Aus dem Institut für Community Medicine
(Direktor: Univ.- Prof. Dr. Wolfgang Hoffmann)
der Universitätsmedizin der Ernst-Moritz-Arndt-Universität Greifswald

Thema:

Prostate volume estimation in MR images for epidemiological and clinical studies – new methods

Inaugural - Dissertation

zur

Erlangung des akademischen

Grades

Doktor der Wissenschaften in der Medizin

(Dr. rer. med.)

der

Universitätsmedizin

der

Ernst-Moritz-Arndt-Universität

Greifswald

2014

vorgelegt von:

Mohamad Habes

geb. am: 05.03.1983

in: Ruhaibeh

Dekan: Prof. Dr. Reiner Biffar

1. Gutachter: Prof. Dr. Wolfgang Hoffmann, MPH

2. Gutachter: Prof. Dr. Norbert Hosten

3. Gutachter: PD. Dr. Thomas Wittenberg

Ort, Raum: Greifswald, ICM Hörsaal Ellernholzstraße 1/2

Tag der Disputation: 30.10.2014

Contents

CONTENTS

1 Introduction

1.1 Background

Benign prostatic hyperplasia (BPH) prevalence reaches 50% among 50-year-old men, and prostatic enlargement probability is 90% among 80-year-old men [Espinosa 2013]. BPH is associated with different symptoms. Often, the patients suffer from bladder outlet obstruction and lower urinary tract symptoms (LUTS or so-called prostatism [Abrams 1994]). A possible explanation for the obstruction is the pressure caused by BPH on the urethra. In general, the relationship between clinical symptoms and BPH is still not well understood [Hald 1989, Oelke 2007]. BPH can compromise the quality of life, requires long-term medication and, in many cases, ultimately surgery [Abrams 1994]. For BPH diagnosis, the prostate volume (PV) and serum prostate-specific antigen (PSA) levels should be measured. Prostatic enlargement showed correlation with serum PSA values [Stamey 1987] and age [Vesely 2003] [Tanguay 2009].

Despite its high prevalence and clinical significance, there are a number of open questions in BPH. The literature reports various cut-off values for pathological enlargement of the prostate. While Kim et al. defined a PV ≥ 25 ml as the criterion for an enlarged prostate [Kim 2011], Herbert et al. defined enlarged prostate as PV ≥ 40 ml [Herbert 2004]. PV measurement in large population-based studies would allow deriving more objective reference values and a more valid early BPH diagnosis.

In the clinical context, measurement of PV is important for treatment response monitoring in the clinical applications for BPH management research. Ditonno et al. [Ditonno 2005] evaluated the efficacy of oral lonidamine treatment for subjects with BPH symptoms. The subjects in the treatment group experienced a decrease of 11.2% in PV. Boyle et al. used PV to predict the outcome of BPH treatment with finasteride [Boyle 1996]. Silvero et al. analyzed whether the effect of further combination with additional medications could offer an advantage in BPH treatment, potentially reducing the time required to achieve its symptomatic effect [Di Silvero 2005]. Accurate PV assessment is therefore an essential step for similar studies. PV assessment is currently achieved using transrectal ultrasound (TRUS) as the gold standard in clinical applications. By convention, TRUS assumes a prostate gland shape to be ellipsoid [Milonas 2003]. Extensive enlargement caused by

BPH, however, can cause additional form variations, rendering that strict shape assumption inaccurate for precise assessment of PV. Nathan et al. revealed that PV estimation using the ellipsoid formula often underestimates the real PV and argued that all measured dimensions contain errors [Nathan 1996]. Rodriguez et al. confirmed the consistent underestimation of the actual gland size with the ellipsoid formula [Rodriguez 2008]. They concluded that the width in TRUS and not the length is the least reliable factor.

Different groups have carried out research on the topic of PV prediction from serum PSA levels [Roehrborn 1999, Morote 2000, Park 2013]. Similar prediction models have the potential to estimate PV without using the TRUS modality, which can be an advantage especially in cases of patients with contraindications against TRUS, e.g., patients with anal fissures. In addition, the shape assumption used as the standard in TRUS can underestimate the PV. In their research, Morote.et al. and Park et al. used the ellipsoid formula for PV estimation. More exact prediction models could be developed using methods for PV estimation more reliable than the ellipsoid formula.

Although MRI is more costly in clinical routines than TRUS, regular MR examination could be completely non-invasive. Hence, MRI provides an additional option for PV estimation, but its usability and precision has not yet been assessed in sufficient detail. To validate PV measurement based on MRI for clinical use requires a characterized sample of the general population, in which participants have undergone a standardized MRI exam.

1.2 The Study of Health in Pomerania (SHIP)

The Study of Health in Pomerania (SHIP) is a general adult population-based prospective cohort study in the region of Western Pomerania in northeastern Germany [Hegenscheid 2009, Voelzke 2011]. The aim of SHIP is to assess general health in the community and measure the prevalence of common risk factors for preclinical diseases to allow better understanding of the cause of disease and ultimately improve prevention.

Data are collected in SHIP using interview-based surveys as well as non-invasive methods. The second follow-up of SHIP included a whole-body Magnetic Resonance Imaging (WB-MRI) standardized protocol for all participants, representing the general

health status of the population. The third follow-up examination of SHIP participants (also including WB-MRI) is scheduled for the year 2014.

1.3 Objectives of the thesis

The first objective of this thesis was to develop a method for automated prostate segmentation in a WB-MRI context of a population-based study. Automation should reduce interobserver variation and human resources in image reading. The second objective was to evaluate a new, accurate MR-guided approach for assessment of PV without any shape assumption and suitable for clinical and clinical epidemiological applications. Both objectives made use of kernel-based methods, which were not applied previously for PV estimation in the MR modality. In particular, the possibility of integrating Support Vector Machines methods in epidemiological and clinical epidemiological applications is to be evaluated.

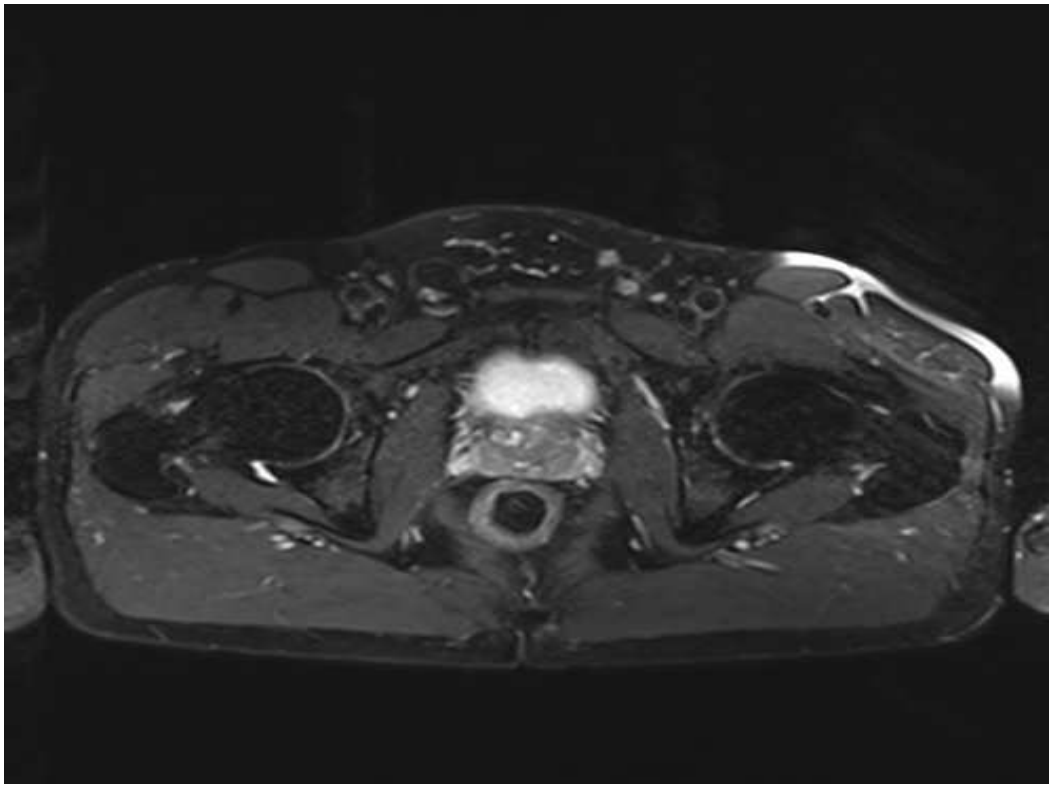
This cumulative thesis is based on two peer-reviewed original articles recently published in international journals [Habes 2013b, Habes 2013c]. The material of both papers is attached to the thesis in Appendix I [Habes 2013b] and Appendix II [Habes 2013c]. In the following section, the basic approach, data and instruments used are summarized. The core results of both papers are presented in the results section. In the discussion section, the present work is put in the context of recent published literature and a detailed outlook toward future research aspects on the topic is provided.

2 Materials and Methods

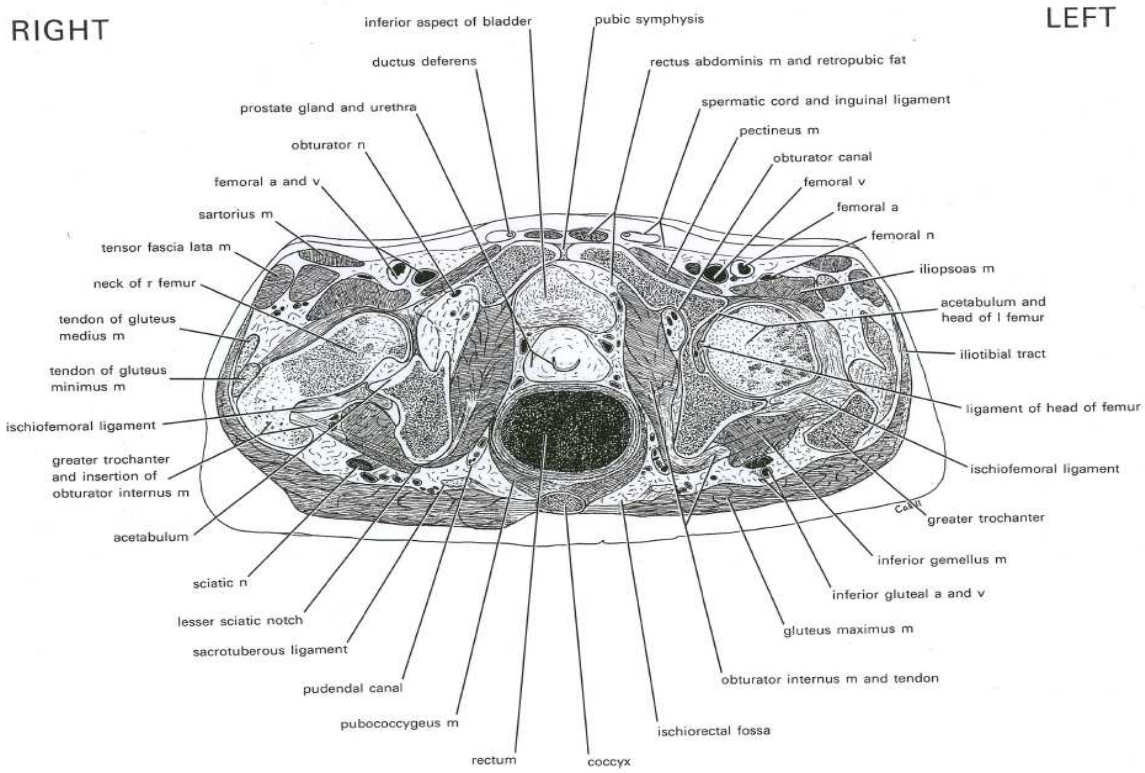
2.1 MR Images

All Study participants underwent MRI scans using a 1.5-T MR scanner (Avanto; Siemens Medical Systems, Erlangen, Germany) [Hegenscheid 2009, Voelzke 2011]. As described in the SHIP protocol, the axial Proton Density Fast Spin Echo Fat Saturated Sequence was used for pelvis visualization [Hegenscheid 2009]. Figure 1 illustrates an example of the sequence used for one SHIP participant and the corresponding anatomical structures of the pelvis. Image parameters of this sequence were: TR 3230 ms, TE 34 ms, flip angle 90°, voxel size (0.8 × 0.8 × 3 mm: 0.9 mm gap). In this thesis, MR images of the pelvis region in two different groups of randomly selected SHIP participants were used.

MATERIALS AND METHODS



a)



b)

Figure 1: a) an axial slice example of the sequence used for visualization of pelvic structures in SHIP for one participant, b) anatomical image of the male pelvis [Cahill 1995]. *This material is reproduced with permission of John Wiley & Sons, Inc*

The first group – consisting of sixteen participants' scans (males between 40 and 60 years old) – was used to develop a fully automatic algorithm, which was validated against randomly selected manual segmentations by two different observers. Following Tingelhoff et al, the decision was made in this thesis to have two expert observers (A and B) assess a limited number of MRI scans rather than have only one observer assess the maximum number of screens in order to parameterize the fully automated prostate segmentation algorithm [Tingelhoff 2008]. However, resources of two manual readings for the evaluation were limited for only sixteen subjects, bearing in mind that this fully automated method is intended to reduce the need for human resources in reading MR scans. The second group – consisting of fifty-three participants' scans (males between 35 and 70 years old) – was used to develop a semi-automatic algorithm for applications in a clinical context. Those measurements and the clinical standard formula for PV estimation (ellipsoid) were compared to the manual delineation of observer A (as the closest to the real PV), who measured the PV with the ellipsoid formula and the semi-automated method developed. This semi-automated method is to be employed as an adjunct tool in clinical studies, and is not intended to replace the human expert. Slice by slice, the experts performed the manual prostate delineation on the scans by visual inspection without further algorithmic support, using the open source Biomedical Image Analysis Package ImageJ, version 1.45 [Schneider 2012].

The ethics committee of the University of Greifswald approved the general SHIP project. The SHIP personnel obtained written informed consent from all participants, which included MRI studies for quantitative evaluation in a large variety of research areas, also including the prostate. All algorithms developed during this thesis used ImageJ, version 1.45 [Schneider 2012] as the developing framework.

2.2 Prostate volume estimation methods

During this Ph.D. thesis, binary class Support Vector Machines (C SVM) algorithms were initially developed for automated segmentation of human structures in ultra-short echo time (UTE) MR images, including the skull as well as cavities, in tasks of different projects [Habes 2013a]. The pipelines developed in that study showed that binary SVM methods can be used for automated segmentation from double components produced simultaneously by the UTE sequence. Both components could help in tissue

discrimination, since they contained complementary information. However, the pelvic sequence used in SHIP – to visualize general pelvic organs – produces only one component. This makes prostate delineation using SVM methods in this sequence more challenging. The integration of C SVM in the fully automated algorithm for PV estimation required additional automatically generated features for prostate tissue discrimination, as well as consideration of a more complex voxel neighborhood [Habes 2013b]. The application of single-class Support Vector Machines (S SVM) required a histogram-based normalization step to automate the skull segmentation [Habes 2013a]. However, in this thesis, the application of S SVM for accurate PV estimation made use of its advantage in training with just one class and with a limited sample [Habes 2013c]. In the following, both algorithms developed for prostate volume estimation are explained.

2.4.1 Method 1: C SVM for epidemiological studies

Cortes and Vapnik presented the binary Support Vector Machines (the term C SVM is used here as suggested in kernel-based methods literature) for binary class classification [Cortes 1995], in which vectors of two classes are labelled and used as the training set. C SVM finds the optimal hyperplane, which separates the two classes in the training set, and uses this for new data classification. Figure 2 illustrates the principle of C SVM.

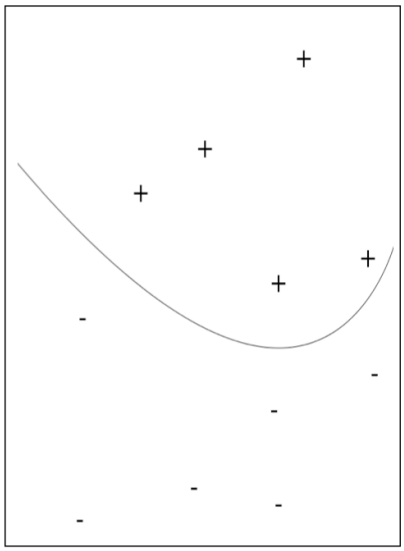


Figure 2: Optimal hyperplane separating the vectors of the training set, which consists of two classes (“+”,”-”).

For non-linearly separable cases, the data are transformed in higher dimension feature space F using a kernel function.

In this thesis, a new algorithm based on C SVM for prostate segmentation in the whole-body MRI context was developed [Habes 2013b]. In this algorithm, the segmentation task is considered a binary voxel classification problem, where every voxel is assigned to either prostate tissue or background. The classification vectors in this algorithm are generated from 3D neighboring voxels using a “plus”-symbol-like structure and automatically generated features for prostate description. The Gaussian Kernel was used for mapping in higher feature space. The features used for appropriate prostate description are: median, gradient, anisotropic diffusion and eigenvalues of the structure tensor. Pre-isolation of urological organs – to reduce the number of voxels to be classified – was achieved using a maximal-entropy thresholding step. This developed method is designed for epidemiological studies, since the algorithm does not require any human interaction for automated prostate volume delineation or volume estimation, which reduces human error and ensures reproducibility. In terms of epidemiological studies, full automation is the first requirement for the developed algorithm, although more accurate methods are available (see results section).

2.4.2 Method 2: S SVM for clinical studies

Schölkopf et al. introduced the single-class Support Vector Machines (S SVM) [Schloelkopf 2001a, Schloelkopf 2001b] for classification using a training set consisting of just one class as the training sample. Figure 3 illustrates the principle of S SVM. Mapping the training set in the higher dimensional space F is achieved via a kernel function. An optimal subset of F must be then sought to separate the training examples from the origin of F. In this thesis project, a new algorithm based on S SVM for prostate volume estimation in a whole-body MRI context was developed [Habes 2013c]. The prostate volume estimation is considered 3D object reconstruction by the S SVM classification technique. Using manually seeded semi-landmarks on the contour of the prostate, the S SVM is capable of whole-prostate reconstruction. Every landmark is a vector in the image coordinates. Mapping in the feature space was established using the Gaussian kernel. This newly developed method is designed for clinical studies, since the resulting PV is more accurate than both method 1 and the ellipsoid formula (see results section). In terms of clinical studies, accuracy is the first requirement for the developed algorithm, although human interaction in some steps is required.

