Shreyan Wankavala 112634232

# Global Registration and Segmentation Framework With Application to Magnetic Resonance Prostate Imagery

Written By: Yi Gao*, *Student Member, IEEE*, Romeil Sandhu, *Student Member, IEEE*, Gabor Fichtinger, *Member, IEEE*, and Allen Robert Tannenbaum, *Fellow, IEEE*

Prostate cancer ranks among one of the most widespread of all cancers in the U.S. male population. Finding a more efficient method to extract prostates from MR imagery would then be a great contribution, as it is normally a very challenging and important task for medical analysis and surgical planning. That is exactly what the authors of this paper sought out to do, as they developed a unified shape-based framework to aid in the process.

Their contribution with this paper was considered three-fold. Normally, shape-based segmentation is a two part problem; one must properly align a set of training shapes first, so that any variation in shape is not due to pose. Only then, can the segmentation of the images be performed. This issue of registration will be addressed by representing the shapes of a training set as point clouds, later referred to as point-sets. By doing this, the more global aspects of registration will be addressed through a certain particle filtering process. Next, the cost functional used by the authors used the image statistics localized in a banded region around the contours of an image. It will aid segmentation both globally and locally, without adding any extra computation when compared to only global processes. Lastly, the two processes listed will be combined in a shape-based segmentation framework and the prostates will be successfully extracted from MR imagery. In this paper, the act of representing images as point-sets will first be thoroughly explained. These images will then be registered under a particle filtering framework. The shape prior will be constructed, and will be combined with local image statistics to perform the segmentation. This process will be used in a real experiment on clinical data sets to highlight its efficiency in dealing with the arduous task of prostate registration and the overall segmentation task.

To begin the registration process, the image needs to be represented as a set of random samples rather than as a discrete function. In the paper, this is done by assuming that the image is represented by a continuous function $f$ on the compact image domain $\Omega$. We will also define it to be a nonnegative function. The point-set representation (PSR) for images is chosen over the

discrete function representation (DFR). This is because in a discrete function representation, a given image would be represented as a discrete function defined on a uniform grid. Each value is associated with each spatial sampling location, more specifically the image intensity. This is not advantageous, as using the PSR gives the same outcome as the DFR, while handling certain registration issues, like naturally handling the case of large translations between the two given images, with a much shorter computation time (under DFR, the traversing of the domain grids could be very time consuming. In contrast, PSR represents the image by far fewer points lowering the computation time). Both methods are fully equivalent, so no information will be lost. In the paper, the function used is given below. This normalization allows

$$\begin{cases} \tilde{f}(\boldsymbol{x}) \geq 0, \quad \forall \boldsymbol{x} \in \Omega \\ \int_{\Omega} \tilde{f}(\boldsymbol{x}) d\boldsymbol{x} = 1 \end{cases}$$

us to treat the function *f* as a PDF defined on the domain *Ω*. This is done, so we can then represent *f* by drawing samples from the distribution. A rejection sampling algorithm is adopted to acquire the corresponding set of points for our PSR. *P* and *G* are shown. Another advantage of

$$\mathcal{P} = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_M : \boldsymbol{p}_i = (p_i^x, p_i^y, p_i^z) \in \Omega\}$$

$$\mathcal{Q} = \{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N : \boldsymbol{q}_i = (q_i^x, q_i^y, q_i^z) \in$$

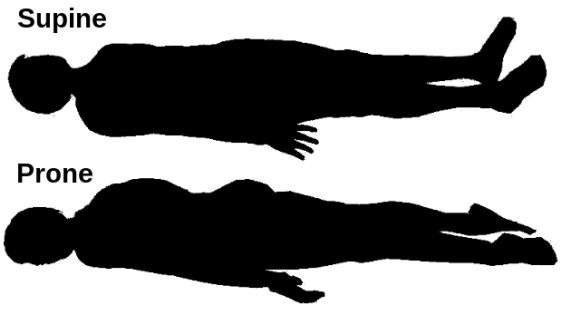
using the PSR is also apparent, as we can now represent the image as purely 2D or 3D points, unlike DFR, where a real number or integer is associated with each spatial position.The two images of *P* and *G* are then registered by aligning their corresponding point-sets using the following:

