visual Analytics solution for the exploration of variability, in an ensemble of alternative radiotherapy dose plans. Our approach provides capabilities for the exploration and analysis of the impact of adjustments from the previous steps of the pipeline on the final dose planning outcome. Chapter 8 is based on the short paper:

■ Silva, P., Raidou, R.G., and Vilanova, A., 2016. Visualization of Variability in Radiotherapy Dose Planning. Proceedings of the 10th MedViz Conference (MedViz 2016), pp.63–66 [252].

8.1. Abstract

As already mentioned at the beginning of this dissertation, radiotherapy is currently able to offer improved tumor control and minimized radiation-induced toxicity. However, parameterizations, choices and assumptions during the earlier steps of radiotherapy planning pipeline may have an impact on the resulting dose plan. Although it is not known a priori which of these assumptions or parameter settings lead to better results, it is valuable for clinical researchers to understand their effect on the final dose plan.

In this chapter, we present an initial Visual Analytics framework, for the interactive exploration and analysis of the variability in an ensemble of possible dose plans. The proposed framework consists of multiple linked views, to allow for detailed, user-driven variability assessment. A usage scenario is employed to illustrate the usefulness of the framework for understanding the effect of adjustments in the previous steps of the pipeline to the final planning outcome. In addition to this, an initial discussion with domain experts provided us with feedback on the developed framework.

8.2. Introduction

To ensure the success of radiotherapy treatment, dose planning needs to be performed in advance. In the steps of the radiotherapy planning pipeline before the actual dose planning, different assumptions, choices and/or parameterizations can be made, resulting in different dose plans. In clinical practice, one particular strategy is chosen based on clinical knowledge and guidelines, and a single dose plan is performed and followed. However, in clinical research it is necessary to analyze and evaluate all possibilities. As it is not known a-priori which of the assumptions or parameter settings lead to better results, it is valuable for clinical researchers to understand how the different assumptions and parameterizations affect the final result. Therefore, there is a need for a solution that would enable clinical researchers to identify the relationship between adjustments of parameters or potential choices in the previous steps of the radiotherapy pipeline and the impact on the planning outcome.

Although multiple studies have been performed to assess the variability at certain steps of the pipeline [45, 260, 312], only few approaches explore the impact of this variability on the final dose plan [278], mainly due to a lack of suitable tools. In the present chapter, we present a visual analysis framework that enables clinical researchers, working in radiotherapy dose planning, to interactively explore and analyze the variability in an ensemble of possible dose plans, each resulting from a different adjustment in the planning pipeline. With the aid of visualization, the user is able to gain insight into the effect of these adjustment on the variability across multiple dose plans at two levels: firstly, based on the radiotherapy dose iso-contours across the dose plans, also referred to as *iso-doses*, and, secondly, directly at a voxel level.

The remainder of this chapter is organized, as follows: Section 8.3 includes all the necessary background information. Section 8.4 provides an overview on previous work, and Section 8.5 is the core of this chapter, where our proposed approach is presented and discussed. Section 8.6 presents the results of a usage scenario and the feedback we received from an informal discussion with our intended users. Finally, Section 8.7 concludes the chapter with a discussion and propositions for future work.

8

8.3. Clinical Background

A radiotherapy dose plan indicates how much radiation dose should be delivered, at which location of the body of the patient. It consists of a dose distribution, shaped according to the size of the target volume of the tumor to be treated. A common way to display and evaluate a radiation dose distribution is through the use of the so-called *iso-dose lines*, which represent points or zones in the body of the patient that receive equal doses of radiation [300]. An example of a radiotherapy dose plan and the employed iso-doses is presented in Figure 8.1.
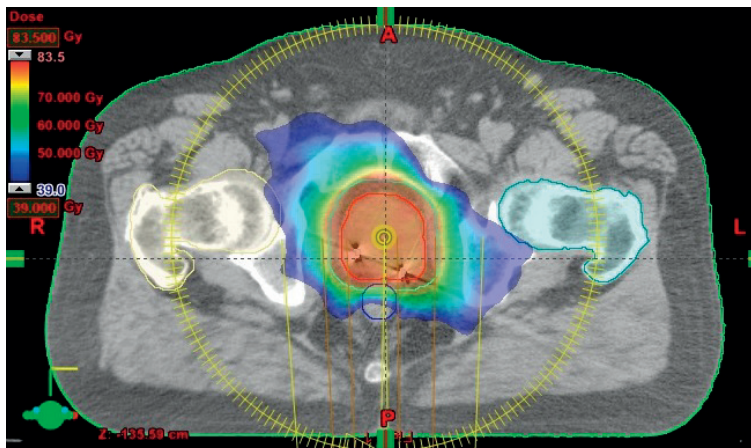
Figure 8.1: An example of a radiotherapy dose plan and the employed iso-doses, indicating how much radiation dose should be delivered, at which location of the body of the patient.

Outside of the radiation therapy application field, an equivalent of an iso-dose line is an *iso-contour*. An iso-contour, or simply contour, is a curve connecting points that have the same particular value of a function [242]. Inside the volume of the data of the patient, an iso-contour can be defined as:

$$f(x, y, z) = c \tag{8.1}$$

where c is the iso-contour value, iso-value or iso-dose level, and (x, y, z) is a particular position in the volume. This equation divides the volume into inside and outside regions of the contour. In the former, the values are higher than a specific iso-dose level ($f(x, y, z) > c$), while the latter contains values lower ($f(x, y, z) < c$) than the iso-dose level.

The process of creating a dose plan can be subjective to variability, due to the different assumptions, parameter settings and choices that can be made along the entire planning pipeline, as presented in Chapter 2. Among all these, the most significant would be imaging inaccuracies, errors in dose calculations, different parameter values in image analysis, target misalignment during image acquisition, and contouring variability [260]. Depending on all these assumptions and choices, the final dose plan might be substantially different. Therefore, these assumptions and choices might have an impact on the final treatment of

the patient. In this work, we will refer to the resulting set of possible dose plans as *ensemble of dose plans*, and each dose plan will be called an *ensemble member*.

As mentioned before, clinical researchers are interested in investigating the impact of different choices, assumptions or parameterizations on the resulting dose plan, and especially how these adjustments relate to the induced variability in the dose planning outcomes. They are interested in investigating how the different dose plans vary within one or more iso-dose levels (*contour-based approach*), but they also require more localized information on the variability (*voxel-based approach*).

At a contour-level, clinical researchers need to know whether specific adjustments induce high or low variability in the outcome of the dose planning, whether this variability is consistent throughout the whole plan or more/less prominent for specific dose levels and, also, whether specific adjustments result into significant differences in the outcome. At a voxel-level, it is more interesting to identify regions where the dose plans present higher/lower variability and, also, how the dose distribution is shaped for specific regions of interest. In particular, the following tasks are relevant for the exploration of the variability in an ensemble of radiotherapy dose plans:

- At an iso-dose level (contour based approach):
  - *Exploration of iso-dose variability* – How different are the dose plans in relation to a particular iso-dose? (T1)
  - *Comparison of iso-dose variability* – How does the variability between dose plans changes for different iso-doses? (T2)
  - *Dose plan outlier detection* – Along the iso-doses, are there any dose plans that significantly differ from the rest? (T3)

- At a voxel-level (voxel-based approach):
  - *Global overview* – Where are the (regions of) voxels with higher (dis)agreement between dose plans? (T4)
  - *Localized view* – Given a (region of) voxel(s), what is the distribution of doses among the dose plans? (T5)

8.4. Related Work

In radiotherapy, a common tool to evaluate different dose plans involves the use of dose-volume histograms (DVHs) [300]. A dose-volume histogram is a plot that indicates which percentage of a structure receives a certain dose. As an extension to DVH, Trofimov et al. [278] proposed the concept of DVH bands to enable the visualization of variability in the outcome of a treatment plan. However, all DVH-related approaches are able to provide only global information – not at a contour or voxel level – lacking also in providing spatial information, with respect to the anatomy of the patients. Rudimentary juxtaposition approaches are often used to add spatial information. These are sub-optimal in insight, as they require from users to use their memory to visually compare the different plans.

Outside of the radiation therapy application field, a standard visualization technique for exploring iso-contour distributions in ensemble data employs *spaghetti plots* [70]. In these plots, the contours of all ensemble members are simultaneously rendered in a single representation. This representation may suffer from clutter, while it does not facilitate trend or outlier detection. Enhancements of spaghetti plots have already been presented in Chapter 3 [85, 216, 238], but they are not fully applicable to our case, especially due to their complexity.

To deal with multiple three-dimensional surfaces, Busking et al. [41] and Alabi et al. [5] proposed several methods that allow their comparison. More recently, Demir et al. [67] used screen-space silhouettes instead of solid surfaces to explore the variability in ensembles of iso-surfaces in an interactive way. Other similar approaches have also been discussed in Chapter 7. Still, these approaches are limited to a small number of ensemble members, and can easily suffer from occlusion and visual clutter. Methods, such as the ones presented in Chapter 3 for the probabilistic visualization of iso-contours [203, 209, 210, 211, 212], have also been vastly employed.

Finally, Whitaker et al. [305] presented an approach, which requires the quantification and visualization of statistical properties on the size, position, shape and structure of ensembles of iso-contours or iso-surfaces. Through the use of non-parametric statistical methods, the method quantitatively shows the variability of the contours, in a way that resembles the descriptive statistics of conventional boxplots. This includes the mean, median, outliers, and the envelopes containing the 50% and 100% of the ensemble members. The contour boxplot was extended to 3D, for visualizing the variability in ensembles of iso-surfaces [228].

Overall, several visualization and exploration methods have been developed in other areas of application. Most of them are suitable for the contour-based aspect of our approach, while others are more appropriate for a voxel-based analysis. In our approach, these visualizations will be employed and enhanced to fulfill the specific needs of our dose planning application.

8.5. Visual Analytics for the Exploration of Dose Planning Variability

As already mentioned, the proposed approach comprises two main components, as shown in Figure 8.2: the contour-based and the voxel-based component.

8.5.1. Contour-Based Analysis of Variability

The contour-based approach, which is presented in Figure 8.3, aims at providing means for exploring the variability in an ensemble of dose plans, at one or multiple iso-dose levels. It incorporates the following three sub-components:

■ *Exploration and analysis of iso-dose variability* (T1). At a contour-level, the goal is to be able to identify major trends and outliers in an ensemble of iso-doses. We use an adaptation of the contour boxplot method [305]. Instead of using the original contour boxplots visualization, we use opacity to encode the band depth value. As a result, iso-doses with high band depth are encoded into high opacity, while iso-doses with low probability are encoded into

low opacity. Identically, the same concept is applied to the bands: the 50% band, since it contains the 50% of deepest iso-doses, is encoded into higher opacity, while the 100% band is encoded into lower opacity. Additionally, the median iso-dose is rendered thicker and with a different color, while the outliers are kept dashed, as shown in Figure 8.3 - a. At the same time, the method must be compatible in a situation where multiple iso-doses are displayed simultaneously. In the present case, a maximum number of three simultaneous iso-doses was considered. The decision was to use yellow, red, and blue. Alpha blending is used to show the overlap of different iso-dose bands.

■ *Comparison of variability along iso-doses* (T2). The goal of the second component is to provide an overview of the variability, so that the user can identify and compare, immediately, which iso-doses have higher or lower variability. Also, it is possible to pick interesting doses and interactively explore them, using a bar chart visualization, as shown in Figure 8.3 - b. The horizontal axis represents the discrete iso-dose levels, while the vertical axis represents a probability that is indicative of the variability at every iso-dose. This component is linked with (T1).

■ *Outlier detection* (T3). In this part, the goal is to compare the dose plans through their iso-doses in a global way. In this way, it is possible to identify which dose plans differ significantly from the rest along the range of iso-doses. Every iso-dose of the ensemble has an associated band depth value, which we encode through a heatmap with a gray scale colormap. Iso-doses with high probability are mapped to white, while iso-doses with low probability are encoded to black, as shown in Figure 8.3 - c. When hovering the mouse on the heatmap, tooltips appear with detailed information on the underlying cell.
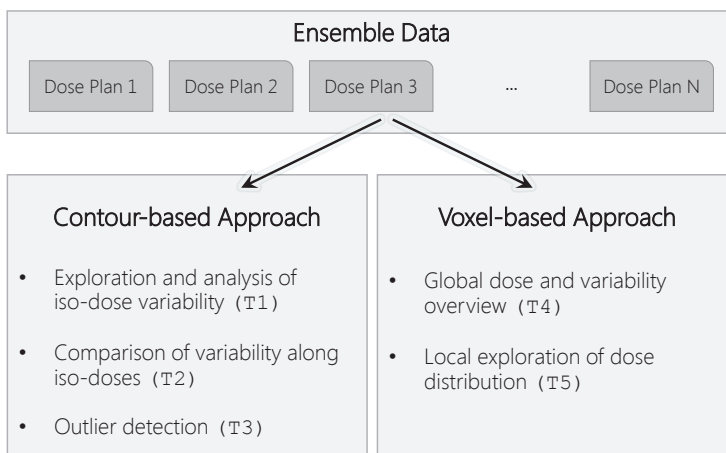
Figure 8.2: The two components (contour-based and voxel-based approach) of our proposed Visual Analytics framework for the exploration of dose planning variability.
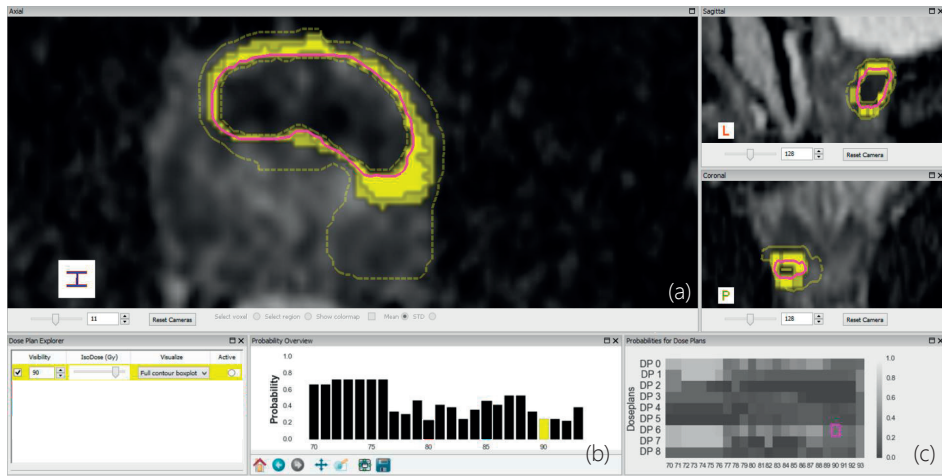
Figure 8.3: Illustration of the contour-based approach at an isodose level of 90 Gy, for an ensemble of 9 simulated plans. (a) Contour boxplot representation: the high opacity yellow band is the 50% band, the low opacity yellow band is the 100% band, the dashed yellow line corresponds to outliers and the magenta line to the median iso-dose (T1). (b) Bar chart visualization, which encodes the variability at every isodose: here, higher probability indicates lower variability (T2). (c) Heatmap visualization, where darker grey corresponds to lower probability, i.e., to outliers (T3).

8.5.2. Voxel-Based Analysis of Variability

The voxel-based approach, which is presented in Figure 8.4, aims at providing means for exploring the variability in an ensemble of dose plans, within one or more anatomical regions, at a localized voxel-level. It incorporates the following two sub-components:

■ *Global dose and variability overview* (T4). At a voxel-level, the goal is to be able to identify voxels or regions with higher or lower variability. To show the dose magnitude, we calculate an average dose plan by computing the mean at every voxel. To show variability, we calculate the standard deviation at every voxel. For a general overview of these two facets of the data we employ a 2D color map, as shown in Figure 8.4 - a.

Also a scatterplot, depicted in Figure 8.4 - d, is employed. Every point in the scatterplot represents a voxel from the data, and is represented by the mean in the horizontal axis, and the standard deviation in the vertical axis. From this representation, voxels with lower/higher variability with respect to a specific dose level can be discovered. To diminish clutter from point overplotting, datapoints in the scatterplot are rendered with higher opacity. Additionally, to help locating dense regions, a bi-variate kernel density plot overlaid on top of the scatterplot was used as an additional visualization method. Still, the scatterplot only is not able to provide any kind of spatial information. To overcome this limitation, we employ brushing and linking from the scatterplot to the anatomy of the patient, illustrated in Figure 8.4 - c.

■ *Local exploration of dose distribution* (T5). Additional visualization methods were incorporated to dynamically explore the distributions at every voxel. Based on 2D slices, we enable
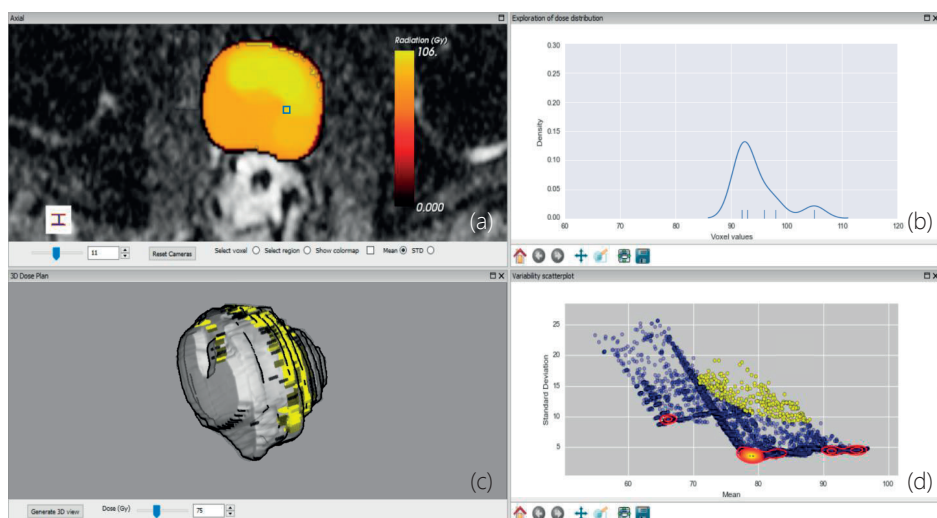
Figure 8.4: Illustration of the voxel-based approach. (a) Color mapping of mean radiation dose value. (b) Inspection of the distribution at a particular voxel position, as selected in (a). (c) Anatomical reference for voxels that were selected in the scatterplot. (d) Scatterplot with overlaid KDE and selected points in yellow. `(T4)` is accomplished with the views (a,c,d), while (b) is related to `(T5)`.

the user to interactively probe the voxels and get a detailed notion of the underlying distribution of values, as shown in Figure 8.4 - b. The resulting visualization depicts a simultaneous combination between a kernel density estimation (KDE) and a rugplot, where the KDE curve provides an indication where the value density is greater. We also enable a region of interest (ROI) selection: instead of considering a single voxel, a ROI can be selected, and the distributions are displayed for every voxel inside the region.

■ We implemented the proposed approach in Python as a DeVIDE module [32], employing the Visualization Toolkit (VTK) [2].

8.6. Evaluation

In this section, we present a usage scenario of our proposed approach with simulated data. The presented usage scenario is meant for demonstration purposes only, and no clinical observations should be inferred from it, as it is based on artificial data. In addition to this, we present the outcome of an initial informal discussion with clinical researchers.

8.6.1. Usage Scenario

The simulated dataset consists of 9 dose plans. One of the dose plans is from an actual prostate cancer patient, depicting a real dose distribution. The remaining dose plans were simulated from the original plan. All dose plans contain the same dimensions ($256 \times 256 \times 24$

voxels) and spacing between voxels ($1mm \times 1mm \times 1mm$). An additional MRI scan data set is used, to provide anatomical reference for the prostate and adjacent organs.

Starting with the *contour-based approach*, the user first can inspect the bar chart overview to get a quick notion of the variability in the iso-doses, as shown in Figure 8.5 - b. At every iso-dose, variability is inferred through the probability or band depth value. Here, for doses ranging from 70 to 76 Gy we denote a higher probability, as iso-contours are relatively consistent, with low variability. On the other hand, probability reaches a minimum value at 80 Gy, which indicates this dose value is associated with higher variability. Additionally, it is interesting to notice the behavior of variability along the iso-doses: variability is lower for lower doses, while higher doses are associated with a higher variability.

After being provided with an overview, the user can pick up three doses in the bar chart, which are directly reflected on the adjacent operation view and on the anatomical views. In this case, the doses selected were 75 Gy, 80 Gy, and 85 Gy, as shown in Figure 8.5 - b. For 75 Gy, the dose plans are relatively consistent, as illustrated in Figure 8.5 - a. The width of the 100% band is relatively small, which indicates there is low variability in the position of the iso-contours. This might indicate that the simulated adjustments that generated the ensemble of the dose plans do not have a big impact on the dose planning outcome for the 75 Gy iso-dose. The median iso-contour, with magenta color, is identified as belonging to the original dose plan (dose plan 0), as the hovering balloon suggests. Thus, in this case, the original dose plan is considered the most representative of the ensemble for a dose of 75 Gy. On the other hand, it can be seen that one of the iso-contours is considered an outlier, depicted through a dashed line, as shown in 8.5 - a. The difference between the median and the outlier can be further explored, using the heatmap.

On the other side, for 90 Gy, the use of the contour boxplot reveals a higher degree of variability among the dose plans, as Figure 8.3 suggests. This might indicate that the simulated adjustments that generated the ensemble of the dose plans result in a dramatic variability in the dose planning outcome for the 90 Gy iso-dose. In the axial plane, one can immediately identify thicker bands, compared to 75 Gy, as shown in Figure 8.5 - a. In contrast with 75 Gy, it is possible to detect a significant outlier. Additionally, the probability values are relatively lower, as it can be seen by the heatmap of Figure 8.5 - c. Overall, the dose plans significantly disagree more in 90 Gy, as opposite to 75 Gy.

Additionally, it is possible to compare the different dose plans along the range of iso-doses. This is accomplished through the heatmap visualization, as shown in Figure 8.3 - c and Figure 8.5 - c. It can be observed that, along the iso-doses, dose plans 0 (original) and 6 are the most representative samples among the ensemble, with higher probability than the rest, as the colormap suggests. Dose plans 1 and 7 are also closer to dose plans 0 and 6, for lower doses. However, as the dose increases, their probability decreases significantly. Lastly, dose plans 2 and 5 are mostly associated with low probabilities along the iso-doses. This means that these dose plans are more likely to be considered outliers, since their respective iso-contours are quite different from the remaining iso-contours in the ensemble. Here, the heatmap view offers the possibility of comparing dose plans along the iso-dose range.

