Selection of Optimal Hyper-Parameters for Estimation of Uncertainty in MRI-TRUS Registration of the Prostate

Petter Risholm*, Firdaus Janoos, Jennifer Pursley, Andriy Fedorov,
Clare Tempany, Robert A. Cormack, and William M. Wells III

Brigham and Women's Hospital, Harvard Medical School
`pettri@bwh.harvard.edu`

Abstract. Transrectal ultrasound (TRUS) facilitates intra-treatment delineation of the prostate gland (PG) to guide insertion of brachytherapy seeds, but the prostate substructure and apex are not always visible which may make the seed placement sub-optimal. Based on an elastic model of the prostate created from MRI, where the prostate substructure and apex are clearly visible, we use a Bayesian approach to estimate the posterior distribution on deformations that aligns the pre-treatment MRI with intra-treatment TRUS. Without apex information in TRUS, the posterior prediction of the location of the prostate boundary, and the prostate apex boundary in particular, is mainly determined by the pseudo stiffness hyper-parameter of the prior distribution. We estimate the optimal value of the stiffness through likelihood maximization that is sensitive to the accuracy as well as the precision of the posterior prediction at the apex boundary. From a data-set of 10 pre- and intra-treatment prostate images with ground truth delineation of the total PG, 4 cases were used to establish an optimal stiffness hyper-parameter when 15% of the prostate delineation was removed to simulate lack of apex information in TRUS, while the remaining 6 cases were used to cross-validate the registration accuracy and uncertainty over the PG and in the apex.

1 Introduction

In conventional trans-rectal ultrasound (TRUS) guided prostate brachytherapy, low dosage radioactive seeds are permanently implanted throughout the prostate gland (PG). However, as the majority of prostate cancer cases are confined to the peripheral zone (PZ), partial gland implants targeting the PZ, which minimize the amount of healthy tissue irradiated, are desirable in treating prostate tumors. However, with TRUS it is difficult to determine the internal prostate substructure [1], and the boundary of the prostate apex (PA) (inferior part of prostate) is difficult to determine because of poor contrast from the constrained TRUS field of view and shadowing artifacts. The limited visibility of the gland in TRUS may lead to less than ideal placement of the radioactive seeds, resulting in iatrogenic complications. MR images obtained with an endorectal coil (ERC) have excellent soft tissue contrast, making them ideal for imaging the gland,

* Corresponding author.

its substructure and adjacent structures. Therefore, non-rigidly registering pre-treatment with the intra-treatment prostate images to enhance the delineation of prostate substructure is clinically important and shown to be feasible with intensity-based [2], model-based [1] and learning-based registration methods [3].

Registration methods typically optimize two terms – a likelihood (data) term and a weighted prior (regularization) term. At poorly visible substructures like the apex, the registration is driven more by the prior model and less by the data. Therefore, it becomes critical to optimize the free hyper-parameters (HPs) (*e.g.* regularization weight) of the registration model in order to minimize the error in their registered delineation [4]. Apart from manually tuning these hyper-parameters, one automated approach is to estimate an optimal value of the HP through minimization of cross-validation error in localizing homologous points in the registered fixed and moving images [4]. However, this approach only uses point-estimates of the registration parameters, and discards the associated uncertainty that a posterior distribution provides, when determining the optimal HP setting.

In contrast, Bayesian registration frameworks [5,6,7] that quantify the posterior distribution on transformation parameters can estimate the uncertainty in registration in terms of credibility intervals or inter-quartile ranges, along with an optimal alignment. For example, Risholm *et al.* [6] estimate the posterior distribution $p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau)$ of the transformation $\mathbf{U}$ using a finite-element model [6] which treats the prostate as a synthetic elastic bio-mechanical object. Here, $\mathcal{P}_{\text{us}}$ and $\mathcal{P}_{\text{mr}}$ are representations of the prostate in TRUS and MR images respectively which are demarcated as part of the standard brachytherapy workflow. Furthermore, assume that $\mathcal{P}_{\text{mr}}$ and $\mathcal{P}'_{\text{us}} = \mathcal{P}_{\text{us}} \cup \mathcal{A}_{\text{us}}$ are the ground truth representations of the prostate, that $\mathcal{A}_{\text{us}}$ represents the missing apex in the TRUS, and that the homologous area of $\mathcal{A}_{\text{us}}$ in the MR image is represented as $\mathcal{A}_{\text{mr}} \subset \mathcal{P}_{\text{mr}}$.