$$E(\boldsymbol{A}, \boldsymbol{t}) := \frac{1}{N} \sum_{i=1}^{N} \|\boldsymbol{A}\boldsymbol{q}_i + \boldsymbol{t} - Cl(\boldsymbol{A}\boldsymbol{q}_i + \boldsymbol{t})\| + \frac{\lambda}{\det^2(\boldsymbol{A})}$$

where in this algorithm, A ∈ R^(3X3), det(A) ≠ 0 is the affine transformation matrix, t ∈ R^3 is the translation vector, and *Cl* : R^3→P maps a point in R^3 to the closest point in P.

We then move onto shape registration via particle filtering to complete the registration process. The PSR is a local optimization procedure. This means that although large translations

are effectively handled, large rotations are not. Large rotations are common especially in prostate registration where the supine and prone views of the prostate need to be registered. We can treat




the registration problem as a system parameter estimation task since supine/prone registration cases commonly have optimal rotations close to 0° or 180°, and we have this as prior, or *a priori* knowledge. In the parameter estimation task, 12 transformational parameters will constitute the state variables in the dynamic system. The particle filtering system will then allow us to easily combine the *a priori* information. To begin the filtering process, we must first denote the state variable at time *t* as *xt* and the observation as *yt.* The objective is to estimate the distribution of *xt* based on all the observations made until time *t.* The process and observation models are shown.

$$\begin{cases} \boldsymbol{x}_{t+1} = f(\boldsymbol{x}_t, u_t) \\ \boldsymbol{y}_t = g(\boldsymbol{x}_t, v_t) \end{cases}$$

*f* and *g* may be nonlinear functions, while *ut* and *vt* are the process and observation noises. We will also assume that they are both independent in time. The recursive estimation of *p* consists of two steps, which we can call the *prediction* and *update* steps. The *prediction* step gives the prior PDF of *xt* at time *t. δ* denotes the Dirac function here.

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) = \int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1}) d\boldsymbol{x}_{t-1}$$

where

$$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) = \int \delta\left(\boldsymbol{x}_t - f(\boldsymbol{x}_{t-1}, u)\right) p(u) du.$$

At time *t* after the observation *yt* is available, it can then be used to *update* the estimation to obtain the posterior PDF using the function given.

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t)p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})}{p(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})}$$

where

$$p(\boldsymbol{y}_t|\boldsymbol{x}_t) = \int \delta(\boldsymbol{y}_t - h(\boldsymbol{x}_t, v))\, p(v) dv.$$

In the case that *f* and *g* are nonlinear, the analytical result for *p* is rarely available. In this scenario, one can expect to find a numerical approximation to the PDF.

The next step is for us to use affine image registration. The reason for doing this is so we can address it as the parameter estimation task from earlier, and solve it using particle filters. The affine image registration by particle filtering allows for the shearing and scaling of an image, in addition to rotation and translation. In affine registration, the state space *S* is 12D, where the first nine dimensions are for the affine matrix, and the last three dimensions are for translation. Denoting the state vector *x,* as $x \in S$, the model will take the form shown below. Here, the

$$\boldsymbol{x}_{t+1} = \mathcal{R}(\boldsymbol{x}_t + \boldsymbol{u})$$

operator $R : S \rightarrow S$ takes *xt* as the initial configuration, goes on with a few steps of deterministic affine registration, and returns the resulting parameter estimated as *x(t+1)*. The observation model is shown below. The operator *ε* gives the cost function under the state *xt,* and both the

$$\boldsymbol{y}_t = \mathcal{E}(\boldsymbol{x}_t) + \boldsymbol{v}$$

process and observation models are very nonlinear. Lastly, on top of the process and observation models, we also need the prior distribution of *p(x0).* This is normally assumed to be Gaussian without any prior knowledge, but since we know the supine and prone views need to be registered, this will be used to formulate the design of *p(x0).* A summary of the algorithm is shown below, as well as (11) and (12) respectively.