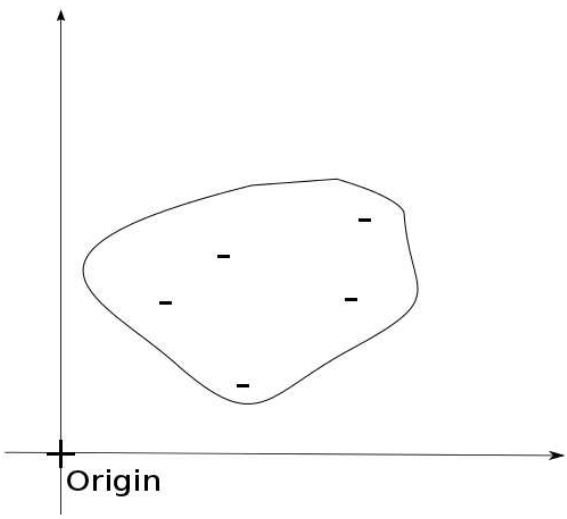


Figure 3: Separating the single training sample ("-") from the origin ("+") of the feature space is similar to the binary case, in which the second class consists of just one vector (the origin).

MRI provided an opportunity for validation of the systematic underestimation of the ellipsoid shape assumption, which is widespread in clinical PV assessment using the TRUS modality. PV based on the ellipsoid formula $V = H \times W \times L \times \pi/6$, where H is the height, W the width, and L the length of the prostate in the 3-D MR image, was calculated. It was then compared to S SVM-based PV [Habes 2013c]. The manual prostate delineation was used here as the reference (the closest to real PV). The dimensions H and W were set on the axial view of the prostate central slice. L was set on the coronal view.

2.5 Evaluation Strategy

For a complete overview of the statistical analysis and the evaluation strategy used for both methods developed here, see Habes 2013b and 2013c. All statistical analyses were calculated using the open-source software for statistical computing R, version 2.15.2. Differences were considered to be statistically significant at a level of $p<0.05$.

Using Bland-Altman plots, the mean difference (MD) and the limits of agreement within 95% confidence intervals (CI) were calculated [Bland 1986]. Furthermore, the strength of the association between automated and reference PV was calculated using Spearman's rank coefficients. The results of Bland-Altman plots and Spearman's rank correlation coefficient analysis compared with manual readings of observer A and B are listed in Table 1 for method 1.

3 Results

For a complete overview of results, see the results sections in Habes 2013b and 2013c.

Evaluation strategy	Observer A	Observer B
Spearman's rank (ρ)	0.936*	0.859*
Bland-Altman Plot	MD = 3 CI: −3.1 to 9.2	MD = 1.9 CI: −7.1 to 10.8

Table 1: Results evaluation of PV estimated using method 1 vs. manual volume estimation of two human experts (observers A and B).
* Significant differences are indicated: $p < 0.05$.
MD: Mean Difference
CI: Limits of agreement within 95% confidence intervals

Evaluation strategy	Method 2	Ellipsoid Formula
Spearman's rank (ρ)	0.965*	0.873*
Bland-Altman Plot	MD = −0.05 CI: −3.8 to 3.7	MD = 8.6 CI: 1 to 16.2

Table 2: Results evaluation of PV estimated using method 2 and the ellipsoid formula method against manual volume estimation of one human expert.
* Significant differences are indicated: $p < 0.05$.
MD: Mean Difference
CI: Limits of agreement within 95% confidence intervals

The results of Bland-Altman plots and Spearman's rank correlation coefficients analysis are listed in Table 1 for method 1 and Table 2 for method 2 with respect to the corresponding reference.

4 Discussion

4.1 Summary

The quantification of prostate enlargement is a significant clinical challenge. WB-MRI provides an innovative option, which was obtained in this thesis. Using kernel-based methods, the first aim of this thesis research was to fully automate prostate segmentation and volume estimation based on WB-MRI in the context of a population-based epidemiological study. This method reduces the need for human resources in image readings and thus allows processing a large number of MR scans with a reasonable amount time and effort. The second aim was to develop a new, accurate, and reliable method without any prostate shape assumption. Although this method requires involvement of an expert, it provides accurate PV estimation. This method was additionally compared to the ellipsoid shape assumption for the prostate [Habes 2013c]. The quantitative evaluation of the prostate shape assumption – similar to the shape assumption used as the standard in clinical diagnostics with the TRUS modality – in the sample of this thesis yielded systematically lower PV compared to method 2 [Habes 2013c] and manual prostate delineation by the expert (8.6 ml underestimation with the ellipsoid formula compared to 0.05 ml overestimation with method 2). A similar comparison was not provided for method 1, since the ellipsoid formula plays no role in epidemiological studies. This cumulative thesis is based on two peer-reviewed journal publications [Habes 2013b, Habes 2013c], which are attached in the Appendix and cover both aims separately.

For an overview of current published research on automated prostate volume estimation based on MR and TRUS images, see Habes 2013b. A comparison between S SVM and C SVM for the segmentation of anatomical structures was provided in a previous study [Habes 2013a] using different kernels for the segmentation. However, the aim of this study was not to directly compare C SVM and S SVM, but rather to develop dedicated methods for specific tasks in epidemiological and clinical studies.

4.2 Limitations

One limitation of both studies is the lack of direct comparison between MR- and TRUS- (the contemporary clinical gold standard) based PV and a validation against prostatectomy specimen-based PV. Both were obviously not possible in the context of the population-based SHIP cohort protocol. Non-invasive methods only could be integrated in the protocol for all participants for ethical reasons in the context of a large prospective cohort study. Furthermore, prostatectomy is performed only in cases of pathologically proven prostatic cancer. In a recent study, Turkbey et al. showed that MR-based PV is smaller than the real prostatectomy specimen-based PV. They revealed a strong positive correlation between prostatectomy specimen-based PV and that derived from manual segmentation of MR scans ($R = 0.89–0.91$, $p < 0.0001$) [Turkbey 2013]. It must be mentioned that prostatectomy specimens in their experiment included the seminal vesicles and variable amounts of adjacent tissue; this may be a systematic reason for underestimation by MRI, as the authors reported. To date, the literature contains no direct comparison of MR-based PV with subsequent prostatectomy specimens without seminal vesicles and without shape assumption. Turkbey et al. also found that the ellipsoid formula could underestimate the PV, which confirms results reported in this thesis.

The literature also contains relatively little data on comparisons between MR- and TRUS-based PV. Al-Rimawi et al. reported a strong correlation between TRUS- and MR-based PV [Al-Rimawi 1994]. Weiss et al. reported similar correlations between TRUS- and endorectal MR-based PV [Weiss 2012]. Both authors, however, used the ellipsoid formula for corresponding PV estimation.

4.3 Outlook

Different steps must still be achieved in order to validate, improve and extend the methods developed during this thesis research. In the following paragraph, some core research plans are summarized:

4.3.1 Direct MR-based prostate volume validation against clinical gold standards

Our research group has currently developed a study protocol to directly compare PV assessment accuracy between TRUS and MRI without shape assumption and prostatctomy specimens. Recently, our research group has received approval from the ethics committee of the University of Greifswald for this analysis to recruit a clinical cohort.

4.3.2 Further optimization of the developed methods

Method 1:

One improvement will be to reduce the time required (currently an average of 10 min). Different strategies could help reduce the necessary generated features for prostate discrimination, e.g., integrating further sequences such as the Dixon sequence from the SHIP protocol used for the abdominal area and with lower resolution compared to the pelvis sequence. In the Dixon sequence, discriminating between the bladder and prostate tissue based on intensity information is possible using the resulting components. Integrating the Dixon sequence in the segmentation pipeline would first require rigid co-registration with the pelvis sequence.

Method 2:

One improvement for this method will be to make it completely automated. Here, the semi-landmarks will be found automatically. From the expert manual segmentations of the prostate, an atlas can be generated. Non-rigid registration of the atlas to a new scan could deliver a prostate probability (p) map. The highest probability voxels could be used for prostate reconstruction with S SVM. Receiver operating curve (ROC) analysis could facilitate finding the optimal p value. It is important to note that the non-rigid registration requires considerable calculation time; therefore, detailed research on the suitable algorithms needs to be carried out.

4.3.3 Prostate sub-region segmentation

McNeal defined the prostate histologically [McNeal 1968, McNeal 1981]. He divided the prostate in four regions: the peripheral zone (PZ, over 70% of the prostate gland), the

central zone (CZ, 25% of the prostate gland), the preprostatic region, which contains the transition zone (TZ), and the anterior fibro-muscular stroma. This histological differentiation between the zones has been limited to biological differences.

The prostatic anatomy is distinguishable using the MRI modality, but the signal intensities of the CZ and the TZ are similar [Hricak 1987, Verma 2011]. Therefore, they are defined in the radiological literature as the central gland (CG) [Verma 2011]. Figure 4 illustrates the zonal anatomy of a 60-year-old SHIP participant.

It is widely accepted that the TZ is the site of origin of BPH [McNeal 1981]. BPH prevalence reaches 50% among 50-year-old men [Espinosa 2013]. Corinca et al. observed a positive correlation between aging and CZ volume [Corica 1999]. A recent study by Turkbey et al. using MRI modality confirmed that CG volume is associated with age and with changes in lower urinary tract symptoms, while PZ volume had no correlation with age [Turkbey 2012]. Zhang et al. confirmed the positive association between TZ volume and aging using the TRUS modality [Zhang 2013]. Hence, the association between prostatic zone volumes and further epidemiological parameters remain an open field of research.

On the other hand, the anatomical regions of the prostate are associated with prostate cancer. It is widely believed in the urological community that the site of origin of most prostatic carcinomas is located in the PZ [McNeal 1981]. Reissigl et al. confirmed that most prostatic cancer originated in the PZ (67% of the patients) and only 28% in the TZ [Reissigl 1997]. Erbersdobler et al. reported that TZ carcinomas are less malignant than PZ carcinomas [Erbersdobler 2002]. Newton et al. revealed that smaller prostate size could predict high-grade prostate cancer [Newton 2010]. To date, the literature contains nothing on the association between volumes of prostate anatomy regions and carcinoma.

Zlotta et al. concluded that TZ volume measurements based on TRUS modality are more accurate than PV assessment using the prolate ellipsoid method [Zlotta 1999], which could make the TZ volume a more reliable clinical biomarker. A similar study on an MR basis would be helpful to further evaluate this possibility. Automation of prostate zonal anatomy segmentation in SHIP will allow studying further associations with clinically relevant epidemiological parameters beyond aging (e.g., smoking).

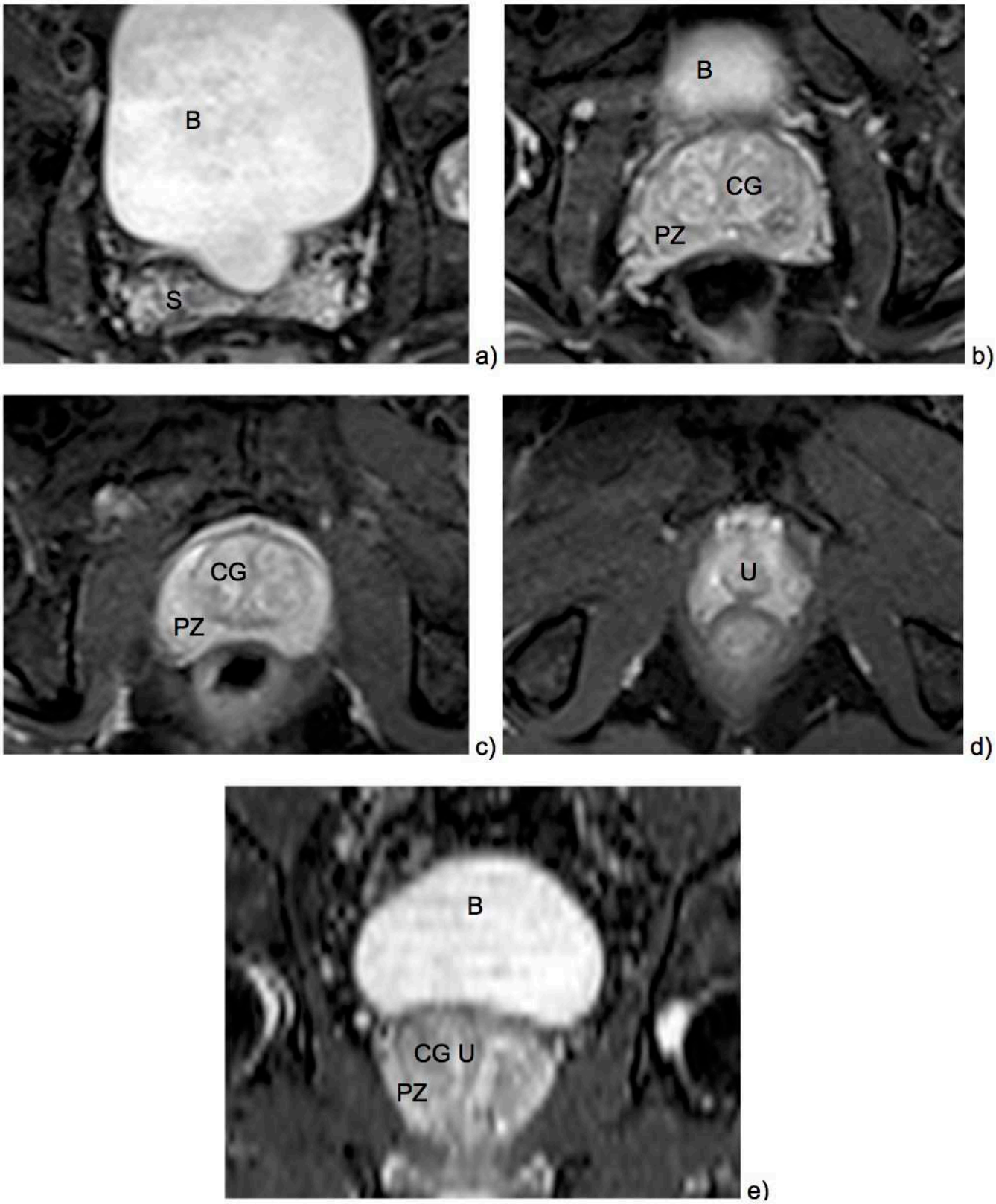


Figure 4: Zonal anatomy of a 60-year-old SHIP participant with prostatic enlargement (PV = 51.3 ml). Axial images (a-d) and coronal image e) represent zonal anatomy: S = seminal vesicle, B = bladder, CG = central gland (central zone and transition zone), PZ = peripheral zone, U = urethra

Finally, prediction models at the prostatic region level, similar to those derived for PV estimation based on serum PSA values are a promising research field. A method combining S SVM and C SVM in the same framework in which prostatic contour is detected could allow more accurate prostatic sub-region segmentation for clinical purposes.

5 Conclusion

The increasing epidemiological need for fully automated methods of prostate volume (PV) estimation in WBI scans requires algorithms with high segmentation quality and without human interaction, which reduces human error and ensures reproducibility. The binary Support Vector Machines (C SVM)-based method developed in this thesis showed PV estimation accuracy comparable to human experts in MR reading and is suitable for integration in epidemiological studies.

In the clinical context, accuracy is the first requirement for any computerized method, although human interaction in some steps is necessary. The widespread ellipsoid formula used in clinical diagnostics showed systematic underestimation of PV in MRI in this thesis. Alternatively, the single-class Support Vector machines (S SVM)-based method agreed excellently with the reference PV. The promising results with respect to accuracy indicate considerable potential for clinical application. Based on the results of this thesis, it is recommended to increase the use of accurate computerized methods in clinics for PV estimation based on MR or TRUS.

Abstract

Benign prostatic hyperplasia (BPH) is one of the most widespread diseases among men older than 50 years. The literature provides various cut-off values for pathological enlargement of the prostate. Prostate volume (PV) measurement in large population-based studies would allow deriving more objective reference values and a more valid early BPH diagnosis. A fully automated method is therefore required. In the clinical context, the measurement of the PV is important for treatment response monitoring in the clinical applications for BPH management research, and an accurate method for PV is essential.

Magnetic Resonance Imaging was used for PV estimation. Two methods based on the Support Vector Machines (SVM) were developed: the binary Support Vector Machines (C SVM)-based method for epidemiological studies and the single-class Support Vector Machines (S SVM)-based method for clinical studies. The second method was additionally compared to the ellipsoid formula for PV estimation, which is widespread in the clinic.

The comparison between volume measurement of the C SVM-based method and manual delineation of observers A and B yielded a strong correlation (Spearman's rank correlation coefficients ρ of 0.936 [$p < 0.001$] and 0.859 [$p < 0.001$], respectively). Comparing the C SVM-based method and the two manual delineations by observers A and B shows an agreement with a mean difference of 3.0 ml (95% confidence interval of -3.1 to +9.2 ml) and 1.9 ml (95% confidence interval of −7.1 to +10.8 ml), respectively.

The S SVM-based method and the reference PV (manual delineation of observer A) show excellent correlation (Spearman’s rank correlation coefficient $\rho = 0.965$, $p < 0.001$), while the ellipsoid formula is less well correlated with the reference PV (Spearman’s rank correlation coefficient $\rho = 0.873$, $p < 0.001$). The mean difference between S SVM and the reference PV was −0.05 ml (95% confidence interval of −3.8 to +3.7 ml); on the other hand, the mean difference between the ellipsoid formula and the reference PV was much greater, with 8.6 ml (95% confidence interval of +1 to +16.2 ml).

The C SVM-based method has considerable potential for integration in epidemiological studies. The prostate volumes obtained by the S SVM-based method agreed excellently with the reference and would be clinically useful for urologists in prostate volumetric analysis.

Zusammenfassung

Die benigne Prostatahyperplasie (BPH) ist eine der am meisten verbreiteten Erkrankungen bei Männern, die älter als 50 Jahre sind. Für diese Vergrößerung der Prostata sind unterschiedliche Volumen-Schwellenwerte aus der Literatur bekannt. Eine Messung des Prostatavolumens (PV) sollte daher in einer populationsbasierten Studie objektive Referenzwerte liefern und eine frühzeitige valide BPH-Diagnostik ermöglichen. Dafür ist eine genaue vollautomatische Methode zur Volumenbestimmung notwendig. Im klinischen Kontext spielt die PV-Messung eine große Rolle als Behandlungsüberwachungsparameter in der klinischen Anwendung sowie für die Forschung im BPH-Management. Eine genaue Methode für die PV-Messung ist hier essentiell.