Boltzmann's distribution is used to convert the elastic energy of a deformation into a prior $p(\mathbf{U} \mid \tau)$, where τ is a temperature HP that can be interpreted as an overall inverse pseudo stiffness of the underlying tissue. The value of this temperature HP affects both the spread/uncertainty of the distribution as well as the mode, especially in areas with weak data like the apex. One strategy [5,8] for dealing with HPs in a Bayesian setting is to equip them with a maximum entropy hyper-prior $p(\tau)$ and marginalize them out from the posterior distribution: $p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}) = \int p(\tau)p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau)d\tau$. However, the drawback with marginalization under an over-dispersed hyper-prior is that it may result in an over-dispersed posterior, as compared to a hyper-prior concentrated around an appropriate value of the HP. This effectively implies increased parameter uncertainty due to HP uncertainty, and this effect will be strong in areas with weak data (*e.g.* $\mathcal{A}_{\text{us}}$).

The main contribution of this paper is a technique for HP selection under uncertainty, where we maximize the likelihood of the temperature τ under the posterior predictive distribution $p(\partial\mathcal{A}_{\text{us}} \mid \partial\mathcal{A}_{\text{mr}}, \tau)$ of the apex boundary $\partial\mathcal{A}_{\text{us}}$ in the TRUS image conditioned on its boundary $\partial\mathcal{A}_{\text{mr}}$ in the MR image. This

distribution is obtained from the posterior distribution of the transformation $p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau)$. Consequently, our method identifies a specific value of the temperature HP that maximizes the ability of the posterior distribution in predicting poorly determined structures like the apex while mitigating the effect of HP related uncertainty. The proposed method for estimating optimal HPs is not restricted to MRI-TRUS registration, but is applicable in any situation where we have ground truth homologous correspondences in the two image spaces.

Because the apex is difficult to delineate in TRUS, the method was trained and validated on a data-set of 10 corresponding pre- and intra-treatment prostate MRI, where accurate ground truth delineation of the prostate gland boundary was possible. With the intra-treatment MRI acting as a surrogate for the TRUS, the apex was systematically removed so that the prior model was the main predictor of the apex location. Temperature estimation was carried out on 4 cases, while the remaining 6 cases were used to cross validate the model's ability to predict the location of the apex. These results are presented in Section 3 and are representative for MRI-TRUS registration results assuming that the deformations seen in the pseudo TRUS (*e.g.* intra-treatment MRI) are good approximations of the deformations seen in real TRUS images.

2 Methods and Materials

2.1 Bio-mechanical Prostate Model

As in Risholm *et al.* [6], we model the prostate with a bio-mechanical finite-element (FE) mesh that assumes the prostate $\mathcal{P}_{\text{mr}}$ to be an elastic material with unknown stiffness and compressibility parameters. The tetrahedral FE-mesh consists of N_e tetrahedra and N_v vertices $\mathbf{V} = [\mathbf{v}_1^\top, \dots, \mathbf{v}_{N_v}^\top]$. The boundary vertices $\mathbf{v}_i = [x_i, y_i, z_i]^\top$ are identified as $\mathbf{V}_b = [\mathbf{v}_1^\top, \dots, \mathbf{v}_{N_b}^\top]$ with $\mathbf{v}_i \in \partial\mathcal{P}_{\text{mr}}$, where $N_b \leq N_v$. Each boundary vertex is associated with a displacement vector $\mathbf{u}_i = [u_i, v_i, w_i]^\top$ such that $\mathbf{v}_i[u_i] = \mathbf{v}_i + \mathbf{u}_i \in \partial\mathcal{P}_{\text{us}}$. Once the deformation of the boundary $\mathbf{U} = [\mathbf{u}_1^\top, \dots, \mathbf{u}_{N_b}^\top]^\top$ that aligns $\partial\mathcal{P}_{\text{mr}}$ with $\partial\mathcal{P}_{\text{us}}$ is determined, the displacement of the internal nodes under the linear elastic model is determined by solving a linear system of equations. Associated with each pair of vertices $\mathbf{v}_i$ and $\mathbf{v}_j$ is a stiffness matrix $\mathbf{K}_{ij}$ which depends on Young's modulus ($E = 1/\tau$) and Poisson's ratio (ρ) of the underlying tissue. Because no boundary forces are applied, E is a unit-less pseudo stiffness. The linear elastic (band) energy for the deformation of the prostate boundary is defined as $E_{\text{el}}(\mathbf{U}) = \sum_{i=1}^{N_b} \sum_{j \in \mathcal{B}_i} \mathbf{u}_i \mathbf{K}_{ij} \mathbf{u}_j$, where $\mathcal{B}_i$ is the set of vertices in the Markov neighborhood of vertex i .