Regarding the *voxel-based approach*, a different kind of insight is provided. Using the color encoding in the slice viewer, it is possible to quickly locate regions of higher dose or lower dose. Yellow regions that are corresponding to high radiation dose, are located in the right lateral lobe of the prostate, as depicted in Figure 8.4 - a. Concerning the distribution plot

8
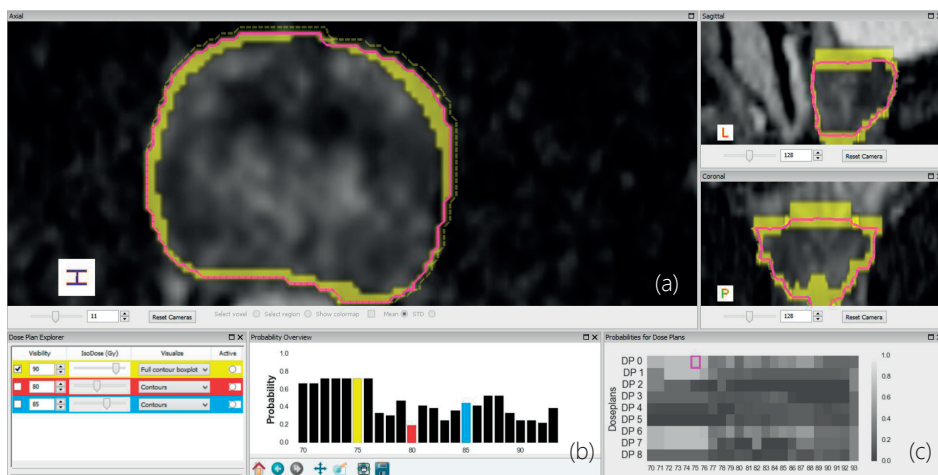
Figure 8.5: Illustration of the contour-based approach at an isodose level of 75 Gy, for an ensemble of 9 simulated plans. (a) Contour boxplot representation (`T1`). (b) Bar chart visualization (`T2`). (c) Heatmap visualization (`T3`).

of Figure 8.4 - b, the voxel highlighted with the blue box in Figure 8.4 - a, has a more spread out distribution, ranging between 90 and 110 Gy approximately, as shown in Figure 8.4 - b.

In the scatterplot, the overlaid KDE shows a high density of doses around 80 Gy, as presented in Figure 8.4 - d. Brushing voxels with higher standard deviation, highlighted with yellow, belong mostly to the boundaries of the dose plan, as denoted in Figure 8.4 - a, c. This suggests that such locations should be analyzed thoroughly, as these are the locations of higher variability.

8.6.2. Informal Discussion with Clinical Researchers

An initial informal discussion with domain experts was conducted in order to get feedback for the developed solution. The group of participants included three medical physicists, working in radiotherapy dose planning. Two of the participants work at the Netherlands Cancer Institute, while the third participant is with the Aarhus University Hospital.

At the beginning, the participants were given a small demonstration of the framework, where the main functionalities and components were shown interactively. The participants observed the demonstration and started a useful discussion with respect to the demonstrated visualizations.

Regarding the contour-based approach, they commented that it is an *interesting method for visualizing variability across treatment plans*. One participant appreciated the presence of interactivity for exploring the variability along the doses. Although the discussion participants do not currently incorporate this kind of analysis in their work, they mentioned it can be useful in a situation where it is interesting to *tune parameters and see how they affect the dose planning result*. Furthermore, all participants suggested ideas that could be incor-

porated in the future, such as employing Pareto fronts [300] or the DVH [278] as additional views on the variability. Concerning the voxel-based approach, all participants agreed that currently there are no means of performing a voxel-level analysis, and the proposed approach employs interesting methods, with the potential of providing insight into more details.

Overall, the participants agreed that the proposed framework provides an interesting tool to study variability. Currently, this aspect has not been extensively explored and incorporated into the dose planning process. Therefore, it might be an interesting approach to be considered in the future. The main strong point of the framework lies on the fact that it enables an interactive exploration of variability at a contour-level as well as a voxel-level, which was not possible before.

8.7. Conclusions and Future Work

We introduced a visual framework for the interactive exploration and analysis of variability in radiotherapy dose planning. It allows to visually assess the variability across multiple possible dose plans, as a result of different assumptions, parameter settings and choices that can be taken during the planning pipeline. The core aspect of the framework is the ability of providing insight on the variability at two different levels: through the iso-contours across the dose plans and directly at a voxel level. An initial, informal discussion with domain experts resulted in a positive feedback for the developed framework, who considered the integration of both perspectives an enrichment to the analysis process and a more complete perception of the underlying variability.

Points for future work include a more thorough evaluation, which would validate the initial observations of the domain experts. In addition to this, the incorporation of assumption or parameter sensitivity analysis directly into the framework would be an interesting improvement, which would enable to connect the previous steps planning pipeline with the dose planning variability visualization process. For example, it would be interesting to link the pharmacokinetic modeling step (Chapter 5) to see how different modeling choices affect the outcome of the dose planning, or how adjustments in the segmentation step (Chapter 7) propagate to the end of the pipeline. Such enhancements would require, though, to adapt several of the visualization techniques used. In this case, more scalable techniques or other approaches that would reduce the dimensionality of the involved data are required.

8

9

Visual Analytics for the Exploration of Tumor Control Probability Modeling

Almost all human life depends on probabilities.

Voltaire (1694 - 1778)

In this chapter, we propose a visual tool that enables clinical researchers working on TCP modeling to explore the information provided by their models, to discover new knowledge and to confirm or generate hypotheses within their data. Chapter 9 is based on the paper:

■ Raidou, R.G., Casares-Magaz, O., Muren, L.P., van der Heide, U.A., Rørvik, J., Breeuwer, M. and Vilanova, A., 2016. Visual Analysis of Tumor Control Models for Prediction of Radiotherapy Response. In Computer Graphics Forum (CGF), Vol. 35, No. 3, pp. 231-240 [221].

9.1. Abstract

As already described in Chapter 2, statistical models are often employed in radiotherapy research in order to quantify the probability that a tumor is effectively treated with a given radiation dose. These statistical models are called tumor control probability (TCP) models. Recently, TCP models started incorporating additional information from imaging modalities. In this way, patient-specific properties of tumor tissues are included, improving radiobiological accuracy. Yet, the employed imaging modalities are subject to uncertainties, which may have a significant impact on the modeling outcome. At the same time, TCP models are sensitive to a number of parameter assumptions. Currently, uncertainty and parameter sensitivity are not incorporated in the analysis, mainly due to time, human and computational resource constraints.

In this chapter, we present a visual tool that enables clinical researchers working on TCP modeling, to explore the information provided by their models, to discover new knowledge and to confirm or generate hypotheses within their data. Our approach incorporates the following four main components. First of all, it supports the exploration of *uncertainty* and its effect on TCP modeling. Moreover, it facilitates *parameter sensitivity analysis* to common assumptions and enables the identification of *inter-patient response variability*. Finally, it allows starting the analysis from the desired treatment outcome, to *identify treatment strategies* that achieve it. To evaluate our proposed approach, we conducted a study with nine clinical researchers. All participants agreed that the proposed visual tool provides better understanding and new opportunities for the exploration and analysis of TCP modeling.

9.2. Introduction

During radiotherapy planning, different treatment strategies can be followed, each considering several assumptions or choices. Examples of these assumptions include dose escalation, uniformity or non-uniformity of tumor irradiation, the amount of the received dose, and eventual fractionation of the treatment, which is the division of the total radiation into smaller doses per session over a period of time. In clinical practice, one specific treatment strategy is chosen among these alternatives, based on clinical experience and guidelines. Still, in clinical research, the goal is to thoroughly evaluate all possibilities. In this way, more targeted treatments can be designed and provided to clinical practice. To simulate and evaluate the effects of a specific strategy, clinical researchers employ Tumor Control Probability (TCP) models [302].

Conventional TCP models are statistical models that quantify the probability that a tumor is effectively controlled or treated, given a specific radiation dose. In plain words, TCP models aim at answering the question: *What is the probability Y that a tumor is treated with this strategy, given a dose X?* For example, Figure 9.1 depicts three different outcomes of a specific TCP model, each from a specific radiotherapy treatment strategy. In this case, by providing a total treatment dose of 77 Gy, the first strategy results in **88%** probability of treating the whole tumor, while the other two have a lower treatment probability response of **63%** for the second strategy, and **31%** for the third strategy.

In the last years, patient-specific tissue characteristics from imaging modalities started being included in planning. This has influenced also TCP modeling, where additional per

9
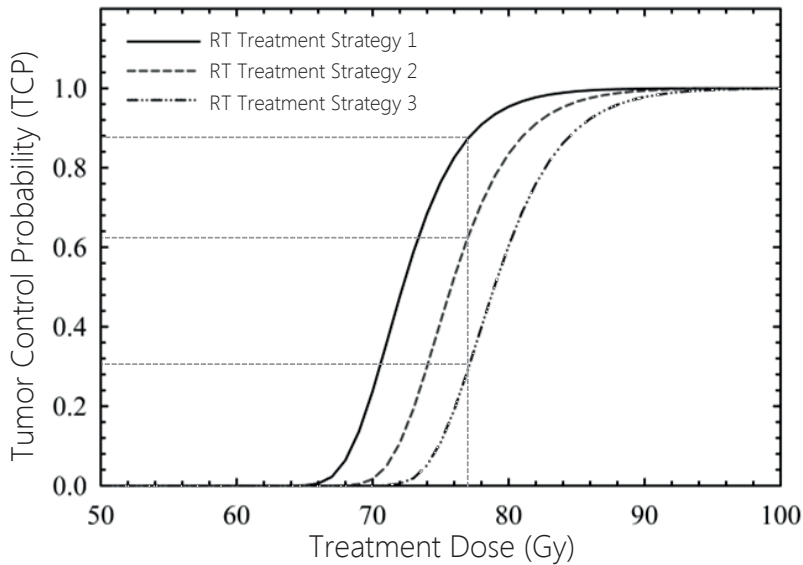
Figure 9.1: An example of the resulting TCP curves for a given dose, with three different radiotherapy treatment strategies. By providing a total dose of 77 Gy, the first strategy results in **88%** probability of treating the whole tumor, while the other two have a lower treatment probability response of **63%** and **31%**, respectively.

voxel information – in essence, properties indicative of tumor characteristics – are being incorporated [266]. In this way, clinical researchers can predict more accurately the tumor treatment probability at a voxel level, by adding radiobiological information in their statistical models, for example, from Diffusion Weighted (DW) MRI.

So far, several interesting aspects of TCP modeling are not incorporated in clinical research, due to complexity, lack of human and computational resources and time constraints. First of all, imaging modalities are subject to *uncertainties* with significant impact on the model outcome and the simulated treatment response [140]. Additionally, there are many different TCP models and different parameter assumptions in each one of them [258, 302]. Usually, these assumptions are educated guesses, and awareness on the *sensitivity* of the models is important.

Moreover, TCP modeling is often applied to entire patient cohorts, to investigate the *inter-patient response variability*. This knowledge can help clinical researchers to design more robust treatment strategies. Finally, clinical researchers are interested in exploring and analyzing their data in a *reverse* manner. Practically, they need to be able to answer also the question: *given a target treatment outcome for a tumor, identify the radiotherapy strategy(-ies) to achieve it?* In this work, we introduce a visual analytics approach to extend the exploration of TCP modeling, to cover also these topics that are currently not possible to be incorporated in the analysis.

9

Our *contribution* is the design and implementation of a visual tool that enables clinical researchers to explore the information provided by their TCP models, to discover knowledge and to confirm or generate hypotheses within their data. As far as we know, there is no other tool to serve this purpose. Our approach incorporates the following four main components:

- It supports the quantification and exploration of imaging-induced uncertainty and its propagation to TCP modeling.

- It facilitates exploring and analyzing the sensitivity of TCP models to different assumptions and parameter variations.

- It enables identifying and exploring inter-patient response variability, within cohorts.

- It allows, given a targeted treatment outcome, to identify the treatment strategies or parameters that would achieve it.

The remainder of this chapter is organized, as follows: Section 9.3 includes all the necessary background information. Section 9.4 provides an overview on previous work, and Section 9.5 is the core of this chapter, where our proposed approach is presented and discussed. Section 9.6 presents the results of our evaluation. Finally, Section 9.7 concludes the chapter with a discussion and propositions for future work.

9.3. Clinical Background

In the last years, radiotherapy research aims at designing more effective and better targeted treatments, to be applied in clinical practice. For the simulation and evaluation of all different treatment strategies, TCP models are being built [302]. Conventional TCP models are usually regression models that summarize empirical knowledge about the effect of radiation to tumors, representing the probability that a tumor is effectively treated with a specific dose [258, 302].

To achieve a more targeted treatment, tailored to the patient-specific tumor tissue characteristics, information from imaging modalities was recently incorporated to TCP modeling [266]. In this way, properties indicative of tissue characteristics were included to improve the radiobiological accuracy of modeling at a voxel-level [266]. In this work, we employ a novel TCP model that involves DW-MRI [48]. This in-vivo imaging technique measures quantitatively the water diffusion per voxel, from apparent diffusion coefficient (ADC) maps, and is employed to identify high-density tissue like tumors [25].

This ADC-based TCP model is subject to uncertainties [140], often due to calculation restrictions in the clinical setting or due to magnetic field inhomogeneities [25]. Although we are considering a specific TCP model, uncertainties are present in all modalities and our proposed approach could be extended also to them. These uncertainties need to be quantified and propagated into modeling to identify their effect on the prediction outcome. More details about the source, quantification and propagation of uncertainty in the employed TCP model are discussed in Section 9.5.

Additionally, all TCP models, including ADC-based ones, incorporate a number of different parameter assumptions. For example, in the explored TCP model, researchers make

9

assumptions for the amount of dose or fractionation, when quantifying the per voxel cell density from ADC maps, or when selecting values for parameters that model the survival and death of tumor cells after irradiation [48]. Still, it is not known which choices lead to better results, as well as what is the effect of different alternatives. Thus, the parameter sensitivity of the model needs to be incorporated in the analysis. In state-of-the-art clinical research, the ADC-derived uncertainty and the model sensitivity to parameter assumptions are not considered yet, as they cannot be explored with the existing tools. This is further obstructed by the fact that the evaluation and analysis of TCP models is usually applied to cohorts of patients, to account for inter-patient response variability.

Finally, the current TCP modeling workflow is based on the question: *What is the probability that a tumor is controlled, given a specific dose?* Yet, clinical researchers have not managed to find an easy and insightful way to answer the inverse: *Which radiotherapy delivery strategy can achieve a specific target treatment?*

After an extensive discussion with clinical researchers working on TCP modeling, we defined together the most relevant open tasks for their research, which are also summarized in Figure 9.2:

- Quantification and interactive exploration of the ADC-induced uncertainty and its propagation to TCP modeling (`T1`).

- Exploration and analysis of the assumption-induced TCP model sensitivity (`T2`).

- Identification of inter-patient variability to radiotherapy treatment response (`T3`).

- A new bi-directional TCP modeling workflow (radiotherapy strategy ↔ Predicted or desired outcome) (`T4`).
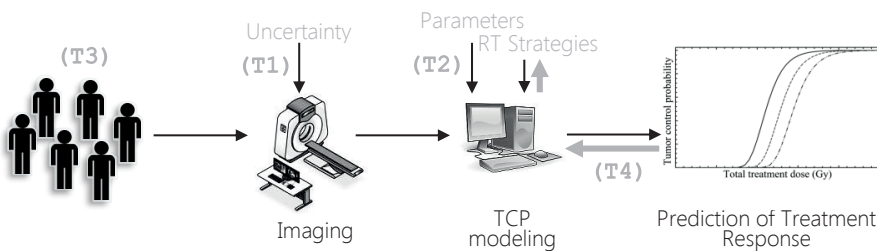
Figure 9.2: The proposed visual analytics approach for the prediction of radiotherapy treatment response in TCP Modeling. With grey, we denote the four requirements (`T1`-`T4`) described in Section 9.3, which are our contributions to the workflow employed in clinical research.

9.4. Related Work

There are several frameworks that cover topics similar to ours. To the best of our knowledge, there is none for the exploration of TCP modeling. In this section, we review the literature,

related to the tasks mentioned in Section 9.3.

■ *Visualizing Uncertainty.* Uncertainty visualization literature is vast [27, 136]. It can be roughly divided into the following main categories: visualizations using visual variables, such as color [105, 115], brightness [71], fuzziness [160], or texture [30]; visualizations that adapt the basic geometry to represent uncertainty [105, 316] or surrounding volume [202, 209, 203]; visualizations with additional graphical variables, such as glyphs [215, 237, 243]; and visualizations employing animations [168]. The selection or combination of these approaches is not limitless and must be done in regard to the data, avoiding clutter. In our case, we need to visualize not only the inherent uncertainty of the imaging data itself, but also how it propagates and affects the outcome of the TCP model. Therefore, several of the previously mentioned approaches need to be carefully adapted to suit our application.

■ *Analyzing Parameter Sensitivity.* Parameter sensitivity is often connected to forecasting or prediction models. A conceptual framework for parameter sensitivity analysis was presented by Sedlmair et al. [246]. Other examples of systems for exploring multi-dimensional parameter spaces are the Ensemble-Vis [216], Noodles [238], OVis [120, 121] and the approach of Berger et al. [20]. Visualizations for parameter sensitivity analysis were also proposed for medical applications [33, 219, 274]. Most of them employ multiple views in an interactive environment, where linking and brushing enables exploration and analysis. Yet, none of these frameworks can be used directly for our purposes.

■ *Studying Cohorts.* In many cases, patients are not analyzed individually. Previous work in cohort visualization mainly focuses on the comparative analysis of shape variability [42, 118, 241]. Recently, Steenwijk et al.[261], Zhang et al.[318] and Klemm et al.[149, 150] proposed interactive visual analysis of cohorts that goes also beyond shape analysis. However, these methods assume that the structure of interest has spatial correspondence between patients and can be compared after matching. This is not valid for tumors. In our case, we need to treat each tumor in the cohort as an entity that we can compare to the rest, while still considering and visualizing the within-cohort heterogeneity.

■ *Redesigning the workflow.* Several visualizations for redesigning the usual workflow in a specific application field have been proposed [37, 55, 69]. Inspired by these strategies, we adapted their approaches to fit our requirements.

9.5. Visual Analytics for the Exploration of Tumor Control Probability Models

The proposed visual tool aims at satisfying the specific exploratory needs of clinical researchers working on TCP modeling, as described in Section 9.3. Our visual tool consists of the four main components `(T1-T4)`, illustrated in Figure 9.2.

9.5.1. Quantification and Interactive Exploration of Uncertainty and its Propagation to TCP Modeling

The ADC-based TCP model [48] aims at incorporating cell density (CD) information. This is a common measure in tumor tissue characterization, referring to the number of tumor cells within a volume. The first step in the ADC-based TCP model requires the calculation of CD

from the ADC values [48]. Among all approaches employed to quantify CD from ADC values, our clinical collaborators chose two for visualization:

- The sigmoid approach, where CD is modeled as a sigmoid function of the ADC values [48].

- The Gibbs approach [97], which is based on experimental data. In this approach, a number of prostate samples were scanned. Subsequently, their ADC values were retrieved from imaging, while their respective CD values were identified from histopathology, as shown in Figure 9.3. From these (ADC, CD) data points, Gibbs et al. established a linear relationship between the two dimensions, as shown in Figure 9.3, with the red dotted line. This linear relationship was described by Gibbs et al., as:

$$CD = \frac{2.1 \cdot 10^{-}3 - ADC}{3 \cdot 10^{-}5} \qquad (9.1)$$

Both approaches – the sigmoid and the Gibbs approach – are affected by uncertainty. In the first case, only the uncertainty of the ADC, calculated below, is present, while in the second case, there is an additional uncertainty in the experimental set-up that was used to determine the relationship between ADC and CD [97].
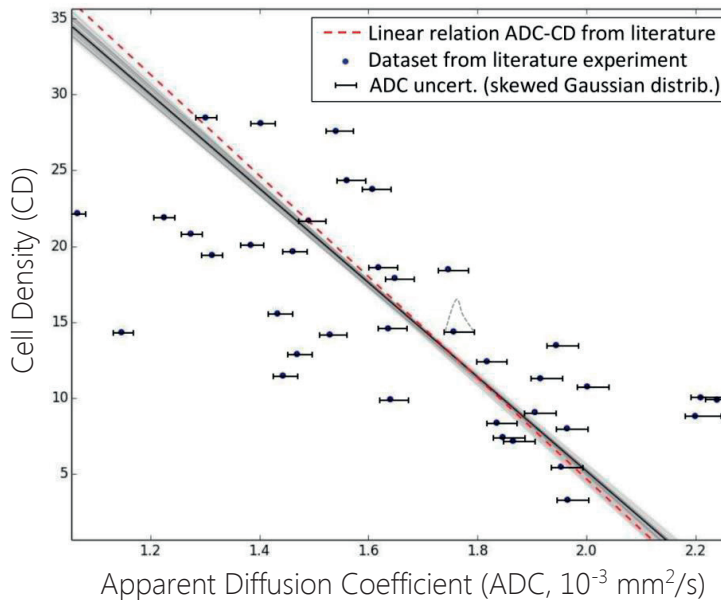
Figure 9.3: Data set extracted from the experiment of Gibbs et al. [97]. The dataset is illustrated with the points, and the linear relation between ADC and CD is shown with the red line. This relationship is without uncertainty. Incorporating ADC uncertainty results in a set of linear fits, shown with the grey scale. Here, dark grey denotes higher probability, and light denotes lower probability.

Uncertainty in ADC maps is usually measured through experimentation in each clinical institution, specifically for their specific scanning set-up [140]. Our clinical collaborators from the Netherlands Cancer Institute and from Aarhus University Hospital have modeled this uncertainty by experiments, as the probability that an ADC value m is measured in imaging, given a quantitative real value r. This is given by a probability density function $p(m|r)$, which is a Gaussian distribution with a standard deviation σ dependent on the real value r. We are interested, though, in quantifying the probability $p(r|m)$ that the real ADC value r has occurred given m. From Bayes' rule we obtain:

$$p(r|m) = \frac{p(r) \cdot p(m|r)}{p(m)} = \frac{p(r) \cdot p(m|r)}{\int_r p(r) \cdot p(m|r) dr} \tag{9.2}$$

where $p(r)$ is the prior probability of the value r. This is assumed to be uniform: $p(r) = \frac{1}{R}$, where R is the range of possible values. Since the standard deviation of $p(m|r)$ depends on the value r, the calculation of $p(r|m)$ is not trivial and was approximated analytically, using Taylor expansion. The analytical calculation was done in `MATLAB` and it resulted to be a skewed Gaussian, dependent on the measured value m:

$$p(r|m) = Q(m) \cdot \frac{1}{r} \cdot \exp\left(-\frac{(m - (1 + \mu) \cdot r)^2}{2 \cdot (\sigma \cdot r)^2}\right) \tag{9.3}$$

where $Q(m)$ is a polynomial of 12$^{\text{th}}$ degree of the measured value m, resulting from the approximation. For generalization, $p(r|m)$ could also have been calculated numerically.