Für die PV-Messung wurden MRT-Bilddaten des Prostata-Areals verwendet. Zwei „*Support Vector Machines*“-basierte Methoden wurden für die Volumenbestimmung entwickelt: die „*binary Support Vektor Machines*“ (C SVM)-basierte Methode für epidemiologische Studien und die „*Single Class Support Vector Machines*“ (S SVM)-basierte Methode für klinische Studien. Die zweite Methode wurde mit der klinisch weitverbreiteten Ellipsen-Formel für PV-Messung verglichen.

Der Vergleich zwischen der PV-Messung der C SVM-basierten Methode und der manuellen Abgrenzung der Prostata durch zwei Beobachter zeigt eine starke Korrelation (Spearman's Korrelationskoeffizient $\rho = 0.936$, $p < 0.001$ und $= 0.859$, $p < 0.001$). Der Bland-Altman-Plot zeigt eine Übereinstimmung mit dem Mittelwert der Differenzen von 3.0 ml (95 % Konfidenzintervall: -3.1 bis +9.2 ml) und 1.9 ml (95 % Konfidenzintervall: −7.1 bis +10.8 ml).

Die S SVM-basierte Methode besitzt eine exzellente Korrelation mit der Referenz-PV –manuelle Abgrenzung vom Beobachter– (Korrelationskoeffizient $\rho = 0.965$, $p < 0.001$), während das bekannte Ellipsen-Modell eine deutlich geringere Korrelation mit der Referenz-PV (Korrelationskoeffizient $\rho = 0.873$, $p < 0.001$) aufweist. Der Mittelwert der Differenzen zwischen der S SVM-basierten Methode und der Referenz-PV war −0.05 ml (95 % Konfidenzintervall: -3.8 bis +3.7 ml). Der Mittelwert der Differenzen zwischen der Ellipsen Formel und der Referenz-PV war mit 8.6 ml deutlich höher (95 % Konfidenzintervall: +1 bis +16.2 ml).

Die C SVM-basierte Methode hat somit ein deutliches Potential für die Integration in epidemiologischen Studien zur BPH. Die S SVM-basierten Prostatavolumen stimmen mit der Referenz exzellent überein. Die Methode ist daher nützlich für klinisch-urologische Prostata-volumetrische Analysen.

Abbreviations

Benign prostatic hyperplasia	BPH
Prostate volume	PV
Transrectal ultrasound	TRUS
The Study of Health in Pomerania	SHIP
Whole-body Magnetic Resonance Imaging	WB-MRI
Binary-class Support Vector Machines	C SVM
Single-class Support Vector Machines	S SVM
Mean difference	MD
Confidence intervals	CI
The peripheral zone	PZ
The central zone	CZ
The transition zone	TZ
The central gland	CG

References

[Abrams 1994] Abrams, P. "New words for old: lower urinary tract symptoms for" prostatism." BMJ: British Medical Journal 308, no. 6934 (1994): 929.

[Al-Rimawi 1994] Al-Rimawi M., Griffiths D.J., Boake R.C., Boake, D., and Johnson, M. "Transseptal ultrasound versus magnetic resonance imaging in the estimation of prostatoc volume." BJU Int 74 no 5 (1994) :596–600

[Boyle 1996] Boyle, P., Lawrence Gould, A., and Roehrborn, C. G. "Prostate volume predicts outcome of treatment of benign prostatic hyperplasia with finasteride: meta-analysis of randomized clinical trials." Urology 48, no. 3 (1996): 398-405.

[Cahill 1995] Cahill D. R., Orland M. J. and Miller G. M. (1995) “Atlas of Human Cross-Sectional Anatomy: With CT and MR Images” John Wiley and sons.

[Cortes 1995] Cortes, C. and Vapnik, V. Support-Vector Networks: Machine Learning.1995: 273-297

[Corica 1999] Corica, F.A., Jacobsen, S.J., King, B.F., David G., Bostwick, D. J., Girman, C. J., and Lieber, M. M. "Prostatic central zone volume, lower urinary tract symptom severity and peak urinary flow rates in community dwelling men." The Journal of urology 161, no. 3 (1999): 831-834.

[Di Silverio 2005] Di Silverio, F, Bosman, C, Salvatori, M, Albanesi, L., Proietti Pannunzi, L., Ciccariello, M., Cardi, A., Salvatori, G., and Sciarra, A.. "Combination therapy with rofecoxib and finasteride in the treatment of men with lower urinary tract symptoms (LUTS) and benign prostatic hyperplasia (BPH)." European urology 47, no. 1 (2005): 72-79.

[Ditonno 2005] Ditonno, P, Battaglia, M, Selvaggio, O, Garofalo, L., Lorusso, V., and Selvaggi, F. P. "Clinical evidence supporting the role of lonidamine for the treatment of BPH." Reviews in Urology 7, no. Suppl 7 (2005): S27.

[Espinosa 2013] Espinosa, G, Esposito, R, Kazzazi, A, and Djavan, B. "Vitamin D and benign prostatic hyperplasia-a review." The Canadian journal of urology 20, no. 4 (2013): 6820-6825.

[Erbersdobler 2002] Erbersdobler, A., Fritz, H., Schnöger, S., Graefen, M., Hammerer, P., Huland, H., and Henke, R. "Tumour grade, proliferation, apoptosis, microvessel density, p53, and bcl-2 in prostate cancers: differences between tumours located in the transition zone and in the peripheral zone." European urology 41, no. 1 (2002): 40-46.

[Habes 2013a] Habes, M., Rota Kops, E., Bahr, J., Kühn, J. P., Hoffmann, W., Lipinski, H. G., and Herzog, H. "Automated Skull and Cavity Segmentation from Ultra Short TE Sequence Images." Current Medical Imaging Reviews 9, no. 2 (2013): 120-128.

[Habes 2013b] Habes, M., Schiller, T., Rosenberg, C., Burchardt, M., and Hoffmann, W. "Automated prostate segmentation in whole-body MRI scans for epidemiological studies. " Phys. Med. Biol. 58 no 17 (2013): 5899-5916.

[Habes 2013c] Habes, M., Bahr, J., Schiller, T., Kühn, J. P., Hoppe, L., Burchardt, M., and Hoffmann, W. "New technique for prostate volume assessment." World journal of urology (2013): (accepted, in press).

REFERENCES

[Hegenscheid 2009] Hegenscheid, K., Kühn, J.P., Völzke, H., Biffar, R., Hosten, N., and Puls, R. Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study Rofo 181 no 8 (2009): 748-759

[Hald 1989] Hald, T. "Urodynamics in benign prostatic hyperplasia: a survey." The prostate 15, no. S2 (1989): 69-77.

[Herbert 2004] Herbert, L. "Evaluating men with benign prostatic hyperplasia." Reviews in Urology 6, no. Suppl 1 (2004): S8.

[Hricak 1987] Hricak, H., Dooms, G.C., McNeal, J.E., Mark, A. S., Marotti, M., Avallone, A., Pelzer, M., Proctor, E. C., and Tanagho, E. A. "MR imaging of the prostate gland: normal anatomy." American journal of roentgenology 148, no. 1 (1987): 51-58.

[Kim 2011] Kim, JM, Song, PH, Kim, HT and Moon, K. H. “Effect of obesity on prostate-specific antigen, prostate volume, and international prostate symptom score in patients with benign prostatic hyperplasia.” Korean J Urol.; 52 no. 6 (2011): 401-405

[Martin Bland 1986] Martin Bland, J. and Altman, D. Statical Methods for Assessing agreement between two methods of clinical measurement: Lancet, 327 no. 8476 (1986): 307-310

[McNeal 1968] McNeal, J.E. "Regional morphology and pathology of the prostate." American journal of clinical pathology 49, no. 3 (1968): 347.

[McNeal 1981] McNeal, J.E. "The zonal anatomy of the prostate." The Prostate 2, no. 1 (1981): 35-49.

[Milonas 2003] Milonas, D., Trumbeckas, D. and Juska, P. The importance of prostatic measuring by transrectal ultrasound in surgical management of patients with clinically benign prostatic hyperplasia Medicina (Kaunas). 2003; 39(9): 860-866

[Morote 2000] Morote J., Encabo G., Lopez M., and de Torres, I. M. “Prediction of prostate volume based on total and free serum prostate–specific antigen: is it reliable?” Eur Urol 38 no. 1 (2000):91–95

[Nathan 1996] Nathan, M. S., Seenivasagam, K., Mei, Q., Wickham, J. E. A., and Miller, R. A. "Transrectal ultrasonography: why are estimates of prostate volume and dimension so inaccurate?." British Journal of Urology 77, no. 3 (1996): 401-407.

[Newton 2010] Newton, M. R., Phillips, S., Chang, S. S., Clark, P. E., Cookson, M. S., Davis, R., Fowke, J. H., Herrell, S. D., Baumgartner, R., Chan, R., Mishra, V., Blume, J. D., Smith Jr., J. A., and Barocas, D. A. "Smaller prostate size predicts high grade prostate cancer at final pathology." The Journal of urology 184, no. 3 (2010): 930-937.

[Oelke 2007] Oelke, M., Höfner, K., Jonas, U, Laval, K. U., and Tunn, U. "Terminologie und Diagnostik des benignen Prostatasyndroms." Dtsch Arztebl 104, no. 33 (2007): 2261-7.

[Park 2013] Park T., Chae J. Y., Kim J. W. Kim, J. W., Oh, M. M., Yoon, C. Y., and Moon, D. G. “Prostate-specific antigen mass and free prostate-specific antigen mass for predicting the prostate volume of Korean men with biopsy-proven benign prostatic hyperplasia.” Korean J Urol 54 no. 9 (2013):609–614

[Reissigl 1997] Reissigl, A., Pointner, J., Strasser, H., Ennemoser, O., Klocker, H., and Bartsch, G. "Frequency and clinical significance of transition zone cancer in prostate cancer screening." The Prostate 30, no. 2 (1997): 130-135.

[Rodriguez 2008] Rodriguez JE, Skarecky, D, Narula, N, and Ahlering, T. E. "Prostate volume estimation using the ellipsoid formula consistently underestimates actual gland size." The Journal of urology 179, no. 2 (2008): 501-503.

REFERENCES

[Roehrborn 1999] Roehrborn C.G., Boyle P., Gould A.L. and Waldstreicher, J. “Serum prostate-specific antigen as a predictor of prostate volume in men with benign prostatic hyperplasia.” Urology 53 no. 3 (1999):581–589

[Schneider 2012] Schneider, CA, Rasband, WS and Eliceiri, KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods, 2012; 9(7), 671-675.

[Schoelkopf 2001a] Schölkopf, B and Smola, AJ. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.MIT Press, 2001

[Schoelkopf 2001b] Schölkopf B, Platt JC, Shawe-Taylor J, Smola, A. J., and Williamson, R. C.. “Estimating the Support of a High-Dimensional Distribution” Neural Comput. 13 no. 7 (2001):1443–1471.

[Stamey 1987] Stamey, T.A., Yang, N., Hay, A.R., McNeal, J. E., Freiha, F. S., and Redwine, E. “Prostate- specific antigen as a serum marker for adenocarcinoma of the prostate” N Engl J Med 317 no. 15 (1987): 909-916

[Tanguay 2009] Tanguay, S., Awde, M., Brock, G., Casey, R., Kozak, J., Lee, J., Nickel, J. C., and Saad, F., "Diagnosis and management of benign prostatic hyperplasia in primary care." Canadian Urological Association Journal 3, no. 3 Suppl 2 (2009): S92.

[Tingelhoff 2008] Tingelhoff K, Eichhorn K W, Wagner I, Kunkel M E, Moral A I, Rilk M E, Wahl F M and Bootz F “Analysis of manual segmentation in paranasal CT images” Eur. Arch. Otorhinolaryngol. 265 (2008): 1061–70

[Turkbey 2012] Turkbey, B., Huang, R., Vourganti, S., Trivedi, H., Bernardo, M., Yan, P., Benjamin, C., Pinto, P. A., and Choyke, P. L. “Age related changes in prostate zonal volumes as measured by high resolution magnetic resonance imaging (MRI): a cross sectional study in over 500 patients.” BJU Int.110 no. 11 (2012): 1642-1647

[Turkbey 2013] Turkbey, B., Sergei V., Fotin, R.J., Yin, Y., Daar, D., Aras, O., Bernardo, M., Garvey, B. E., Weaver, J., Haldankar, H., Muradyan, N., Merino, M. J., Pinto, P. A., Periaswamy, S., and Choyke, P. L. "Fully Automated Prostate Segmentation on MRI: Comparison With Manual Segmentation Methods and Specimen Volumes." American Journal of Roentgenology 201, no. 5 (2013): W720-W729.

[Verma 2011] Verma, S, and Arumugam R. "A clinically relevant approach to imaging prostate cancer: review." American Journal of Roentgenology 196, no. 3_supplement (2011): S1-S10.

[Vesely 2003] Vesely, S., Knutson, T., Damber, J.E., Dicuio, M., and Dahlstrand, C. “Relationship between age, prostate volume, prostate-specific antigen, symptom score and uroflowmetry in men with lower urinary tract symptoms” Scand J Urol Nephrol. 37 no. 4 (2003): 322-328

[Voelzke 2011] Völzke, H, Alte, D, Schmidt, CO, Radke, D., Lorbeer, R., Friedrich, N., Aumann, N., Lau, K., Piontek, M., Born, G., Havemann, C., Ittermann, T., Schipf, S., Haring, R., Baumeister, S. E., Wallaschofski, H., Nauck, M., Frick, S., Arnold, A., Jünger, M., Mayerle, J., Kraft, M., Lerch, M. M., Dörr, M., Reffelmann, T., Empen, K., Felix, S. B., Obst, A., Koch, B., Gläser, S., Ewert, R., Fietze, I., Penzel, T., Dören, M., Rathmann, W., Haerting, J., Hannemann, M., Röpcke, J., Schminke, U., Jürgens, C., Tost, F., Rettig, R., Kors, J. A., Ungerer, S., Hegenscheid, K., Kühn, J. P. ,Kühn, J., Hosten, N., Puls, R., Henke, J., Gloger, O., Teumer, A., Homuth, G., Völker, U., Schwahn, C., Holtfreter, B., Polzer, I., Kohlmann, T., Grabe, H. J., Rosskopf, D., Kroemer, H. K., Kocher, T. Biffar, R., John, U. and Hoffmann, W.
Cohort profile: the study of health in Pomerania Int J Epidemiol. 40 no. 2 (2011): 294-307

[Weiss 2012] Weiss B.E., Wein A.J., Malkowicz S.B. and Guzzo, T. J. "Comparison of prostate volume measured by transrectal ultrasound and magnetic resonance imaging: is transrectal ultrasound suitable to determine which patients should undergo active surveillance? " Urol Oncol 31 no. 8 (2013):1436–1440

REFERENCES

[Zhang 2012] Zhang, S. J, Qian, H. N., Zhao, Y., Sun, K., Wang, H. Q., Liang, G. Q., Li, F. H. and Li, Z. "Relationship between age and prostate size." Asian journal of andrology. 15 no. 1 (2012): 116-120.

[Zlotta 1999] Zlotta, A. R., Djavan, B., Damoun, M., Roumeguere, T., Petein, M., Entezari, K., Marberger, M. and Schulman, C. C. "The importance of measuring the prostatic transition zone: an anatomical and radiological study." BJU international 84 (1999): 661-666.

Appendix I: Automated prostate segmentation in whole-body MRI scans for epidemiological studies

Appandix I is based on:

Mohamad Habes, Thilo Schiller, Christian Rosenberg, Martin Burchardt and Wolfgang Hoffmann, "Automated prostate segmentation in whole-body MRI scans for epidemiological studies. " Phys. Med. Biol. 2013; 58 : 5899-5916, DOI:10.1088/0031-9155/58/17/5899

The final version can be downloaded from:

http://iopscience.iop.org/0031-9155/58/17/5899/article#

I Abstract

The whole prostatic volume (PV) is an important indicator for benign prostate hyperplasia. Correlating the PV with other clinical parameters in a population-based prospective cohort study (SHIP-2) requires valid prostate segmentation in a large number of whole-body MRI scans. The axial proton density fast spin echo fat saturated sequence is used for prostate screening in SHIP-2. Our automated segmentation method is based on support vector machines (SVM). We used three-dimensional neighborhood information to build classification vectors from automatically generated features and randomly selected 16 MR examinations for validation. The Hausdorff distance reached a mean value of 5.048 ± 2.413, and a mean value of 5.613 ± 2.897 compared to manual segmentation by observers A and B. The comparison between volume measurement of SVM-based segmentation and manual segmentation of observers A and B depicts a strong correlation resulting in Spearman's rank correlation coefficients (ρ) of 0.936 and 0.859, respectively. Our automated methodology based on SVM for prostate segmentation can segment the prostate in WBI scans with good segmentation quality and has considerable potential for integration in epidemiological studies.

I. 1. Introduction

Benign prostatic hyperplasia (BPH) is the leading cause of enlargement of the prostate. Clinical indicators for BPH are prostate volume (PV), prostate-specific antigen (PSA) levels and maximum flow rate on uroflowmetry. The radiological indicator for BPH is PV, which correlates with PSA serum values (Stamey *et al* 1987, Vesely *et al* 2003). To date, the gold standard to assess PV is trans-rectal ultrasound (TRUS) (Bates *et al* 1996). MRI provides an alternative for non-invasive prostate visualization. In addition to a strong prostatic enlargement, BPH can cause form variations, which means that strict form assumptions, such as the ellipsoid formula widely used in combination with TRUS, can be inaccurate for PV quantification.