2.2 Bayesian Estimation of Boundary Conditions

Under the posterior distribution $p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau) \propto p(\mathcal{P}_{\text{us}} \mid \mathbf{U}, \mathcal{P}_{\text{mr}})p(\mathbf{U} \mid \tau)$, the likelihood term corresponds to the distance of the deformed boundary $\partial\mathcal{P}_{\text{mr}} \circ \mathbf{U}$ from the TRUS boundary $\partial\mathcal{P}_{\text{us}}$, and is a normal distribution on the ℓ_2 distance from $\partial\mathcal{P}_{\text{us}}$:

$$p(\mathcal{P}_{\text{us}} \mid \mathbf{U}, \mathcal{P}_{\text{mr}}) = \mathcal{N}(\|\partial\mathcal{P}_{\text{mr}} \circ \mathbf{U} - \partial\mathcal{P}_{\text{us}}\|_2^2 \mid 0, \sigma^2), \quad (1)$$

where σ^2 models the uncertainty of the position of the boundary $\partial\mathcal{P}_{\text{us}}$.

The prior distribution on deformations uses Boltzmann's distribution on the linear elastic energy $E_{\text{el}}(\mathbf{U})$ parametrized by the temperature τ so the log posterior becomes

$$\log p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau) = -\frac{1}{\sigma^2} \sum_{i=1}^{N_b} \Delta^2(\mathbf{v}_i[u_i]; \partial\mathcal{P}_{\text{us}}) - \frac{E_{\text{el}}(\mathbf{U})}{\tau} + \text{const}, \quad (2)$$

where Δ is the (signed) ℓ_2 distance of vertex $\mathbf{v}_i[u_i]$ from $\partial\mathcal{P}_{\text{us}}$.

The Boltzmann's temperature controls the overall stiffness of the elastic material (*i.e.* it uniformly modulates all stiffness matrices $\mathbf{K}_{ij}$) and is, in general, a priori unknown. Higher temperatures correspond to lower overall stiffness, increases the spread of both the prior and hence the posterior distribution, while lower temperatures indicate stiffer materials with tight spreads of the distributions. Furthermore, this temperature also affects the location of the mode of the distributions.

The posterior distribution is characterized non-parametrically using Metropolis Hastings (MH) Markov Chain Monte Carlo (MCMC) sampling. Starting from an initial deformation $\mathbf{U}_0$, a proposal boundary deformation $\mathbf{U}_*$ is generated from a multi-variate normal proposal distribution and accepted ($\mathbf{U}_n = \mathbf{U}_*$) or rejected ($\mathbf{U}_n = \mathbf{U}_{n-1}$) according to the MH-criterion. After convergence of the MCMC chain, assessed with the Geweke criterion, the collection of posterior samples $\mathcal{U} = \{\mathbf{U}_1, \dots, \mathbf{U}_{N_s}\}$ characterizes the posterior distribution in Eqn. (2).

2.3 Temperature Estimation

This section describes how to estimate the temperature HP through maximization of the posterior predictive distribution given a representative training data-set with ground truth markers, either in the form of homologous points, contours or surfaces. In this paper, these markers are the manually demarcated boundaries of the prostate apex, $\partial\mathcal{A}_{\text{us}} \subset \partial\mathcal{P}'_{\text{us}}$ and $\partial\mathcal{A}_{\text{mr}} \subset \partial\mathcal{P}_{\text{mr}}$, extracted from pre- and intra-treatment prostate MRI, where the intra-treatment MRI is used as a surrogate for TRUS. Note that the apex boundary $\partial\mathcal{A}_{\text{us}}$ is not used to estimate the posterior distribution on deformations, but is used only post-registration in the temperature estimation step. Therefore, the distribution of the location of the apex boundary $\partial\mathcal{A}_{\text{mr}}^{\text{us}}$ as *predicted by* the MR image is:

$$p(\partial\mathcal{A}_{\text{mr}}^{\text{us}} \mid \partial\mathcal{A}_{\text{mr}}, \tau) = \int p(\partial\mathcal{A}_{\text{mr}}^{\text{us}} \mid \mathbf{U}, \partial\mathcal{A}_{\text{mr}}) p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau) d\mathbf{U}, \quad (3)$$

where $p(\partial\mathcal{A}_{\text{mr}}^{\text{us}} \mid \mathbf{U}, \partial\mathcal{A}_{\text{mr}}) = \delta(\|\partial\mathcal{A}_{\text{mr}} \circ \mathbf{U} - \partial\mathcal{A}_{\text{mr}}^{\text{us}}\|_2^2)$ is modeled with a Dirac delta which is zero everywhere except where, after applying a transformation $\mathbf{U}$, the apex boundary from the MR image $\partial\mathcal{A}_{\text{mr}}$ overlaps with the candidate apex boundary $\partial\mathcal{A}_{\text{mr}}^{\text{us}}$. The posterior probability of a particular deformation $p(\mathbf{U} \mid \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau)$ is given by the set of deformation samples generated through MCMC (§Section 2.2) using non-parametric Kernel Density Estimation (KDE).

Therefore, integrating out $\mathbf{U}$ in Eqn. (3), and through the sifting property of the Dirac delta function, the kernel estimate of predictive density becomes:

$$p(\partial\mathcal{A}_{\text{mr}}^{\text{us}} \mid \partial\mathcal{A}_{\text{mr}}, \tau) = \frac{1}{N_s} \sum_{s=1}^{N_s} k_h \left(\|\partial\mathcal{A}_{\text{mr}} \circ \mathbf{U}_s - \partial\mathcal{A}_{\text{mr}}^{\text{us}}\|_2^2 \right), \quad (4)$$

where k is a symmetric positive kernel with bandwidth h , mean 0 and integral 1, and $\mathbf{U}_s$ is the s -th sample from the MCMC chain of N_s samples. Now, given a training data-set of D_t pairs of ground-truth prostate delineations from pre- and intra-treatment images, and assuming independence conditioned on the temperature, the *maximum likelihood estimate* (MLE) of the temperature HP from the posterior predictive distribution is:

$$\tau^* = \underset{\tau}{\operatorname{argmax}} \prod_{d=1}^{D_t} p(\mathcal{A}_{\text{mr}}^{\text{us}} = \mathcal{A}_{\text{us}}^{(d)} \mid \mathcal{A}_{\text{mr}}^{(d)}, \tau), \quad (5)$$

where $\mathcal{A}_{\text{us}}^{(d)}$ is the manually demarcated boundary of the apex in the surrogate TRUS image. Hence, we choose the model temperature τ such that the ground truth position of the apex has highest predictive probability.

2.4 Patient Data

In this evaluation, intra-treatment MRI was used as a proxy for TRUS because it permitted an accurate demarcation of the apex boundary. The patient data, collected from $D = 10$ patients undergoing prostate biopsy, was acquired as part of a prospective clinical research protocol which is HIPPA compliant and approved by the local institutional review board. For each patient, a diagnostic MRI ($512 \times 512 \times 30$ with a pixel size of $0.3 \times 0.3 \times 3 \text{ mm}^3$) and an intra-treatment MRI ($320 \times 320 \times 40$ with pixel size $0.5 \times 0.5 \times 3 \text{ mm}^3$) used for biopsy guidance were acquired. There is significant change in the prostate between the two images due to the use of an ERC when acquiring the diagnostic MRI and different patient positioning during imaging. The PGs were segmented for each patient on both the pre- and intra-treatment MR images by an abdominal radiologist with over 10 years of experience in prostate MR interpretation.

3 Results

The label-maps of the pre- and intra-treatment segmentations were rigidly registered, and the manually contoured boundaries were smoothed through anti-aliasing followed by Gaussian smoothing with variance of 1 mm^2 . For each of the D cases, a tetrahedral FE-mesh of $\mathcal{P}_{\text{mr}}^{(d)}$ was created. We also removed the apex $\mathcal{A}_{\text{us}}^{(d)}$, corresponding to approximately 15% of the prostate volume, from the ground truth label-map of the surrogate TRUS image such that $\mathcal{P}_{\text{us}}^{(d)} = \mathcal{P}_{\text{us}}^{(d)'} \setminus \mathcal{A}_{\text{us}}^{(d)}$, in order to evaluate the ability of the model to correctly predict the location of the apex in the total absence of image contrast.