The previously calculated uncertainty in ADC maps is propagated in the calculation of the CD. The two approaches for CD quantification from ADC mentioned above, are influenced by ADC uncertainty in a different way:

- In the sigmoid approach, where CD is modeled as a sigmoid function of the ADC, the uncertainty of the ADC is directly propagated to the CD calculation. Hence, the CD uncertainty $CD(r|m)$ is a function of $p(r|m)$.

- In the initial experiment of Gibbs et al. [97] no uncertainty was taken into account. In this empirical approach, the CD values are obtained from histopathology and, hence, have no uncertainty. However, ADC uncertainty is present and needs to be included, affecting also the relationship between ADC and CD. To quantify this, we randomly sample the ADC uncertainty distribution of each data point of Figure 9.3. Then, we take 2 million sets of samples and we calculate all fits, as well as their respective probabilities. These are given by the product of the probability functions of the data point samples. All generated fits can be seen in Figure 9.3, where the grayscale colormap denotes the probability of each one of the fitted lines, where white is the least probable and black the most probable. From the generated fits, we obtain the CD probability function $CD(r|m)$.

The remaining steps of the TCP model are mathematical equations [48], which do not include additional uncertainties and use the CD value as input. Therefore, for the *sigmoid* approach, the uncertainty in the TCP model will depend directly on the ADC uncertainty $p(r|m)$, while for *empirical Gibbs* approach, it will depend on the CD uncertainty $CD(r|m)$.

For the interactive exploration of the uncertainty, our users are initially interested in having a global overview on the regions of the prostate that are most subject to ADC and CD uncertainty. To simultaneously explore the two uncertainties, we employ a 2D colormap [35]. In this visual representation, we encoded the per voxel difference between the most probable real ADC value r and the measured value m, as shown in Figure 9.4:

$$ADC_{diff} = argmax\{p(r|m)\} - m \qquad (9.4)$$

In the present case, this difference is always positive and its magnitude depends on the measured ADC values. Therefore, we decided to map it to the luminance dimension of the colormap, as shown in Figure 9.4. Also, we encoded the per voxel difference between the most probable real CD value – after the propagation of the ADC uncertainty – and the value measured as proposed by Gibbs et al. in the literature, as illustrated in Figure 9.4:

$$CD_{diff} = argmax\{CD(r|m)\} - CD_{Gibbs} \qquad (9.5)$$

This difference is mapped to a divergent hue dimension of the colormap, as both positive and negative values are possible. No transparency is employed in the colormap. Other approaches, such height fields, were considered, but a discussion with the users showed that the colormap was easier to understand and use.

In addition to color-encoding, we enable users to probe the prostate and interactively explore the entire probability density distributions for the ADC and CD values per voxel, depicted in Figure 9.4. With this dual visualization, the user has an overview on the uncertainty at a prostate level, and locally at the voxel level. Finally, when the user performs TCP modeling, the uncertainty is propagated also to the model outcome, as described before, and visualized on the resulting TCP curve as a density band, presented in the zoomed view of Figure 9.4.

9.5.2. Exploration and Analysis of the Assumption-induced TCP Model Sensitivity

In this part, we use two main components. First, the clinical researcher adds a finite number of combinations of TCP parameter sets, for the calculation of the respective TCP models. This is consistent with the traditional way of exploration of TCP modeling, where one or more TCP models are compared to each other. In this case, each combination of TCP parameter sets is encoded to the visualizations depicted in Figure 9.5 - a, which we call *pianola plots*, inspired by the scrolls used by the musical instrument.

In the example of Figure 9.5, the clinical user has added three parameter sets, depicted by the three white planes. Each row of a pianola plot is a parameter. The first parameter cd is the cell density calculation approach, which is a categorical variable that defines which approach is used for the calculation of CD. It can either take the value *empirical* or *sigmoid*, or others, if available. This is encoded with a dot located in the middle or the end of the first row, respectively. The rest of the parameters are radiobiological parameters of the model with continuous values, which can take values in known and pre-defined ranges. For these,

9

the selected value for each parameter is encoded in the location of each one of the scribbled lines.

The user can also explore the effect of varying one or more parameters continuously, through a range. This is denoted with a box instead of a line, the width of which depicts the desired range of values. An example of this case is given for parameter α in the second set of Figure 9.5 - a. To intuitively link these sets to the respective TCP curves, the scribbles of each pianola plot are assigned a different hue. In this way, the user can easily detect the effect of different parameter choices on the TCP curves, as illustrated in Figure 9.5 - b.
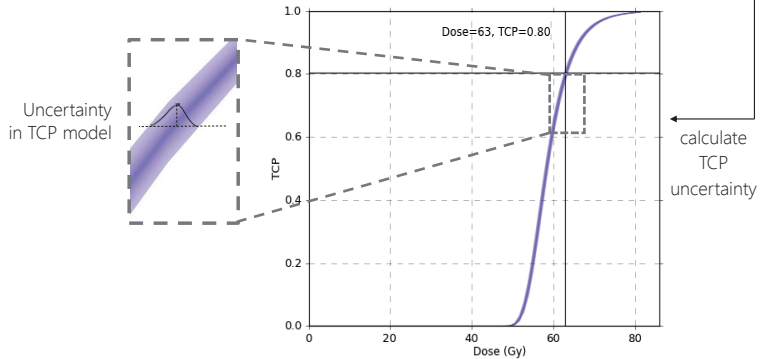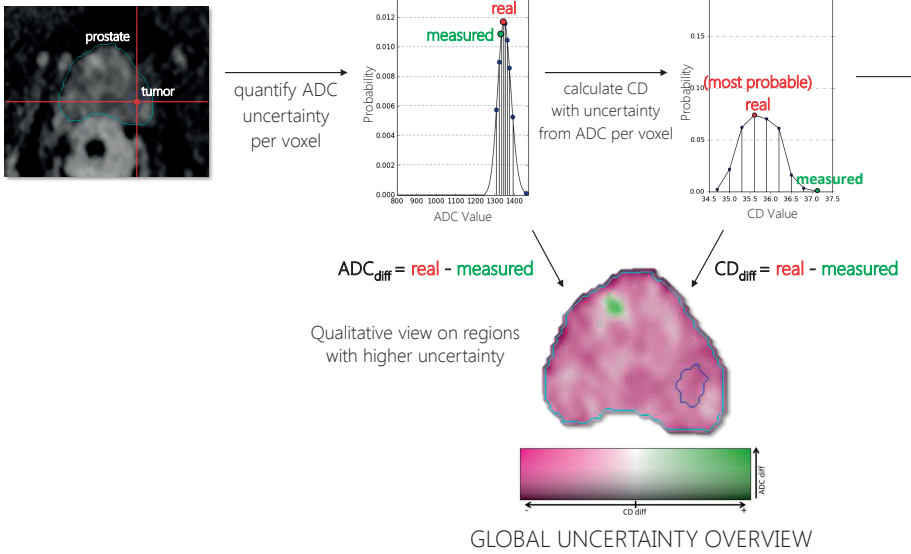
Figure 9.4: Component of our proposed approach, which enables the quantification and interactive exploration of the ADC-induced uncertainty and its effect on TCP modeling (T1).
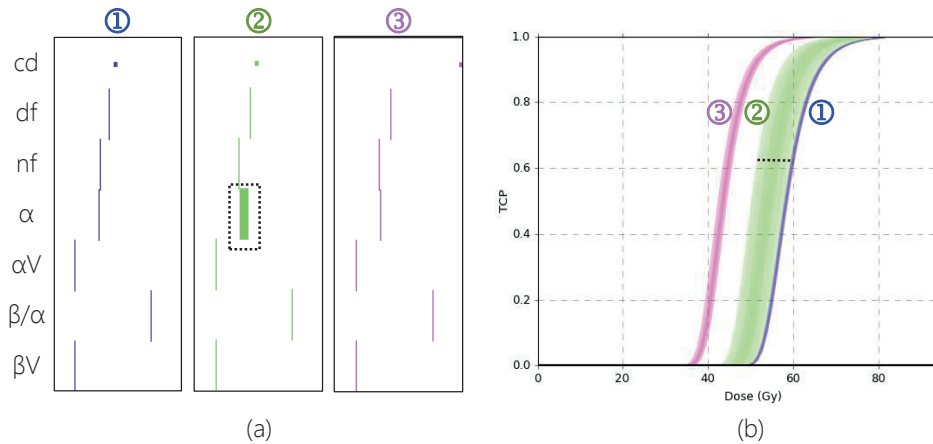
Figure 9.5: Component of our proposed approach, which enables the exploration of the parameter-induced TCP model sensitivity (T2), for three different approaches (①-③). (a) The pianola plots are employed, to intuitively link the chosen parameter sets to the respective TCP curves. (b) The respective curves of each parameter set.

With the current workflow, TCP models can be explored only globally. Although TCP curves can be extracted per voxel, clinical researchers currently calculate the expected average response of the whole tumor to a given dose, and they only analyze the whole tumor TCP curve. They are not able to perform a voxel-based exploration, to detect whether there are specific parts of the tumor that behave differently than the rest, and to analyze why this happens. To enable this, we provide a functionality to probe the TCP curve. This can be either conducted for a TCP value, and identify the linked required dose per voxel, as presented in Figure 9.6, or for a specific dose and identify the linked achieved TCP per voxel. The latter relates also to task (T4).

To provide direct anatomical context, the linked variable is encoded with a heated-body colormap on the imaging slices of the patient, as shown in Figure 9.6. When the user has employed several TCP modeling approaches, we visualize also the variability in the respective dose or TCP value, due to the effect of these alternatives, using circular glyphs [28]. An example of the employed glyphs is illustrated in Figure 9.6. The size of the glyphs denotes the per voxel variability and the blue color is chosen to be complementary to the underlying heated-body colormap. Circular glyphs were chosen, as they preserve visibility on the underlying color-encoded values. This design also helps identifying the relation between values and variability.

9.5.3. Identification of Inter-patient Variability to Treatment Response

The exploration and analysis of the performance of a specific radiotherapy strategy is usually evaluated in a cohort of patients. For example, it is interesting for clinical researchers to know how much the per voxel achieved TCP, or the required dose of their patients varies within a cohort. Within-cohort variability is important, as it can determine whether a treatment

9
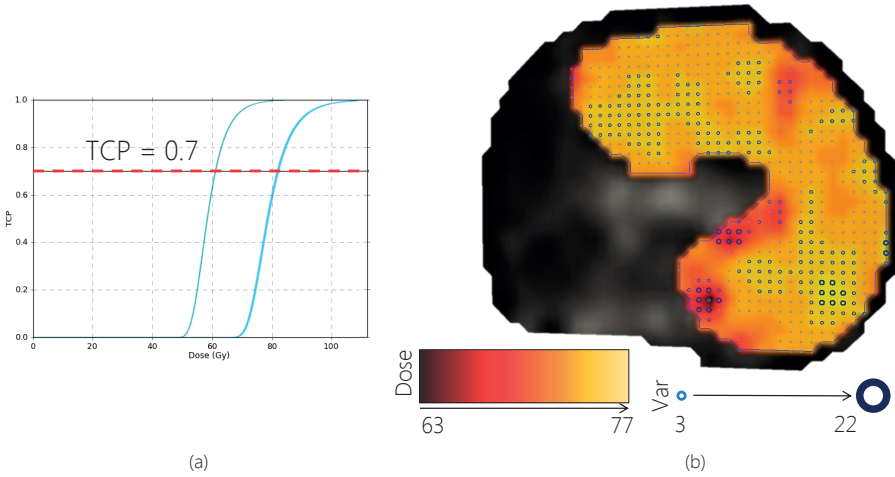
Figure 9.6: Component of our proposed approach, which enables to probe the model curve at a specific TCP value, to inspect the required dose per voxel, depicted with the heated-body colormap, and the respective variability at each voxel position, encoded to the size of the circular glyphs (T2).

strategy is robust enough to treat different patients, and whether it can aid the design of better treatment strategies.

The identification of inter-patient variability to treatment response is linked to (T2): the user probes the TCP curve for either the TCP response or the dose, as shown in Figure 9.6, and the other variable is calculated for the whole cohort, also for multiple TCP modeling approaches, as described in (T2). We provide functionality, with which the users can explore the distributions of the calculated dose or TCP, or multiple sets of these, from multiple TCP modeling approaches. Subsequently, they can partition the cohort of patients, to identify patients that behave similarly throughout different TCP modeling approaches.

To illustrate our approach for the partitioning, we employ the example depicted in Figure 9.7. Here, distributions of dose, required for achieving a specific TCP level, have been calculated for four patients, through three different parameter settings in TCP modeling. We are interested to form groups of patients with similar response patterns, along different parameter settings. For this, we cluster the patients based on the spreads, determined by the distribution dispersion, in each set of distributions. We quantify the dispersion of each distribution, using the median absolute deviation (MAD). This is a robust measure of dispersion [235] and can be described as:

$$MAD = median_i \left(\left| X_i - median_i(X_i) \right| \right) \tag{9.6}$$

where X_i is the distribution of the dose for a specific TCP level, or the distribution of TCP values for a specific dose level. In essence, the MAD is the median of the absolute deviations of the distribution data X_i, from their median. After calculating all MAD measures of

the distributions, we employ a k-means clustering algorithm on the vectors of MADs, calculated per person. This specific clustering algorithm was chosen, due to its simplicity and computational efficiency.

In our approach, the user interactively selects the number of clusters k. To aid the adequate selection of the number of clusters, we employ an additional cluster analysis view. For this, we use a visualization employed in our previous work [226]. In this work, the goal of the visual cluster analysis view was to help users decide whether the visual clusters are well-defined. This is similar to our present goal and we decided to adopt again the same strategy in our system. In this approach, every cluster is mapped to a sphere. For each cluster, we provide internal validity information on the cohesion and separation [226], but also on the inter-patient and inter-assumption variability.

Cohesion is a measure of intra-cluster similarity, while separation is a measure of inter-cluster dissimilarity. For these two measures, we employ the same encoding, as in our previous work [226]: small and opaque spheres describe high cohesion within a cluster, while large and transparent spheres depict low cohesion. Also, thin arrows denote well-separated clusters giving the illusion of distance, while thick arrows stand for less separated clusters. The inter-patient and inter-assumption variability are encoded in the size of the two dimensions of a box, located at the core of each sphere. These encodings are shown in the legend of Figure 9.7.

With the cluster analysis view, the users interactively change the number of clusters, while following the graphical changes on the glyphs, and decide the most satisfactory result, based on the visual optimization of the cluster view. The users interactively partition the patient cohort, inspecting the achieved TCP response, while at the same time, they can identify how much the sub-cohorts of responses vary. An automatized initial selection of a good cluster size or number of clusters would be an interesting future extension.

9.5.4. Bi-Directional Design of TCP Modeling Workflow

With the introduction of `(T4)`, we enable clinical researchers to start their workflow from the desired outcome, to identify and compare the strategies that achieve it. For this, the user defines an acceptability range for the desired TCP outcome, by sketching it on a canvas, as shown in Figure 9.8. Then, all the acceptable parameter combinations that can achieve this are computed, using a brute-force search. The user is presented with these combinations, using a heatmap matrix, depicted in Figure 9.8. In this matrix, every column corresponds to an acceptable combination, and every row to one of the parameters. In the last row, we present also the quantified uncertainty that is introduced by each one of these combinations, calculated in the same way, as in `(T1)`.

The colormap denotes the range of values for each parameter. The user can interact with the matrix and threshold values that are not plausible or interesting for the analysis, or even select combinations based on their uncertainty, as illustrated in Figure 9.8. Probing and linking is employed for the inspection of the TCP curve of each combination. The functionality of `(T4)` is expected to open new ways of exploration and analysis for clinical researchers, as up to now the workflow was done in one direction. Now, the inverse is possible too.

■ *Implementation.* We implemented the visual tool in Python as a DeVIDE module[32], using the Visualization Toolkit (VTK) [2], numpy, scipy, matplotlib and scikit-learn.
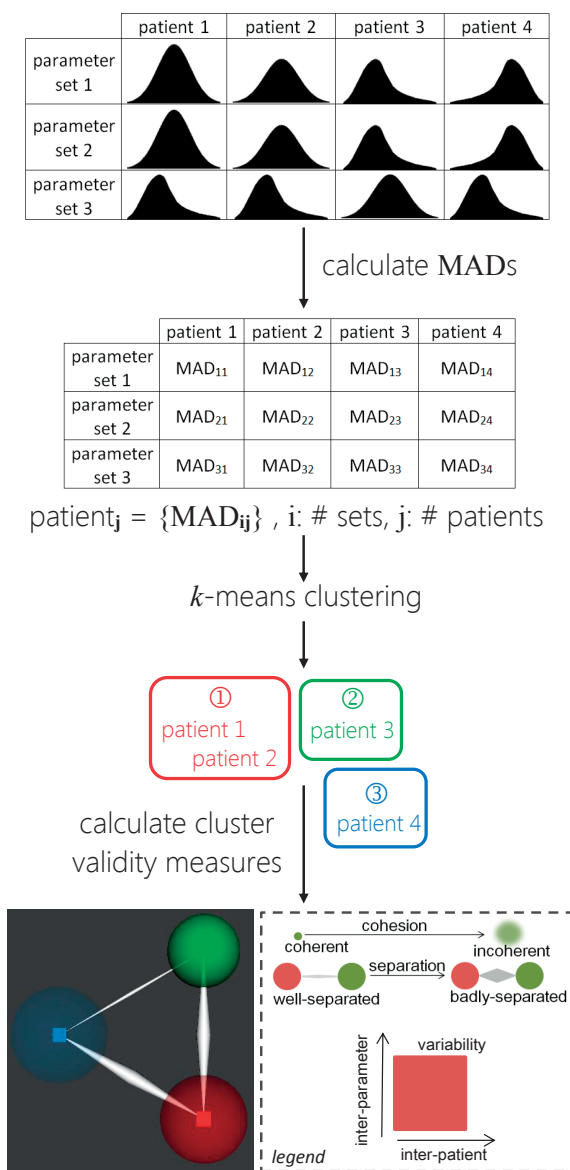
Figure 9.7: Component of our proposed approach, which enables to partition a patient cohort based on TCP treatment response (`T3`). The cluster analysis view (bottom panel), adapted from [226], is used for the visual optimization of clustering.
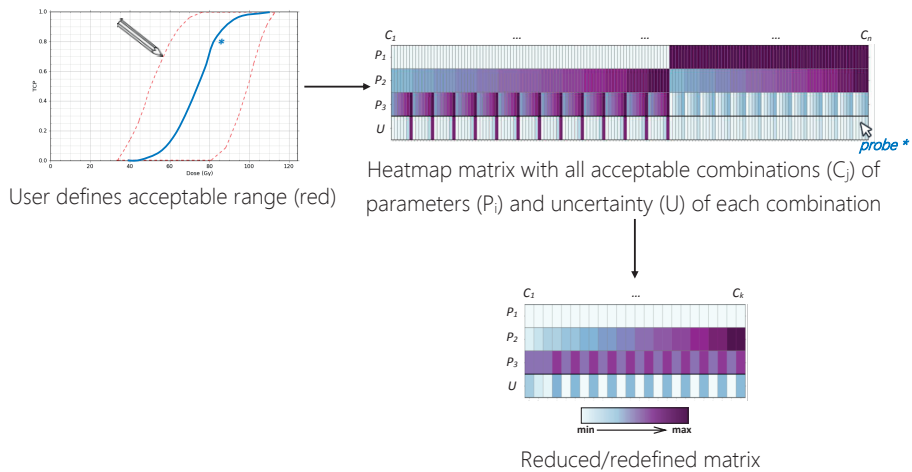
Figure 9.8: Component of our proposed approach, which enables to reverse the workflow in TCP modeling (`T4`). In this case, we start from the desired outcome, to define which treatment strategies can achieve it.

9.6. Evaluation

To assess the value of our visual tool, we performed an evaluation inspired by the paper of Lam et al. [159]. The evaluation was performed with nine domain experts from two clinical institutions (Netherlands Cancer Institute and Aarhus University Hospital). The group of participants included three physicists, five medical physicists and one biomedical engineer. Their field experience varies from medium level (<5 yrs) to very high level (>10 yrs). Two of the participants were actively involved in the design of our tool. Both of them have a very high level of experience in the field of TCP modeling.

All participants were involved in the *first part* of the evaluation. We demonstrated the visual tool, where we showed the main components, simulating the visual environment for the exploration and analysis a TCP modeling workflow. The evaluation participants observed the demonstration and were involved in an active discussion about the various visualizations. Then, they completed a questionnaire.

The *second part* was conducted only with the two participants involved in the design of the tool and the analysis was performed with data already familiar to them. For a deeper understanding on the insights that the tool provides, we performed a case study with hands-on exploration. Each of the four tasks of Section 9.3 was performed with the thinking-out-loud method, as the clinical researchers explained and reasoned on findings in the data. At the end, we asked them to complete again the same questionnaire as before, to see whether their opinion was affected by interacting with the tool.

9

9.6.1. Interviews with Clinical Researchers

During the interviews, the participants completed a questionnaire. The first questions were related to the main tasks of Section 9.3. Each question required an open answer, but also grading using Likert scales $(1-5)$ for the perceived effectiveness, efficiency and satisfaction. To avoid compromising the results, we separated in our analysis the two people involved in the design from the other seven. Half of the tasks were graded higher by the first group and the other half by the second, but overall the two groups had comparable results without significant difference. Also, we separated our analysis based on level of experience, as it possibly indicates different user categories, performing different tasks. Again, the results were comparable among the different groups.

All measured variables received high scores, as shown in Figures 9.9 and 9.10, with a minimum average grade of 4. Uncertainty (T1) and Sensitivity (T2) received high grades. Partitioning (T3) received lower grades, but not lower than 3, as presented in Figures 9.9 and 9.10. This was explained by the fact that participants wanted to see additional information on the data, when partitioning their cohorts. After the case study of the following section, the two participants involved in the design recompiled the questionnaire. For the efficiency of (T1) and effectiveness of (T3), the grades improved. This is interesting, as these two participants consist half of our group with very high experience, who had initially graded (T3) lower than all the others. This could be an indication that after hands-on exploration, this task became clearer to them. Overall, the results between the two rounds are consistently high.