The study of health in Pomerania (SHIP) is a population-based epidemiological study of adults in the region of western Pomerania in northeastern Germany (Völzke *et al* 2011, Hegenscheid *et al* 2009). The focus in SHIP is utilizing non-invasive methods to investigate common risk factors, preclinical disease states and manifest diseases. In SHIP-2, a standardized imaging protocol is applied to all participants. In the context of a large population study, it is important to develop a method for automatic prostate segmentation. A fully automatic segmentation method is important to reduce the influence of inter-observer variation and to improve reproducibility. In epidemiological applications, it is important to develop a method with minimal human interaction. A variety of strategies has been previously suggested for prostate segmentation in CT images. Mazonakis proposed an algorithm based on region growing (Mazonakis *et al* 2001). Region-growing based methods are at a disadvantage, since they heavily depend on image quality. Furthermore, Mazonakis' approach is still semi-automated, i.e., it requires manual setting of three seed points. Lee and Chung introduced a fuzzy-inference-based radial basis function (RBF) for the segmentation task in CT images (Lee and Chung 2004). This approach, like other neural networks, requires costly system training, which can limit its wider practical application. An automated segmentation method was proposed by Chen *et al* (2011). The authors implemented a segmentation cost function based on a Bayesian framework with the incorporation of anatomical constraints. However, for the prostate, the reliability of anatomical constraints can be limited, since the position of the prostate changes due to dynamic filling situations of the bladder or rectum. There are also

approaches for CT images which depend upon large deformation 3D registration, such as in (Foskey *et al* 2005) for prostate segmentation. Conceptually, registration-based approaches are very sensitive to form or intensity variations, which may limit their usefulness (Costa *et al* 2007). Recently, the use of MRI for prostate visualization and related clinical diagnosis has become commonplace, a diagnostic trend which has driven the need to develop new techniques for this type of images. An atlas-based approach for prostate segmentation was presented by Klein *et al* (2008). That method does not distinguish between the prostate tissues and the adjacent seminal vesicles, which we however consider an essential requirement for clinical prostate research. Klein *et al* 's method also includes an intensity-based registration step, which can fail in cases of intensity inhomogeneities or strong form variation. Martin *et al* described a spatially constrained deformable model for segmentation (Martin *et al* 2010). Deformable models usually require a large amount of prior information for an adequate training process. Including all of the possible prostate shape variations would make the human interaction step very time consuming and thus costly. Deformable models are also known to fail in situations with inordinate noise or poor image resolution, as may be the case with WBI using an MR sequence protocol that is not selectively optimized for prostate imaging. Again it is important to note that with BPH, the prostate can undergo strong shape variation. Pasquier *et al* (2007) introduced a model-based method for prostate contour delineation in T2-weighted MRI images, which does not require the costly training step (Pasquier *et al* 2007). However, this is a semi-automatic approach which requires user interaction to initialize the model position from a selected target in the MRI volume. As described in the standard SHIP examination protocol, Hegenscheid *et al* used the axial proton density fast spin echo fat saturated (Ax PD TSE FS) sequence, which showed the best performance in discriminating between structures in the pelvic region (Hegenscheid *et al* 2009).

No work has been published to date on segmenting the prostate in WBI scans or in the Ax PD TSE FS sequence, either with a semi- or a fully automatic method. In the Ax PD TSE FS sequence, it is still unclear how well prostatic tissues can be distinguished from the surrounding vessels and seminal vesicles. The prostate in this sequence cannot be considered a continuous organ, since it is occasionally barely distinguishable from the seminal vesicles. Therefore, there is a need for a new segmentation method that can be

utilized in the framework of WBI scans. In the literature, kernel-based methods have not yet been used for whole deformable organ segmentation. Our choice is motivated by the fact that SVM needs only a small number of vectors to train the system effectively. In this paper, we propose an SVM-based algorithm that allows automated detection of the prostate in Ax PD TSE FS MRI data, and suggest suitable features for the accurate assessment of the prostate. Automated image analysis results are compared with the results of manual assessments by two urological radiology experts.

I. 2. Materials and methods

The MRI volumes used in this study were acquired with a 1.5-T MR imager (MAGNETOM Avanto; Siemens Medical Systems, Erlangen, Germany). All subjects in this study were among the participants of the SHIP (Völzke et al 2011, Hegenscheid et al 2009).

Participants were informed in detail about the study by trained personnel (radiology assistants and radiologists). Oral and written informed consent was obtained from all participants. The ethics committee of University of Greifswald approved the study protocol. The participants were placed in the MRI scanner in supine position. Five stations were used to create the WBI image (head, neck, abdomen, pelvis and lower extremities). The reconstruction of the WBI volume can be done as a post-processing step (Rummeny et al 2006). The pelvic region image volumes were used for the prostate segmentation.

We used the Ax PD TSE FS sequence. Typical parameters of this sequence are: TR 3230 ms, TE 34 ms, flip angle 90°, voxel size (0.8 × 0.8× 3 mm: 0.9 mm gap), 1 average used for pelvis imaging. The bore size was 60 cm, horizontal. An SQ engine gradient system with a maximum gradient field strength of 45 mT m^{-1} was available with a slew rate of up to 200 T m^{-1} s^{-1}. 512 phase-encoding steps were used; the phase-encoding direction was anterior to posterior. The acquisition time for one volume was 2:43 min. Each dataset had a resolution of 512 × 512 × 60 voxels. The multi-station technique was used to establish a WBI scan.

The experiments in this paper are based on 16 of randomly selected participants images. Since both bladder and prostate lie near the center of the volume, it is possible to reduce the calculation room for the prostate segmentation by applying a window of size 176 × 256 × 60 voxels in the center of the volume.

Figure 1 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

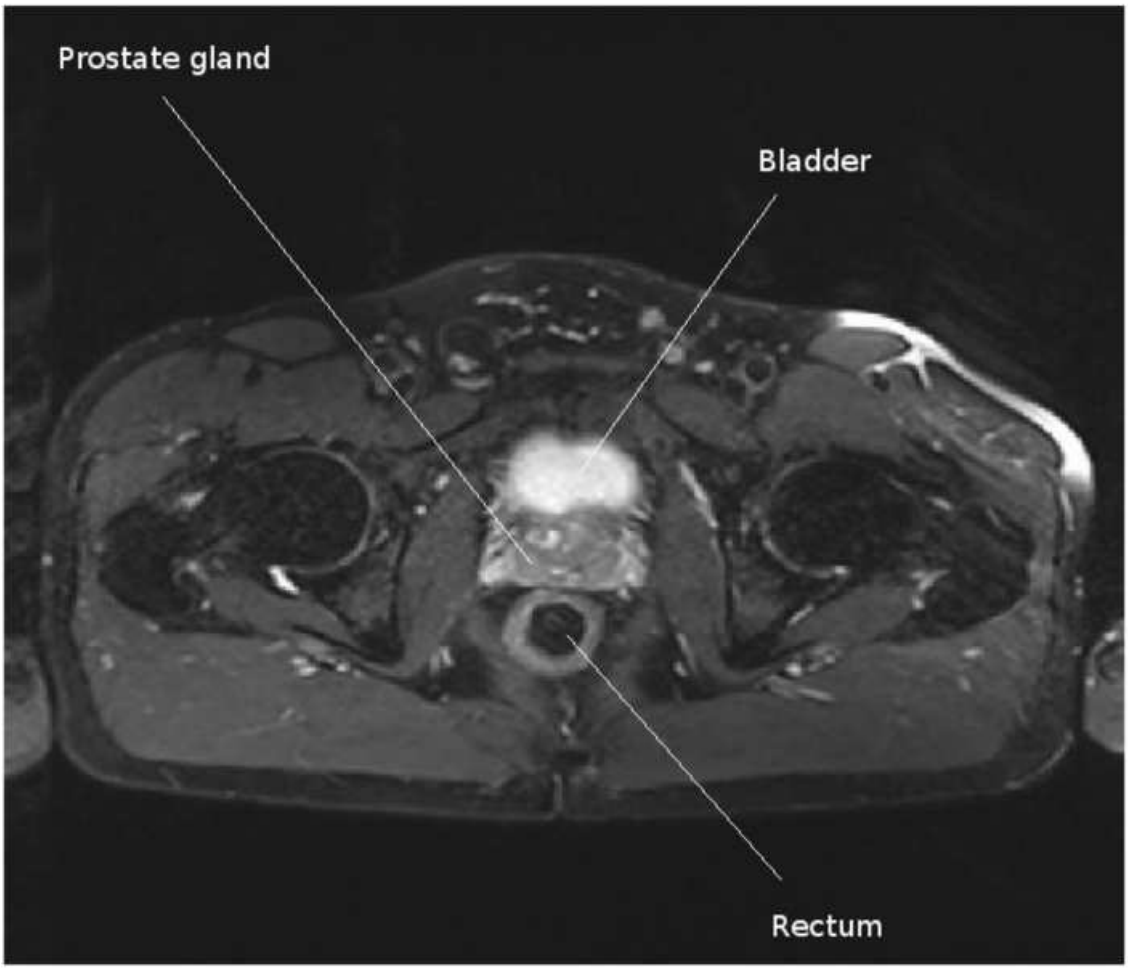


Figure 1. Axial slice example of the Ax PD TSE FS sequence. Axial slice example of the Ax PD TSE FS sequence used in the SHIP study for one participant. This sequence is used to visualize pelvic organs.

Figure 1 illustrate an axial slice example of the Ax PD TSE FS sequence for one participant in the SHIP study and the related anatomical structures. To segment the prostate, we designed and developed a method based on support vector machines (SVM).

I. 2.1. SVM-based segmentation

The consecutive steps of our approach are explained in detail in the following subsections. A flowchart explaining the whole process of the segmentation as well as the training pipeline is illustrated in figure 2

.

Figure 2 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

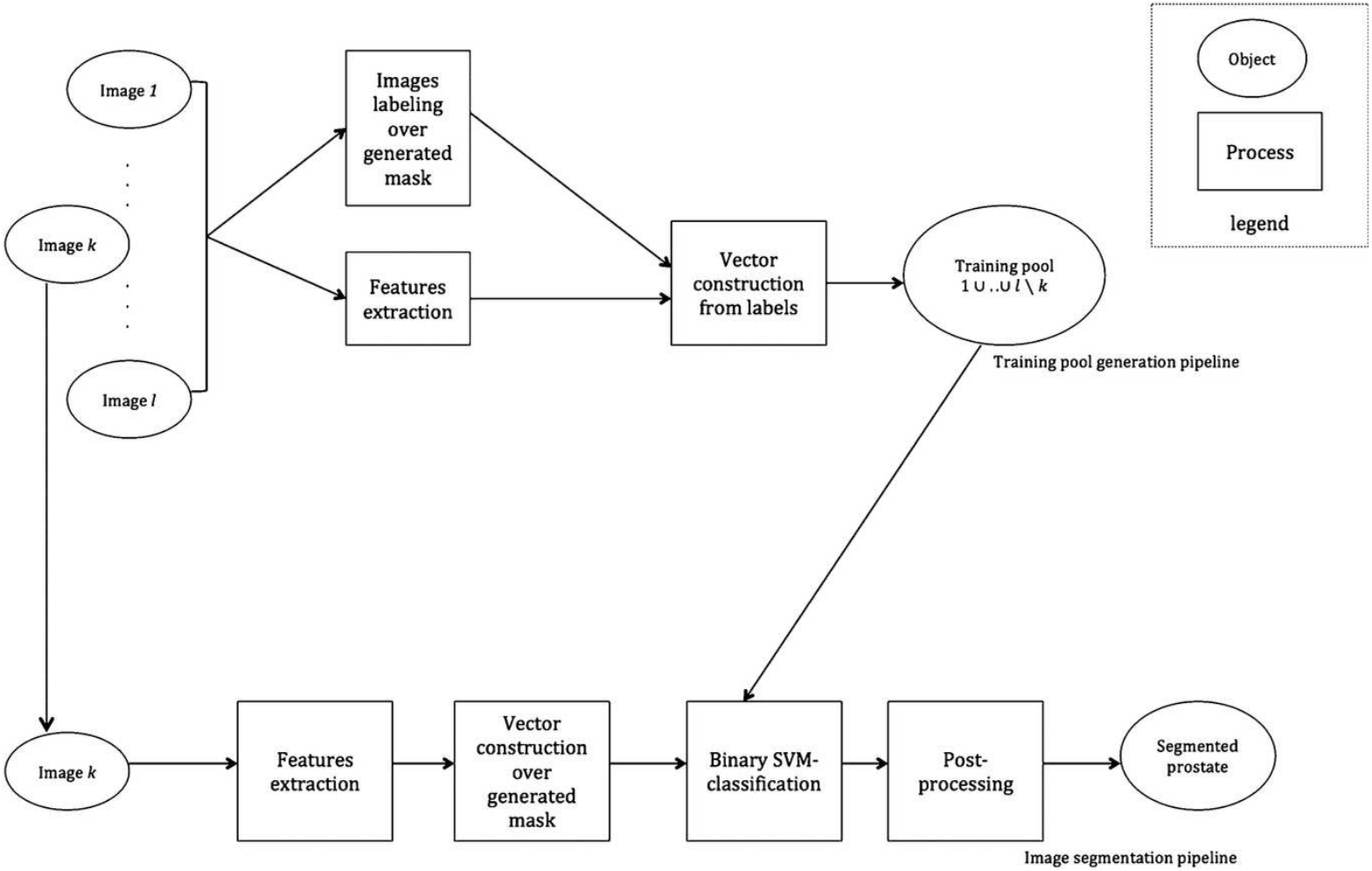


Figure 2. Flowchart of our SVM-based developed method. Flowchart exploring the SVM-based method developed here for prostate segmentation in WBI scans.

This flowchart shows that after system training, a new image can be segmented fully automatically. The steps of the algorithm are explained in detail in the following.

I. 2.1.1. Images labeling and system training

The system was trained with 16 randomly selected image volumes from the SHIP-2 database. The labeling process was done by one urological radiology expert. He labeled the prostate tissues in the images 'by hand'. The radiology expert was asked to use all prostate slices in the volume for each subject. Labeling was done over a masked pelvis image. When one label is set on one voxel, the label is drawn simultaneously for control. The prostate tissue was labeled with 100 labels and another 100 labels were used for the background. Figure 3 illustrates the labeling process led by one radiology expert. Every label is expressed with one symbol drawn simultaneously for control.

Figure 3 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

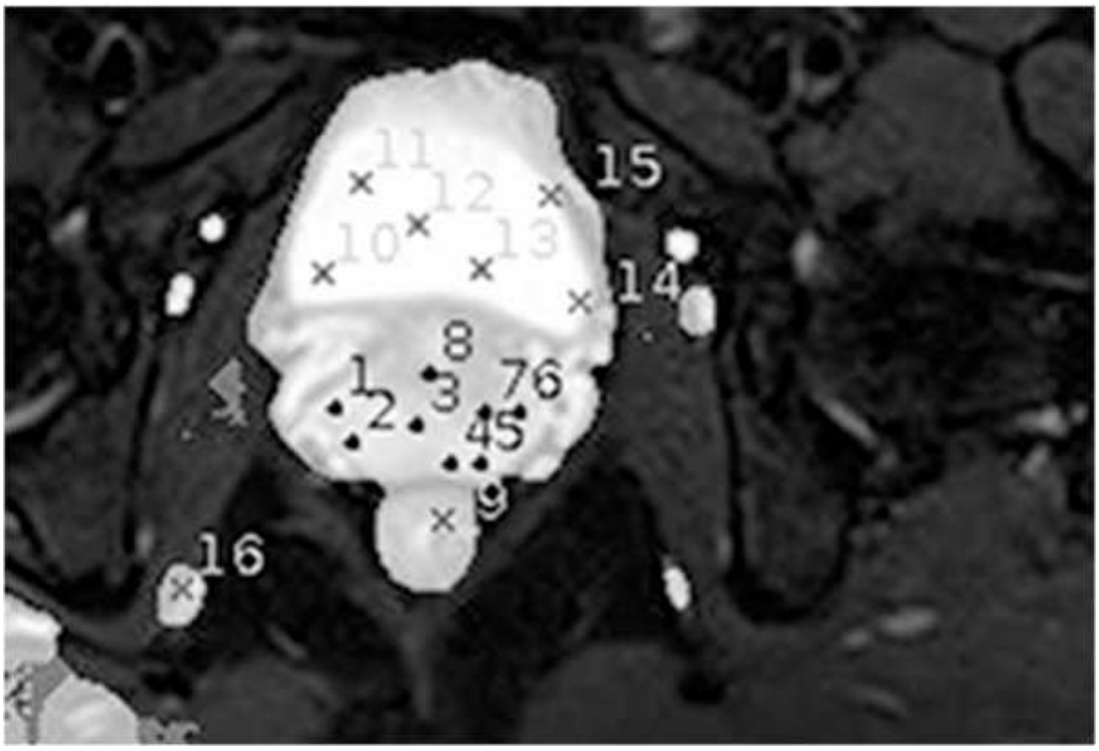


Figure 3. Prostate labeling. Axial slice example for generated mask superimposed on prostate used in this study during the training process led by the expert (dark points correspond to the prostate class and 'x-symbols' to the background class. Sample of eight labels for every class isllustrated; for every subject, 100 vectors from the whole volume were used for every class).

In this figure, samples of eight labels for every class are illustrated. The leave-one-out method was used to generate training pools. This procedure iteratively leaves out the

training set of one definite subject (*k*). The system is trained on remaining training sets for segmentation purposes and then image of *k* is segmented. This method ensures generalization of the segmentation method to independent test images not present in the training pool.

I. 2.1.2. Mask generation

The main aim of this mask is to reduce the amount of the input data. This is essential to reduce the necessary data for the SVM training step.

