SIAM J. IMAGING SCIENCES
Vol. 8, No. 2, pp. 1007–1029



# A Kalman Filtering Perspective for Multiatlas Segmentation*

Yi Gao†, Liangjia Zhu‡, Joshua Cates§, Rob S. MacLeod§, Sylvain Bouix¶, and Allen Tannenbaum‖

**Abstract.** In multiatlas segmentation, one typically registers several atlases to the novel image, and their respective segmented label images are transformed and fused to form the final segmentation. In this work, we provide a new dynamical system perspective for multiatlas segmentation, inspired by the following fact: The transformation that aligns the current atlas to the novel image can be not only computed by direct registration but also inferred from the transformation that aligns the previous atlas to the image together with the transformation between the two atlases. This process is similar to the global positioning system on a vehicle, which gets position by inquiring from the satellite and by employing the previous location and velocity—neither answer in isolation being perfect. To solve this problem, a dynamical system scheme is crucial to combine the two pieces of information; for example, a Kalman filtering scheme is used. Accordingly, in this work, a Kalman multiatlas segmentation is proposed to stabilize the global/affine registration step. The contributions of this work are twofold. First, it provides a new dynamical systematic perspective for standard independent multiatlas registrations, and it is solved by Kalman filtering. Second, with very little extra computation, it can be combined with most existing multiatlas segmentation schemes for better registration/segmentation accuracy.

**Key words.** multiatlas segmentation, Kalman filter, registration, dynamical systems

**AMS subject classifications.** 65D19, 68T45, 62H35, 37Mxx

**DOI.** 10.1137/130933423

## 1. Introduction.

Multiatlas segmentation is a widely adopted method due to its accuracy, ease of user intervention, and robustness. The basic idea underpinning the atlas-based segmentation methodology is to drive segmentation by registration; in order to segment a novel image, one registers an already segmented image (training image, atlas image) to this novel image, and utilizes the resulting transformation to deform the corresponding segmentation (training label image) to the space of the novel image. We should note that several

*Received by the editors August 19, 2013; accepted for publication (in revised form) February 12, 2015; published electronically April 30, 2015. This research was supported in part by grants from the National Center for Research Resources (P41-RR-013218), the National Institute of Biomedical Imaging and Bioengineering (P41-EB-015902) of the National Institutes of Health, by NIH grants R01 MH82918 and 1U24CA18092401A1, and by AFOSR grants FA9550-09-1-0172 and FA9550-15-1-0045.

http://www.siam.org/journals/siims/8-2/93342.html

†Department of Biomedical Informatics and Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794 (gaoyi.cn@gmail.com).

‡Department of Computer Science, Stony Brook University, Stony Brook, NY 11790 (liangjiazhu@gmail.com).

§Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, UT 84112 (cates@sci.utah.edu, macleod@sci.utah.edu).

¶Department of Psychiatry, Harvard Medical School, Boston, MA 02215 (sylvain@bwh.harvard.edu).

‖Department of Computer Science and Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794 (arobertan@gmail.com).

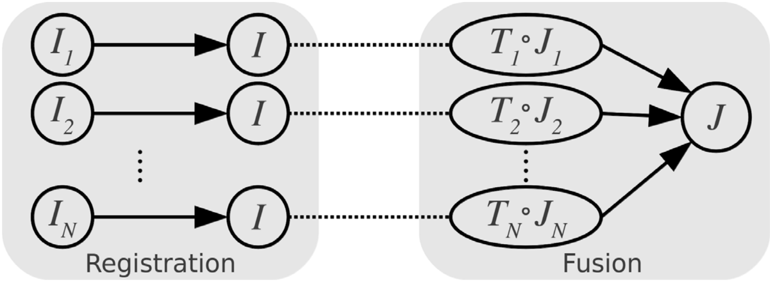


**Figure 1.** *Sketch of multiatlas segmentation. The label images $J_i$ are deformed by the transformations $T_i$ and fused to form the segmentation.*

authors also use the term "atlas" for the *pair* consisting of the training image together with its corresponding label image.

Ideally, registering a single training image to the novel image and deforming the label image using the derived transformation should suffice to produce a segmentation of the novel image. However, since the registration is never perfect, multiple training images are employed where each of their respective deformed label images will provide a consensus on the final segmentation. Such a multiatlas technique, which is briefly reviewed in the next section, has been shown to be more accurate and robust [46, 26, 5, 38, 49, 55].

**1.1. Multiatlas framework.** Typically, the process of multiatlas segmentation consists of two steps. First, several training images are independently registered to the novel image and the transformations, usually nonlinear, are recorded. Then, all the segmented label images are transformed and fused to form the final segmentation. This is shown in Figure 1. The efforts for improving the segmentation accuracy naturally focus on finding either better registration techniques or better label fusion methods.