The nine participants were also asked to compare the visual tool to what they are currently using and to evaluate the overall usefulness of our tool. They commented that they currently, *do not have any other means of analysis, apart from looking at individual graphs*.
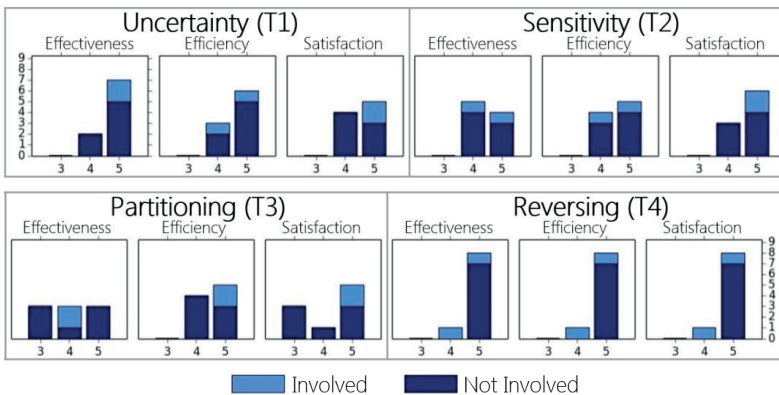
Figure 9.9: Schematic representation of the evaluation results, for each one of the tasks of Section 9.3. Here, the evaluation participants have been split into two groups: the ones that were involved in the design of the visual analytics tool and the others that were not. The horizontal axis represents the grade received in each case, while the vertical axis represents the number of participants that assigned the specific grade.

9
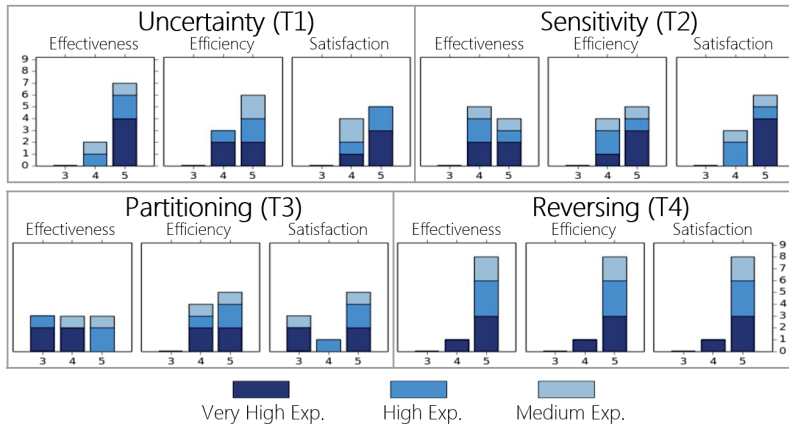
Figure 9.10: Schematic representation of the evaluation results, for each one of the tasks of Section 9.3. Here, the evaluation participants have been split into three groups, based on their experience. The horizontal axis represents the grade received in each case, while the vertical axis represents the number of participants that assigned the specific grade.

For them, the framework requires training and a level of familiarization, but it *removes a significant overhead from the analysis*, giving *important input*. More specifically, the uncertainty part (`T1`) provides *directly understandable and quantitative feedback*, while with sensitivity analysis (`T2`) they can *perceive directly the influence of the dose prescription*. For the cohort partitioning part (`T3`), there were mixed opinions. According to the evaluation participants, it *raises a lot of questions about the subgroups of the cohorts*. It could be the *most important clinical application*, but it should be done also *based on other variables*, or also *for intra-tumor regions*. Reversing the workflow (`T4`) can have *great potential*.

All participants agreed that the visual tool is overall understandable and useful. The strong features of the visual tool are the ability to perform a voxel-based analysis, especially, the probing and linking functionality in the TCP curves and the view on the variability from the different modeling approaches (`T2`), as well as the workflow reversing task (`T4`). Improvement proposals were mostly related to cohort partitioning (`T3`).

9.6.2. Case Study

For the case study, ADC data from a cohort of 11 locally advanced prostate tumor patients was used. The ADC maps were derived with a *b*-value of 1000 and have a size of $256 \times 256 \times 24$ voxels and a resolution of $0.97 \times 0.97 \times 3.6$.

During the task of uncertainty (`T1`), it was noticed that CD might be overestimated in literature, as visualized by the dominant purple color in the slice view of Figure 9.4. Some voxels, highlighted with the green color, have been noticed to be sometimes mis-delineations of voxels that belong to the bladder or to the urethra. In the rest of cases, like in Figure 9.4, these are locations in the prostate that should be checked more thoroughly. Less uncertainty is expected within tumors, due to lower ADC values. The effect of the uncertainty on the TCP

was also found to be interesting: it reaches almost 5% of the TCP for the *empirical* approach, at $D_{50\%}$, i.e., the dose required for achieving 50% control only in the tumor location. This is denoted in Figure 9.11 - a, with the purple curve. The TCP uncertainty is 2.5% for the *sigmoid* approach, as shown in Figure 9.11 - a, with the orange curve.

During the exploration of sensitivity (T2), four examples of parameter sets were explored. The first two are *empirical* approaches with the same radiotherapy strategy, but they differ in the αVar parameter, which is the inter-patient variability in the survival rate of tumor cells. The third is a *sigmoid* approach with the same radiotherapy strategy. The last is an additional *empirical* approach with a range of α between 0.17 and 0.21, which is a radiobiological parameter indicating the survival rate of tumor cells. In the resulting TCP curve graph, there are indications that the *sigmoid* model might predict tumor control with a lower dose, i.e, the curve is more to the left, than the respective *empirical*, depicted in Figure 9.11 - b. Also, when the αVar is neglected, then the model suffers less from uncertainty, as shown in Figure 9.11 - b. For ranging α, the effect on the TCP is more prominent. In this case, probing the TCP at 70% shows that a dose ranging between 43 and 82 Gy is required within the tumor only. The variability between the four models, though, is large for the whole tumor, as illustrated in Figure 9.11 - b.

For cohort partitioning (T3), after probing the TCP at 70%, we obtain the dose distributions per patient, as presented in Figure 9.11 - c. Patient 8 seems to have a different behavior. His TCP curve, shown in Figure 9.11 - c, with the dotted TCP curve, is on the right side of the average curve, which means that to achieve a 70% TCP, he requires a higher dose, as he is a patient with a much larger tumor. After interactive clustering, the visually optimal cluster analysis view is achieved with two clusters, where patients behave similarly in terms of TCP curves, e.g., the ones in the blue cluster are all on the right side of the average curve.

In reversing the workflow (T4), patient 8 was explored to check whether a more satisfying strategy can be identified. A wide range of acceptable TCP is drawn, as in Figure 9.8. More than 200 different combinations are identified, as seen by the columns of the heatmap matrix. On first sight, it seems that the *sigmoid* approach (Figure 9.8 - heatmap, first row, purple section) may be less sensitive to changes in parameters than *empirical* (Figure 9.8 - heatmap, first row, white section), as less combinations are computed for the *sigmoid*. Also, this approach may suffer less from uncertainty (Figure 9.8 - heatmap, last row), as the range of uncertainties does not go up to the maximum value of uncertainty, indicated by the deep purple color. By redefining the acceptable limits for the parameters and the uncertainty, only 24 different combinations are preserved, as shown in Figure 9.8. According to the evaluators, (T4) functionality could be helpful to determine the suitability of this patient for a specific therapy. However, for this, no conclusions can be made, as it would also require the involvement of oncologists, and a more extensive study. The examined cases are meant to demonstrate the use of the visual tool, not as an actual analysis with direct clinical inferences.

9.7. Conclusions and Future Work

In this work, we proposed a visual tool to enable clinical researchers to explore and analyze different aspects of the TCP modeling workflow. We tackled the quantification and interactive exploration of uncertainty and its propagation to TCP modeling, parameter sensitivity

9
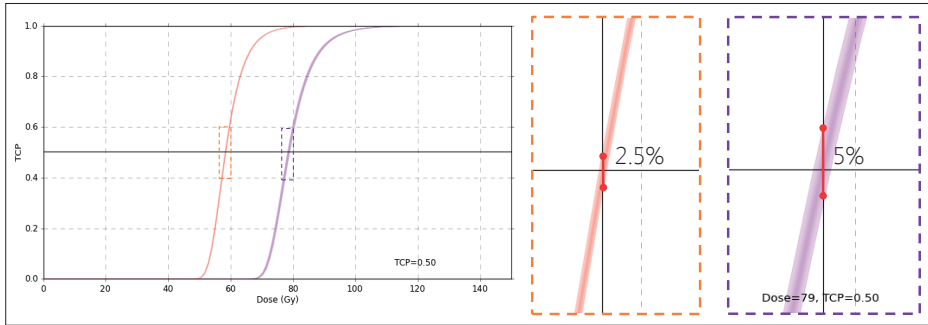
(a)
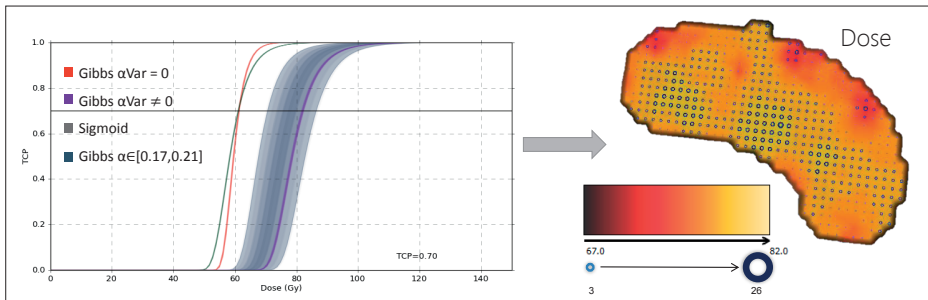
(b)
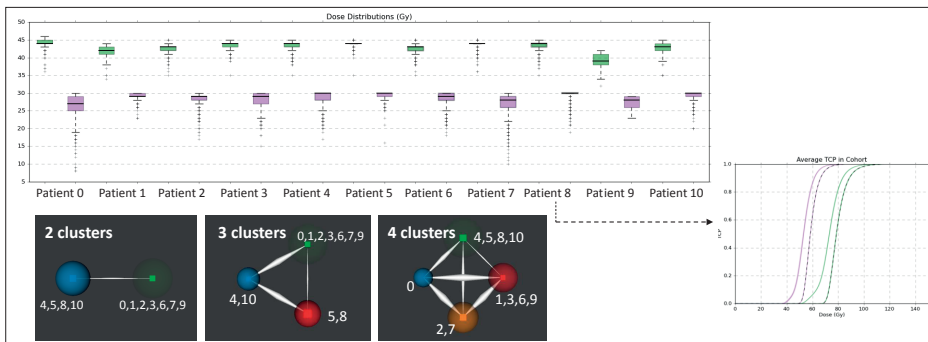
(c)

Figure 9.11: Results from the case study with the two clinical researchers that were involved into the design of our approach.

9

analysis of TCP models, cohort partitioning based on treatment response and a novel functionality for enabling also a reverse workflow. Nine clinical researchers evaluated and confirmed the usefulness of the visual tool, as it opens new possibilities and provides access to new insight in the data. We illustrated this also with a case study.

A direction for future work includes improving the partitioning of the cohorts to enable clustering also based on other attributes, and also linking to intra-tumor tissue characteristics [226]. In this way, more meaningful inter-patient analysis can also be performed. The proposed visual tool is a promising basis for clinical researchers to gain more knowledge on their complex TCP modeling processes, to explore the data from the models in a more insightful way and to generate and confirm new hypotheses.

9

10

Conclusion

It seems impossible, until it's done.

Nelson Mandela (1918 - 2013)

10.1. Summary

As radiotherapy technology focuses on providing a more personalized treatment based on the specific intra-tumor tissue characteristics of each patient, the integration of all available patient- and tumor-specific knowledge from the distinct steps of the radiotherapy pipeline, is required. Current limitations in the exploration and analysis of the involved information can be overcome, by employing solutions from the field of Visual Analytics.

The goal of this dissertation was to present visualization solutions, which are able to integrate the information from all steps of the radiotherapy planning pipeline, in order to fulfill the need of clinical researchers to interactively explore and analyze their data, to derive and understand information from them, and to formulate or confirm hypotheses, with respect to their findings. Each of the previous chapters addressed one step of the tumor-tailored radiotherapy pipeline, providing strategies to tackle the specific needs of clinical researchers with respect to their data and processes.

The present work reached two main advancements. First, in radiotherapy research, we accomplished to provide solutions and tools for clinical researchers, which address their exploratory needs at all steps of the tumor-tailored radiotherapy planning pipeline. Involving clinical experts through visual analysis and interaction in the workflow of the entire radiotherapy pipeline had not been tackled before, and our proposed Visual Analytics solutions address specific tasks that could not be conducted with previously existing tools. The development of these new interactive and investigative tools has now empowered clinical researchers with new exploratory possibilities.

In visualization research, we managed to expand the field of Visual Analytics to support a new clinical domain. As discussed in previous chapters, although there are numerous ways of visualizing multi-dimensional and complex data, none of them was fully applicable to the data and processes of radiotherapy planning. Our contribution to the field of Visual Analytics is a compound of applications tailored to a specific step of the pipeline, to the involved data and to the requirements of clinical researchers. With our work, we presented our user-centered viewpoint on how visualization applications should be designed, implemented, and later evaluated. All proposed approaches promote the integration and combination of the strengths of human perception for exploration and analysis, together with semi-automated methods, with the purpose to increase the understanding in otherwise complex data and processes.

More in particular, Chapter 4 proposed the Orientation-Enhanced Parallel Coordinates Plots (OPCPs), a novel method for the representation and exploration of multi-variate and multi-dimensional data. The proposed technique contributes towards the enhancement of the – otherwise cluttered – display of the data, emphasizing the underlying patterns and improving the discernibility of outliers, or data structures that may be obstructed by noise or other data patterns. The OPCPs technique can be used in several steps of the pipeline, but also outside of this domain in other broader applications.

In Chapter 5, we presented a visualization solution that aids clinical researchers to explore and understand how different choices in modeling may affect the parameter space derived from modeling DCE-MRI data. As part of this Visual Analytics tool, we designed a new interactive representation, the Cocoon, that allows users to explore all required DCE-MRI data aspects in a single combined view. The interactive features of the tool facilitate the ex-

ploration and interpretation of the data and, especially, the correspondence of observations from the feature space to the patient anatomy.

Tissue characterization has been addressed in Chapter 6, where we contributed with a Visual Analytics system that enables the easy exploration and analysis of the high-dimensional feature space of imaging-derived tissue characteristics. This visualization solution enables the identification of distinct intra-tumor regions, the understanding of tumor heterogeneity, new knowledge discovery within the data, as well as hypothesis generation and confirmation, with respect to reference data used in clinical research. Although this work has been developed around the specific topic of intra-tumor tissue characterization, we additionally demonstrated a generalized use of the proposed system, namely aiding the design of brain tissue characterization classifiers and understanding their behavior and outcomes.

In Chapter 7, we addressed the segmentation step of the radiotherapy planning pipeline. In particular, we addressed the need for a Visual Analytics solution that allows the exploration and assessment of potential errors in the segmentation of the prostate and the surrounding organs at risk. With this solution, we enabled the visual analysis of errors in the model-based segmentation of the involved pelvic structures, and made the first step towards understanding the prediction of the performance of the employed segmentation algorithms.

Chapter 8 presented a Visual Analytics solution for the exploration and analysis of the variability in an ensemble of radiotherapy dose plans. Several assumptions or different parameterizations in the previous steps of the pipeline may result into different dose plans. In order to examine the impact of those assumptions or parameterizations on the final planning outcome, and to visually assess the variability across multiple possible dose plans, we designed a framework that enables the exploration and visual analysis of dose variability at an iso-contour level and at a voxel level.

Finally, the exploration and visual analysis of tumor control probability (TCP) modeling, often employed for the evaluation and prediction of the outcome of a radiotherapy treatment, was tackled in Chapter 9. Here, we proposed a visual tool to enable clinical researchers to explore and analyze different aspects of the TCP modeling workflow: the quantification and interactive exploration of uncertainty and its propagation to TCP modeling, parameter sensitivity analysis of TCP models, cohort partitioning based on treatment response and a novel functionality for enabling also a reverse workflow.

All in all, our work is a promising basis, which offers new exploratory possibilities for researchers working on the complex processes and data, present at all steps of the patient- and tumor-tailored radiotherapy pipeline. Nevertheless, it has also revealed a multitude of topics for future research, which are discussed in the upcoming section.

10.2. Discussion and Directions for Future Research

A thorough discussion and specific recommendations for topics that might be interesting for future research have already been provided at the end of each previous chapter. In the present section, we would like to present our own perspective, with respect to a number of general lessons learned in the course of this work, as well as several general points that could be considered for future research.

With respect to the representation and exploration of multi-variate and multi-dimensional

data, it is clear that our approach addresses a specific topic: the reduction of clutter in a specific multi-dimensional representation, the PCPs. Yet, there is a multitude of multi-dimensional representations, which can and should be carefully employed, depending on the tasks that need to be fulfilled and the data aspects that need to be brought forward. Each one of the existing multi-dimensional representations has its strength and limitations. However, as already discussed in many instances of this dissertation, scalability, concerning the number of data points and the number of data dimensions, is a quite often reoccurring problem, in all of them. In our opinion, strategies to address the scalability issue should be devised. A future direction could include the investigation of simple enhancement solutions around existing representations, such as the one proposed within the OPCPs (Chapter 4), opposed to the usual tendency of developing complicated solutions.

In our specific application, as well as in the majority of clinical applications, the often exploding number of features that can be extracted from imaging modalities makes it impeding to consider additional strategies, which employ dimensionality reduction and data mining. Such strategies have been successfully exploited in this dissertation, in particular for tissue characterization. Visual Analytics strategies, such as the framework proposed in Chapter 6, can be a very promising basis for a number of clinical applications, supporting clinical researchers to explore and analyze their data. Yet, the results of the incorporated dimensionality reduction techniques may be not entirely understood, or it may not be feasible to achieve results *on-the-fly*. The latter implies that the users might be required to wait until the completion of the employed algorithm, every time that they need to redefine it, which is not always optimal for the analysis workflow. To this end, the field of *Progressive Visual Analytics*, where partial results of an algorithm can be produced and interactively explored and analyzed, may be beneficial. Future research towards this direction would be an interesting enhancement to our current work.

Designing and employing Progressive Visual Analytics solutions can additionally enable the incorporation and exploration of (im-)precision of the employed processes into the analysis. This, together with the integration of the uncertainty, induced by the parametrization of the chosen dimensionality reduction techniques, would lead to more reliable observations within the data. For example, this could be interesting for the tumor tissue characterization step or for the exploration of the effect of adjustments in the pipeline on the final dose planning outcome. This kind of solutions may, furthermore, allow the extension of a number of applications to enable meaningful follow-up or inter-patient analysis. In essence, we expect that this research direction can help clinical researchers to explore, analyze and deliberate, with respect to the progression of a disease, or to treatment response, also regarding different groups of individuals. From this, diagnosis, prognosis and treatment may significantly benefit. From a visualization point of view, follow-up and inter-patient analysis would be a very challenging topic, due to the implicated dimensionality and complexity of the data, and due to the fact that anatomic correspondences are not ensured. Smart strategies to address these two key-points need to be devised.

Moving forwards, from the imaging and characterization step of the radiotherapy pipeline to the next steps of registration and segmentation, several points have emerged. Concerning the former, our dissertation did not extend to any aspects of the registration processes involved in the radiotherapy planning pipeline. However, the registration aspect is crucial, as it might compromise the subsequent steps of the pipeline. The propagation of potential uncertainty from the registration phase to the rest of the pipeline, as well as the exploration of

the impact of inaccuracies to the final outcome, should definitely be taken into consideration in future research. This should be conducted together with the exploration of the effect of inaccuracies or errors of the segmentation phase to the final dose planning result. In the segmentation step, it is also important to find solutions for the investigation of the relation of the shapes and sizes of the various organs to the algorithm performance.

In the tumor control probability modeling phase, we foresee that the proposed visual tool can be a promising basis for clinical researchers to gain more knowledge on their complex modeling processes. In particular, it opens new directions for clinical research, by providing new ways of exploring and interacting with the data, for example by allowing clinical researchers to start from the desired outcome to explore the feasibility of the strategies required to achieve it. Using this tool, in retrospective, to evaluate treatment prognosis in a big cohort study, would be an interesting future enrichment. To this end, more meaningful inter-patient analysis would be required, which can be achieved by associating this step to the specific intra-tumor characteristics of each group of patients. From a visualization aspect, this would also be a very demanding and worthwhile topic for future research.

In general, the integration of the different solutions, to explore and analyze the accumulation and propagation of the uncertainty from one step of the radiotherapy pipeline to the final outcome, is an additional important next step. In our current work, our contribution to that was the incorporation of the imaging-induced uncertainty into tumor control probability modeling, but there is still a lot to be done towards this direction. This should also take into consideration the uncertainties that may occur during the actual delivery of the treatment, in order to provide a complete overview of the potential inaccuracies in the final dose plan and their potential impact on the treatment and prognosis of the patients. At this point, additional non-imaging patient-specific characteristics of the patients should also be incorporated in the analysis, which would be very interesting for future visualization research.

To conclude, we have proposed a number of Visual Analytics solutions, meant to be employed by clinical researchers. Still, visualization solutions for clinical users, such as dosimetrists or radiologists, should also be designed. This kind of solutions has a fundamental difference with the proposed approaches: they are not meant for exploration, knowledge discovery and hypothesis generation, but for hypothesis confirmation and decision making, instead. From this point of view, we foresee two important aspects that can positively contribute to successfully develop this kind of visualization solutions. First, compliance with tools that the users are already familiar with should be achieved. To this end, the standardization and automatization of the analytic processes would be required, and the amount of required interaction would need to be carefully evaluated and, eventually, minimized. Secondly, as these solutions would mainly aim at supporting – not replacing – decision making, they should also be carefully designed to be adequate for routine use, to be task-oriented and intended to improve the accuracy, efficiency and effectiveness of the involved clinical decision process. Adequate evaluation of these systems would also be required.

This dissertation has described solutions from the field of Visual Analytics, which aim at incorporating the information from the distinct steps of the radiotherapy planning pipeline, along with eventual sources of uncertainty, into comprehensible visualizations. These approaches contribute towards the interactive exploration, visual analysis and understanding of the involved data and processes at each step of the radiotherapy planning pipeline, creating fertile ground for future research in radiotherapy planning, but also in Visual Analytics.