Figure 4 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

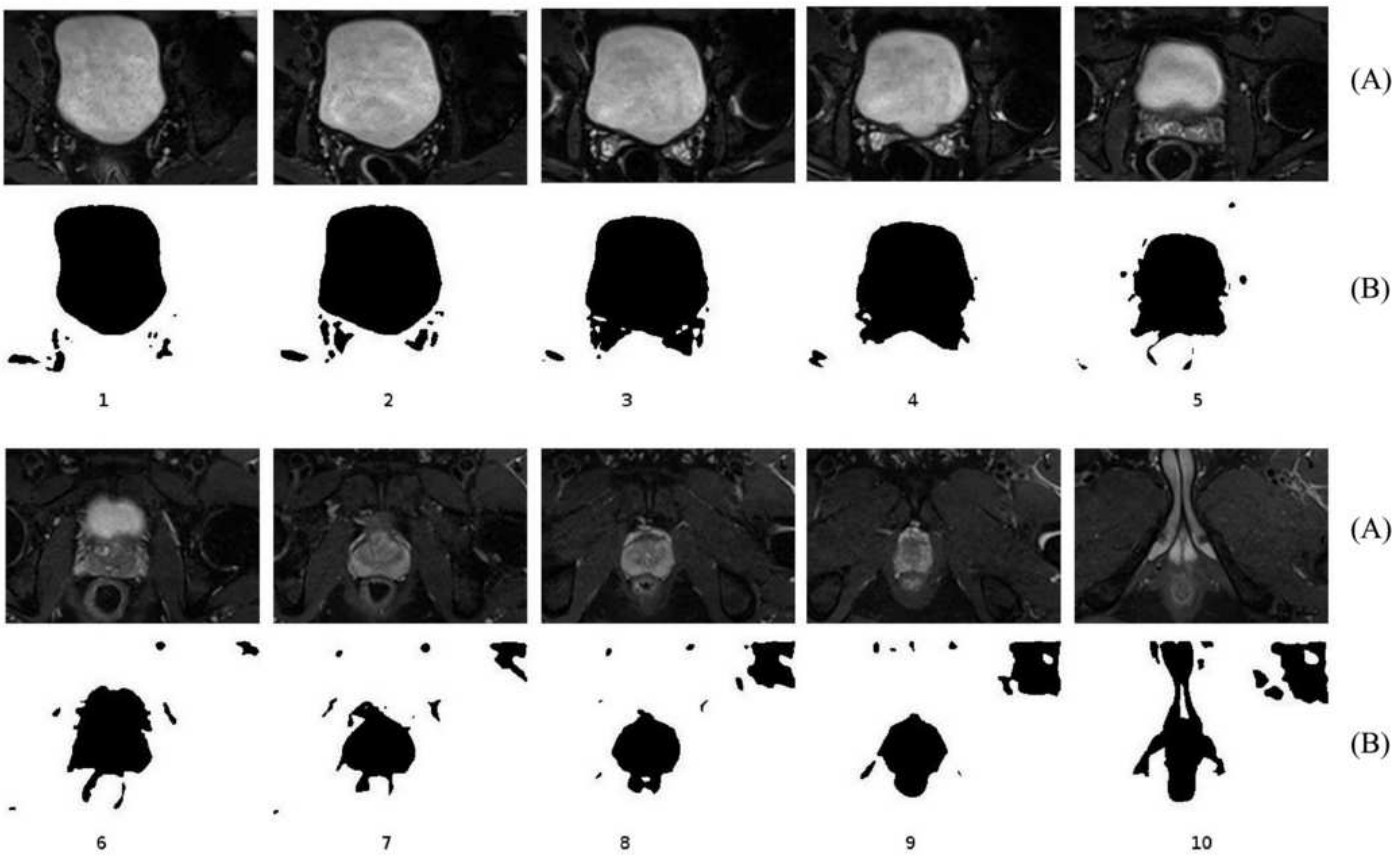


Figure 4. Example slices of the maximal-entropy thresholding. Example slices of the maximal-entropy thresholding result of the pelvis for a random SHIP-2 participant. Since the urological system in the Ax PD TSE FS sequence is always located in the middle of the image, a window in the center of the field of view was applied (A = the original image / axial, B = thresholding results / axial. Tissues with high water content such as prostate, bladder, intestine and penis are thresholded with this technique. The urological system is included in the largest three-dimensional component in the binary volume. We applied a region-growing-based algorithm to isolate it. The results shown are after applying this algorithm).

Use is made of an established thresholding technique based on the maximal-entropy of the distribution of the intensities in the volume to binarize the input image (Kapur et al 1985). In this thresholding technique two types of classes are considered, foreground and background. The entropy of each class is calculated based on the following formulas:

$$H_{\text{foreground}}(T) = -\sum_{g=0}^{T} \frac{p_g}{p_T} ln \frac{p_g}{p_T} \tag{1}$$

$$H_{\text{background}}(T) = -\sum_{g=T+1}^{G} \frac{p_g}{p_T} ln \frac{p_g}{p_T}, \tag{2}$$

where G is the maximum gray-level value in the volume, p_g is the cumulative probability function of the gray-level g, and T is the thresholding value in the histogram of the MRI volume. The thresholding problem can then be considered as an optimization problem. The optimal threshold of the image is defined as:

$$T = \max(H_{\text{foreground}}(T) + H_{\text{background}}(T)), \tag{3}$$

and the image is optimally binarized when the sum of the foreground and background's entropy is maximal. As illustrated in figure 4, the result of this thresholding technique includes soft tissues with a large proportion of water as 'prostate, bladder, intestine and penis'. To isolate the urological system, a binary region-growing-based algorithm is applied to find the largest three-dimensional, connected region (Burger and Burge 2008).

I. 2.1.3. Features extraction

In this step, the features required for the definition of the prostate are generated. Various features have been well studied in the traditional two dimensions in computer vision. In the field of medical imaging, however, organ description should be based on three-dimensional features. The features used in this study are 3D median, 3D gradient, two based on 3D anisotropic diffusion and two 3D based on the eigenvalue of structure tensor.

The 3D filters were applied to achieve best discrimination between prostate and surrounding tissues. In general in MRI, different tissue types can share the same intensity interval (Gonzalez et al 2009); hence, to enable valid discrimination for SVM, further information is needed. Anisotropic diffusion features discriminate between prostate and surrounding vesicles. Median features discriminate between prostate and seminal vesicles. Eigenvalues of structure tensors and gradient features discriminate between prostate and bladder.

3D anisotropic diffusion

Smoothing is a common technique to reduce noise in medical images. Filters such as the Gauss filter are often used in computer vision. In medical imaging, however, it is very important not to affect vascular structures through the noise reduction filtering process. Perona (1990) introduced the anisotropic diffusion filter to reduce noise without removing significant structures from the image (Perona and Malik 1990). Smoothing in this method is formulated as a diffusion process that can be controlled at the boundaries by selection-adaptive diffusion strengths. The formulation of the anisotropic smoothing process can also be used in MRI data. Extension to three dimensions can be done easily (Gerig et al 1992). When the filtering process makes use of all three dimensions, the noise reduction is more effective, since the nature of the edges will also be three dimensional and the useful neighborhood is larger. Anisotropic diffusion can be mathematically formulated as:

$$\frac{\delta I}{\delta t} = \mathrm{div}(c(x, y, z, t)\nabla I) \tag{4}$$

The diffusion function *c(x, y, z, t)*, which controls diffusion strength, is a function of the voxel intensity gradient magnitude. The coordinates *(x, y, z)* represents the spatial coordinates of the image in the three-dimensional MRI volume set, *t* is the diffusion process order parameter and is used in the discrete implementation as the iteration step. This analysis employs two types of diffusion functions:

$$c(x,y,z,t) = \mathrm{e}^{-\left(\frac{\nabla I}{\kappa}\right)^2} \tag{5}$$

$$c(x,y,z,t) = \frac{1}{1 + (\frac{\nabla I}{\kappa})^2}, \tag{6}$$

where κ represents the diffusion constant, ∇ denotes the gradient operator. The two functions 5 and 6 define two different scale spaces. The first scale space enhances object edges during the diffusion process over low-contrast cases. This is important for better object discrimination through the object boundaries. The second function enhances object regions over smaller regions, which allows better enhancement of the prostate region over small regions in the created mask. Empirical analysis led to using the function parameter κ = 70, t = 10 for our experiments.

Eigenvalues of structure tensor

The structure tensor includes first- and second-order intensity-related information. This kind of information has been used by different authors for different tasks. For instance, Rao and Schunk (1991), analyzed flow-like texture information (Rao and Schunck 1991), Gülich et al (1987) and Nitzberg et al (1992) detected corners (Nitzberg and Shiota 1992).

Eigenvalue decomposition of a 3 × 3 symmetric matrix as the structure tensor results in three corresponding eigenvectors (*e1, e2, e3*). The eigenvector *e1* stand for the largest eigenvalue, and the eigenvectors *e2, e3* for middle and smallest eigenvalues, respectively. The largest and smallest eigenvalues of the structure tensors were used as features for the present system.

I. 2.1.4. Vector construction

We are dealing with three-dimensional neighborhood information for the construction of vectors. To reduce classification room dimensions, the conventional three-dimensional bounding box surrounding labeled voxels was not used. The values of the target voxel and its three-dimensional (*n-1*) neighbors in the automatically generated features are used for

the vector construction. The neighborhood structure employed is three dimensional plus symbol-like structure. The vector is defined from one feature, depending on this three-dimensional neighborhood structure, and is thus termed 'feature vector'. For a given radius r, the n dimensions of the feature vector from the target voxel and its three dimensional (n − 1) neighbors can be formulated as:

$$n = (1 + 4 \cdot r) \cdot (1 + 2 \cdot r), \text{where} r = 1,2, \dots ,5 \tag{7}$$

The feature vector can be defined as

$$F_{(x,y,z)} = (f_1, f_2 \dots, f_n), \tag{8}$$

where *x, y* and *z* are the spatial coordinates of the target voxel *v*, and $f_1, f_2 \dots f_n$ are the feature values from the three-dimensional neighborhood information. We obtained the best segmentation in our investigations with the radius r = 2 and resulting feature vectors with n = 45. The classification vectors were built into our segmentation algorithm using the feature vectors. The classification vector can be formulated as:

$$V_{\text{classification}} = (F_{1,(x,y,z)}, F_{2,(x,y,z)} \dots F_{u,(x,y,z)}), \tag{9}$$

where u is the number of features. In our experiments, u has the value 6 (i.e., 2 structure tensor based-, 2 anisotropic diffusion-, 1 median- and 1 gradient images).

I. 2.1.5. Binary SVM classification

The SVM method is used for the binary classification (Cortes and Vapnik 1995, Vapnik 1999). For the training pool P, consists of m vectors x_i from an $\Re^d$ space and labeled in two different classes such as $y_i = -1$ for background and $y_i = +1$ for prostate tissue. The SVM finds the best hyperplane wx−b=0, which divides the space into two classes, and where *w* represents the normal vector and b the bias value. These parameters *w* and *b*

characterize the optimal hyperplane. In this context the following objective function should be minimized:

$$\Phi(w,\xi) = \frac{1}{2}(w \cdot w) + C\left(\sum_{i=1}^{m} \quad \xi_i\right), \tag{10}$$

under the following constraints:

$y_i((w \cdot x_i) - b) \geq 1 - \xi_i \forall i = 1, \dots, m$ and $\xi_i \geq 0 \forall i = 1, \dots, m$, where the parameter ξ represents a slack variable indicating how much the ith instance extends onto the wrong side of the optimal hyperplane, and *C* is the penalty misclassification parameter. The kernel trick is used to map the data, in a nonlinear separable case, to a more suitable space via a kernel function *K* (Scholkopf and Smola 2001). The decision function to classify the test vector *x* not included in the training set using Lagrange multipliers a_i, and kernel function *K* is:

$$f_{\mathrm{svm}}(x) = \mathrm{sgn}\left(\sum_{\mathrm{supportvectors}} \quad y_i a_i K(x, x_j) - b\right) \tag{11}$$

The kernel used here is the RBF:

$$K(x_i, x_j) = \mathrm{e}^{(-\gamma \cdot \|x_i - x_j\|^2)} \tag{12}$$

The use of other kernels such as the polynomial kernel were investigated, but performed less effectively than the RBF kernel. Empirical analysis led to using following parameter (formula 10: *C* = 100), (formula 12: *γ* = 1). Finding the optimal hyperplane separating the two classes of training vectors (formula 10) took on average not more than 2 s for a training pool consisting of training data of 15 scans (100 × 100 vectors × 15). The classification task required considerably more time, since many more vectors needed to be applied on formula 11. This task took on average 8 min in our implementation (CPU: Intel Xeon 2.4 GHz, RAM: 8 GiB, OS: Linux Ubuntu 10.10, programming language: Java). The result of the classification is a binary volume.

I. 2.1.6. Post processing

This step in our method is used to fill in gaps resulting from voxels of cysts, which are not detected as prostate tissue. Cysts in the prostate gland are prevalent pathological findings (Nghiem et al 1990), which can be observed in 7.6% of the population (Ishikawa et al 2003). Cysts appear much lighter than the surrounding tissue in the sequence used. For such cases a binary region-growing algorithm (Burger and Burge 2008) was applied to fill in gaps of the segmented prostate as a post-processing step, since voxels of the cysts should be also considered in the PV assessment. In two subjects, cysts in the prostate were present. The binary region-growing algorithm is integrated in the automatic segmentation pipeline.

I. 2.2. 3D level sets-based segmentation

For segmentation with this method, the Fiji implementation of level sets (Sethian 1999) was used as a semi-automated computerized reference. Fiji is an open-source imaging package freely available (Schindelin et al 2012). This implementation is based on the sparse-field method (Yoo 2004). An initial contour is required to segment the prostate, which will be considered as the initial curve. It expands while trying to find the prostate boundaries. We applied this segmentation technique on the 3D anisotropic filtered image (formula 5).

I. 2.3. Evaluation criteria

To validate the developed method, two different observers segmented the prostate manually (observer A, observer B). The manual segmentations were done slice by slice and the contours of the prostate were defined without any further algorithmic support. The sensitivity (S_n) and specificity (S_p) of whole segmented prostate were used as evaluation criteria. These measurements are widely used for segmentation evaluation (Jian et al 2012, Hu et al 2012) and can be formulated as:

$$S_n = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{13}$$

$$S_p = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}, \tag{14}$$

where true positive (*TP*), false positive (*FP*), true negative (*TN*), and false negative (*FN*) values between the segmentations are calculated.

Besides sensitivity and specificity, the shape-based measurement was also used to evaluate the segmentation performance. A robust, well-known measurement based on the shape is the Hausdorff distance (*HD*) (Huttenlocher et al 1993), which can be formulated as (Böttger et al 2007, Hu et al 2012):

$$\mathrm{HD}(X, Y) = max(d(X, Y), d(Y, X)), \tag{15}$$

where *X* and *Y* are sets of surface points of the two binary segmentation objects. *d(X, Y)* is the Euclidean distance between two points. The *HD* results in the maximum error of the binary objects.

PV was quantified by summing all of the resulting segmentation voxels multiplied by the voxel size. The Bland–Altman method was subsequently used to evaluate the disagreements in PV measurements. The Bland–Altman plot is commonly used to evaluate the agreement between two methods, usually in the context of clinical measurements (Bland and Altman 1986). Spearman's rank correlation coefficient analysis was performed to show the strength of the association between the volumes estimated. All previous evaluation criteria were calculated between the binary segmentations (automated versus manual or manual versus manual) as 3D objects.

I. 2.4. Statistical analysis

Statistical significance for PV measurements was assessed using R 2.15.1, a free software environment for statistical computing and graphics. The Shapiro–Wilk test (Zar 1984) was used to test for the normality of data. Further, the F-test was performed to prove the equality of two variances. To test for the equality of the means of any two samples, a paired t-test was used by assuming normality and equality of variances. In

cases of non-normal distribution, the Wilcoxon rank sum test was used to test the equality of the means of the two examples. Differences were considered to be statistically significant at a level of $p < 0.05$.

I. 3. Results

The prostate segmentation with the method presented here was automatically performed with no user interaction. Figure 5 illustrates the differences between automatically segmented prostate and the corresponding two manual segmentations by the radiology experts for one subject.

Figure 5 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

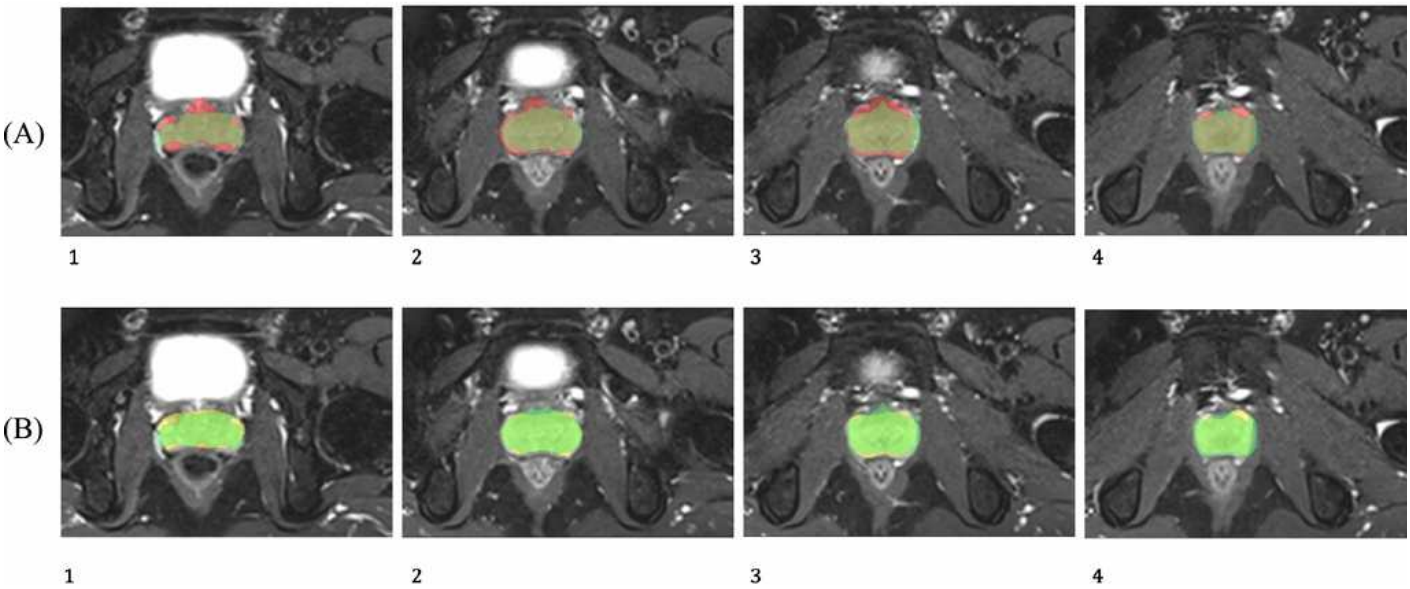


Figure 5. Segmentation results of the SVM-based approach. Segmentation results of the SVM-based approach (green) superimposed on manual segmentations by urological radiology experts in one sequence. (A = superimposed over segmentation by observer A (red); B = superimposed over segmentation by observer B (yellow)).

I. 3.1. SVM-based segmentation evolution experiment

A complete overview of the results obtained from the different evolution strategies (sensitivity, specificity, *HD*) used to evaluate SVM-based method developed here and the level sets-based method are listed in table 1. The evaluation was performed between the SVM-based method and the level sets-based method against the manual segmentation of observers A and B. Values are expressed as means ± SD. Results of statistical analysis

for comparison between SVM-based and level sets-based are also listed in the same table.

I. 3.2. PV agreement experiment

A comparison of the PV between the computerized and the two manual measurements is shown in figure 6, which illustrates the box-plot of resulting PV values.