1: Sample from $p(\boldsymbol{x}_0)$ to get $\{\boldsymbol{x}_0(i) : i = 1, \ldots, N\}$

2: **for** $i = 1, 2, \ldots, t$ **do**

3: Obtain the prior samples $\{\boldsymbol{x}_t^*(i)\}$ by the process model.

4: Evaluate the likelihood $q_i$'s using (11)

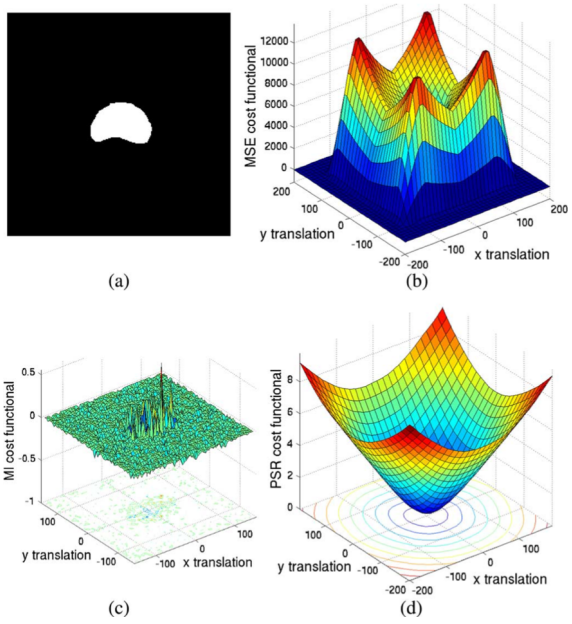
5: Resample to get posterior $\{\boldsymbol{x}_t(i)\}$ using (12)

6: **end for**

$$q_i = \frac{p\left(\boldsymbol{y}_t | \boldsymbol{x}_t^*(i)\right)}{\sum_{j=1}^{N} p\left(\boldsymbol{y}_t | \boldsymbol{x}_t^*(j)\right)}$$

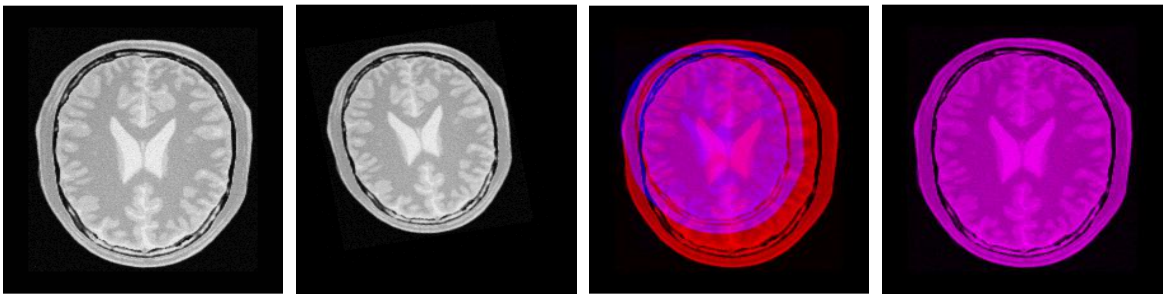
$$\sum_{j=0}^{M-1} q_j < w_j \leq \sum_{j=0}^{M} q_j$$

The researchers perform a series of experiments now that registration and particle filtering are complete. They provide experiments to demonstrate the behavior of different registration cost functionals, the 'robustness' of the proposed method to initialization, the supine/prone prostate registration, and the computer efficiency of their algorithm. The plots of different cost functional values are made with respect to various 2D translations. The test prostate binary image, mean square error (MSE), mutual information (MI), and PSR are all graphed.