References

[1] Institute for Information Systems at the TU Braunschweig (Germany).

[2] The Visualization Toolkit. `http://vtk.org/`.

[3] three.js: Javascript 3D library. `https://threejs.org/`.

[4] XmdvTool Datasets. `http://davis.wpi.edu/xmdv/datasets.html`.

[5] O. S. Alabi, X. Wu, J. M. Harter, M. Phadke, L. Pinto, H. Petersen, S. Bass, M. Keifer, S. Zhong, and C. Healey. Comparative visualization of ensembles using ensemble surface slicing. In *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2012.

[6] American Cancer Society. Cancer Facts and Figures 2016. American Cancer Society, 2016.

[7] American Cancer Society. Prostate cancer risk factors. American Cancer Society, 2016.

[8] American College of Radiology. PI-RADS, Prostate Imaging Reporting and Data System. ACR, American College of Radiology, 2015.

[9] D. E. Amir, K. L. Davis, M. D. Tadmor, E. F. Simonds, J. H. Levine, S. C. Bendall, D. K. Shenfeld, S. Krishnaswamy, G. P. Nolan, and D. Pe'er. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature biotechnology*, 31(6):545–552, 2013.

[10] P. Anbeek, K. L. Vincken, M. J. Van Osch, R. H. Bisschops, and J. Van Der Grond. Probabilistic segmentation of white matter lesions in mr imaging. *NeuroImage*, 21(3):1037–1044, 2004.

[11] D. F. Andrews. Plots of high-dimensional data. *Biometrics*, pages 125–136, 1972.

[12] K. Andrews. Evaluation comes in many guises. In *AVI Workshop on BEyond time and errors (BELIV) Position Paper*, pages 7–8, 2008.

[13] M. Ankerst, S. Berchtold, and D. A. Keim. Similarity clustering of dimensions for an enhanced visualization of multidimensional data. In *Information Visualization, 1998. Proceedings. IEEE Symposium on*, pages 52–60. IEEE, 1998.

[14] A. O. Artero. Uncovering clusters in crowded parallel coordinates visualizations. In *Proceedings of the IEEE Symposium on Information Visualization 2004*, pages 81–88. IEEE Computer Society, 2004.

[15] D. Asimov. The Grand Tour: a tool for viewing multidimensional data. *SIAM journal on scientific and statistical computing*, 6(1):128–143, 1985.

[16] J. O. Barentsz, J. Richenberg, R. Clements, P. Choyke, S. Verma, G. Villeirs, O. Rouviere, V. Logager, and J. J. Fütterer. ESUR prostate MR guidelines 2012. *European radiology*, 22(4):746–757, 2012.

[17] S. Barlowe, T. Zhang, Y. Liu, J. Yang, and D. Jacobs. Multivariate visual explanation for high dimensional datasets. In *Visual Analytics Science and Technology, 2008. VAST'08. IEEE Symposium on*, pages 147–154. IEEE, 2008.

[18] R. A. Becker and W. S. Cleveland. Brushing scatterplots. *Technometrics*, 29(2):127–142, 1987.

[19] S. M. Bentzen. Dose painting and theragnostic imaging: towards the prescription, planning and delivery of biologically targeted dose distributions in external beam radiation oncology. In *Radiation Oncology Advances*, pages 40–61. Springer, 2008.

[20] W. Berger, H. Piringer, P. Filzmoser, and E. Gröller. Uncertainty-aware exploration of continuous parameter spaces using multivariate prediction. In *Computer Graphics Forum*, volume 30, pages 911–920. Wiley Online Library, 2011.

[21] E. Bertini, L. Dell'Aquila, and G. Santucci. Springview: Cooperation of radviz and parallel coordinates for view optimization and clutter reduction. In *Coordinated and Multiple Views in Exploratory Visualization, 2005.(CMV 2005). Proceedings. Third International Conference on*, pages 22–29. IEEE, 2005.

[22] C. M. Bishop. *Pattern recognition and machine learning*. New York: Springer„ 2006.

[23] J. Blaas, C. P. Botha, and F. H. Post. Interactive visualization of multi-field medical data using linked physical and feature-space views. In *Computer Graphics Forum*, pages 123–130, 2007.

[24] G. H. Bol, A. N. Kotte, U. A. van der Heide, and J. J. Lagendijk. Simultaneous multi-modality ROI delineation in clinical practice. *Computer methods and programs in biomedicine*, 96(2):133–140, 2009.

[25] D. Bonekamp, M. A. Jacobs, R. El-Khouli, D. Stoianovici, and K. J. Macura. Advancements in MR imaging of the prostate: from diagnosis to interventions. *Radiographics*, 31(3):677–703, 2011.

[26] D. Bonekamp and K. J. Macura. Dynamic contrast-enhanced magnetic resonance imaging in the evaluation of the prostate. *Topics in Magnetic Resonance Imaging*, 19(6):273–284, 2008.

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

[28] R. Borgo, J. Kehrer, D. H. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward, and M. Chen. Glyph-based visualization: Foundations, design guidelines, techniques and applications. *Eurographics State of the Art Reports*, pages 39–63, May 2013.

[29] Bostock, M. D3.js: Data-Driven Documents. `https://d3js.org/`.

[30] R. P. Botchen, D. Weiskopf, and T. Ertl. Interactive visualization of uncertainty in flow fields using texture-based techniques. In *Proc. Intl. Symp. on Flow Visualization*, 2006.

[31] C. Botha, B. Preim, A. Kaufman, S. Takahashi, and A. Ynnerman. From individual to population: challenges in medical visualization. In *Scientific Visualization*, pages 265–282. Springer, 2014.

[32] C. P. Botha and F. H. Post. Hybrid scheduling in the DeVIDE dataflow visualisation environment. In *SimVis*, pages 309–322, 2008.

[33] R. Brecheisen, A. Vilanova, B. Platel, and B. ter Haar Romeny. Parameter sensitivity visualization for DTI fiber tracking. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1441–1448, 2009.

[34] S. Bremm, T. Von Landesberger, M. Heß, and D. Fellner. PCDC-On the highway to data-a tool for the fast generation of large synthetic data sets. In *EuroVis Workshop on Visual Analytics*, pages 7–11, 2012.

[35] C. A. Brewer. Guidelines for use of the perceptual dimensions of color for mapping and visualization. In *IS&T/SPIE 1994 International Symposium on Electronic Imaging: Science and Technology*, pages 54–63. International Society for Optics and Photonics, 1994.

[36] K. Brodlie, R. A. Osorio, and A. Lopes. A review of uncertainty in data visualization. In *Expanding the frontiers of visual analytics and visualization*, pages 81–109. Springer, 2012.

[37] S. Bruckner and T. Möller. Result-driven exploration of simulation parameter spaces for visual effects design. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1468–1476, 2010.

[38] P. Bryant. A brief history and review of accelerators. *CERN European Organization for Nuclear Research - Reports CERN*, pages 1–1, 1994.

[39] D. L. Buckley, C. Roberts, G. J. Parker, J. P. Logue, and C. E. Hutchinson. Prostate Cancer: Evaluation of Vascular Characteristics with Dynamic Contrast enhanced T1 weighted MR Imaging - Initial Experience. *Radiology*, 233(3):709–715, 2004.

[40] A. Buja, J. A. McDonald, J. Michalak, and W. Stuetzle. Interactive data visualization using focusing and linking. In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on*, pages 156–163. IEEE, 1991.

[41] S. Busking, C. P. Botha, L. Ferrarini, J. Milles, and F. H. Post. Image based rendering of intersecting surfaces for dynamic comparative visualization. *The visual computer*, 27(5):347–363, 2011.

[42] S. Busking, C. P. Botha, and F. H. Post. Dynamic Multi-View Exploration of Shape Spaces. In *Computer Graphics Forum*, volume 29, pages 973–982. Wiley Online Library, 2010.

[43] D. Bystrov, N. Schadewaldt, H. Schulz, J. Peters, et al. A new method for robust organ positioning in CT images. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 338–341. IEEE, 2012.

[44] C. Brewer, M. Harrower and The Pennsylvania State University. ColorBrewer 2.0 - Color advice for cartography. `http://colorbrewer2.org/`.

[45] L. Caravatta, G. Macchia, G. Mattiucci, A. Sainato, N. Cernusco, G. Mantello, M. Di Tommaso, M. Trignani, A. De Paoli, G. Boz, M. Friso, V. Fusco, M. Di Nicola, A. Morganti, and D. Genovesi. Inter-observer variability of clinical target volume delineation in radiotherapy treatment of pancreatic cancer: a multi-institutional contouring experience. *Radiation Oncology*, 9(1):1–9, 2014.

[46] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[47] D. B. Carr and R. J. Littlefield. Scatterplot matrix techniques for large N. *Journal of the American Statistical Association*, 82(398):424–436, 1987.

[48] O. Casares-Magaz, U. van der Heide, J. Rørvik, P. Steenbergen, and L. Muren. A Tumor Control Probability model for Radiotherapy of prostate cancer using MRI-based Apparent Diffusion Coefficient maps. *Radiotherapy and Oncology*, 119(1):111–116, 2016.

[49] A. Cedilnik and P. Rheingans. Procedural annotation of uncertain information. In *Visualization 2000. Proceedings*, pages 77–84. IEEE, 2000.

[50] J. M. Chambers. *Graphical methods for data analysis*. 1983.

[51] Y. J. Choi, J. K. Kim, N. Kim, K. W. Kim, E. K. Choi, and K.-S. Cho. Functional MR Imaging of Prostate Cancer. *Radiographics*, 27(1):63–75, 2007.

[52] J. Choo, H. Lee, J. Kihm, and H. Park. iVisClassifier: An interactive visual analytics system for classification based on supervised dimension reduction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 27–34. IEEE, 2010.

[53] P. Cignoni, M. Corsini, and G. Ranzuglia. MeshLab: An open-source 3D mesh processing system. *Ercim news*, 73(45-46):6, 2008.

[54] A. Cockburn, A. Karlson, and B. B. Bederson. A review of overview+ detail, zooming, and focus+ context interfaces. *ACM Computing Surveys (CSUR)*, 41(1):2, 2009.

[55] D. Coffey, C.-L. Lin, A. G. Erdman, and D. F. Keefe. Design by dragging: An interface for creative forward and inverse design with simulation ensembles. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2783–2791, 2013.

[56] C. Collins and S. Carpendale. VisLink: Revealing relationships amongst visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1192–1199, 2007.

[57] A. Coninx, G.-P. Bonneau, J. Droulez, and G. Thibault. Visualization of uncertain scalar data fields using color scales and perceptually adapted noise. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, pages 59–66. ACM, 2011.

[58] K. A. Cook and J. J. Thomas. Illuminating the path: The research and development agenda for visual analytics. Technical report, Pacific Northwest National Laboratory (PNNL), Richland, WA (US), 2005.

[59] M. J. Costa, H. Delingette, S. Novellas, and N. Ayache. Automatic segmentation of bladder and prostate using coupled 3D deformable models. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 252–260. Springer, 2007.

[60] E. Coto, S. Grimm, S. Bruckner, E. Gröller, A. Kanitsar, and O. Rodriguez. *MammoExplorer: an advanced CAD application for breast DCE-MRI*. Citeseer, 2005.

[61] G. O. Cron, C. Foottit, T. E. Yankeelov, L. I. Avruch, M. E. Schweitzer, and I. Cameron. Arterial input functions determined from MR signal magnitude and phase for quantitative Dynamic Contrast-Enhanced MRI in the human pelvis. *Magnetic resonance in medicine*, 66(2):498–504, 2011.

[62] J. Crook, Y. Raymond, D. Salhani, H. Yang, and B. Esche. Prostate motion during standard radiotherapy as assessed by fiducial markers. *Radiotherapy and Oncology*, 37(1):35–42, 1995.

[63] B. M. Dale, J. A. Jesberger, J. S. Lewin, C. M. Hillenbrand, and J. L. Duerk. Determining and optimizing the precision of quantitative measurements of perfusion from dynamic contrast enhanced MRI. *Journal of Magnetic Resonance Imaging*, 18(5):575–584, 2003.

[64] A. Dasgupta, M. Chen, and R. Kosara. Conceptualizing Visual Uncertainty in Parallel Coordinates. *Computer Graphics Forum*, 31(3pt2):1015–1024, 2012.

[65] T. J. Davis and C. P. Keller. Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences*, 23(4):397–408, 1997.

[66] G. Delaney, S. Jacob, C. Featherstone, and M. Barton. The role of radiotherapy in cancer treatment. *Cancer*, 104(6):1129–1137, 2005.

[67] I. Demir, J. Kehrer, and R. Westermann. Screen-space Silhouettes for Visualizing Ensembles of 3D Isosurfaces. In *Proc. IEEE Pacific Visualization Symp. (Visualization Notes)*, 2016.

[68] P. Di Giovanni, C. Azlan, T. Ahearn, S. Semple, F. Gilbert, and T. Redpath. The accuracy of pharmacokinetic parameter measurement in DCE-MRI of the breast at 3 T. *Physics in medicine and biology*, 55(1):121, 2009.

[69] A. Diehl, L. Pelorosso, C. Delrieux, C. Saulo, J. Ruiz, M. E. Gröller, and S. Bruckner. Visual analysis of spatio-temporal data: Applications in weather forecasting. *Computer Graphics Forum*, 34(3):381–390, 2015.

[70] P. Diggle, P. Heagerty, K.-Y. Liang, and S. Zeger. Analysis of longitudinal data. *Oxford University Press*, 2002.

[71] S. Djurcilov, K. Kim, P. Lermusiaux, and A. Pang. Visualizing scalar volumetric data with uncertainty. *Computers & Graphics*, 26(2):239–248, 2002.

[72] H. Doleisch. SimVis: Interactive visual analysis of large and time-dependent 3D simulation data. In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, pages 712–720. IEEE Press, 2007.

[73] H. Doleisch, M. Gasser, and H. Hauser. Interactive feature specification for Focus+Context visualization of complex simulation data. In *VisSym*, volume 3, pages 239–248, 2003.

[74] H. Doleisch, P. Muigg, and H. Hauser. Interactive visual analysis of hurricane Isabel with SimVis. In *IEEE Visualization*, 2004.

[75] P. Drapikowski. Surface modeling-uncertainty estimation and visualization. *Computerized Medical Imaging and Graphics*, 32(2):134–139, 2008.

[76] O. Ecabert, J. Peters, M. J. Walker, T. Ivanc, C. Lorenz, J. von Berg, J. Lessick, M. Vembar, and J. Weese. Segmentation of the heart and great vessels in CT images using a model-based adaptation framework. *Medical image analysis*, 15(6):863–876, 2011.

[77] S. B. Edge and C. C. Compton. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Annals of surgical oncology*, 17(6):1471–1474, 2010.

[78] G. Ellis and A. Dix. Enabling automatic clutter reduction in parallel coordinate plots. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):717–724, 2006.

[79] G. Ellis and A. Dix. An explorative analysis of user evaluation studies in information visualisation. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, pages 1–7. ACM, 2006.

[80] G. Ellis and A. Dix. A taxonomy of clutter reduction for information visualisation. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6):1216–1223, 2007.

[81] N. Elmqvist, P. Dragicevic, and J.-D. Fekete. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE transactions on Visualization and Computer Graphics*, 14(6):1539–1148, 2008.

[82] J. I. Epstein, L. Egevad, M. B. Amin, B. Delahunt, J. R. Srigley, P. A. Humphrey, G. Committee, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: definition of grading patterns and proposal for a new grading system. *The American journal of surgical pathology*, 40(2):244–252, 2016.

[83] E. Fanea, S. Carpendale, and T. Isenberg. An interactive 3D integration of parallel coordinates and star glyphs. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pages 149–156. IEEE, 2005.

[84] Z. Fang, T. Möller, G. Hamarneh, and A. Celler. Visualization and exploration of time-varying medical image data sets. In *Proceedings of Graphics Interface 2007*, pages 281–288. ACM, 2007.

[85] F. Ferstl, K. Bürger, and R. Westermann. Streamline variability plots for characterizing the uncertainty in vector field ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):767–776, Jan 2016.

[86] Fisher, R.A. Iris Data Set. `http://archive.ics.uci.edu/ml/datasets/Iris`, 1936.

[87] J. Folkman. Role of angiogenesis in tumor growth and metastasis. In *Seminars in oncology*, volume 29, pages 15–18. Elsevier, 2002.

[88] V. S. Fonov, A. C. Evans, R. C. McKinstry, C. Almli, and D. Collins. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47:S102, 2009.

[89] FP7 DR-THERAPAT. The Digital Radiation Therapy Patient. `http://drtherapat.eu/`, 2 2013.

[90] B. Fröhler, T. Möller, and C. Heinzl. GEMSe: Visualization-Guided Exploration of Multi-channel Segmentation Algorithms. *Computer Graphics Forum (Proceedings of the Eurographics Conference on Visualization (EuroVis) 2016*, 35(3), June 2016.

[91] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner. Hierarchical parallel coordinates for exploration of large datasets. In *Proceedings of the conference on Visualization'99: celebrating ten years*, pages 43–50. IEEE Computer Society Press, 1999.

[92] Z. Geng, Z. Peng, R. S. Laramee, J. C. Roberts, and R. Walker. Angular histograms: Frequency-based visualizations for large, high dimensional data. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2572–2580, 2011.

[93] A. Gerbaulet, R. Potter, J.-J. Mazeron, H. Meertens, and E. V. Limbergen. The GEC ESTRO handbook of brachytherapy. 2002.

[94] G. Gerig, M. Jomier, and M. Chakos. Valmet: A new validation tool for assessing and improving 3D object segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001*, pages 516–523. Springer, 2001.

[95] N. D. Gershon. Visualization of fuzzy data using generalized animation. In *Proceedings of the 3rd Conference on Visualization '92*, VIS '92, pages 268–273. IEEE Computer Society Press, 1992.

[96] A. Geurts, G. Sakas, A. Kuijper, M. Becker, and T. von Landesberger. Visual comparison of 3D medical image segmentation algorithms based on statistical shape models. In *Digital Human Modeling. Applications in Health, Safety, Ergonomics and Risk Management: Ergonomics and Health*, pages 336–344. Springer, 2015.

[97] P. Gibbs, M. D. Pickles, and L. W. Turnbull. Repeatability of echo-planar-based diffusion measurements of the human prostate at 3T. *Magnetic resonance imaging*, 25(10):1423–1429, 2007.

[98] S. Glaßer, K. Lawonn, and B. Preim. Visualization of 3D Cluster Results for Medical Tomographic Image Data. In *In Proc. of Conference on Computer Graphics Theory and Applications (VISIGRAPP/GRAPP)*, pages 169–176, 2014.

[99] S. Glaßer, U. Preim, K. Tönnies, and B. Preim. A visual analytics approach to diagnosis of breast DCE-MRI data. *Computers & Graphics*, 34(5):602–611, 2010.

[100] S. Glaßer, P. Saalfeld, P. Berg, N. Merten, and B. Preim. How to Evaluate Medical Visualizations on the Example of 3D Aneurysm Surfaces. In *Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)*, pages 153–162, 2016.

[101] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2002.

[102] M. Graham and J. Kennedy. Using curves to enhance parallel coordinate visualisations. In *Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on*, pages 10–16. IEEE, 2003.

[103] D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. WEAVE: A system for visually linking 3-D and statistical visualizations, applied to cardiac simulation and measurement data. In *Proceedings of the conference on Visualization'00*, pages 489–492. IEEE Computer Society Press, 2000.

[104] H. Griethe and H. Schumann. Visualizing uncertainty for improved decision making. SimVis, 2006.

[105] G. Grigoryan and P. Rheingans. Point-based probabilistic surfaces to show surface uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 10(5):564–573, 2004.

[106] G. Groenendaal, A. Borren, M. R. Moman, E. Monninkhof, P. J. van Diest, M. E. Philippens, M. van Vulpen, and U. A. van der Heide. Pathologic validation of a model based on diffusion-weighted imaging and dynamic contrast-enhanced magnetic resonance imaging for tumor delineation in the prostate peripheral zone. *International Journal of Radiation Oncology, Biology, Physics*, 82(3):e537–e544, 2012.

[107] G. Groenendaal, M. R. Moman, J. G. Korporaal, P. J. van Diest, M. van Vulpen, M. E. Philippens, and U. A. van der Heide. Validation of functional imaging with pathology for tumor delineation in the prostate. *Radiotherapy and Oncology*, 94(2):145–150, 2010.

[108] E. H. Grubbe. Priority in the therapeutic use of X-rays. *Radiology*, 21(2):156–162, 1933.

[109] L. Harrison, F. Yang, S. Franconeri, and R. Chang. Ranking visualizations of correlation using Weber's law. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1943–1952, 2014.

[110] J. A. Hartigan. Printer graphics for clustering. *Journal of Statistical Computation and Simulation*, 4(3):187–213, 1975.

[111] H. Hauser, F. Ledermann, and H. Doleisch. Angular brushing of extended parallel coordinates. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, pages 127–130. IEEE, 2002.

[112] J. Heinrich and D. Weiskopf. State of the art of parallel coordinates. *STAR Proceedings of Eurographics*, 2013:95–116, 2013.

[113] E. Henderson, B. K. Rutt, and T.-Y. Lee. Temporal sampling requirements for the tracer kinetics modeling of breast disease. *Magnetic resonance imaging*, 16(9):1057–1073, 1998.

[114] R. J. Hendley, N. S. Drew, A. M. Wood, and R. Beale. Case study. Narcissus: visualising information. In *Information Visualization, 1995. Proceedings.*, pages 90–96. IEEE, 1995.

[115] T. Hengl. Visualisation of uncertainty using the HSI colour model: computations with colours. In *7th International Conference on GeoComputation*, 2003.

[116] A. Hennemuth, S. Behrens, C. Kuehnel, S. Oeltze, O. Konrad, and H.-O. Peitgen. Novel methods for parameter-based analysis of myocardial tissue in mr images. In *Medical Imaging*, pages 65111N–65111N. International Society for Optics and Photonics, 2007.

[117] M. Hermann, A. C. Schunke, and R. Klein. Semantically steered visual analysis of highly detailed morphometric shape spaces. In *Biological Data Visualization (BioVis), 2011 IEEE Symposium on*, pages 151–158. IEEE, 2011.

[118] M. Hermann, A. C. Schunke, T. Schultz, and R. Klein. A visual analytics approach to study anatomic covariation. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 161–168. IEEE, 2014.