Figure 6 from Mohamad Habes et al 2013 Phys. Med. Biol. 58 5899

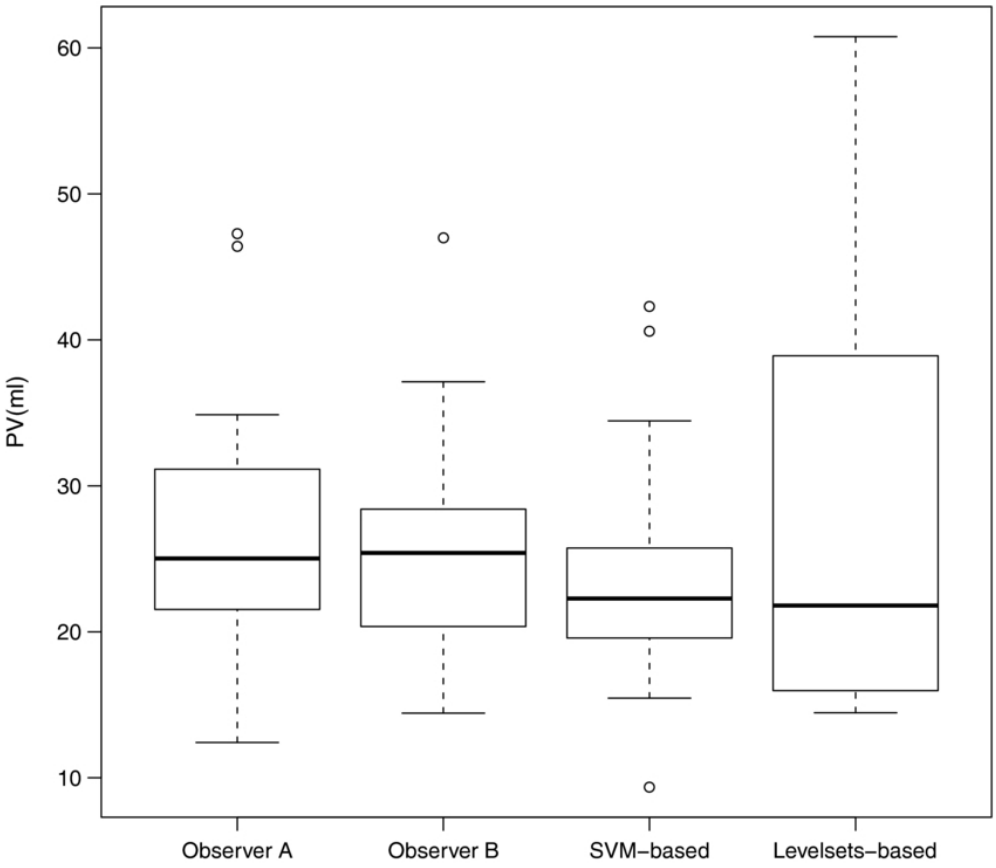


Figure 6. Box-plot for resulted PV values. PVs were determined manually by observers A and B and by use of two different methods 1: our developed SVM-based, 2: level sets-based.

The mean manual volume measurement by observer A was 26.996 ml with a standard deviation of 9.717 ml. The mean manual volume measurement by observer B was 25.809 ml with a standard deviation of 8.217 ml, while the mean volume measurement by the SVM-based method developed here was 23.950 ml with a standard deviation of 8.721 ml and by level sets-based method was 28.613 ml with a standard deviation of 15.057 ml. No statistically significant difference in the mean values between PV of observer A and B measurements was found.

The comparison between volume measurement of SVM-based segmentation and manual segmentation of observers A and B depicts a strong correlation, resulting in a Spearman's rank correlation coefficient (ρ) as listed in table 1.

Bland–Altman plot results between the measurements of our automatic SVM-based versus the level sets-based approach as well as the measurements by the two observers A and B are listed in table 1. The results of our SVM-based method indicate that the mean difference compared to observers A and B as well as the limits of agreement within the 95% confidence interval (CI) were small enough to show good agreement.

I. 3.3. Inter-observer variations

PV measurements showed relevant inter-observer variation. Observer A measurement of the PV reached a mean difference of 1.2 ± 3.6 ml (mean ± standard deviation) and a 95% CI −6.1 to 8.5 versus observer B's measurement.

	Observer A versus			Observer B versus		
Evaluation strategy	**SVM-based**	**Level sets-based**	**p-value**	**SVM-based**	**Level sets-based**	**p-value**
Sensitivity	0.800 ± 0.066	0.656 ± 0.137	0.0006	0.788 ± 0.077	0.655 ± 0.139	0.0040
Specificity	0.998 ± 0.0004	0.998 ± 0.0001	0.2611	0.998 ± 0.0006	0.998 ± 0.0001	0.3332
Hausdorff distance	5.048 ± 2.413	5.564 ± 2.528	0.2522	5.613 ± 2.897	5.945 ± 2.518	0.5619
Spearman's rank (ρ)	0.936*	0.664*		0.859*	0.624*	
Bland–Altman Plot	md = 3 ± 3.1 CI: −3.1 to 9.2	md = −1.6 ± 13 CI: −27.5 to 24.3		md = 1.9 ± 4.5 CI: −7.1 to 10.8	md = −2.8 ± 11.8 CI: −26.3 to 20.7	

Table 1. Segmentation evaluation results comparing the fully automatic SVM-based method and the level sets-based method. Evaluation was performed against the manual segmentation of observers A and B.
*Significant differences are indicated: $p < 0.05$. md: mean difference.

I. 4. Discussion and conclusion

I. 4.1. Comparison to clinical PV standards

PV is an important indicator for BPH. The definition of pathological enlargement of the prostate differs in the literature. In (Kim et al 2011) BPH was defined as PV > 25 ml and an international prostate symptom score > 8. PV measurement in large population-based studies would make it possible to derive more objective threshold values (e.g. the 95% quantile of the PV distribution).

The gold standard method to quantify PV in urological practice is TRUS. The most common way to calculate PV is by using the prolate ellipsoid formula: $PV = H \times W \times L \times \pi/6$, where: H is height, W is width and L is length of the prostate in the ultrasound image (Bates et al 1996). Terris et al showed that this method has a high correlation with the actual PV as measured after radical prostatectomy (Terris and Stamey 1991). Using TRUS, Bates et al reported considerable inter-observer variations in PV quantification (Bates et al 1996). They reported a mean difference of 0.3 ± 5.3 ml (mean ± standard deviation) with a 95% CI of −10.9 to 10.9. To our knowledge, inter-observer variations of manual PV quantification of 3D objects in MRI images has not been studied systematically. In our analysis, manual segmentations by two urological radiology experts showed slightly better PV quantification accuracy in MRI with mean difference of 1.2 ± 3.6 ml (95% CI −6.1 to 8.5) compared to Bates et al results. Note that the comparison with published values was not direct, because different databases were used.

I. 4.2. Evaluation of the new methodology

To our knowledge, no algorithm to segment the prostate in WBI scans or the Ax PD TSE FS has been published to date, neither for semi- nor fully-automatic segmentation. Conventional computerized segmentation techniques, such as fast marching (Sethian 1999, Yoo 2004), did not yield acceptable segmentation results. Kallenber et al used neural networks to automate breast density segmentations (Kallenberg et al 2011). With limited training efforts, we developed a new prostate segmentation technique for Ax PD TSE FS, which can be used in the context of epidemiological studies, since it requires no user interaction and hence can be readily implemented in large scale studies.

Automatic segmentation of the prostate in 3D MRI volumes can be regarded as a binary classification problem, where each voxel can be assigned to either prostate tissue or the background. To our knowledge, SVM has not been previously used to segment deformable organs like the prostate in MRI images. The SVM approach developed here depends upon automatically generated 3D features of the volume as well as 3D neighborhood structures. Median, gradient, anisotropic, and eigenvalues of the structure tensor features were automatically generated for the classification task. These features can give a full description of the prostate. The dimensions of the volume to be distinguished were reduced through maximal-entropy thresholding.

PV automatically measured with the method developed here has a CI width of −3.1 to 9.2 versus observer A, and −7.1 to 10.8 versus observer B, which is comparable to previously published values for inter-observer variation in TRUS (−10.9 to 10.9)(Bates et al 1996). Mean PV differences of our automated approach compared to human observers (3 ± 3.1 ml versus observer A, and 1.9 ± 4.5 ml versus observer B) are also close to the published values for TRUS (0.3 ± 5.3 ml, gold standard) (Bates et al 1996). Note that these comparisons with published values were not direct, because different databases were used.

I. 4.3. Comparison to level sets-based segmentation

Even the results obtained by the level sets-based segmentation in some evaluation strategies are similar: specificity (S_p) = 0.998 ± 0.0001 versus observer A and S_p = 0.998 ± 0.0001 versus observer B, *HD* = 5.564 ± 2.528 versus observer A and *HD* = 5.945 ± 2.518 versus observer B. Other evaluation strategies show that the SVM-based segmentation is more accurate than the level sets-based one for prostate segmentation in whole-body MRI scans: sensitivity (S_n) = 0.656 ± 0.137 < 0.800 ± 0.066 versus observer A and S_n = 0.655 ± 0.139 < 0.788 ± 0.077 versus observer B. Spearman's rank (ρ) compared to volumes estimated by observer A (ρ) = 0.664 < 0.936 and (ρ) = 0.624 < 0.859 compared to volumes estimated by observer B. Bland–Altman plots show CI width of −27.5 to 24.3 versus observer A, and −26.3 to 20.7 versus observer B, which are obviously worse than those of the SVM-based method (−3.1 to 9.2 versus observer A and −7.1 to 10.8 versus observer B).

I. 4.4. Validation through literature review

Direct comparisons of our computerized algorithm with previously developed methods in the literature are limited because different databases, protocols and evaluation criteria were used. Furthermore, to our knowledge, this is the first study segmenting the prostate in MR modality with the Ax PD TSE FS.

Recent works in MR modality, such as that presented by Toth et al , reported similar results: S_n = 0.81 and S_p = 0.99 (Toth et al 2011b). Langerak et al , also reported a similar S_p value of 0.99 value (Langerak et al 2010). Gao et al , reported an HD of 10.22 ± 4.03 mm (Gao et al 2010), which are worse than the results achieved in the present study. Although the sensitivity obtained by using our SVM-based method is high, it does not reach the one reported by Martin et al , (S_n = 0.86) (Martin et al 2010). Note that the above comparisons were not direct, because different databases were used. Some researchers used the root mean square distance, e.g., Zhu et al , who obtained a value of 5.4811 ± 2.9 mm (Zhu et al 2007). Other authors used the mean absolute distance as evaluation strategy, Allen et al , achieved a value of 2.8 ± 0.82 mm (Allen et al 2006) and Toth et al , reported a value of 5 mm (Toth et al 2011b). Toth et al , using volumetric ratio for evaluation, obtained a value 1.05 ± 0.21 (Toth et al 2011a). Dowling et al , used median Dice coefficient for evaluation and obtained a value of 0.86 (Dowling et al 2011).

Other recent methods have been developed to automate prostate segmentation in TRUS modality such as that by Diaz et al , who achieved results close to ours; they reported an S_n value of 80% (Diaz and Castaneda 2008). Garnier et al , also presented a similar mean HD value (4.79 ± 1.62) (Garnier et al 2011). Some researchers used volume error as an evaluation criterion; for instance, Mahdavi et al , obtained volume errors of 6.63 ± 0.9% (Mahdavi et al 2011), while other authors validated their methods with yet other evaluation strategies. Medina et al , used the contour mean distance and reported a value of 3.58 ± 1.49 pixels (Medina et al 2006). Yan et al , evaluated their method with the mean absolute distance and reported a value of 2.01 ± 1.02 mm (Yan et al 2010), while Ghose et al , evaluated their method using Dice coefficients and reported a value of 0.95 ± 0.2 (Ghose et al 2012).

I. 4.5. Advantages of the SVM-based segmentation

Besides the full automation and the accurate segmentation proven from our results, our approach has the advantage of being able to highly accurately segment the prostate with a limited training set (200 vectors × 15 images).

Methodologies based on shape for segmentation require large number of masks in the case of the prostate. The reason is that prostate shape differs strongly between individuals especially in the case of BPH. The training set of such methodologies should include all possible prostate shape variations to enable accurate prostate detection. Such methodologies could segment the prostate accurately once their systems are sufficiently trained. Chen et al 's methodology for prostate segmentation in CT images presented in (Chen et al 2011) needed 184 scans for training. Klein et al 's method requires masks generated through manual segmentation by experts, which include all possible shape variations of the prostate (Klein et al 2008). Depending on radiological expertise, manual segmentation of the prostate for evolution purposes in this study took an average of 10 min, while labeling of vectors on the prostate gland in our study took an average of about 1 min.

Our developed methodology however is classification-based. Furthermore SVM can make good generalization from a limited number of training sets (Shao and Lunetta 2012) and can achieve very good classification accuracy. Following Tingelhoff et al we decided having two expert observers rather than the maximum number of screens with only one observer (Tingelhoff et al 2008). Resources of two manual readings for our evaluation were available for only randomly selected 16 subjects. These were sufficient however for our developed algorithm to segment the prostate highly accurately.

I. 4.6. Limitations and outlook

Nonetheless, some limitations should be considered regarding the results of the present study. First of all, MRI-based PV determination for the subjects participating in this study could not be directly compared to corresponding TRUS data due to the ethical consent standards of the SHIP study. Second, the Bland–Altman plots indicate very similar PV estimation results compared to the manual segmentation by observer A, but less similar compared to observer B. Therefore, the system in its current status may need a visual

radiological control of its results for clinical use. The optimization of our algorithm will increase accuracy in discrimination between prostate and seminal vesicles or penis-related tissue.

Implementation optimization is possible using more processing units for classifying vectors simultaneously (e.g., GPU-based implementation). We assume that this will improve the performance time considerably (calculation time could be reduced to less than 1 s). In the context of epidemiological studies and considering the fact that manual prostate segmentation by our human observer took 10 min on average, the performance of our implementation is acceptable, since it completely lacks human interaction.

An algorithm was developed for the automatic segmentation of the liver in whole-body MRI scans (Gloger et al 2010). This method, however, cannot be adapted for prostate segmentation in one component sequence, since the Dixon sequence (Dixon 1984) used in that work produces four image components. It was also possible to automate skull segmentation using the ultra-short echo time (UTE) sequence, which produces two image components (Rota Kops et al 2011, Habes et al 2013). The present authors intend to optimize the system by considering more information acquired about the prostate in the SHIP protocol to improve the segmentation itself and to expand our system with a decision-support unit. Moreover, we plan to investigate the possibilities of further integration of more SHIP sequences for automatic segmentation and PV quantification in SHIP-2 to improve our segmentation results.

I. 5. Conclusion

Our automatic methodology based on SVM for prostate segmentation in Ax PD TSE FS can segment the prostate in WBI scans with high segmentation quality. Our SVM-based method compared to that based on level sets showed better segmentation accuracy. The results show that the methodology we developed here is comparable to urological radiology expert prostate readings in WBI scans and it is suited for epidemiological studies.

I. References

Allen P D, Graham J, Williamson D C and Hutchinson C E 2006 Differential segmentation of the prostate in MR images using combined 3D shape modelling and voxel classification ISBI: IEEE Int. Symp. on Biomedical Imaging pp 410–3

Bates T, Reynard J, Peters T and Gingell J 1996 Determination of prostatic volume with transrectal ultrasound: a study of intra-observer and interobserver variation J. Urol. 155 1299–300

Bland J M and Altman D G 1986 Statistical methods for assessing agreement between two methods of clinical measurement Lancet 327 307–10

Böttger T, Grunewald K, Schöbinger M, Fink C, Risse F, Kauczor H U, Meinzer H P and Wolf I 2007 Implementation and evaluation of a new workflow for registration and segmentation of pulmonary MRI data for regional lung perfusion assessment Phys. Med. Biol. 52 1261

Burger W and Burge M J 2008 Digital Image Processing (Berlin: Springer)

Chen S, Lovelock D M and Radke R J 2011 Segmenting the prostate and rectum in CT imagery using anatomical constraints Med. Image Anal. 15 1–11

Cortes C and Vapnik V 1995 Support-vector networks Machine Learning vol 20 (Boston: Kluwer) pp 273–97

Costa M J, Delingette H, Novellas S and Ayache N 2007 Automatic segmentation of bladder and prostate using coupled 3D deformable models Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007 vol 4791 (Berlin: Springer) pp 252–60

Diaz K and Castaneda B 2008 Semi-automated segmentation of the prostate gland boundary in ultrasound images using a machine learning approach Proc. SPIE 6914 69144A

Dixon W T 1984 Simple proton spectroscopic imaging Radiology 153 189–94

Dowling J A, Fripp J, Chandra S, Pluim J P W, Lambert J, Parker J, Denham J, Greer P B and Salvado O 2011 Fast automatic multi-atlas segmentation of the prostate from 3D MR images Proc. 2011 Int. Conf. on Prostate Cancer Imaging: Image Analysis and Image-Guided Interventions (Berlin: Springer) pp 10–21

Foskey M, Davis B, Goyal L, Chang S, Chaney E, Strehl N, Tomei S, Rosenman J and Joshi S 2005 Large deformation three-dimensional image registration in image-guided radiation therapy Phys. Med. Biol. 50 5869–92

Gao Y, Sandhu R, Fichtinger G and Tannenbaum A R 2010 A coupled global registration and segmentation framework with application to magnetic resonance prostate imagery IEEE Trans. Med. Imaging 29 1781–94

Garnier C, Bellanger J J, Wu K, Shu H, Costet N, Mathieu R, de Crevoisier R and Coatrieux J L 2011 Prostate segmentation in HIFU therapy IEEE Trans. Med. Imaging 30 792–803

Gerig G, Kubler O, Kikinis R and Jolesz F A 1992 Nonlinear anisotropic filtering of MRI data IEEE Trans. Med. Imaging 11 221–32

Ghose S, Oliver A, Marti R, Llado X, Freixenet J, Mitra J, Vilanova J C, Comet-Batlle J and Meriaudeau F 2012 Statistical shape and texture model of quadrature phase information for prostate segmentation Int. J. Comput. Assist. Radiol. Surg. 7 1–13

Gloger O, Kuhn J, Stanski A, Volzke H and Puls R 2010 A fully automatic three-step liver segmentation method on LDA-based probability maps for multiple contrast MR images Magn. Reson. Imaging 28 882–97

Gonzalez F A, Romero E, Gonzalez F A and Romero E 2009 Biomedical Image Analysis and Machine Learning Technologies: Applications and Techniques 1st edn (Hershey, PA: IGI Publishing)