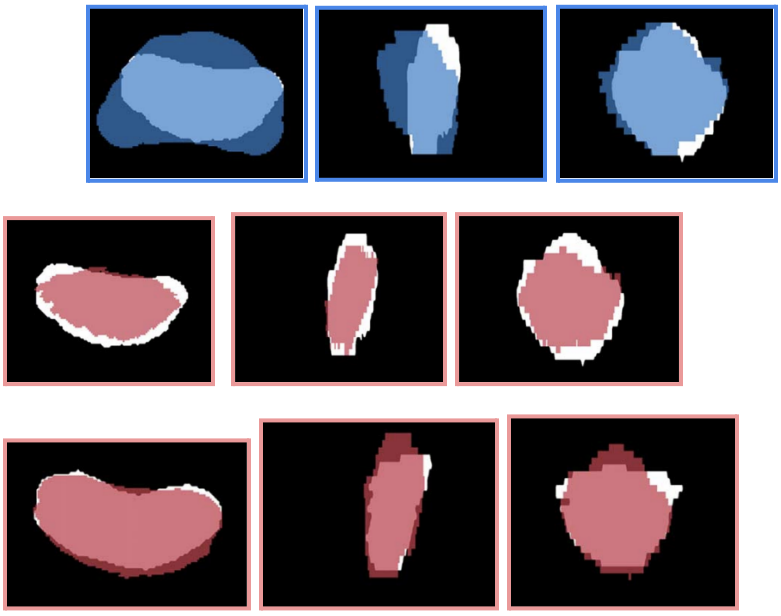



The researchers determine that the MSE is an inadequate cost functional for stochastic probing based on global image registration. This is because not only is the region of convergence of the MSE small relative to that of the PSR, the cost functional values drop to zero when the transitions are large. When looking at the MI, we see that apart from being noisy, the cost functional is flat for most of the regions. This makes the registration process very sensitive to the initialization and the optimization step size. Finally, when looking at the PSR, we see that the minimum of the proposed cost functional is located at the correct position and smoothly grows monotonically outward as translation increases. We can conclude that no matter how large the translation is, the registration process is able to drive it to its correct place.

Examples of affine image registration by particle filtering are shown. The image is first sheared, scaled, rotated, and translated. The mean image before and after registration shows how the image is aligned when it was not before.



In the paper, the picture process of supine-prone prostate registration is shown. A challenging problem is to register these prostates. The moving prostate 3D image in blue is overlaid on the fixed image in white. The results generated by the proposed algorithm are shown by the red images. Here, the axial, sagittal, and coronal views are presented.



Recall the *a priori* knowledge when formulating the rotation for the supine and prone views. The prior distribution can be defined as follows. An important point to note is that in the case such

$$p(\theta_0) = \frac{1}{2\sqrt{2\pi}\sigma}\left(e^{-\frac{\theta^2}{2\sigma^2}} + e^{-\frac{(\theta-\pi)^2}{2\sigma^2}}\right)$$

prior knowledge is not available, a uniform distribution can be used.

The shape prior must be constructed before we can segment the image now that we have registered the shapes. We use a modification of a signed distance function (SDF) via a transformation of the form shown below, to provide a better representation for shapes.

$$s(\boldsymbol{x}) = \mathcal{T}(\mathrm{SDF}(\mathbf{x}))$$

Particularly, we want to choose a mapping of $\tau$ that preserves the zero level set of the SDF and eliminates the large values and variances far from that level set. We can then apply the hyperbolic tangent function to the SDF to represent the shapes.

$$s(\boldsymbol{x}) = \tanh\left(\mathrm{SDF}(\boldsymbol{x})\right) = \frac{e^{2\mathrm{SDF}(\boldsymbol{x})} - 1}{e^{2\mathrm{SDF}(\boldsymbol{x})} + 1}$$

A standard principal component analysis (PCA) is adopted to learn the shapes. To obtain the mean shape, we use the formula below.

$$\bar{s}(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N} s_i(\boldsymbol{x})$$