Indeed, any improvement in the registration technique would improve the performance of the procedure indicated by each horizontal arrow in the left panel of Figure 1. The study of registration constitutes its own challenges and is not the subject of the present study. After performing registration, the label images are deformed using their respective transformations, and then they are fused to form the final segmentation. Most fusion techniques can be categorized as statistical frameworks for classifier combination; see [34] and the references therein. Some researchers choose training images according to their similarity with the novel image [46, 2]. Weighted averaging is also a commonly adopted strategy [26, 47]. Furthermore, instead of assigning a weight for an entire (transformed) label image, local weighting schemes many times achieve better accuracy [5, 27, 10, 48]. In particular, Sabuncu et al. introduced a generative model and showed that several previous voting schemes were special cases of such a model [48]. More recently, Wang et al. proposed a regression-based fusion technique [55, 56]. One of the key insights of the latter study is that the correlated redundancy in the training images can be reduced by the negative weighting obtained from regression analysis.

Some researchers first build single or several representative images out of all the training images and perform multiatlas segmentation with this new set of atlases [17, 31, 3, 7, 48]. This approach not only reduces the computation time, but may also improve the overall accuracy due to the fact that some outliers are avoided in these representatives.

From the above brief review and referring to Figure 1, we can observe that most previous research has attempted to improve the fusion step, or each individual registration in the registration panel. In this study, we propose a new perspective for the registration panel by adopting the insights of dynamical system estimation [52], an approach that is motivated in the next section.

**1.2. Our contributions.** The core novelty of our approach is that the relationship between the training images has been utilized in order to reduce the registration time [13]. We provide a new dynamical system perspective for multiatlas segmentation, inspired by the following fact: There are (at least) two solutions for the registration between the current training image to the novel image. The first is by direct optimization; the second can be computed as the composition of the two registration transformations, one of which is between the previous training image and the novel image, and the other is between the previous and current training images. Unfortunately, due to various factors, neither transformation produces a perfect match of the two images. However, we can combine them to get a better solution.

This approach is similar to that of a global positioning system (GPS) on a vehicle. In order to get the current location, it gathers two pieces of information: directly inquiring from the satellite and via inference from the previous location and velocity. Similar to the two registration solutions, neither of the two positions is 100% accurate. To solve such a problem in GPS, the dynamics of the vehicle are modeled as a linear system and an optimal trajectory in the least square sense is estimated by combining those pieces of information with the Kalman filter [52]. In the image analysis field, it has also been shown that incorporating the concept of dynamical systems into a deterministic process has improved the accuracy and robustness for both registration [39, 43, 51, 21] and segmentation [22].

In this work, a dynamic model of the multiatlas segmentation procedure is constructed by exploiting the relationships among the training images in the transformation group. As shown in Figure 2, we propose to treat the registration panel as an entire system.

Most previous literature divide the entire registration into two stages: a linear (affine) registration step performs the global alignment, which is followed by a deformable step that addresses the local mismatching. In theory, a full filtering approach that stabilizes both the affine and the deformable stages of the registration is possible. However, in the current settings we only perform the filtering in the affine step for the following reasons.

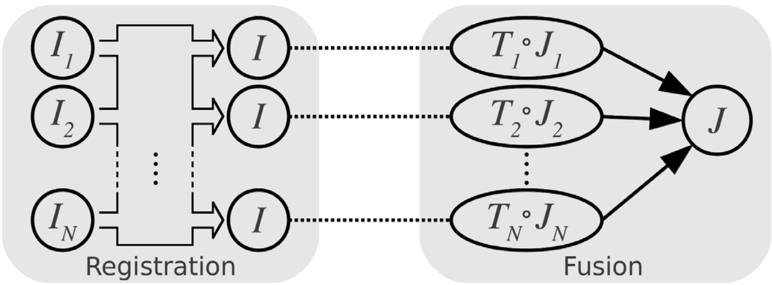


**Figure 2.** *The conceptual diagram of the Kalman multiatlas. Note that the only difference is we treat the multiple registration as an entire dynamical system. The fusion panel is the same as in Figure 1.*