Habes M, Rota Kops E, Bahr J, Kühn J P, Hoffmann W, Lipinski H G and Herzog H 2013 Automated skull and cavity segmentation from ultra short TE sequence images Curr. Med. Imaging Rev. 9 120–8

Hegenscheid K, Kühn J P, Völzke H, Biffar R, Hosten N and Puls R 2009 Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study Rofo 181 748–59

Hu Y C, Grossberg M D, Wu A, Riaz N, Perez C and Mageras G S 2012 Interactive semiautomatic contour delineation using statistical conditional random fields framework Med. Phys. 39 4547

Huttenlocher D P, Klanderman G A and Rucklidge W J 1993 Comparing images using the hausdorff distance IEEE Trans. Pattern Anal. Mach. Intell. 15 850–63

Ishikawa M, Okabe H, Oya T, Hirano M, Tanaka M, Ono M, Kawamura K, Fujimoto N and Sakurada K 2003 Midline prostatic cysts in healthy men: incidence and transabdominal sonographic findings Am. J. Roentgenol. 181 1669–72

Jian W, Sun X and Luo S 2012 Computer-aided diagnosis of breast microcalcifications based on dual-tree complex wavelet transform Biomed. Eng. Online 11 96

Kallenberg M G, Lokate M, van Gils C H and Karssemeijer N 2011 Automatic breast density segmentation: an integration of different approaches Phys. Med. Biol. 56 2715

Kapur J N, Sahoo P K and Wong A K C 1985 A new method for graylevel picture thresholding using the entropy of the histogram Comput. Vis. Graph. Image Process. 29 273–85

Kim J M, Song P H, Kim H T and Moon K H 2011 Effect of obesity on prostate-specific antigen, prostate volume, and international prostate symptom score in patients with benign prostatic hyperplasia Korean J. Urol. 52 401–5

Klein S, van der Heide U A, Lips I M, van Vulpen M, Staring M and Pluim J P W 2008 Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information Med. Phys. 35 1407–17

Langerak R, van der Heide U A, Kotte A N T J, Viergever M A, van Vulpen M and Pluim J P W 2010 Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE) IEEE Trans. Med. Imaging 29 2000–8

Lee C C and Chung P C 2004 Identifying abdominal organs using robust fuzzy inference model Proc. IEEE Int. Networking, Sensing and Control Conf. pp 1289–94

Mahdavi S S, Chng N, Spadinger I, Morris W J and Salcudean S E 2011 Semi-automatic segmentation for prostate interventions Med. Image Anal. 15 226–37

Martin S, Troccaz J and Daanenc V 2010 Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model Med. Phys. 37 1579–90

Mazonakis M, Damilakis J, Varveris H, Prassopoulos P and Gourtsoyiannis N 2001 Image segmentation in treatment planning for prostate cancer using the region growing technique Br. J. Radiol. 74 243–8 PMID: 11338100

Medina R, Bravo A, Windyga P, Toro J, Yan P and Onik G 2006 A 2-D Active appearance model for prostate segmentation in ultrasound images Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society pp 3363–6

Nghiem H T, Kellman G M, Sandberg S A and Craig B M 1990 Cystic lesions of the prostate Radiographics 10 635–50 PMID: 1696019

Nitzberg M and Shiota T 1992 Nonlinear image filtering with edge and corner enhancement IEEE Trans. Pattern Anal. Mach. Intell. 14 826–33

Pasquier D, Lacornerie T, Vermandel M, Rousseau J, Lartigau E and Betrouni N 2007 Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy Int. J. Radiat. Oncol. Biol. Phys. 68 592–600

Perona P and Malik J 1990 Scale-space and edge detection using anisotropic diffusion IEEE Trans. Pattern Anal. Mach. Intell. 12 629–39

Rao A and Schunck B G 1991 Computing oriented texture fields CVGIP: Graph. Models Image Process. 53 157–85

Rota Kops E, Habes M, Kaffanke J, Tellmann L, Lipinski H G, Herzog H and Shah J 2011 Attenuation maps for brain PET-MR scanners based on segmentation of UTE images J. Nucl. Med. Meet. Abstr. 52 1991 (Abstract)

Rummeny E J, Reimer P and Heindel W (ed) 2006 Ganzkörper-MR-Tomographie (Stuttgart: Thieme)

Schindelin J et al 2012 Fiji: an open-source platform for biological-image analysis Nature Methods 9 676–82

Scholkopf B and Smola A J 2001 Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Cambridge, MA: MIT Press)

Sethian J A 1999 Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science vol 3 (Cambridge: Cambridge University Press)

Shao Y and Lunetta R S 2012 Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points ISPRS J. Photogramm. Remote Sens. 70 78–87

Stamey T A, Yang N, Hay A R, McNeal J E, Freiha F S and Redwine E 1987 Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate New. Engl. J. Med. 317 909–16

Terris M K and Stamey T A 1991 Determination of prostate volume by transrectal ultrasound J. Urol. 145 984–7 PMID: 2016815

Tingelhoff K, Eichhorn K W, Wagner I, Kunkel M E, Moral A I, Rilk M E, Wahl F M and Bootz F 2008 Analysis of manual segmentation in paranasal CT images Eur. Arch. Otorhinolaryngol. 265 1061–70

Toth R, Bloch B N, Genega E M, Rofsky N M, Lenkinski R E, Rosen M A, Kalyanpur A, Pungavkar S and Madabhushi A 2011a Accurate prostate volume estimation using multifeature active shape models on T2-weighted MRI Acad. Radiol. 18 745–54

Toth R, Tiwari P, Rosen M, Reed G, Kurhanewicz J, Kalyanpur A, Pungavkar S and Madabhushi A 2011b A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation Med. Image Anal. 15 214–25

Vapnik V 1999 The Nature of Statistical Learning Theory (Berlin: Springer)

Vesely S, Knutson T, Damber J E, Dicuio M and Dahlstrand C 2003 Relationship between age, prostate volume, prostate-specific antigen, symptom score and uroflowmetry in men with lower urinary tract symptoms Scand. J. Urol. Nephrol. 37 322–8

Völzke H et al 2011 Cohort profile: the study of health in Pomerania Int. J. Epidemiol. 40 294–307

Yan P, Xu S, Turkbey B and Kruecker J 2010 Discrete deformable model guided by partial active shape model for TRUS image segmentation IEEE Trans. Biomed. Eng. 57 1158–66

Yoo T S 2004 Insight into Images: Principles and Practice for Segmentation, Registration, and Image Analysis (Massachusetts: AK Peters Ltd/CRC Press)

Zar J H 1984 Biostatistical Analysis 2nd edn (Englewood Cliffs NJ: Prentice-Hall)

Zhu Y, Williams S and Zwiggelaar R 2007 A hybrid ASM approach for sparse volumetric data segmentation Pattern Recognit. Image Anal. 17 252–8

Appendix II: New technique for prostate volume assessment.

Appandix II is based on:

Mohamad Habes, Jeanette Bahr, Thilo Schiller, Jens-Peter Kühn, Laura Hoppe, Martin Burchardt, Wolfgang Hoffmann, "New technique for prostate volume assessment." World journal of urology (2013): (accepted, in press). DOI:10.1007/s00345-013-1220-2

The final published version can be downloaded from:

http://link.springer.com/article/10.1007%2Fs00345-013-1220-2/fulltext.html

II. Abstract

Purpose

The prostate-specific antigen density (PSAD) helps distinguish between benign prostatic hyperplasia (BPH) and prostate cancer. Accurate prostate volume (PV) assessment is necessary for PSAD calculation and both BPH diagnosis and treatment response monitoring; therefore, accurate PV measurement is increasingly becoming an essential step in the urology.

Methods

Magnetic resonance imaging was used for PV estimation. A new technique based on single-class support-vector machines (S SVM) for accurate PV estimation was realized. Three estimation methods were compared; method 1: planimetry (reference), method 2: S SVM based, and method 3: prolate ellipsoid.

Results

Method 1 and method 2 depict a strong correlation (Spearman's rank correlation coefficient $\rho = 0.965$, $p < 0.001$). The interrater reliability for method 1 and method 2 readings as expressed by the intraclass correlation coefficient (ICC) was 0.975 ($p < 0.001$). Comparison between method 3 and the two other methods shows $\rho = 0.873$ ($p < 0.001$), and $\rho = 0.795$ ($p < 0.001$), respectively. ICC was 0.54 and 0.505, respectively. The mean difference between method 1 and method 2 was −0.05 ml. The limits of agreement

with the 95 % confidence interval were −3.8 to 3.7 ml. Comparing method 3 and the two other methods shows a worse agreement with mean difference of 8.6 ml (95 % confidence interval of 1.0–16.2 ml) and 8.6 ml (95 % confidence interval of −0.7 to 18.0 ml), respectively.

Conclusions

The prostate volumes obtained by our technique agreed excellently with the planimetry (reference) method. This new technique would be clinically useful for urologists in prostate volumetric analysis.

II. Introduction

Benign prostatic hyperplasia (BPH) is the leading cause of enlargement of the prostate [1]. BPH can cause compression of the intraprostatic urethra, which in turn often leads to urinary flow obstruction. This common disease often compromises the quality of life, requires long-term medication, and sometimes ultimately surgical intervention.

Accepted indicators for BPH are the prostate volume (PV), prostate-specific antigen (PSA) levels, and the maximum flow rate on uroflowmetry (MFR). The most important radiological indicator for BPH is PV, which has been shown to correlate with clinical PSA serum values [2, 3] and reduction in MFR. The definition of pathological enlargement of the prostate is variable and differs in the literature. Kim et al. [4].defined a PV ≥ 25 ml as the cutoff value for BPH. It was shown that not only the transition zone volume but also the whole prostate volume (PV) correlates with the volume and weight of resected tissue in patients after transurethral resection of the prostate due to BPH [5].

Furthermore, the measurement of the PV can be used as both a diagnostic and a monitoring parameter in the clinical management of BPH [1, 2]. Therefore, a robust and device-independent technique to determine the true PV is necessary to assess reliable volume. At this time, the standard of reference to evaluate the PV is transrectal ultrasound using the ellipsoid formula [5]. In addition to strong prostatic enlargement, BPH can cause form variations, which means that strict shape assumptions can be inaccurate for assessment of PV based on transrectal ultrasound with the ellipsoid formula. The purpose of this study was to evaluate a new MR-guided approach for assessment of PV based on pattern recognition methodology without any shape assumption.

II. Materials and methods

In the population-based prospective cohort Study of Health in Pomerania (SHIP), healthy volunteers underwent whole-body magnetic resonance imaging using a 1.5-T MR scanner (Avanto; Siemens Medical Systems, Erlangen, Germany) [6, 7]. Hegenscheid et al. [7] used the axial proton density fast spin echo fat-saturated sequence for the pelvis region visualization in axial slice orientation. Image parameters of this sequence were the following: TR 3,230 ms, TE 34 ms, flip angle 90 °, and voxel size (0.8 × 0.8 × 3 mm: 0.9 mm gap). In this study, MR images of the prostate of randomly selected SHIP participants were used to determine the prostate volume by means of three different estimation methods. Fifty-three images of volunteers (with PV < 60 ml) were assessed. In this paper, we describe the three estimation methods that were used to calculate the prostate volume and compare their results.

The Ethics Committee of the University of Greifswald approved the general SHIP study project, and this associated project. Written informed consent was obtained from all SHIP-cohort participants.

II. Prostate volume estimation methods

II. Method 1: planimetry

This method is based on the manual tracing of the prostate by an urologist with radiological expertise. By manual delineation, the prostate was outlined slice by slice upon visual inspection without further algorithmic support. A binary 3-D object was generated from all slices. PV was then calculated from the number of voxels in the object multiplied by the voxel size of the scan. For manual delineation, the open-source software Image J for image analysis was used [8]. Since planimetry-based assessment of PV is closest to the actual prostate shape, we used the planimetry-based assessment as the reference (or “gold-standard”) in this study.

II. Method 2: single-class support-vector machines-based technique

The support-vector machines (SVM) method is well established in pattern recognition for binary classification tasks [9, 10]. Using training data of two classes, SVM can find the optimal hyperplane separating them. The optimal hyperplane will then be used for

classification of new data that were not included in the training set. We used recently binary SVM for prostate segmentation [11]. However, in this study, we developed PV assessment algorithm based on the single-class SVM (S SVM) method, which is suggested by Schölkopf et al. [12] for classification tasks. Here, the SVM training data are sufficient to be labeled with just one class. The training set in this case should have probability distribution function P in the feature space F. An optimal subset S of F must be sought that separates the whole training set from the origin of F.

Using the S SVM, six landmarks on the contour of the prostate in every slice were manually seeded. Every landmark represents a vector in the spatial coordinates of the image. Separating the landmarks with S SVM from the origin of F reconstructs a 3-D prostate object. Prostate volume estimation in this method is then similar to that applied to the manually delineated slices.

II. Method 3: prolate ellipsoid formula

The prostate volume was determined by a urologist with radiological expertise using the widely clinically recognized prolate ellipsoid formula: $V = H \times W \times L \times \pi/6$, where H is height, W is width, and L is length of the prostate in the 3-D image. This calculation assumed that the prostate has an ellipsoid shape. Dimensions of the prostate were calculated using the software Image J [8]. Three views of the three-dimensional orthogonal projections were used. The dimensions H and W were set on the axial view of the prostate center slice. L was set on the coronal view.

II. Evaluation strategy

The Wilcoxon rank sum test was used to test the equality of the means of the two examples. Multiple tests were performed by applying Bonferroni corrections. PV data are given as mean ± SD. To demonstrate the width of the confidence interval of the estimated PV, Bland–Altman plots with 95 % confidence intervals are shown. This plot is commonly used to evaluate the agreement between two methods in the context of clinical measurements [13].

Scatter plots and Spearman's rank correlation coefficient analysis were calculated to show the strength of the association between the volumes estimated with various methods. To show the interrater reliability and agreement with quantitative data, the measurements

were evaluated by using the intraclass correlation coefficient (ICC) [14, 15]. We used the R package irr, which included estimation of the intraclass correlation coefficient (ICC) for one-way and two-way models. This ICC computes single-score or average-score ICCs as an index of interrater reliability of quantitative data. Additionally, the F test and confidence intervals were computed. We used the two-way model with type of agreement and single-score ICC (2, 1). Further, we illustrated a figure with PV estimation error (PV of reference measurement—PV of estimation method) for all volunteers. Statistical analyses were performed using R 2.15.2, a free software environment for statistical computing and graphics. Differences were considered to be statistically significant at a level of $p < 0.05$.

II. Results

Figure 1 illustrates an example of one slice of resulted prostate contour for one subject via planimetry and S SVM.

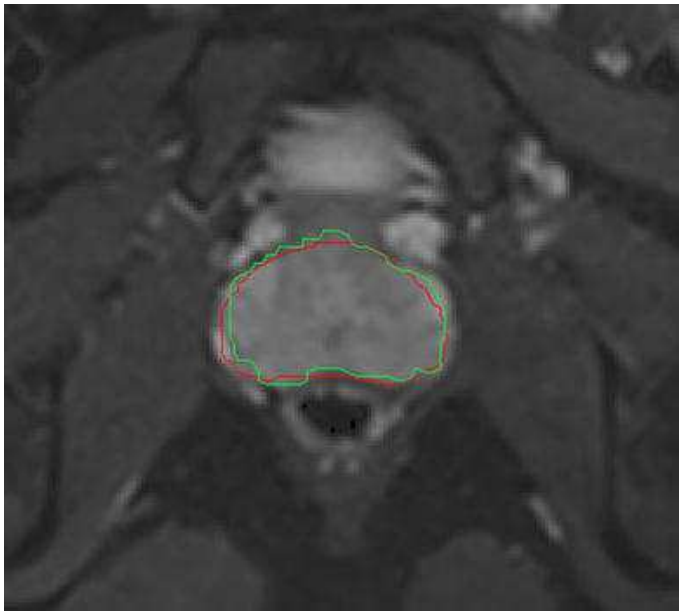


Fig. 1: Prostate contouring via two different methods (method 1: planimetry [green] and method 2: S SVM [red])

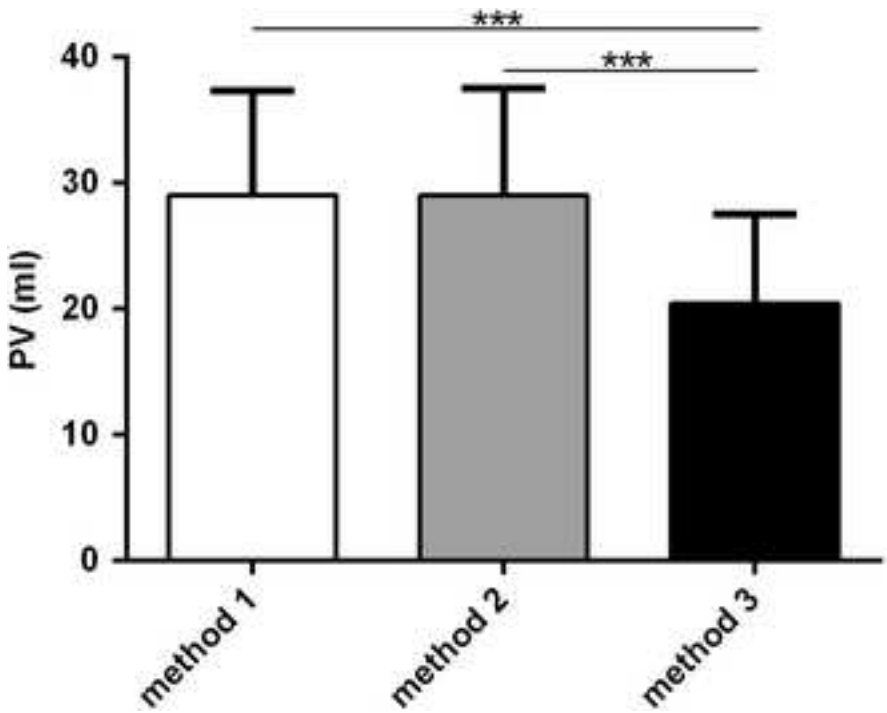


Fig. 2: Prostate volumes were determined by use of three different methods (method 1 planimetry, method 2 S SVM, and method 3 prolate ellipsoid). Values are expressed as mean ± SD. Significant differences are indicated: ***p < 0.001