The mean shape is then subtracted from each shape. We can form the covariance matrix below:

$$C = \frac{1}{N}(\eta_1, \ldots, \eta_N)(\eta_1, \ldots, \eta_N)^T$$

And finally, the singular value decomposition gives the following:

$$C = G\Lambda G^T$$

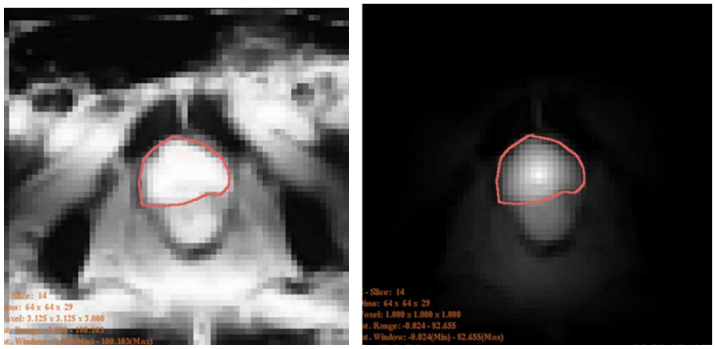
In this formula, $\Lambda$ is a diagonal matrix containing the eigenvalues and the columns of $G$ to store the eigenvectors. These are all reshaped by the original image size and are denoted by $g(x)$. It is kept in mind that besides the shapes of the prostate, the mean and variance of the image intensity within the prostate could also be learned if the training shapes and their corresponding original images are both present.

The registration of the images has been completed, and the shape prior has been found. We can now segment the image under a Bayesian framework to highlight the region of interest. After that, a variational scheme based on local regional information will be used to extract the prostate from the posterior image. Given an image $J(\mathrm{x})$, the likelihood $lJ(\mathrm{x})$ is computed is shown below, where $\mu$ is the mean and $\sigma$ is the standard deviation of the object intensity.

$$l_J(\mathbf{x}) := \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(J(\mathbf{x}) - \mu)^2}{2\sigma^2}\right)$$

The values for the mean and the standard deviation could either be provided by the user or learned during the learning process. To calculate the posterior, the prior term is still needed. While a uniform prior can be used (in this case, the posterior would be normalized), constructing

our own will make the posterior much more convincing as well as the segmentation easier. Below, we can see the computed directional distance map (DDM) of the image. We can see why this is beneficial when comparing the posterior image of a slice using a uniform prior versus using a DDM as a prior. We can differentiate the prostate fairly well, and even though the bright



region below it is difficult to exclude, the shape-based segmentation successfully recognizes it as the background. Now, we will use the image content based directional distance in the prior. What this means is that both the image content and the distance to the cert of the object are considered when calculating the distance. Given *J(x)*, we can construct

$$d(\mathbf{x}, \phi) = |J(\mathbf{x} - J(\mathbf{x} + \phi))|$$

where $\Phi \in A$ *{(0,0,+-1),(0,+-1,0),(+-1,0,0)}*. The DDM, which was shown in the images above, can be calculated by solving the *Hamilton-Jacobi-Bellman* equation.

$$\begin{cases} 0 = \inf\limits_{\phi \in A} \{d(\mathbf{x}, \phi) + \langle \phi, \nabla \psi(\mathbf{x}) \rangle\} \\ \psi(\mathbf{p}_c) = 0. \end{cases}$$

To segment the posterior image, a local regional information based segmentation method is proposed, where the segmentation curve is used to maximize the difference between the average posterior within a region inside and outside of the curve. We aim to minimize the function given below as:

$$
\begin{aligned}
E &:= -\frac{1}{2}(u-v)^2 \\
&= -\frac{1}{2}\left(\frac{\int_\Omega \mathcal{H}_B\left(h(\boldsymbol{x})\right) K(\boldsymbol{x}) d\boldsymbol{x}}{\int_\Omega \mathcal{H}_B\left(h(\boldsymbol{x})\right) d\boldsymbol{x}}\right. \\
&\qquad \left. -\frac{\int_\Omega \mathcal{H}_B\left(-h(\boldsymbol{x})\right) K(\boldsymbol{x}) d\boldsymbol{x}}{\int_\Omega \mathcal{H}_B\left(-h(\boldsymbol{x})\right) d\boldsymbol{x}}\right)^2
\end{aligned}
$$