[119] C. M. Hoeks, J. O. Barentsz, T. Hambrock, D. Yakar, D. M. Somford, S. W. Heijmink, T. W. Scheenen, P. C. Vos, H. Huisman, I. M. van Oort, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology*, 261(1):46–66, 2011.

[120] T. Höllt, A. Magdy, G. Chen, G. Gopalakrishnan, I. Hoteit, C. D. Hansen, and M. Hadwiger. Visual analysis of uncertainties in ocean forecasts for planning and operation of offshore structures. In *Visualization Symposium (PacificVis), 2013 IEEE Pacific*, pages 185–192. IEEE, 2013.

[121] T. Höllt, A. Magdy, P. Zhan, G. Chen, G. Gopalakrishnan, I. Hoteit, C. D. Hansen, and M. Hadwiger. Ovis: A framework for visual analysis of ocean forecast ensembles. *Visualization and Computer Graphics, IEEE Transactions on*, 20(8):1114–1126, 2014.

[122] T. Höllt, N. Pezzotti, V. van Unen, F. Koning, E. Eisemann, B. P. Lelieveldt, and A. Vilanova. Cytosplore: Interactive immune cell phenotyping for large single-cell datasets. *Computer Graphics Forum (Proc. of EuroVis)*, 35(3):171–180, june 2016.

[123] D. Holten and J. J. Van Wijk. Evaluation of cluster identification performance for different PCP variants. In *Computer Graphics Forum*, volume 29, pages 793–802. Wiley Online Library, 2010.

[124] H. Hricak, P. L. Choyke, S. C. Eberhardt, S. A. Leibel, and P. T. Scardino. Imaging prostate cancer: a multidisciplinary perspective. *Radiology*, 243(1):28–53, 2007.

[125] P. A. Humphrey. Gleason grading and prognostic factors in carcinoma of the prostate. *Modern pathology*, 17(3):292–306, 2004.

[126] C. B. Hurley and R. Oldford. Pairwise display of high-dimensional information via eulerian tours and hamiltonian decompositions. *Journal of Computational and Graphical Statistics*, 2012.

[127] S. Ingram, T. Munzner, V. Irvine, M. Tory, S. Bergner, and T. Möller. Dimstiller: Workflows for dimensional analysis and reduction. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on*, pages 3–10. IEEE, 2010.

[128] A. Inselberg. The plane with parallel coordinates. *The visual computer*, 1(2):69–91, 1985.

[129] A. Inselberg and B. Dimsdale. Parallel coordinates for visualizing multi-dimensional geometry. In *Computer Graphics 1987*, pages 25–44. Springer, 1987.

[130] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013.

[131] D. Jaffray, D. Yan, and J. Wong. Managing geometric uncertainty in conformal intensity-modulated radation therapy. In *Seminars in Radiation Oncology*, volume 9, pages 4–19. Elsevier, 1999.

[132] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.

[133] D. H. Jeong, C. Ziemkiewicz, B. Fisher, W. Ribarsky, and R. Chang. iPCA: An Interactive System for PCA-based Visual Analytics. In *Computer Graphics Forum*, volume 28, pages 767–774. Wiley Online Library, 2009.

[134] J. Johansson, P. Ljung, M. Jern, and M. Cooper. Revealing structure within clustered parallel coordinates displays. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pages 125–132. IEEE, 2005.

[135] S. Johansson and J. Johansson. Interactive dimensionality reduction through user-defined combinations of quality metrics. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):993–1000, 2009.

[136] C. R. Johnson and A. R. Sanderson. A next step: Visualizing errors and uncertainty. *IEEE Computer Graphics and Applications*, 23(5):6–10, 2003.

[137] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.

[138] D. K. Jones. Determining and visualizing uncertainty in estimates of fiber orientation from diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49(1):7–12, 2003.

[139] J. F. Kallehauge, K. Tanderup, C. Duan, S. Haack, E. M. Pedersen, J. C. Lindegaard, L. U. Fokdal, S. M. I. Mohamed, and T. Nielsen. Tracer kinetic model selection for dynamic contrast-enhanced magnetic resonance imaging of locally advanced cervical cancer. *Acta Oncologica*, 53(8):1064–1072, 2014.

[140] J. F. Kallehauge, K. Tanderup, S. Haack, T. Nielsen, L. P. Muren, L. Fokdal, J. C. Lindegaard, and E. M. Pedersen. Apparent diffusion coefficient (ADC) as a quantitative parameter in diffusion weighted MR imaging in gynecologic cancer: Dependence on b-values used. *Acta Oncologica*, 49(7):1017–1022, 2010.

[141] M. R. Kaus, K. K. Brock, V. Pekar, L. A. Dawson, A. M. Nichol, and D. A. Jaffray. Assessment of a model-based deformable image registration approach for radiation therapy planning. *International Journal of Radiation Oncology Biology Physics*, 68(2):572–580, 2007.

[142] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon. Visual analytics: Definition, process, and challenges. In *Information visualization*, pages 154–175. Springer, 2008.

[143] D. A. Keim. Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[144] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann. *Mastering the information age-solving problems with visual analytics*. Florian Mansmann, 2010.

[145] E. Kerkhof, R. Van Der Put, B. Raaymakers, U. Van Der Heide, M. Van Vulpen, and J. Lagendijk. Variation in target and rectum dose due to prostate deformation: an assessment by repeated MR imaging and treatment planning. *Physics in medicine and biology*, 53(20):5623, 2008.

[146] D. W. Kerst. The acceleration of electrons by magnetic induction. *Physical Review*, 60(1):47, 1941.

[147] J. H. Kim, J. K. Kim, B.-W. Park, N. Kim, and K.-S. Cho. Apparent diffusion coefficient: prostate cancer versus noncancerous tissue according to anatomical region. *Journal of Magnetic Resonance Imaging*, 28(5):1173–1179, 2008.

[148] S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. Pluim. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Medical physics*, 35(4):1407–1417, 2008.

[149] P. Klemm, K. Lawonn, M. Rak, B. Preim, K. D. Tönnies, K. Hegenscheid, H. Volzke, and S. Oeltze. Visualization and analysis of lumbar spine canal variability in cohort study data. In *VMV*, pages 121–128, 2013.

[150] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Volzke, and B. Preim. Interactive visual analysis of image-centric cohort study data. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):1673–1682, 2014.

[151] J. G. Korporaal, C. A. van den Berg, C. R. Jeukens, G. Groenendaal, M. R. Moman, P. Luijten, M. van Vulpen, and U. A. van der Heide. Dynamic Contrast-enhanced CT for Prostate Cancer: Relationship between Image Noise, Voxel Size, and Repeatability. *Radiology*, 256(3):976–984, 2010.

[152] J. G. Korporaal, C. A. van den Berg, M. J. van Osch, G. Groenendaal, M. van Vulpen, and U. A. van der Heide. Phase-based arterial input function measurements in the femoral arteries for quantification of dynamic contrast-enhanced (DCE) MRI and comparison with DCE-CT. *Magnetic resonance in medicine*, 66(5):1267–1274, 2011.

[153] J. G. Korporaal, M. van Vulpen, C. A. van den Berg, and U. A. van der Heide. Tracer kinetic model selection for dynamic contrast-enhanced computed tomography imaging of prostate cancer. *Investigative radiology*, 47(1):41–48, 2012.

[154] J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

[155] H. J. Kuijf, C. M. Tax, L. K. Zaanen, W. H. Bouvy, J. de Bresser, A. Leemans, M. A. Viergever, G. J. Biessels, and K. L. Vincken. The added value of diffusion tensor imaging for automated white matter hyperintensity segmentation. In *Computational Diffusion MRI*, pages 45–53. Springer, 2014.

[156] J. Kurhanewicz, D. Vigneron, P. Carroll, and F. Coakley. Multiparametric magnetic resonance imaging in prostate cancer: present and future. *Current opinion in urology*, 18(1):71, 2008.

[157] G. J. Kutcher, G. S. Mageras, and S. A. Leibel. Control, correction, and modeling of setup errors and organ motion. In *Seminars in Radiation Oncology*, volume 5, pages 134–145. Elsevier, 1995.

[158] F. Ladstädter, A. K. Steiner, B. C. Lackner, G. Kirchengast, P. Muigg, J. Kehrer, and H. Doleisch. SimVis: An Interactive Visual Field Exploration Tool Applied to Climate Research. In *New Horizons in Occultation Research*. Springer, 2009.

[159] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *Visualization and Computer Graphics, IEEE Transactions on*, 18(9):1520–1536, 2012.

[160] C. H. Lee and A. Varshney. Representing thermal vibrations and uncertainty in molecular surfaces. In *Electronic Imaging 2002*, pages 80–90. International Society for Optics and Photonics, 2002.

[161] A. Leemans, B. Jeurissen, J. Sijbers, and D. Jones. ExploreDTI: a graphical toolbox for processing, analyzing, and visualizing diffusion MR data. In *17th Annual Meeting of Intl Soc Mag Reson Med*, volume 209, page 3537, 2009.

[162] S. A. Leibel, Z. Fuks, M. J. Zelefsky, S. L. Wolden, K. E. Rosenzweig, K. M. Alektiar, M. A. Hunt, E. D. Yorke, L. X. Hong, H. I. Amols, et al. Intensity-modulated radiotherapy. *The Cancer Journal*, 8(2):164–176, 2002.

[163] B. Levin, D. A. Lieberman, B. McFarland, R. A. Smith, D. Brooks, K. S. Andrews, C. Dash, F. M. Giardiello, S. Glick, T. R. Levin, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *CA: a cancer journal for clinicians*, 58(3):130–160, 2008.

[164] M. Lichman. UCI machine learning repository. `http://archive.ics.uci.edu/ml`, 2013.

[165] C. C. Ling, J. Humm, S. Larson, H. Amols, Z. Fuks, S. Leibel, and J. A. Koutcher. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *International Journal of Radiation Oncology Biology Physics*, 47(3):551–560, 2000.

[166] L. Linsen, T. Van Long, and P. Rosenthal. Linking multidimensional feature space cluster visualization to multifield surface extraction. *Computer Graphics and Applications, IEEE*, 29(3):85–89, 2009.

[167] S. K. Lodha, A. Pang, R. E. Sheehan, and C. M. Wittenbrink. UFLOW: Visualizing uncertainty in fluid flow. In *Visualization'96. Proceedings.*, pages 249–254. IEEE, 1996.

[168] C. Lundström, P. Ljung, A. Persson, and A. Ynnerman. Uncertainty visualization in medical volume rendering using probabilistic animation. *IEEE transactions on visualization and computer graphics*, 13(6):1648–1655, 2007.

[169] P. Maillard, O. Carmichael, D. Harvey, E. Fletcher, B. Reed, D. Mungas, and C. DeCarli. FLAIR and diffusion MRI signals are independent predictors of white matter hyperintensities. *American Journal of Neuroradiology*, 34(1):54–61, 2013.

[170] M. M. Malik, C. Heinzl, and M. E. Gröller. Comparative visualization for parameter studies of dataset series. *Visualization and Computer Graphics, IEEE Transactions on*, 16(5):829–840, 2010.

[171] K. G. Manton, I. Akushevich, and J. Kravchenko. *Cancer mortality and morbidity patterns in the US population: an interdisciplinary approach*. Springer Science & Business Media, 2008.

[172] L. B. Marks, E. D. Yorke, A. Jackson, R. K. Ten Haken, L. S. Constine, A. Eisbruch, S. M. Bentzen, J. Nam, and J. O. Deasy. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology, Biology, Physics*, 76(3):S10–S19, 2010.

[173] A. R. Martin and M. O. Ward. High dimensional brushing for interactive exploration of multivariate data. In *Proceedings of the 6th Conference on Visualization'95*, page 271. IEEE Computer Society, 1995.

[174] K. Matkovic, M. Jelovic, J. Juric, Z. Konyha, and D. Gracanin. *Interactive visual analysis and exploration of injection systems simulations*. IEEE, 2005.

[175] K. T. McDonnell and K. Mueller. Illustrative parallel coordinates. In *Computer Graphics Forum*, volume 27, pages 1031–1038. Wiley Online Library, 2008.

[176] A. McKenzie, M. van Herk, and B. Mijnheer. Margins for geometric uncertainty around organs at risk in radiotherapy. *Radiotherapy and oncology*, 62(3):299–307, 2002.

[177] A. M. Mendrik, K. L. Vincken, H. J. Kuijf, M. Breeuwer, W. H. Bouvy, J. De Bresser, A. Alansary, M. De Bruijne, A. Carass, A. El-Baz, et al. MRBrainS Challenge: Online evaluation framework for brain image segmentation in 3T MRI scans. *Computational intelligence and neuroscience*, 2015.

[178] S. Miksch and W. Aigner. A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290, 2014.

[179] P. Muigg, J. Kehrer, S. Oeltze, H. Piringer, H. Doleisch, B. Preim, and H. Hauser. A Four-level Focus+ Context Approach to Interactive Visual Analysis of Temporal Features in Large Scientific Data. In *Computer Graphics Forum*, volume 27, pages 775–782. Wiley Online Library, 2008.

[180] T. Munzner. *Fundamentals of Computer Graphics*, chapter Visualization, pages 675–707. Eds. AK Peters, third edition, 2009.

[181] T. Munzner. A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics*, 15(6):921–928, 2009.

[182] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[183] National Cancer Institute (NCI). Understanding Cancer. `https://www.cancer.gov/about-cancer/understanding/what-is-cancer`, 2015.

[184] National Institute of Health, National Cancer Institute. Cell Biology and Cancer, Understanding Cancer. `https://science.education.nih.gov/supplements/nih1/cancer/guide/understanding1.html`, 9 2013.

[185] T. S. Newman and W. Lee. On visualizing uncertainty in volumetric data: techniques and their evaluation. *Journal of Visual Languages & Computing*, 15(6):463–491, 2004.

[186] K. T. Nguyen, A. Bock, A. Ynnerman, and T. Ropinski. Deriving and Visualizing Uncertainty in Kinetic PET Modeling. In *Proceedings of the EG Visual Computing for Biology and Medicine*, pages 107–114, 2012.

[187] M. Noirhomme-Fraiture. Visualization of large data sets: the Zoom Star solution. *International Electronic Journal of Symbolic Data Analysis*, pages 26–39, 2002.

[188] M. Novotny and H. Hauser. Outlier-preserving Focus+Context visualization in parallel coordinates. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):893–900, 2006.

[189] S. Oeltze, H. Doleisch, H. Hauser, P. Muigg, and B. Preim. Interactive visual analysis of perfusion data. *IEEE transactions on visualization and computer graphics*, 13(6):1392–1399, 2007.

[190] E. A. O'Flynn and N. M. deSouza. Functional magnetic resonance: biomarkers of response in breast cancer. *Breast Cancer Research*, 13(1):204, 2011.

[191] W. K. Oh, M. Hurwitz, A. V. D'Amico, J. P. Richie, and P. W. Kantoff. Biology of prostate cancer. *Holland-Frei Cancer Medicine, 6th edition.*, 2003.

[192] J. S. Olson. *The history of cancer: an annotated bibliography*. Number 3. Greenwood, 1989.

[193] A. R. Padhani, G. Liu, D. Mu-Koh, T. L. Chenevert, H. C. Thoeny, T. Takahara, A. Dzik-Jurasz, B. D. Ross, M. Van Cauteren, D. Collins, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia*, 11(2):102–125, 2009.

[194] A. Pang and A. Freeman. Methods for comparing 3D surface attributes. In *Electronic Imaging: Science & Technology*, pages 58–64. International Society for Optics and Photonics, 1996.

[195] A. T. Pang, C. M. Wittenbrink, and S. K. Lodha. Approaches to uncertainty visualization. *The Visual Computer*, 13(8):370–390, 1997.

[196] L. Pantoni. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *The Lancet Neurology*, 9(7):689–701, 2010.

[197] G. J. Parker and D. L. Buckley. Tracer kinetic modelling for T1-Weighted DCE-MRI. In *Dynamic Contrast-Enhanced Magnetic Resonance Imaging in Oncology*, Medical Radiology, pages 81–92. Springer Berlin Heidelberg, 2005.

[198] G. J. M. Parker, C. Roberts, A. Macdonald, G. Buonaccorsi, S. Cheung, D. L. Buckley, A. Jackson, Y. Watson, K. Davies, and G. C. Jayson. Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for Dynamic Contrast-Enhanced MRI. *Magnetic resonance in medicine*, 56(5):993–1000, 2006.

[199] D. Pasquier, T. Lacornerie, M. Vermandel, J. Rousseau, E. Lartigau, and N. Betrouni. Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy. *International Journal of Radiation Oncology, Biology, Physics*, 68(2):592–600, 2007.

[200] S. Pelengaris and M. Khan. *The molecular biology of cancer: A bridge from bench to bedside*. John Wiley & Sons, 2013.

[201] W. Peng, M. O. Ward, and E. A. Rundensteiner. Clutter reduction in multi-dimensional data visualization using dimension reordering. In *Information Visualization, 2004. IN-FOVIS 2004. IEEE Symposium on*, pages 89–96. IEEE, 2004.

[202] T. Pfaffelmoser, M. Mihai, and R. Westermann. Visualizing the variability of gradients in uncertain 2d scalar fields. *Visualization and Computer Graphics, IEEE Transactions on*, 19(11):1948–1961, 2013.

[203] T. Pfaffelmoser, M. Reitinger, and R. Westermann. Visualizing the positional and geometrical variability of isosurfaces in uncertain scalar fields. In *Computer Graphics Forum*, volume 30, pages 951–960. Wiley Online Library, 2011.

[204] Pickover, C.A. and Tewksbury, S. (eds.). *Frontiers of Scientific Visualization*. Wiley, 1994.

[205] C. Plaisant. The challenge of information visualization evaluation. In *Proceedings of the working conference on Advanced visual interfaces*, pages 109–116. ACM, 2004.

[206] M. D. Plumlee and C. Ware. Zooming versus multiple window interfaces: Cognitive costs of visual comparisons. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(2):179–209, 2006.

[207] J. Poco, D. M. Eler, F. V. Paulovich, and R. Minghim. Employing 2D projections for fast visual exploration of large fiber tracking data. In *Computer Graphics Forum*, volume 31, pages 1075–1084. Wiley Online Library, 2012.

[208] T. Porter and T. Duff. Compositing digital images. In *ACM Siggraph Computer Graphics*, volume 18, pages 253–259. ACM, 1984.

[209] K. Pothkow and H.-C. Hege. Positional uncertainty of isocontours: Condition analysis and probabilistic measures. *IEEE Transactions on Visualization and Computer Graphics*, 17(10):1393–1406, 2011.

[210] K. Pöthkow and H.-C. Hege. Nonparametric models for uncertainty visualization. In *Computer Graphics Forum*, volume 32, pages 131–140. Wiley Online Library, 2013.

[211] K. Pöthkow, C. Petz, and H.-C. Hege. Approximate level-crossing probabilities for interactive visualization of uncertain isocontours. *International Journal for Uncertainty Quantification*, 3(2), 2013.

[212] K. Pöthkow, B. Weber, and H.-C. Hege. Probabilistic marching cubes. In *Computer Graphics Forum*, volume 30, pages 931–940. Wiley Online Library, 2011.

[213] K. Potter, H. Hagen, A. Kerren, and P. Dannenmann. Methods for presenting statistical information: The box plot. *Visualization of Large and Unstructured Data Sets, s*, 4:97–106, 2006.

[214] K. Potter, J. Kniss, R. Riesenfeld, and C. R. Johnson. Visualizing summary statistics and uncertainty. In *Computer Graphics Forum*, volume 29, pages 823–832. Wiley Online Library, 2010.

[215] K. Potter, P. Rosen, and C. R. Johnson. From quantification to visualization: A taxonomy of uncertainty visualization approaches. In *Uncertainty Quantification in Scientific Computing*, pages 226–249. Springer, 2012.

[216] K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, and C. R. Johnson. Ensemble-vis: A framework for the statistical visualization of ensemble data. In *2009 IEEE International Conference on Data Mining Workshops*, pages 233–240. IEEE, 2009.

[217] B. Preim and C. P. Botha. *Visual Computing for Medicine: Theory, Algorithms, and Applications*. Newnes, 2013.

[218] B. Preim, S. Oeltze, M. Mlejnek, E. Gröller, A. Hennemuth, and S. Behrens. Survey of the visual exploration and analysis of perfusion data. *IEEE Transactions on Visualization and Computer Graphics*, 15(2):205–220, 2009.

[219] A. Pretorius, M.-A. P. Bray, A. E. Carpenter, R. Ruddle, et al. Visualization of parameter space for image analysis. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):2402–2411, 2011.

[220] J. L. Prince and J. M. Links. *Medical imaging signals and systems*. Pearson Prentice Hall Upper Saddle River, NJ, 2006.

[221] R. G. Raidou, O. Casares-Magaz, L. P. Muren, U. A. van der Heide, J. Rørvik, M. Breeuwer, and A. Vilanova. Visual Analysis of Tumor Control Models for Prediction of Radiotherapy Response. *In Computer Graphics Forum (CGF)*, 35(3):231–240, 2016.

[222] R. G. Raidou, M. Eisemann, M. Breeuwer, E. Eisemann, and A. Vilanova. Orientation-enhanced parallel coordinate plots. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of the Information Visualization 2016)*, 22(1):589–598, 2016.

[223] R. G. Raidou, H. J. Kuif, N. Sepasian, N. Pezzotti, W. H. Bouvy, M. Breeuwer, and A. Vilanova. Employing visual analytics to aid the design of white matter hyperintensity classifiers. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Proceedings, Part II, Springer International Publishing*, pages 97–105, 2016.

[224] R. G. Raidou, F. J. J. Marcelis, M. Breeuwer, E. Gröller, A. Vilanova, and H. M. M. van de Wetering. Visual Analytics for the Exploration and Assessment of Segmentation Errors. In *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine*, pages 193–202. The Eurographics Association, 2016.

[225] R. G. Raidou, M. P. Moreira, W. van Elmpt, M. Breeuwer, and A. Vilanova. Visual analytics for the exploration of multiparametric cancer imaging. *In Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*, pages 263–264, 2014.

[226] R. G. Raidou, U. A. van der Heide, C. V. Dinh, G. Ghobadi, J. Kallehauge, M. Breeuwer, and A. Vilanova. Visual analytics for the exploration of tumor tissue characterization. *Computer Graphics Forum*, 34(3):11–20, 2015.

[227] R. G. Raidou, U. A. van der Heide, P. J. van Houdt, M. Breeuwer, and A. Vilanova. The iCoCooN: Integration of Cobweb Charts with Parallel Coordinates for Visual Analysis of DCE-MRI Modeling Variations. In *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine*, pages 11–20. The Eurographics Association, 2014.