The likelihood variance was set to the maximum pixel spacing (3mm) to model the boundary label uncertainty, and the prostate was modeled as a nearly incompressible material ($\rho = 0.45$). The average number of FE-vertices over the 10 cases was $\bar{N}_v = 415$. In all MCMC runs, 2×10^6 samples were generated, which after thinning with a factor of 10 and discarding the first 25% of the chain for burn-in, resulted in $N_s = 1.5 \times 10^5$ remaining samples. The total running time for one chain was ≈ 4 hrs.

3.1 Temperature Estimation

We first performed MLE (§Section 2.3) for the temperature HP on $D_t = 4$ randomly chosen data-sets. To optimize Eqn. (3), the temperature τ was discretized into 10 bins over the range $20 \leq \tau \leq 380$ and the predictive density (§Eqn. (4)) was estimated for each temperature and data-set in parallel (4×10 parallel computations). Because of the high number of samples, the particular choice of kernel and bandwidth in Eqn. (4) has a negligible effect on the result – we used a standard normal kernel and found the optimal temperature to be $\tau^* = 180$. Figure 1 shows the marginal posterior distribution of the deformed prostate in relation to the ground truth at three different temperature settings and shows that uncertainty plays an important role in selecting HPs.

Fig. 1. Coronal slices of the marginal posterior probability $p(\partial\mathcal{P}_{\text{mr}}^{\text{us}} | \phi, \theta)$ of $\partial\mathcal{P}_{\text{mr}}^{\text{us}}$ crossing a ray, with spherical coordinates ϕ and θ , originating from the prostate center at three different temperatures for one of the 4 training cases. The region below the horizontal line denotes $\mathcal{A}_{\text{us}}$, and the black outline delineates $\partial\mathcal{P}_{\text{us}}'$. For each temperature we compare the predicted log probability $\Xi(\tau) = \log p(\partial\mathcal{A}_{\text{us}} | \partial\mathcal{A}_{\text{mr}}, \tau)$ with the conventional Expected Mean Squared Error (EMSE): $\Lambda(\tau) = \frac{1}{N_b} \sum_{i=1}^{N_b} \int \Delta^2(\mathbf{v}_{u_i}, \partial\mathcal{A}_{\text{us}}) p(\mathbf{U} | \mathcal{P}_{\text{us}}, \mathcal{P}_{\text{mr}}, \tau) d\mathbf{U}$. **Fig. (a):** $\Xi(20) \approx -\infty$ and $\Lambda(20) = 22.7$. **Fig. (b):** $\Xi(180) = -69.8$ and $\Lambda(180) = 37.9$. **Fig. (c):** $\Xi(340) = -73.1$ and $\Lambda(340) = 66.3$. The trend, which is confirmed by the 3 other training data-sets, is that EMSE is overly optimistic about distributions with low temperatures, and that these distributions, as shown in Fig. (a), are implausible when evaluated on the training data. Conversely, the MLE chooses the temperature ($\tau^* = 180$) that maximizes the posterior predictive distribution on the training data.

3.2 Evaluation of Method

The accuracy and precision of the HP estimation method was evaluated using the 6 remaining cases. The predictive distribution of both the apex boundaries and the remaining prostate gland were characterized for each case at τ^* , and error in boundary prediction as well as the inter-quartile range (IQR) on the distance errors were computed, as listed in Table 1. Figure 2 shows the marginal posterior distribution of the prostate, at optimum temperature, which conveys the mode as well as the corresponding uncertainty and is compatible with the ground-truth delineation of the prostate.

Fig. 2. Results from the evaluation using the optimal temperature ($\tau^* = 180$). Coronal slices of the marginal posterior probability $p(\partial\mathcal{P}_{\text{mr}}^{\text{us}} | \phi, \theta)$ of $\partial\mathcal{P}_{\text{mr}}^{\text{us}}$ crossing a ray, with spherical coordinates ϕ and θ , originating from the prostate center. The ground truth boundary of the prostate $\partial\mathcal{P}_{\text{us}}'$ is overlaid as a black contour. The area below the horizontal line delineates $\mathcal{A}_{\text{us}}$.