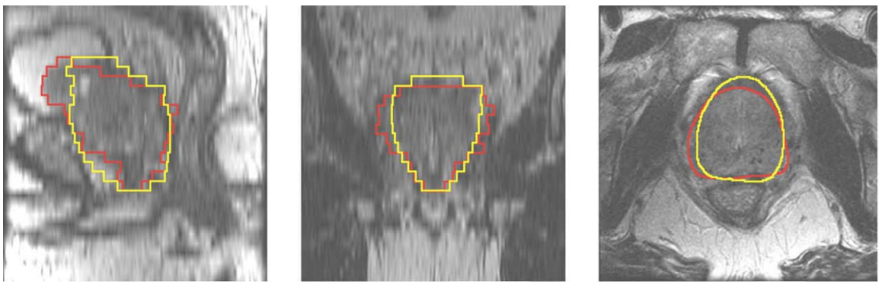
In this function, *h*(x) is the T-SDF representation of the curve, and a banded Heaviside function is also used. Below are the T-SDF and Heaviside functions respectively:

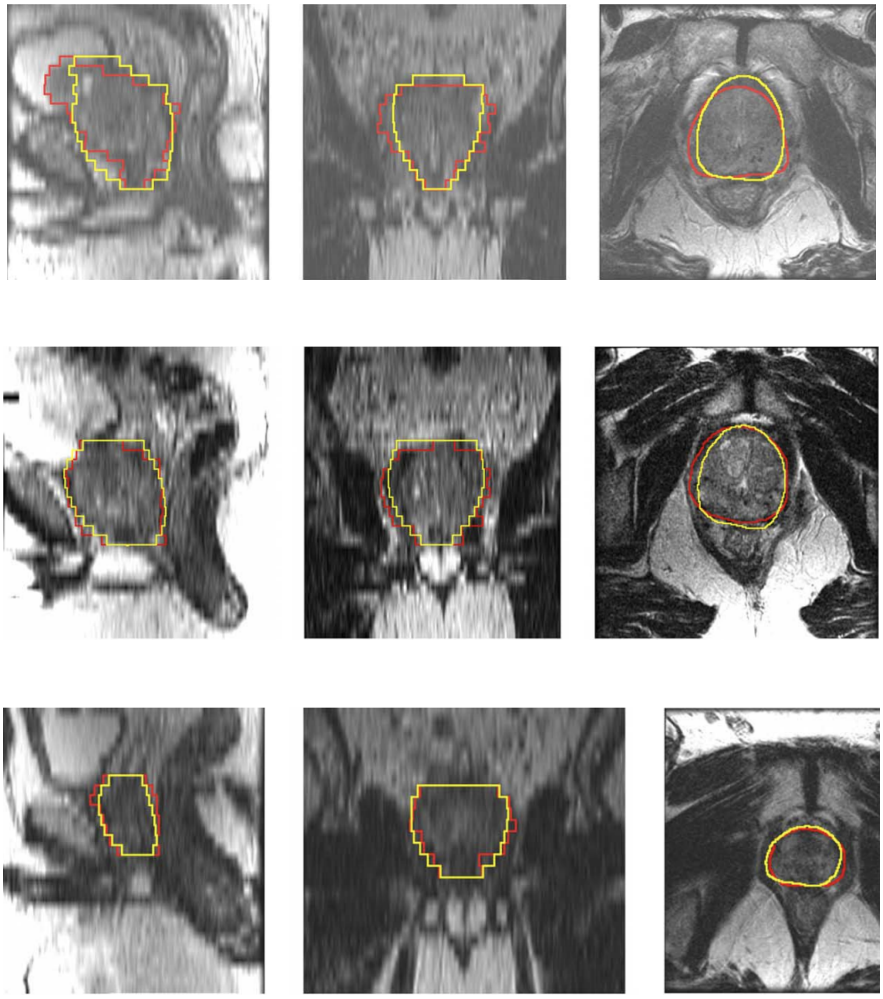
$$
h(\boldsymbol{x}) = \bar{s}(A\boldsymbol{x}+T) + \sum_i^L \omega_i g_i(A\boldsymbol{x}+T)
$$

$$
\mathcal{H}_B(t) = \frac{1}{\pi}\left(\arctan\left(\frac{t}{\epsilon}\right) - \arctan\left(\frac{t-B}{\epsilon}\right)\right)
$$