[228] M. Raj, M. Mirzargar, J. Preston, R. M. Kirby, and R. T. Whitaker. Evaluating alignment of shapes by ensemble visualization. *IEEE Computer Graphics and Applications*, 36(3):60–71, 2016.

[229] P. L. Rheingans. Task-based color scale design. In *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making*, pages 35–43. International Society for Optics and Photonics, 2000.

[230] P. J. Rhodes, R. S. Laramee, R. D. Bergeron, T. M. Sparr, et al. Uncertainty visualization methods in isosurface rendering. In *Eurographics*, volume 2003, pages 83–88, 2003.

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

[232] J. C. Roeske, J. D. Forman, C. F. Mesina, T. He, C. A. Pelizzari, E. Fontenla, S. Vijayakumar, and G. T. Chen. Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy. *International Journal of Radiation Oncology, Biology, Physics*, 33(5):1321–1329, 1995.

[233] R. E. Roth, A. W. Woodruff, and Z. F. Johnson. Value-by-alpha maps: An alternative technique to the cartogram. *The Cartographic Journal*, 47(2):130–140, 2010.

[234] V. Rudat, P. Schraube, D. Oetzel, D. Zierhut, M. Flentje, and M. Wannenmacher. Combined error of patient positioning variability and prostate motion uncertainty in 3D conformal radiotherapy of localized prostate cancer. *International Journal of Radiation Oncology, Biology, Physics*, 35(5):1027–1034, 1996.

[235] D. Ruppert. *Statistics and Data Analysis for Financial Engineering*. Springer Texts in Statistics, 2011.

[236] A. Saad, T. Möller, and G. Hamarneh. Probexplorer: Uncertainty-guided exploration and editing of probabilistic medical image segmentation. In *Computer Graphics Forum*, volume 29, pages 1113–1122. Wiley Online Library, 2010.

[237] J. Sanyal, S. Zhang, G. Bhattacharya, P. Amburn, and R. J. Moorhead. A user study to compare four uncertainty visualization methods for 1D and 2D datasets. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1209–1218, 2009.

[238] J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn, and R. Moorhead. Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1421–1430, 2010.

[239] N. Schadewaldt, D. Bystrov, T. Vik, H. Schulz, J. Peters, A. Franz, C. Buerger, and K. Bzdusek. Robust initialization of multi-organ shape models. *In MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications*, 2013.

[240] J. Schmidt, M. E. Gröller, and S. Bruckner. VAICo: Visual analysis for image comparison. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2090–2099, 2013.

[241] J. Schmidt, R. Preiner, T. Auzinger, M. Wimmer, M. E. Gröller, and S. Bruckner. YMCA – your mesh comparison application. In *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on*, pages 153–162. IEEE, 2014.

[242] W. J. Schroeder and K. M. Martin. Overview of visualization. In *Visualization Handbook*, chapter 1, pages 3–35. Elsevier, 2005.

[243] T. Schultz, L. Schlaffke, B. Schölkopf, and T. Schmidt-Wilcke. HiFiVE: a hilbert space embedding of fiber variability estimates for uncertainty modeling and visualization. In *Computer Graphics Forum*, volume 32, pages 121–130. Wiley Online Library, 2013.

[244] A. Sciarra, J. Barentsz, A. Bjartell, J. Eastham, H. Hricak, V. Panebianco, and J. A. Witjes. Advances in magnetic resonance imaging: how they are changing the management of prostate cancer. *European urology*, 59(6):962–977, 2011.

[245] M. Sedlmair. Design study contributions come in different guises: Seven guiding scenarios. In *Proc. of the IEEE VIS Workshop Beyond Time and Errors: Novel Evaluation Methods for Visualization (BELIV)*. ACM Press, 2016.

[246] M. Sedlmair, C. Heinzl, S. Bruckner, H. Piringer, and T. Moller. Visual parameter space analysis: A conceptual framework. *Visualization and Computer Graphics, IEEE Transactions on*, 20(12):2161–2170, 2014.

[247] J. Seo and B. Shneiderman. Interactively exploring hierarchical clustering results. *Computer*, 35(7):80–86, 2002.

[248] R. Shimofusa, H. Fujimoto, H. Akamata, K. Motoori, S. Yamamoto, T. Ueda, and H. Ito. Diffusion-weighted imaging of prostate cancer. *Journal of computer assisted tomography*, 29(2):149–153, 2005.

[249] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343. IEEE, 1996.

[250] H. Siirtola. Direct manipulation of parallel coordinates. In *Information Visualization, 2000. Proceedings. IEEE International Conference on*, pages 373–378. IEEE, 2000.

[251] H. Siirtola and K.-J. Räihä. Interacting with parallel coordinates. *Interacting with Computers*, 18(6):1278–1309, 2006.

[252] P. Silva, R. G. Raidou, and A. Vilanova. Visualization of variability in radiotherapy dose planning. *Proceedings of the 10th MedVis Conference*, pages 63–66, 2016.

[253] S. Silva, J. Madeira, and B. S. Santos. PolyMeCo - An integrated environment for polygonal mesh analysis and comparison. *Computers & Graphics*, 33(2):181–191, 2009.

[254] J. M. Slater. From X-rays to ion beams: a short history of radiation therapy. In *Ion Beam Therapy – Fundamentals, Technology, Clinical Applications*, pages 3–16. Springer, 2012.

[255] N. N. Smit and K. Lawonn. An introduction to evaluation in medical visualization. In *Proceedings of EuroRV3: EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization*. Eurographics Digital Library, 2016.

[256] D. M. Somford, J. J. Fütterer, T. Hambrock, and J. O. Barentsz. Diffusion and perfusion MR imaging of the prostate. *Magnetic resonance imaging clinics of North America*, 16(4):685–695, 2008.

[257] S. P. Sourbron and D. L. Buckley. Classic models for dynamic contrast-enhanced MRI. *NMR in biomedicine*, 26(8):1004–1027, 2013.

[258] Å. Søvik, E. Malinen, Ø. S. Bruland, S. M. Bentzen, and D. R. Olsen. Optimization of tumour control probability in hypoxic tumours by radiation dose redistribution: a modelling study. *Physics in medicine and biology*, 52(2):499, 2007.

[259] T. C. Sprenger, R. Brunella, and M. H. Gross. H-BLOB: a hierarchical visual clustering method using implicit surfaces. In *Proceedings of the conference on Visualization'00*, pages 61–68. IEEE Computer Society Press, 2000.

[260] P. Steenbergen, K. Haustermans, E. Lerut, R. Oyen, L. De Wever, L. Van den Bergh, L. Kerkmeijer, F. Pameijer, W. Veldhuis, J. van der Voort van Zyp, F. Pos, S. Heijmink, R. Kalisvaart, H. Teertstra, C. Dinh, G. Ghobadi, and U. van der Heide. Prostate tumor delineation using multiparametric magnetic resonance imaging: Inter-observer variability and pathology validation. *Radiotherapy and Oncology*, 115(2):186–190, 2015.

[261] M. D. Steenwijk, J. Milles, M. Buchem, J. Reiber, and C. P. Botha. Integrated visual analysis for heterogeneous datasets in cohort studies. In *IEEE VisWeek Workshop on Visual Analytics in Health Care*, volume 3, page 3, 2010.

[262] Stewart, B.W. and Wild C.P. *World Cancer Report 2014*. International Agency for Research on Cancer, World Health Organization, 2014. WHO reference number: 17600031.

[263] J. C. Stroom and B. J. Heijmen. Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report. *Radiotherapy and oncology*, 64(1):75–83, 2002.

[264] L. Tabar, A. Gad, L. Holmberg, U. Ljungquist, K. C. P. Group, C. Fagerberg, L. Baldetorp, O. Gröntoft, B. Lundström, J. Månson, et al. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the breast cancer screening working group of the swedish national board of health and welfare. *The Lancet*, 325(8433):829–832, 1985.

[265] P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison Wesley, 2005.

[266] K. Tanderup, D. R. Olsen, and C. Grau. Dose painting: art or science? *Radiotherapy and Oncology*, 79(3):245–248, 2006.

[267] J. Thariat, J.-M. Hannoun-Levi, A. S. Myint, T. Vuong, and J.-P. Gérard. Past, present, and future of radiotherapy for the benefit of patients. *Nature reviews Clinical oncology*, 10(1):52–60, 2013.

[268] H. Theisel. Higher Order Parallel Coordinates. In *VMV*, pages 415–420, 2000.

[269] J. J. Thomas. *Illuminating the path:[the research and development agenda for visual analytics]*. IEEE Computer Society, 2005.

[270] M. Thor, M. Væth, Á. Karlsdottir, and L. P. Muren. Rectum motion and morbidity prediction: Improving correlation between late morbidity and dvh parameters through use of rectum planning organ at risk volumes. *Acta Oncologica*, 49(7):1061–1068, 2010.

[271] P. S. Tofts, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B. Larsson, T.-Y. Lee, N. A. Mayr, G. J. Parker, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. *Journal of Magnetic Resonance Imaging*, 10(3):223–232, 1999.

[272] P. S. Tofts, G. Brix, D. L. Buckley, J. L. Evelhoch, E. Henderson, M. V. Knopp, H. B. W. Larsson, T.-y. Lee, N. A. Mayr, G. J. M. Parker, R. E. Port, J. Taylor, and R. M. Weisskoff. Contrast-Enhanced T1-Weighted MRI of a Diffusable Tracer: Standardized Quantities and Symbols. *Journal of Magnetic Resonance Imaging*, 10:223–232, 1999.

[273] C. Tominski, J. Abello, and H. Schumann. Interactive poster: 3D axes-based visualizations for time series data. In *Poster Compendium of IEEE Symposium on Information Visualization, 2005. INFOVIS 2005.* Citeseer, 2005.

[274] T. Torsney-Weir, A. Saad, T. Möller, H.-C. Hege, B. Weber, J.-M. Verbavatz, and S. Bergner. Tuner: Principled parameter finding for image segmentation algorithms using visual response surface exploration. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12):1892–1901, 2011.

[275] G. J. Tortora and B. H. Derrickson. *Principles of anatomy and physiology*. John Wiley & Sons, 2008.

[276] M. Tory and T. Moller. Human factors in visualization research. *IEEE transactions on visualization and computer graphics*, 10(1):72–84, 2004.

[277] M. Tory and T. Moller. Rethinking visualization: A high-level taxonomy. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*, pages 151–158. IEEE, 2004.

[278] A. Trofimov, J. Unkelbach, T. F. DeLaney, and T. Bortfeld. Visualization of a variety of possible dosimetric outcomes in radiation therapy using dose-volume histogram bands. *Practical Radiation Oncology*, 2(3):164–171, 2012.

[279] E. R. Tufte and P. Graves-Morris. *The visual display of quantitative information*, volume 2. Graphics press Cheshire, CT, 1983.

[280] P. Tukey and J. Tukey. Graphic display of data sets in 3 or more dimensions. *The collected works of John Tukey*, 5:189–288, 1988.

[281] C. Turkay, J. Parulek, N. Reuter, and H. Hauser. Interactive visual analysis of temporal cluster structures. In *Computer Graphics Forum*, volume 30, pages 711–720. Wiley Online Library, 2011.

[282] B. Turkbey, P. S. Albert, K. Kurdziel, and P. L. Choyke. Imaging localized prostate cancer: current approaches and new developments. *AJR. American journal of roentgenology*, 192(6):1471, 2009.

[283] B. Turkbey, D. Thomasson, Y. Pang, M. Bernardo, and P. L. Choyke. The role of Dynamic Contrast-Enhanced MRI in cancer diagnosis and treatment. *Diagnostic and interventional radiology.*, 16(3):186–92, 2010.

[284] R. Twiddy, J. Cavallo, and S. M. Shiri. Restorer: A visualization technique for handling missing data. In *Proceedings of the conference on Visualization 1994*, pages 212–216. IEEE Computer Society Press, 1994.

[285] A. Unwin, C. Volinsky, and S. Winkler. Parallel coordinates for exploratory modelling analysis. *Computational Statistics & Data Analysis*, 43(4):553–564, 2003.

[286] L. Van der Maaten and G. Hinton. Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9(85):2579–2605, 2008.

[287] F. van Ham and J. J. van Wijk. Interactive visualization of small world graphs. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on*, pages 199–206. IEEE, 2004.

[288] S. Verma, B. Turkbey, N. Muradyan, A. Rajesh, F. Cornud, M. A. Haider, P. L. Choyke, and M. Harisinghani. Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management. *American Journal of Roentgenology*, 198(6):1277–1288, 2012.

[289] A. Vilanova, R. G. Raidou, and N. Pezzotti. Visual analysis for hypothesis generation in medical imaging research. *Proceedings of the 9th MedVis Conference*, pages 11–12, 2015.

[290] T. von Landesberger, G. Andrienko, N. Andrienko, S. Bremm, M. Kirschner, S. Wesarg, and A. Kuijper. Opening up the black box of medical image segmentation with statistical shape models. *The Visual Computer*, 29(9):893–905, 2013.

[291] T. von Landesberger, D. Basgier, and M. Becker. Comparative Local Quality Assessment for 3D Medical Image Segmentation with Focus on Statistical Shape Model-based Algorithms. *Visualization and Computer Graphics, IEEE Transactions on*, 2015.

[292] T. von Landesberger, S. Bremm, M. Kirschner, S. Wesarg, and A. Kuijper. Visual Analytics for model-based medical image segmentation: Opportunities and challenges. *Expert Systems with Applications*, 40(12):4934–4943, 2013.

[293] J. Wang, W. Peng, M. O. Ward, and E. A. Rundensteiner. Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *IEEE Symposium on Information Visualization, 2003. INFOVIS 2003.*, pages 105–112, 2003.

[294] M. Q. Wang Baldonado, A. Woodruff, and A. Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces*, pages 110–119. ACM, 2000.

[295] M. O. Ward. XmdvTool: Integrating multiple methods for visualizing multivariate data. In *Proceedings of the Conference on Visualization'94*, pages 326–333. IEEE Computer Society Press, 1994.

[296] M. O. Ward. Creating and manipulating n-dimensional brushes. In *Proceedings of Joint Statistical Meeting*, pages 6–14. Citeseer, 1997.

[297] J. H. Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.

[298] C. Ware. Designing with a 2 1/2-D attitude. *Information Design Journal*, 10(3):258, 2001.

[299] C. Ware. *Information visualization: perception for design*. Elsevier, 2012.

[300] Washington, C.M. and Leaver, D.T. *Principles and Practice of Radiation Therapy*. Elsevier - Health Sciences Division, 2015.

[301] S. Webb. *Intensity-Modulated Radiation Therapy*. CRC Press, Taylor & Francis, 2001.

[302] S. Webb and A. Nahum. A model for calculating tumour control probability in radiotherapy including the effects of inhomogeneous distributions of dose and clonogenic cell density. *Physics in Medicine and Biology*, 38(6):653, 1993.

[303] E. J. Wegman. Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association*, 85(411):664–675, 1990.

[304] C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis. Processing and visualization for diffusion tensor mri. *Medical image analysis*, 6(2):93–108, 2002.

[305] R. T. Whitaker, M. Mirzargar, and R. M. Kirby. Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2713–2722, 2013.

[306] C. M. Wittenbrink, A. T. Pang, and S. K. Lodha. Glyphs for visualizing uncertainty in vector fields. *IEEE transactions on Visualization and Computer Graphics*, 2(3):266–279, 1996.

[307] N. Wong, S. Carpendale, and S. Greenberg. EdgeLens: An interactive method for managing edge congestion in graphs. In *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*, pages 51–58. IEEE, 2003.

[308] P. C. Wong and R. D. Bergeron. Multiresolution multidimensional wavelet brushing. In *Visualization'96. Proceedings.*, pages 141–148. IEEE, 1996.

[309] World Health Organization. Cancer Fact Sheet No.297. WHO Media Center, 2 2014.

[310] Z. Xie, S. Huang, M. O. Ward, and E. A. Rundensteiner. Exploratory visualization of multivariate data with variable quality. In *2006 IEEE Symposium On Visual Analytics Science And Technology*, pages 183–190. IEEE, 2006.

[311] L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T.-F. Li, G.-Y. Kim, G. Luxton, and A. Koong. Overview of image-guided radiation therapy. *Medical Dosimetry*, 31(2):91–112, 2006.

[312] H. Yamazaki, H. Shiomi, T. Tsubokura, N. Kodani, T. Nishimura, N. Aibe, H. Udono, M. Nishikata, Y. Baba, M. Ogita, K. Yamashita, and T. Kotsuma. Quantitative assessment of inter-observer variability in target volume delineation on stereotactic radiotherapy treatment for pituitary adenoma and meningioma near optic tract. *Radiation Oncology*, 6(1):1–6, 2011.

[313] J. Yang, W. Peng, M. O. Ward, and E. A. Rundensteiner. Interactive hierarchical dimension ordering, spacing and filtering for exploration of high dimensional datasets. In *Information Visualization, 2003. INFOVIS 2003. IEEE Symposium on*, pages 105–112. IEEE, 2003.

[314] T. E. Yankeelov, G. O. Cron, C. L. Addison, J. C. Wallace, R. C. Wilkins, B. A. Pappas, G. E. Santyr, and J. C. Gore. Comparison of a reference region model with direct measurement of an AIF in the analysis of DCE-MRI data. *Magnetic resonance in medicine*, 57(2):353–361, 2007.

[315] X. Yuan, P. Guo, H. Xiao, H. Zhou, and H. Qu. Scattering points in parallel coordinates. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1001–1008, 2009.

[316] B. Zehner, N. Watanabe, and O. Kolditz. Visualization of gridded scalar data with uncertainty in geosciences. *Computers & Geosciences*, 36(10):1268–1275, 2010.

[317] C. Zhang, T. Schultz, K. Lawonn, E. Eisemann, and A. Vilanova. Glyph-based comparative visualization for diffusion tensor fields. *Visualization and Computer Graphics, IEEE Transactions on*, 22(1):797–806, 2016.

[318] Z. Zhang, D. Gotz, and A. Perer. Iterative cohort analysis and exploration. *Information Visualization*, page 1473871614526077, 2014.

[319] H. Zhou, W. Cui, H. Qu, Y. Wu, X. Yuan, and W. Zhuo. Splatting the lines in parallel coordinates. In *Computer Graphics Forum*, volume 28, pages 759–766. Wiley Online Library, 2009.

[320] H. Zhou, X. Yuan, H. Qu, W. Cui, and B. Chen. Visual clustering in parallel coordinates. In *Computer Graphics Forum*, volume 27, pages 1047–1054. Wiley Online Library, 2008.

[321] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

List of Figures

List of Tables

List of Abbreviations

AICc	Akaike Information Criterion
ANOVA	Analysis of Variance
ADC	Apparent Diffusion Coefficient
AIF	Arterial Input Function
AD	Axial Diffusivity
BT	Brachytherapy
B/L	Brushing and Linking
CD	Cell Density
CTV	Clinical Target Volume
CC	Cobweb Charts
CRC	Composite Ray Casting
CT	Computed Tomography
CA	Contrast Agent
DTI	Diffusion Tensor Imaging
DW(I)	Diffusion-Weighted Imaging
DVH	Dose-Volume Histogram
DCE-MRI	Dynamic-Contrast Enhanced MR Imaging
EPI	Echo-Planar Imaging
ETM	Extended Tofts Model
EBRT	External Beam Radiation Therapy
F+C	Focus+Context
FA	Fractional Anisotropy
GS	Gleason Score
Gy	Gray
GTV	Gross Tumor Volume
HSD	Tukey's Honest Significant Difference
IGRT	Image-Guided Radiation Therapy
IMRT	Intensity-Modulated Radiation Therapy
IR	Inverse Recovery
IV	Irradiated Volume
KDE	Kernel Density Estimation
LDA	Linear Discriminant Analysis

LAMP	Local Affine Multidimensional Projections
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MIP	Maximum Intensity Projection
MD	Mean Diffusivity
MAD	Median Absolute Deviation
MDS	Multi-Dimensional Scaling
NTCP	Normal Tissue Complication Probability
OAR	Organ At Risk
OPCP	Orientation-Enhanced Parallel Coordinate Plots
PCP	Parallel Coordinates Plot
PK	Pharmacokinetic Modeling
PTV	Planning Target Volume
PET	Positron Emission Tomography
PCA	Principle Component Analysis
PI-RADS	Prostate Imaging Reporting and Data System
PSA	Prostate Specific Antigen
RD	Radial Diffusivity
RT	Radiotherapy or Radiation Therapy
ROI	Region of Interest
SPLOM	Scatterplot Matrix
SVD	Small Vessel Disease
t-SNE	t-Distributed Stochastic Neighbor Embedding
TAC	Time-Activity Curve
TIC	Time-Intensity Curve
TNM	Tumor Node Metastasis (Classification of Tumors)
TM	Tofts Model
TF	Transfer Function
TCP	Tumor Control Probability Modeling
TTC	Tumor Tissue Characterization
TV	Treated Volume
2CXM	Two-Compartmental Exchange Model
UE	User Experience
UP	User Performance
VDAR	Visual Data Analysis and Reasoning
VA	Visual Analytics
VTK	Visualization Toolkit
WMH	White Matter Hyperintensity

Once you have a PhD,
every meeting becomes
a doctor's appointment.

Acknowledgements

After four years, this deeply emotional phase of my life has come to its end. It was a long path, most of the times "sunny", sometimes less fun, but I tried to make the most of it and I enjoyed it to the last bit. After all these years, I learned to cherish the happy moments, as much as the sad ones, I embraced some of the PhD stereotypes, I earned new knowledge about my field – but mostly about myself and my boundaries – and I was taught how to face situations and stand tall in front of them. There are so many people that have contributed to all these, and their contribution has come in so many different ways that it becomes impossible to thank them all enough in a couple of pages. If I forgot someone, thank you and forgive me!