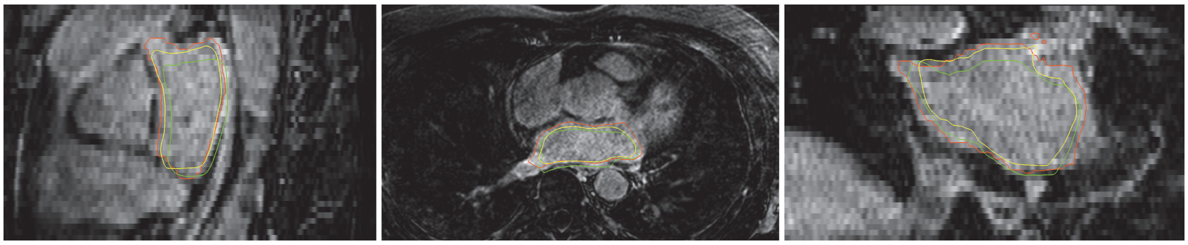

**Figure 3.** *Result of the affine registration. The red contour is the "ground truth" of the left atrium in the heart MR images. The green contour is the result of the proposed method without the Kalman filtering. This demonstrates the global errors after the affine step. The yellow contour is the result of the proposed method with the Kalman filtering, in which the global drifting has been corrected.*

First, affine registration controls the major component of the alignment. Being the first step of the registration pipeline, its performance determines the deformable step and the overall results. A robust estimation of this step would therefore improve the overall segmentation performance. It is observed that while the affine registration in general works quite well for brain images, for chest and pelvic images, unfortunately, huge variances exist. For brain images in which the content is at least a coarse approximation of an ellipsoid, if the two ellipsoids are registered well, then the internal structures roughly align. Unfortunately, for the chest or pelvic images, not only are there huge variances in the content of the images (e.g., leg versus no leg, arm versus no arm, only top of the liver versus more liver region, etc.), but also the organ positions and shapes vary significantly inside the body (e.g., lung/heart shapes in different phases). For example, in Figure 3, there is a global mismatch of the heart after the affine registration step. This mismatch may be due to the large organ position and shape variability inside the body. The Kalman filter effectively addresses such discrepancies in the affine step, giving a much better initial state for the deformable registration.

The second rationale for applying the Kalman filter only on the affine but not on the deformable step is because of nonlinearity and computational cost. The deformable registration step is a nonlinear process. In order to harness the full potential of filtering, the Kalman filter is not sufficient and one needs computationally more intensive approaches. Techniques that address this weakness include, for example, unscented Kalman filtering (UKF) and particle filtering (PF). However, if we denote the dimension of the deformable registration as $D$ (for example $D = 3000$ in the case of $10 \times 10 \times 10$ control nodes in the B-spline registration), then we will perform approximately 6000 extra full deformable registrations if UKF is used, and much more for PF. While the extended Kalman is able to handle a certain level of nonlinearity with moderate extra computation, it does not fit the current scenario since there is no closed form expression for the nonlinear process and therefore the derivative of the process is difficult to evaluate.

Because of these reasons, the filtering is performed only on the affine step. In such a way, the Kalman filter may be utilized with negligible extra computational cost. This situation is different from the approach described in [21, 22] in two ways. First, the goal of the statistical estimation there is to improve a single optimization/registration process. In our approach, the goal is to use filtering to harness the correlation among *different* training images. Moreover, from a computational load perspective, we have a linear process model. As a result, the proposed Kalman filter-based multiatlas has effectively no added computational load.

Therefore, the contribution of the current work is twofold. First, it provides a new systematic perspective for the multiatlas segmentation task. In particular, we treat it as a dynamical system and solve it using Kalman filtering. Second, with very little extra computation, it can be combined with most existing multiatlas segmentation schemes for better registration/segmentation accuracy.

## 2. Method: Kalman multiatlas segmentation.

In this section, we first set up the notation necessary to explain the multiatlas segmentation method and then detail the proposed Kalman filter multiatlas algorithm.

### 2.1. Multiatlas segmentation.

We denote the novel image to be segmented as $I : \Omega \subset \mathbb{R}^3 \to \mathbb{R}^+$. The training images, the *multiatlas*, are denoted as $I_i : \Omega \to \mathbb{R}^+, i = 0, \dots, N-1$, and their respective manually segmented binary images as $J_i : \Omega \to \{0, 1\}; i = 0, \dots, N-1$ with 1 indicating being inside the target. The basic scheme of the original multiatlas segmentation can be divided into two steps: *registration* and *label fusion*. First, each of the $I_i$ images is registered to $I$, and the optimal transformation $\Psi_i : \mathbb{R}^3 \to \mathbb{R}^3$ minimizes the cost function,

$$\Psi_i = \underset{\Psi:\mathbb{R}^3\to\mathbb{R}^3}{\arg\min}\, d(I(\boldsymbol{x}), I_i(\Psi(\boldsymbol{x}))), \tag{2.1}$$

where the (dis-)similarity measurement $d(\cdot, \cdot)$ may be chosen to be the mean square errors (MSE), if the training images are of the same modality of the novel image, or mutual information, when the training images are of different modalities of the novel image.