Table 1. Accuracy and precision of the posterior distribution at optimal temperature value for PA ($\partial\mathcal{A}_{\text{mr}}^{\text{us}}$) and the rest of the PG ($\partial\mathcal{P}_{\text{mr}}^{\text{us}} \setminus \partial\mathcal{A}_{\text{mr}}^{\text{us}}$). **Note:** All values in mm. For each case we estimated the mode $\hat{\mathbf{u}}$ of the posterior distribution. The mean and maximum errors are the mean and maximum absolute distances of boundary vertices deformed by $\hat{\mathbf{u}}$ from $\partial\mathcal{P}_{\text{us}}'$ measured by the Δ metric. Also tabulated is the inter-quartile range (IQR) of the distribution of the distances $\{\Delta(\mathbf{v}_i \circ \mathbf{u}_i^s; \partial\mathcal{P}_{\text{us}}')\}_{i \in \partial\mathcal{P}_{\text{mr}}, s=1, \dots, N_s}$.

Case:	1	2	3	4	5	6
Apex Mean Error	3.9	5.5	5.0	4.7	4.7	4.5
Apex Max Error	8.5	12.9	11.6	12.3	10.7	15.4
Apex IQR	6.6	8.1	7.4	7.4	7.6	8.8
Prostate Mean Error	1.6	1.7	1.8	1.7	1.9	1.6
Prostate Max Error	14.1	12.7	11.5	15.0	12.9	11.4
Prostate IQR	3.8	4.3	4.6	4.1	5.4	3.6

4 Discussion

With a Bayesian method, we estimate the posterior distribution on boundary conditions of an elastic model of the prostate extracted from MRI that will align it with the prostate delineated in TRUS. Because the posterior distribution, and in particular the posterior prediction of the location of the prostate apex, is sensitive to the pseudo stiffness (prior) HP, we maximize the posterior predictive distribution with regards to the HP given homologous ground truth apex delineation. The registration method, with optimal HPs, was evaluated on 6 MRI-MRI datasets where the mean error on the boundary was found to be 1.7mm. Comparable results are reported on homologous landmarks, *e.g.* using intensity-based registration [2] with a mean error of 1.5mm and the model-based approach of Hu *et al.* [1] with median RMS of 2.4mm. However, the advantage of our method is that we prescribe error-bars on the registration results, in terms of IQRs, and that the true prostate boundary is within these error-bars.

Acknowledgments. This work was supported by NIH grants P41EB015898, R01CA111288, P41RR019703, P41RR013218 and P41EB015902.

References

1. Hu, Y., Ahmed, H.U., Taylor, Z., Allen, C., Emberton, M., Hawkes, D., Barratt, D.: MR to ultrasound registration for image-guided prostate interventions. *Med. Image Anal.* 16(3), 687–703 (2012)
2. Fedorov, A., Tuncali, K., Fennessy, F.M., Tokuda, J., Hata, N., Wells, W.M., Kikinis, R., Tempany, C.M.: Image registration for targeted MRI-guided transperineal prostate biopsy. *J. Magn. Reson. Imaging* (2012)
3. Shi, Y., Liao, S., Shen, D.: Learning statistical correlation for fast prostate registration in image-guided radiotherapy. *Med. Phys.* 38(11), 5980–5991 (2011)
4. Yeo, B.T.T., Sabuncu, M.R., Vercauteren, T., Holt, D.J., Amunts, K., Zilles, K., Golland, P., Fischl, B.: Learning task-optimal registration cost functions for localizing cytoarchitecture and function in the cerebral cortex. *IEEE Trans. Med. Imaging* 29(7), 1424–1441 (2010)
5. Simpson, I.J., Schnabel, J.A., Groves, A.R., Andersson, J.L., Woolrich, M.W.: Probabilistic inference of regularisation in non-rigid registration. *NeuroImage* 59(3), 2438–2451 (2012)
6. Risholm, P., Fedorov, A., Pursley, J., Tuncali, K., Cormack, R., Wells, W.M.: Probabilistic non-rigid registration of prostate images: Modeling and quantifying uncertainty. In: *IEEE Int. Symp. Biomed. Imaging*, pp. 553–556 (2011)
7. Gee, J.C., Bajcsy, R.K.: Elastic matching: Continuum mechanical and probabilistic analysis. In: *Brain Warping*, pp. 183–197. Academic Press (1999)
8. Janoos, F., Risholm, P., Wells, W.M.: Robust non-rigid registration and characterization of uncertainty. In: Zhou, K., Duncan, J.S., Ourselin, S. (eds.) *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, vol. 1, pp. 4321–4326 (2012)