First of all, I would like to thank my daily supervisor, dr. Anna Vilanova, and my promotor, prof. Marcel Breeuwer for giving me this terrific opportunity, for trusting me through all these years and for their enormous contribution to the production of this thesis. Your high standards motivated me to work harder and to achieve better results, and also brought me out of my comfort zone and matured me. Dear Anna, you have always been there for me, patient and understanding, supportive and motivating. What I enjoyed most in our collaboration was your directness and critical point-of-view, which allowed me to know at every moment how things were going, and at the same time your positive attitude, which could give me a push even in difficult times – work-wise and at a personal level. *No podia demanar millor supervisor, moltes gràcies per tot!* Dear Marcel, thank you very much for your constant encouragement, guidance and support. Our meetings and also our talks have always been a pleasure, especially when I was seeking for a sanity check in crazy ideas. I benefited so much from your advice, in particular when exploring possibilities for new projects. Your close connection to industry and your experience gave me the opportunity to obtain a different perspective about my work, boosting also my self-confidence. *Hartelijk bedankt voor je enthusiasme en voor je begeleiding!*

I would also like to thank my co-promotor and head of the IMAG/e group at TU Eindhoven, prof. Josien Pluim. Dear Josien, thank you for having me as part of your group and for all the time, support and guidance that you have dedicated to me. From the IMAG/e group, I would also like to thank all my other colleagues who made every Thursday a very *gezellig* and special day of the week. Roy, thank you for being an infinite source of enthusiasm, a huge inspiration and a very pleasant person to work with! You definitely helped me a lot to make a good start at TU/e. Birgit, Neda and Samaneh, our girly talks will always accompany me. It is amazing how an almost girls-only room can make a difference! Jiong, our room-companion, thank you for putting up with us and for all the funny and serious discussions! Mitko, *ti blagodaram* for all the fake Greek-Macedonian quarrels! We can be very proud that we proved how non-sense all this is. Still, my feta is better than yours! Dear Pim, thank you for translating my summary and for all the great time, especially during MICCAI! Dear Behdad, Erik and Maxime thank you for all the fun! Wouter, thank you for being a great collaborator in our student projects! I learned so much working with you. Dear Rina, thank you very much for being always so helpful, eager and willing to lend a hand! It was always a pleasure to be

around such a cheerful person as you! Dear Bart, Andrea, Remco, Laura and Pauly thank you very much. From the Algorithms and Visualization group, I would like to thank prof. Jack van Wijk, first of all, for accepting to be part of my graduation committee and for several helpful discussions we have had during these years. I would also like to thank dr. Huub van de Wetering for our collaboration. Huub, it was great working with you and you helped me so much with our project! Also, I would like to thank Michiel, Stef, Alberto, Bram and Paul for all the great fun we had at conferences! Dear Roeland, I especially thank you for all the coffees and beers and chats and wanderlust talks, we had together! Still, you did not convince me to eat the disgusting pineapple pizza!

During my PhD, I was also part of the Computer Graphics and Visualization group at TU Delft, as a visiting researcher. I would like to thank the head of the group prof. Elmar Eisemann for this great opportunity! Dear Elmar, thank you very much for having me in this *amaaaazing* group for all these years, for all the kind words of support, for your help and motivation, but also for all the funny jokes! I would also like to thank my two special room-mates, Thomas Höllt and Nicola Pezzotti. I am really happy that I shared the biggest part of my days with you and I am sure that I earned two very good friends! Dear Thomas, thank you for your critical feedback and for your advice on my future steps, but also for the outings, for the pizza's and for sharing my coffee and cupcake cravings. I feel that your addition to the group really changed its dynamics and its mood! *Caro Nicola, da quando sei arrivato nel nostro gruppo, ho trovato un amico vero con il quale posso condividere le mie ansie, le gioie e gli scherzi – specialmente per le scelte gastronomiche di Mary (grazie Mary per le risate immense)! Vi ringrazio tutti e due di cuore per essere sempre stati disponibili a sentirmi ed aiutarmi!* Then, I would also like to thank my former colleagues and good friends, Jean Marc and Martin! It has always been a pleasure to have you around and I admit it, I really suffered when both of you left the group! Jean Marc, thank you for being always such a positive per-son, so much passionate about everything and always with your party mood on, *surtout dans mon jardin*! Dear Martin, thank you for all our funny moments, full of jokes and laughter, for all the partying in Rotterdam and for our failed attempts to play badminton and do sports! Apart from these, I was extremely lucky to have worked with you, as well. You taught me so much and I really look forward to doing this again! I would like to thank also all my other colleagues, with whom I also shared so many beautiful moments. Noeska, a special thanks to you, for always being there for me! Changgong and Jingtang, thank you for teaching me so much about the Chinese culture and for the wonderful Chinese dinners! Thomas K. thank you very much for all the funny and inappropriate jokes, you always managed to bring out the best (or worst?) in me! Ricardo, Pedro and Peter Kok, you are among the very first people I had the chance to meet in the group and I really cherish all the moments we had. Special thanks go to Francois, who was also my supervisor and mentor when I was doing my mas-ter thesis in the group! Dear Rafa, thank you for all the serious and less serious talks that we had through all these years! *Obrigada!* My most recent colleagues, Klaus, Niels, Timo-thy, Pablo, Michael, Christopher, Victor, Nestor and Leo, thank you for the nice talks we have had! I would like to thank Ruud and Bart, in particular, who were always welcoming me in their office and helping me always, with a smile on! Also, I would like to thank Frits Post, for all the beautiful advice he gave me throughout these years and for all the motivational conversations we have had. Last, but definitely not least, I would like to thank Charl Botha. Dear Charl, with your unlimited enthusiasm you have really inspired me to work in the field of Visualization! Doing my master thesis with you taught me so much and prepared me really well for my PhD. Thank you very much for everything!

I also want to thank all the students I had throughout these years. Dear Marta, Pedro, Freek and Giannis, you trusted me to do your master project with me and you helped me carry on significant parts of my research. I learned so much with you and you helped me realize how much I like supervising and mentoring students. Thank you very much! I would also like to thank a number of people from the INSY department, who I had the pleasure to meet during these years. Dear Zeki, thank you for all the beautiful talks, for making each other jealous with our trips, for always being full of positivity which made me smile even in difficult days! Dear Babak, Alessio and Raynor, thank you for all the kitchen talks over coffee! Dear Marco, it was a pleasure chatting with you, while raising a glass or two at the /Pub! Dear Boudewijn, thank you very much for your kindness and positivity, but also for the nice discussions we have had all these years!

In the last year of my PhD, I was also invited as a visiting researcher at the Institute of Computer Graphics and Algorithms of TU Wien. This was an amazing experience for me and also a great opportunity, not only for my career, but also for my personal development, as I had the chance to meet a lot of great people! First of all, I would like to thank Ao. Prof. Eduard Gröller (*Meister! oh, Meister!*) and Prof. Werner Purgathofer for offering me this possibility and for trusting me again for our future collaboration. I would also like to thank all the wonderful people that I had the chance to interact with: the always smiling Ivan, Manuela (thank you for your advices!), Peter Mindek (special thanks to you for the amazing drawing that is in my thesis!!), Haichao (with your board games!), Johannes (*Yourrrrr Highness*), Tobias (one more ukulele for the band!), Alexey, Victor, Mathieu, Johanna, David, Nicolas, Gabriel, Sebastian and Jiri. Special thanks to Katja Bühler from the VRVis for all the wonderful and enlightening discussions! Dearest Bara, thank you so much for all the great talks, your advice and the amazing dancing! My dear "Viennese", thank you all so much for helping me to feel at home, for all the fun evenings and for half-destroying my liver during my stay. I am really looking forward to be and work with you again! After all, *you can always check out, but you can never leave!*

My research would have never been possible without the significant contribution of a number of collaborators. I was extremely fortunate to be part of a large, multi-disciplinary and international consortium for my project and to co-operate with a group of bright scientists and researchers with great vision. Working with you has been a great opportunity, which made me broaden my horizons and deepen my knowledge into the field. I would like to thank my collaborators from the Netherlands Cancer Institute, prof. Uulke van der Heide, Cuong Dinh, Peter Steenbergen, Petra van Houdt and Ghazaleh Ghobadi for their contribution to my work and for setting up great projects together! Dear Uulke, thank you very much for always having your door open for me and for providing a critical eye that greatly improved my research! I would also like to thank my collaborating partners from Aarhus University and Aarhus University Hospital, who played a great part in all our projects, but also welcomed me multiple times in cold Denmark! Prof. Ludvig Muren, Oscar Casares, Jesper Kallehauge, Søren Haack, Susanne Rylander and Prof. Kari Tanderup, thank you so much for our co-operation and for your great input for my projects! In particular, dear Ludvig and dear Oscar, it has always been a great pleasure discussing and working with you! Your enthusiasm and feedback was a real driving force for my work, and I always had a great time also outside the working hours! *Tusen takk! Muchisimas gracias por todo!* Special thanks go to my collaborators from Philips Research Hamburg and, in particular, Steffen Reinisch, Nicole Schadewaldt and Sven Prevrhal. I always enjoyed our discussions and I would like to thank you for the ef-

fort and time that you dedicated to our projects. Finally, I would like to thank Hugo Kuijf from Utrecht University, for his great impact to our project, which resulted into a MICCAI paper, of which I am particularly proud! Dear Hugo, thank you for your patience and your feedback, but also thank you for the joyful time we had during the conference.

My closest friends, I would like to address each one of you separately. Δωράκι μου γλυκό, όλα αυτά τα χρόνια που είσαι δίπλα μου στα όμορφα και στα δύσκολα ήσουν ένα πραγματικό δώρο για μένα! Από την πρώτη στιγμή εκεί στο αμφιθέατρο του Ζάχου! Όσα ευχαριστώ και να σου πω δεν θα είναι ποτέ αρκετά! Ηρακλάκο μου, αστέρι μου, προσπαθώ να σου γράψω δύο γραμμές και σκέφτομαι όλα όσα έχουμε περάσει μαζί, ευχάριστα, δυσάρεστα, αστεία και ενίοτε γελοία! Τα περισσότερα δεν περιγράφονται δημόσια, οπότε θα αρκεστώ να σε ευχαριστήσω μέσα από την καρδιά μου για όσες (καθόλου λίγες) φορές που ήσουν δίπλα μου και μου έδωσες κουράγιο και δύναμη, πάντα με το χαμόγελο και την αισιοδοξία σου! Μαριλούδι μου, χάρη σε σένα ξέρω τι πάει να πει να έχεις μια μικρή αδερφή! Ακόμα και τώρα που είμαστε μακριά, ξέρω ότι με νοιάζεσαι και είσαι πάντα δίπλα μου, όπως κι εγώ! Σε ευχαριστώ για το γλυκό σου χαμόγελο και τη στήριξή σου, που τα διορθώνουν όλα με μιας! Βασίλη, πιγκουίνε μου, σε ευχαριστώ για τις απεριόριστες στιγμές γέλιου που μου έχεις χαρίσει, για τα όμορφα ταξίδια που κάναμε μαζί, για τις φορές που ήσουν εκεί για μένα (κυρίως τα πρώτα δύσκολα χρόνια της Ολλανδίας) και για τα αξέχαστα κάλαντα στην πλατεία του χωριού (ξέρεις εσύ)! Εύχομαι να έχουμε άπειρες τέτοιες στιγμές στο μέλλον!

My dear Babbolz, Martina and Daniel, I thank you wholeheartedly for all the years of friendship, for all the late and tipsy evenings, for the dancing, for sharing and discussing all our deep thoughts and fears and hopes for the future! Also, thank you for the large quantities of hummus, *gbejna* and wine that you have often shared with me! *Grazzi!*

Ειρηνάκι και Νικολιό μου (φυσικά και το μικρό Μαριλούκι, ναι!) είστε η πιο γλυκιά νότα στη μουντάδα της Ολλανδίας, τα τελευταία χρόνια! Σας ευχαριστώ πολύ για τις υπέροχες βραδιές τραγουδιού και κεφιού, το ένα και μοναδικό μάθημα γιουκαλίλι (είμαι ανεπίδεκτη, το κατάλαβα), τις βόλτες στην εξοχή, τις στιγμές δημιουργίας στα κεραμικά και την οικογενειακή ζεστασιά που νιώθω κάθε φορά μαζί σας. Κιμωνάκο, Τανιούχο και μικρό Φοιβόνι, μπορεί να είστε οι πιο πρόσφατες προσθήκες στον οικογενειακό μου κύκλο, αλλά νιώθω να σας ξέρω χρόνια! Είμαι πραγματικά ευγνώμων που είχα την τύχη να σας γνωρίσω και σας ευχαριστώ για όλα τα όμορφα που έχουμε περάσει, και για όσες φορές μου ανυψώσατε το ηθικό με το χιούμορ σας και την θετική σας συμπεριφορά (με την Τάνια να αλληθωρίζει)! Σας ευχαριστώ επίσης που μου προσφέρατε ένα σπιτικό σε δύσκολους καιρούς! Leonoraki and Γιώργο, thank you wholeheartedly for all the beautiful moments, full of laughter and joy! My dear Leonor, talking to you has always felt so right and fulfilling – whether it was about *pastéis*, or about deep and troublesome matters!

Meine lieben Angelika und Ivo, vielen Dank für all die schönen Momente, für die Gastfreundschaft, für die Spaziergänge in unserem schönen Wien, für die leckeren, schläfrigen Frühstücke und für die Filmabende und Vorschläge! Mit euch ist es immer lustig und hat mir wirklich geholfen, mich in dieser Großstadt zu Hause zu fühlen! Dearest Corrie, thank you so much for your dancing lessons and for your support and determination! Taking your class has given me confidence, has made me to believe in myself and was a wonderful, fun break from my stressful routine! Thank you for believing in me! *Carissimi Rodrigo, Francesca e Vituccio (amorino!), infinitamente grazie di cuore! È così importante sapere che ci sono persone come voi nella mia vita! Spero di poter vedervi più spesso, ora che saremo vicini! Silvinho,*

obrigada por sempre apreciares os meus cozinhados, pela boa companhia e pelo tempo tão bem passado durante as nossas viagens! Beste Marije, bedankt! Dear Belén, thank you very much for your friendship and your love all those years! Despite ups and downs, you were always there for me! Dear Slavik, thank you! My dear Todor and Milena, my little Tamara, thank you for all the beautiful moments we have lived together, in Holland and in Serbia! I wish we have much more time together in the future! My dear Indian friends, Nishant, Smita, Aswin and Revathi, thank you so much for the unforgettable experiences we had in India, at your wonderful and touching weddings! Thank you for all the years of friendship, for your warm hearts that are always open for everyone, and for teaching me how to remain always positive! Τάσο, όχι που θα σε ξέχναγα! Ευχαριστώ για τα άπειρα κρύα ανέκδοτα, τα οποία μόνο εγώ μπορώ να εκτιμήσω και να γελάσω μαζί τους. Dear Marieke and Eilko, thank you so much for all the great dinners and talks, but especially for caring always so much about me! *Hartelijk bedankt!* Dear Anthony, thank you for being always so enthusiastic and for all the great fun, especially during our NZ trip!

Βαλάντη μου και Γιούλη μου, σας ευχαριστώ μέσα από καρδιάς για όλα αυτά τα χρόνια που είστε δίπλα μου και με στηρίζετε. Μαζί σας νιώθω πως έχω μια δεύτερη οικογένεια – ή καλύτερα – μια κερκίδα! Μπορεί να είστε μακριά, αλλά ξέρω ότι είστε πάντα εκεί για μένα! Κυρία Έλλη μου, σας ευχαριστώ για την αγάπη σας και την φροντίδα σας όλα αυτά τα χρόνια! Επίσης, θα ήθελα να ευχαριστήσω τους πιο παλιούς μου φίλους: Σαρρή, Λένα, Γιώργη και – φυσικά – Λούλη! Είστε η ζωντανή απόδειξη ότι με κάποιους ανθρώπους δένεις για μια ζωή, όσο μακριά και αν είναι, όσα χρόνια κι αν περάσουν. Σας ευχαριστώ!

During all these years, my biggest support and my rock has been my family. Μανούλα μου και μπαμπάκα μου, σας ευχαριστώ πραγματικά για όλα όσα έχετε κάνει για μένα όλα αυτά τα χρόνια. Σας ευχαριστώ πολύ για την συμπαράσταση στις δύσκολες στιγμές και για τα δάκρυα χαράς όταν κάτι πήγαινε καλά. Σας είμαι ευγνώμων που από μικρή με ενθαρρύνατε να κάνω αυτό που μου αρέσει και να δρομολογήσω την πορεία μου. Ποτέ δεν μου αρνηθήκατε τη βοήθειά σας, πάντα ακούγατε τις γκρίνιες μου, πάντα ανεχόσασταν τις παραξενιές μου. Συγγνώμη αν μερικές φορές σας στεναχώρησα με τα παράπονά μου. Μπαμπά μου σε ευχαριστώ για το γονίδιο της παλαβομάρας και της ανεμελιάς που πολλές φορές μου έκανε τη ζωή πιο ευχάριστη και εύκολη, μαμά μου σε ευχαριστώ που ποτέ δε μου χάιδεψες τα αυτιά και με έμαθες να πατάω στα πόδια μου. Δε θα μπορούσα να έχω καλύτερους γονείς, ελπίζω να ανταπόδωσα στο ελάχιστο, κάνοντάς σας περήφανους για μένα! Σας αγαπώ πολύ! *Nonnina mia, ti ringrazio per essere sempre stata la piu dolce nonna del mondo, per essere la mia seconda mamma. Ti voglio un mondo di bene! Zio Franco, a te devo la mia testa dura ed il cuore tenero. Spero che tu sia fiero della tua versione bassa e stonata! Nonno Nicola, grazie per essermi sempre vicino, anche da lassù.*

Ευχαριστώντας όσους ήρθαν στη ζωή μου για να μείνουν, σκέφτομαι όσους έφυγαν στη διαδρομή. Ευχαριστώ λοιπόν κι εκείνους, αφού άλλωστε, "οι άνθρωποι υπάρχουν από τη στιγμή που βρίσκουν μια θέση στη ζωή των άλλων".

Curriculum Vitæ

Renata Georgia Raidou was born on April 9, 1987, in Volos, Greece. She received her secondary education at the 5th Lyceum of Volos, Greece, in June 2005. In September 2005, she started her studies at the Faculty of Electrical and Computer Engineering of the National Technical University of Athens (NTUA), in Greece. In July 2010, she obtained her Diploma (BSc+MSc), with a major in Telecommunications and a minor in Biomedical Engineering.

In September 2010, she started her studies in Biomedical Engineering, at the Faculty of Mechanical, Maritime and Materials Engineering (3mE) at Delft University of Technology, in the Netherlands. She specialized in Medical Imaging, and conducted her master thesis project within the group of Computer Graphics and Visualization (CGV), in collaboration with Leiden University Medical Center (LUMC). Her project was on the *Visualization for the Planning and Guidance of Minimally Invasive Cement Injection in Orthopedics*. It was successfully defended, in October 2012. After her graduation, she started working as a Data Scientist, at the Orthopedics Department of LUMC, the Netherlands.

In March 2013, she started her Ph.D. at the Medical Image Analysis Group (IMAG/e) of Eindhoven University of Technology, in the Netherlands. Her project was conducted as part of the FP7 project *DR THERAPAT – Digital Radiation Therapy Patient*, in a strong collaboration with the Netherlands Cancer Institute (NKI-AvL), Aarhus University Hospital in Denmark, Philips Healthcare Best and Philips Research Hamburg. The results of this project are presented in this dissertation.

During her Ph.D., she was a visiting researcher at the Computer Graphics and Visualization (CGV) Group of Delft University of Technology. Additionally, she performed an externship at the Institute of Computer Graphics and Algorithms of Vienna University of Technology, in Austria, and a short externship at the Department of Medical Physics, Aarhus University Hospital, in Denmark. As part of her Ph.D., she supervised four master thesis projects and undertook educational tasks in visualization courses.

Since February 2017, she is employed as a researcher at the Computer Graphics Group of the Institute of Computer Graphics and Algorithms at Vienna University of Technology, in Austria.

List of Publications

Conference and Journal Papers:

7. **Raidou, R.G.**, Kuijf, H.J., Sepasian, N., Pezzotti, N., Bouvy, W.H., Breeuwer, M., and Vilanova, A., 2016. Employing Visual Analytics to Aid the Design of White Matter Hyperintensity Classifiers. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Proceedings, Part II, Springer International Publishing*, pp. 97-105

6. **Raidou, R.G.**, Marcelis, F.J.J., Breeuwer, M., Gröller, M.E., Vilanova, A., and van de Wetering, H.M.M., 2016. Visual Analytics for the Exploration and Assessment of Segmentation Errors. *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM 2016)*, pp. 193-202.

5. Malan, D.F., van der Walt, S.J., **Raidou, R.G.**, van den Berg, B., Stoel, B.C., Botha, C.P., Nelissen, R.G. and Valstar, E.R., 2016. A fluoroscopy-based planning and guidance software tool for minimally invasive hip refixation by cement injection. *International journal of computer assisted radiology and surgery (JCARS)*, 11(2), pp.281-296.

4. **Raidou, R.G.**, Casares-Magaz, O., Muren, L.P., van der Heide, U.A., Rørvik, J., Breeuwer, M. and Vilanova, A., 2016. Visual Analysis of Tumor Control Models for Prediction of Radiotherapy Response. *In Computer Graphics Forum (CGF)*, Vol. 35, No. 3, pp. 231-240.

3. **Raidou, R.G.**, Eisemann, M., Breeuwer, M., Eisemann, E. and Vilanova, A., 2016. Orientation - Enhanced Parallel Coordinate Plots. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of the Information Visualization 2016)*, 22(1), pp.589-598.

2. **Raidou, R.G.**, Van Der Heide, U.A., Dinh, C.V., Ghobadi, G., Kallehauge, J.F., Breeuwer, M. and Vilanova, A., 2015. Visual analytics for the exploration of tumor tissue characterization. *In Computer Graphics Forum (CGF)*, Vol. 34, No. 3, pp. 11-20.

1. **Raidou, R.G.**, van der Heide, U.A., van Houdt, P.J., Breeuwer, M. and Vilanova, A., 2014. The iCoCooN: Integration of Cobweb Charts with Parallel Coordinates for Visual Analysis of DCE-MRI Modeling Variations. *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM 2014)*, pp. 11-20.

Short Papers:

3. Silva, P., **Raidou, R.G.**, and Vilanova, A., 2016. Visualization of Variability in Radiotherapy Dose Planning. *Proceedings of the 10th MedViz Conference (MedViz 2016)*, pp.63-66.

2. Vilanova, A., **Raidou, R.G.**, and Pezzotti, N., 2015. Visual Analysis for Hypothesis Generation in Medical Imaging Research. *Proceedings of the 9th MedViz Conference (MedViz 2015)*, pp.11-12.

1. **Raidou, R.G.**, Moreira, M.P., van Elmpt, W., Breeuwer, M. and Vilanova, A., 2014. Visual analytics for the exploration of multiparametric cancer imaging. *In IEEE Conference on Visual Analytics Science and Technology (VAST) 2014*, pp. 263-264.

Much to learn you still have,
my young padawan!