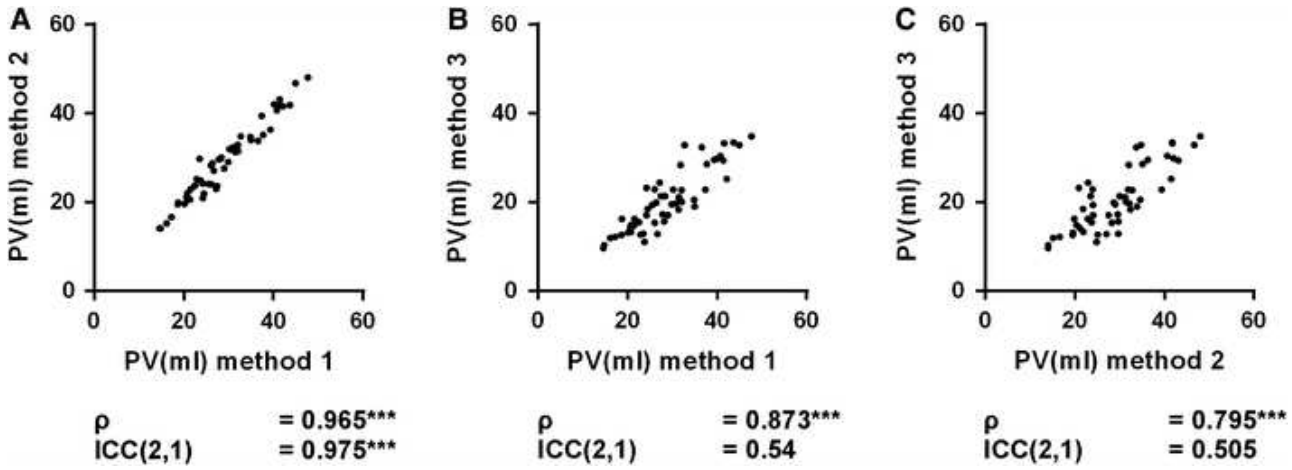


Fig. 3: Scatter plots a, b, and c are used to show the relation between methods. Relationships between prostate volumes calculated with three different methods (method 1 planimetry, method 2 S SVM, and method 3 prolate ellipsoid) were determined by use of Spearman's rank correlation coefficient (ρ) and the intraclass correlation coefficient (ICC). Significant differences are indicated as: ***p < 0.001

As shown in Fig. 2, the PV assessment based on the prolate ellipsoid formula yielded significantly smaller volumes as compared with both planimetry and S SVM. The

relationship between the PVs estimated by all three methods is shown in Fig. 3. The comparison between method 1 and method 2 depicts a strong correlation resulting in a Spearman's rank correlation coefficient (ρ) of 0.965 ($p < 0.001$). The interrater reliability for method 1 and method 2 readings as expressed by the ICC was 0.975 ($p < 0.001$). Further, the comparison between method 3 and the two other methods shows worse correlation and interrater reliability: the Spearman's rank correlation coefficients were 0.873 ($p < 0.001$) and 0.795 ($p < 0.001$), respectively. The intraclass correlation coefficients (ICC) were 0.54 and 0.505, respectively, neither of which were statistically significant.

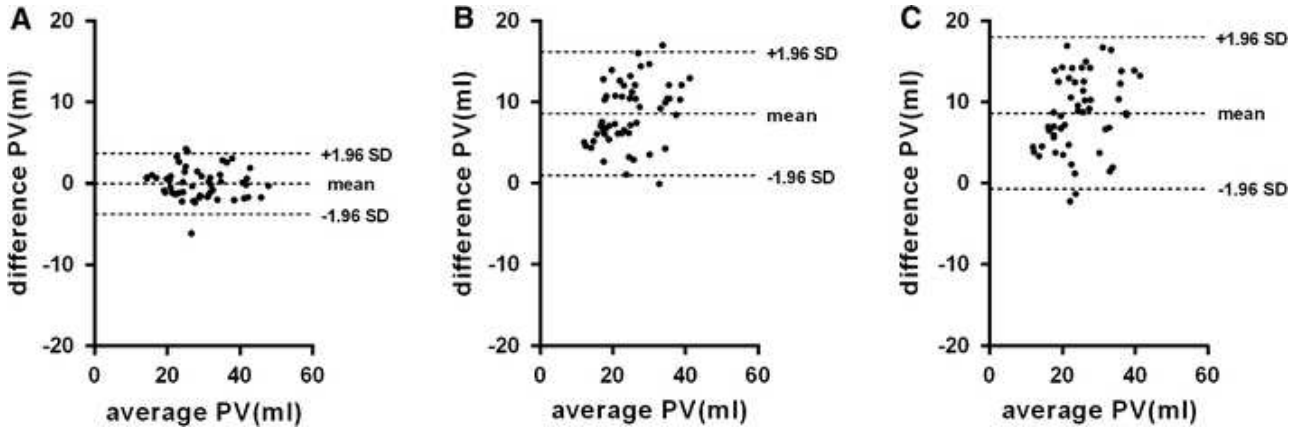


Fig. 4: Bland–Altman plots of prostate volumes calculated with three different methods (method 1: planimetry, method 2: S SVM, and 3: prolate ellipsoid). a method 1 versus method 2, b method 1 versus method 3, and c method 2 versus method 3

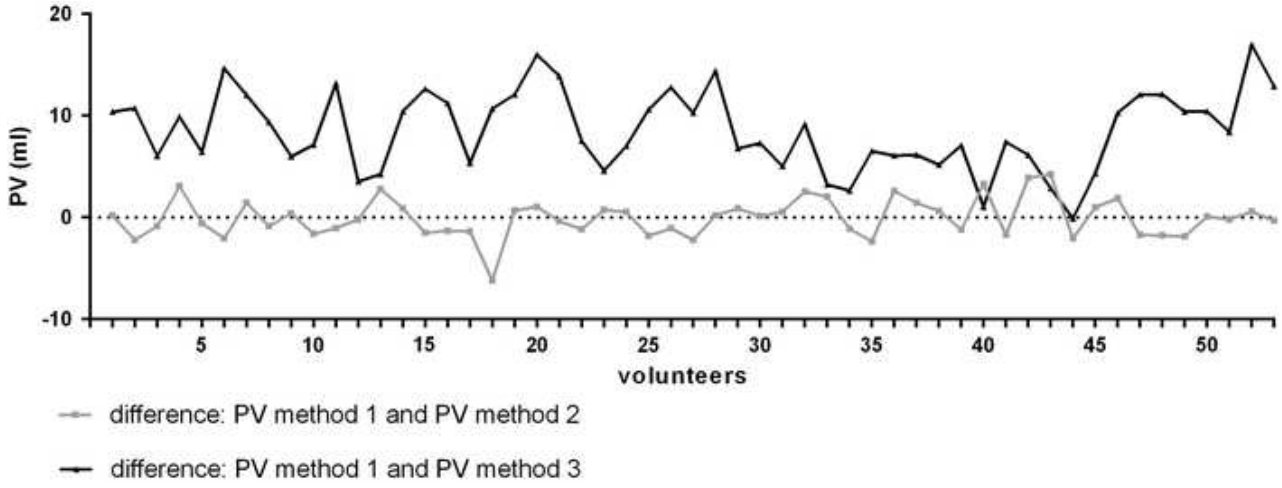


Fig. 5: Prostate volume estimation error values (method 1 planimetry, method 2 S SVM, and method 3 prolate ellipsoid) were determined. Reference is planimetry measurement

The Bland–Altman plot for assessing agreement is also presented in Fig. 4. The mean difference between method 1 and method 2 was −0.05 ml. The limits of agreement with the 95 % confidence interval were −3.8–3.7 ml, which were small enough to show a good agreement between the two volumetric methods. The plot, which compares method 3 and the two other methods, shows a worse agreement with mean difference of 8.6 ml (95 % confidence interval of 1.0–16.2 ml) and 8.6 ml (95 % confidence intervals were −0.7–18.0 ml), respectively. The error values in PV estimation for method 2 and method 3 lie in the range of −6.2–4.2 ml and −0.08–16.99, respectively. Method 3 significantly underestimated PV (Fig. 5).

II. Discussion

Benson et al. [16] introduced the term prostate-specific antigen density (PSAD) and suggested PSAD to help distinguish between BPH and prostate cancer. Recently, Stephan et al. [17] recommended the PSAD, since it can show better performance in prostate cancer detection than the percent free PSA. PSAD is calculated by dividing the serum prostate-specific antigen level by the PV. Additionally, PV has a significant correlation with the serum PSA level and age [3]. Roehrborn et al. [18] revealed an age-dependent log-linear relationship between PV (resulted from either TRUS or MRI modalities) and serum PSA level. Morote et al. [19] predicted PV by free PSA and total PSA. Park et al. predicted PV by PSA mass and free PSA mass [20]. Vesely et al., Morote et al., and Park.et al. used in their studies the ellipsoid formula with the TRUS modality. The assumption of the ellipsoid shape of the prostate can underestimate the PV as shown from our results. Improvements in volume assessment based on either TRUS or MRI modalities could provide more exact relationship and prediction models. Furthermore, the measurement of the PV can be used as both a diagnostic and a treatment response monitoring parameter of BPH [3, 21].

In our present study, we proposed new pattern recognition-based technique for PV calculation. The technique we describe has been developed as post-processing step to estimate the prostate volume based on images obtained in the context of a representative cohort of the general population. In further clinical applications, our method can also be implemented in real-time context and can be integrated in computer-aided diagnosis systems. We assume that the volume derived from manual prostate delineation is closest

to the actual PV, since it is calculated from the actual prostate shape without any shape assumption. We demonstrated that our computerized S SVM-based technique for prostate volume calculation was more accurate than prolate ellipse volume calculation. One possible reason is the poor correlation between traced prostate length in the MRI image (L) and the real prostate length. The PV obtained through our computerized technique had a strong correlation between PV obtained using the planimetry method (Spearman's rank correlation coefficient ρ was 0.965); in contrast, the prolate ellipsoid formula revealed considerably less correlation (Spearman's rank correlation coefficient ρ was 0.795).

Furthermore, the prolate ellipsoid formula compared to planimetry underestimates PV with a mean difference (planimetry–prolate ellipsoid) of 8.6 ml (95 % confidence interval 1.0–16.2 ml). The PV underestimation with the prolate ellipsoid formula has been reported by Mac Mahon et al. [22] for the TRUS modality. Moreover, they obtained limits of agreement similar to those of our MRI study (95 % confidence interval of −12.4–0.6 ml and mean difference prolate ellipsoid–planimetry = −5.9 ml). Note that this comparison was not a direct comparison because different databases, and imaging modalities were used. Kimura et al. [23] introduced a new PV calculation methodology called biplane planimetry, using information of both cross and sagittal sections in TRUS. They reported a volume error of −6.6 ± 8.8 ml (mean ± SD) compared with the gold-standard planimetry [23]. Mac Mahon et al. [22] suggested an alternative formula based on a bullet shape for PV estimation to replace the prolate ellipsoid method for PV estimation in TRUS. The limits of agreement between our computerized volumes and the planimetry volumes were −3.8–3.7 ml, and the mean difference was −0.05 ml. These limits are smaller than the results reported by Mac Mahon et al. [22], where the limits between the newly suggested bullet-shape formula and planimetry volumes were −6.7–9.6 ml, with a mean difference of 1.5. Note that all previous comparisons were not direct comparisons because different databases and imaging modalities were used. There are several parameters to be adjusted in our scheme. They were determined by empirical analysis.

Al-Rimawi et al. [24] reported that PV estimated using the ellipsoid formula correlates well in both TRUS and MRI modalities, but they also reported that MRI gave a significantly larger volume than TRUS, which was due to larger values for the cephalocaudal and anteroposterior diameters. Weiss et al. [25] found that PV estimation with TRUS and endorectal MRI is highly correlated. They also reported differences in average of prostate

volume of only 1.7 ml. The higher volume was measured with TRUS. Turkbey et al. [26] reported that prostate MRI is able to document age-related changes in prostate zonal volumes. Their results suggest a role for MRI in measuring accurate prostate zonal volumes.

One limitation of the present study is the absence of direct comparison between planimetry PV measurement resulted from TRUS and MRI, but this was due to ethical consent restrictions of the SHIP study. The method was initially developed in the context of a population-based cohort. PV assessment based on TRUS is clearly the present clinical standard. Prior to integration in any clinical application, our method needs to be validated in direct comparison with TRUS modality, which is practical and efficient tool for prostate volumetry.

II. Conclusion

The increasing need for valid methods for prostate volume estimation leads to demands for novel accurate approaches. Computerized prostate volumetry may provide an alternative to the present clinically wide spread ellipsoid formula-based method. In this study, we developed a method for prostate volumetry applied to MR images by employing single-class support-vector machines, a technique recently introduced into the field of pattern recognition. Prostate volumes obtained by our technique agreed excellently with the reference (planimetry) method. The promising results with respect to accuracy indicate considerable potential in clinical application, but further validation against TRUS modality is still required.

II. Acknowledgments

The present work has been supported by a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the “Center of Knowledge Interchange” program of Siemens AG.

II. Conflict of interest

No conflict of interest to be declared.

II. References

1. Zoltan E, Lee R, Staskin DR et al (2008) Combination therapy for benign prostatic hyperplasia: does size matter? Curr Bladder Dysfunct Rep 3(2):102–108

2. Stamey TA, Yang N, Hay AR et al (1987) Prostate specific antigen as a serum marker for adenocarcinoma of the prostate. N Engl J Med 317(15):909–916

3. Vesely S, Knutson T, Damber JE et al (2003) Relationship between age, prostate volume, prostate-specific antigen, symptom score and uroflowmetry in men with lower urinary tract symptoms. Scand J Urol Nephrol 37(4):322–328

4. Kim JM, Song PH, Kim HT et al (2011) Effect of obesity on prostate-specific antigen, prostate volume, and international prostate symptom score in patients with benign prostatic hyperplasia. Korean J Urol 52(6):401–405

5. Milonas D, Trumbeckas D, Juska P (2003) The importance of prostatic measuring by transseptal ultrasound in surgical management of patients with clinically benign prostatic hyperplasia. Medicina (Kaunas) 39(9):860–866

6. Völzke H, Alte D, Schmidt CO et al (2011) Cohort profile: the study of health in Pomerania. Int J Epidemiol 40(2):294–307

PubMedCrossRef

7. Hegenscheid K, Kühn JP, Völzke H et al (2009) Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. Rofo 181(8):748–759

8. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9(7):671–675

9. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

10. Schölkopf B, Smola AJ (2001) Learning with Kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Massachusetts (USA)

11. Habes M, Schiller T, Rosenberg C et al (2013) Automated prostate segmentation in whole-body MRI scans for epidemiological studies. Phys Med Biol 58:5899–5916

12. Schölkopf B, Platt JC, Shawe-Taylor J et al (2001) Estimating the support of a high-dimensional distribution. Neural Comput 13(7):1443–1471

13. Bland MJ, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327(8476):307–310

14. Shrout PE, Fleiss JL (1979) Estimating the support of a high-dimensional distribution. Psychol Bull 86(2):420–428

15. Portney LG, Watkins MP (1993) Foundations of Clinical Research: applications to Practice Pearson/Prentice Hal, USA

16. Benson M, Whang I, Pantuck A et al (1992) Prostate septic antigen density: a means of distinguishing benign prostatic hypertrophy and prostate cancer. J Urol 147(3 Pt 2):815

17. Stephan C, Stroebel G, Heinau M et al (2005) The ratio of prostate-septic antigen (PSA) to prostate volume (PSA density) as a parameter to improve the detection of prostate carcinoma in PSA values in the range of <4 ng/ml. Cancer 104(5):993

18. Roehrborn CG, Boyle P, Gould AL, Waldstreicher J (1999) Serum prostate-specific antigen as a predictor of prostate volume in men with benign prostatic hyperplasia. Urology 53(3):581–589

19. Morote J, Encabo G, Lopez M, de Torres IM (2000) Prediction of prostate volume based on total and free serum prostate–specific antigen: is it reliable? Eur Urol 38(1):91–95

20. Park T, Chae JY, Kim JW et al (2013) Prostate-specific antigen mass and free prostate-specific antigen mass for predicting the prostate volume of Korean men with biopsy-proven benign prostatic hyperplasia. Korean J Urol 54(9):609–614

21. Di Silverio F, Bosman C, Salvatori M et al (2005) Combination therapy with rofecoxib and finasteride in the treatment of men with lower urinary tract symptoms (LUTS) and benign prostatic hyperplasia (BPH). Eur Urol 47(1):72–79

22. MacMahon PJ, Kennedy AM, Murphy DT et al (2009) Modified prostate volume algorithm improves transseptal US volume estimation in men presenting for prostate brachytherapy. Radiology 250(1):273–280

23. Kimura A, Kurooka Y, Kitamura T et al (1997) Biplane planimetry as a new method for prostatic volume calculation in transseptal. Int J Urol 4(2):152

24. Al-Rimawi M, Griffiths DJ, Boake RC et al (1994) Transseptal ultrasound versus magnetic resonance imaging in the estimation of prostatoc volume. BJU Int 74(5):596–600

25. Weiss BE, Wein AJ, Malkowicz SB et al (2012) Comparison of prostate volume measured by transrectal ultrasound and magnetic resonance imaging: is transrectal ultrasound suitable to determine which patients should undergo active surveillance? Urol Oncol 31(8):1436–1440

26. Turkbey B, Huang R, Vourganti S et al (2012) Age-related changes in prostate zonal volumes as measured by high-resolution magnetic resonance imaging (MRI): a cross-sectional study in over 500 patients. BJU Int 110(11):1642–1647

Acknowledgement

I would like to thank John Wiley & Sons, Inc and the authors of the book “Atlas of human cross-sectional anatomy” for the permission of reproducing part b) of figure 1.

Danksagung

Bedanken möchte ich mich bei meinem Betreuer und Lehrer Herrn:

Prof. Dr. med. Wolfgang Hoffmann, MPH

für die wissenschaftliche Unterstützung und essentielle Orientierung.

Ich danke auch alle Mitglieder der Prostata AG im Universitätsmedizin Greifswald:

Prof. Dr. Martin Burchardt, Dr. Jens Kühn, Dr. Laura Hoppe, Dr. Thilo Schiller.

Ich möchte auch ALLE Kollegen aus dem Institut für Community Medicine für die schöne Arbeitsatmosphäre danken.

Zuletzt möchte ich bei meiner Familie danken, die mich immer unterstützt hat.