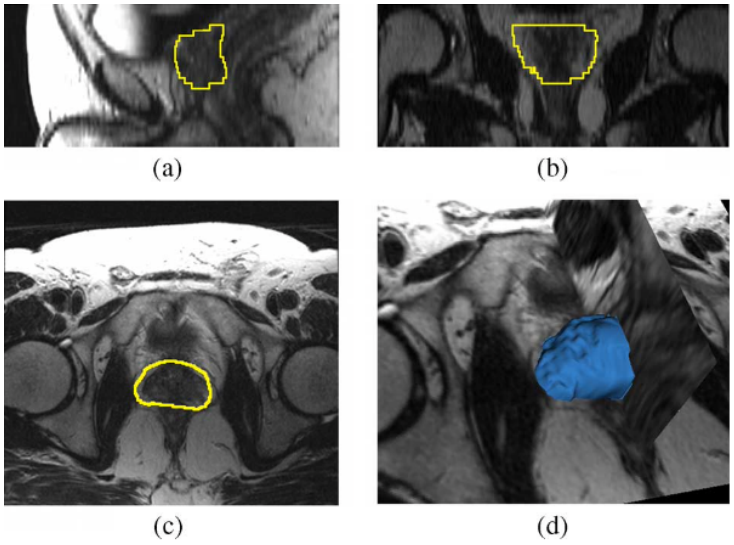
The banded Heaviside function realizes the localized property of the cost functional. From this, we find that the *u* is the mean of the posterior in a banded region *inside* of the object. Just like this, *v* is the mean of the mean of the posterior in a banded region *outside* of the object. We can see that this cost functional is very effective in its influence remote to the curve. By minimizing the segmentation cost functional, the optimal contour and transformation are found. The segmentation is now complete.

In our results, the algorithm is run on a series of clinical data sets to test its effectiveness. The prostate is extracted from the MR imagery of multiple patients. Some of them are shown below:





Above are the segmentation results for multiple patients (sagittal, coronal, and axial views). The orange contour is the expert result, while the yellow is generated by the algorithm. In some regions, the method does not provide perfect results, but in some scenarios like in places of low intensity contrast, it may be difficult even for a clinical expert to define a boundary. Still, it can definitely be said that the method yields a visually excellent segmentation and this indicates that the learned shape prior does have the capacity to represent different shapes, especially when the prostate shape differs from the previous learnt shapes as seen below. It can be seen as more spherical.




We were able to extract the prostate successfully from MR prostate imagery by first, proposing a discrete representation for the images using point sets. Next, a presentation of a training shape-based on the SDF was proposed (constructing the shape prior). Lastly, the images were segmented by using a regional statistical based segmentation process which was focused around the changing contour of the images. By representing an image as a set of points during the registration process, we saw that the registration procedure was much faster, our cost functional behaved better, especially when the translation was large between the two images, and finally by treating the registration problem as a parameter estimation task, global registration of the images was achieved under the particle filtering framework. In the future, this process will be further automized, an example being by better utilizing the learnt centers of the training shapes. In addition to this, the segmentation cost functional will rely on the Bhattacharyya coefficient to better separate the prostate from the background instead of only relying on first order statistics.

## Works Cited

"Affine Registration¶." *Affine Registration - SimpleElastix 0.1 Documentation*, https://simpleelastix.readthedocs.io/AffineRegistration.html.

"A Coupled Global Registration and Segmentation Framework with Application to Magnetic Resonance Prostate Imagery." *IEEE Xplore*, https://ieeexplore.ieee.org/document/5482197.