In practice, the computation of $\Psi_i$ is often achieved in two consecutive steps. First, $\Psi_i$ is decomposed into two transformations, an affine transformation $A_i \in \mathbb{R}^{3\times3}, \boldsymbol{t}_i \in \mathbb{R}^3$ and a nonlinear deformable transformation $\psi_i : \mathbb{R}^3 \to \mathbb{R}^3$, and two optimization problems,

$$A_i = \underset{A:\in\mathbb{R}^{3\times3}, \boldsymbol{t}\in\mathbb{R}^3}{\arg\min}\, d(I(\boldsymbol{x}), I_i(A\boldsymbol{x} + \boldsymbol{t})) \tag{2.2}$$

$$\psi_i = \underset{\psi:\mathbb{R}^3\to\mathbb{R}^3}{\arg\min}\, d(I(\boldsymbol{x}), I_i(\psi(A_i\boldsymbol{x} + \boldsymbol{t}_i))), \tag{2.3}$$

are solved consecutively. Then, each optimal $\Psi_i$ is recorded. Usually, the first of these two steps heavily determines the overall registration quality. Indeed, after the first affine registration step, if the transformed training label image is not covering the target region well, it is very difficult for the subsequent nonlinear step to provide a necessary correction. After the registration, in the second stage of the multiatlas segmentation, each training label image $J_i$ is transformed with $\Psi_i$ and the transformed training label images, $J_i \circ \Psi_i, i = 0, \dots, N-1$, are fused to form the segmentation.

Based on these steps, most of the approaches try to improve the segmentation performance by improving either the individual registration accuracy or the fusion process after individual registration. However, one point that has not been considered in much of the literature is that the registrations of the training images can be correlated. As a result, the *registration* step may be viewed as a dynamical system estimation problem. By adopting this point of view, the dynamical system framework can be utilized to obtain better segmentation results. Indeed, after $\Psi_1$ has been computed by optimizing (2.1), the subsequent $\Psi_i, i \geq 2$, can be computed directly by transformation composition without the optimization. This is because $I_{i+1}$ can be registered to $I_i$, in an offline manner, with the optimal transformation being denoted as $\Psi_{i+1}^i$. As a result, the transformation from $I_{i+1}$ to $I$ can be computed as the composition of $\Psi_i \circ \Psi_{i+1}^i$, which is denoted as $\widetilde{\Psi}_{i+1}$.

Unfortunately, neither $\widetilde{\Psi}_{i+1}$ nor the direct computed transformation $\Psi_{i+1}$ may be suitable. That is, neither $J_{i+1} \circ \widetilde{\Psi}_{i+1}$ nor $J_{i+1} \circ \Psi_{i+1}$ gives the perfect segmentation of the target in $I$. A natural question is, how should the two transformations be combined in a way that the resulting transformation is a better choice?

Fortunately, such a problem has been encountered and tackled in other fields of study. For example, the GPS on a vehicle gets the current location from two sources: by directly inquiring the satellite(s) and by computing the previous position and speed. Neither source gives a 100% accurate position. In order to solve such a problem on the GPS, a dynamical system estimate is crucial to combine them for the final output using the Kalman filter. Inspired by such an idea, the registration of multiple images can be viewed as an estimation of a dynamical system, and Kalman filtering can be employed for better overall results. To aid further discussion, we give a brief review of the Kalman filter in the next section.

**2.2. Kalman filter.** The Kalman filter is a recursive least square estimation method [52]. Given a linear discrete-time dynamical system, whose state variable is $\boldsymbol{\alpha}_k$ and observation variable is $\boldsymbol{\beta}_k$ (both at time $k$), the objective of the Kalman filter is to provide the a posteriori estimate for the state $\boldsymbol{\alpha}_k$, $\hat{\boldsymbol{\alpha}}_k^+$, given the observations $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_k =: \boldsymbol{\beta}_{1:k}$, i.e.,

$$\hat{\boldsymbol{\alpha}}_k^+ = E(\boldsymbol{\alpha}_k | \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_k) = E(\boldsymbol{\alpha}_k | \boldsymbol{\beta}_{1:k}). \tag{2.4}$$

With this goal, the process and observation models are given as

$$\boldsymbol{\alpha}_k = F_{k-1}\boldsymbol{\alpha}_{k-1} + u_{k-1}, \tag{2.5}$$

$$\boldsymbol{\beta}_k = H_k \boldsymbol{\alpha}_k + v_k, \tag{2.6}$$

where the system dynamics are represented by the matrix $F_k$, and the observation is represented by the matrix $H_k$. The system noise $u_k$ and observation noise $v_k$ are white, zero-mean, and uncorrelated. Moreover, their covariance matrices $Q_k$ and $R_k$ of $u_k$ and $v_k$, respectively, are assumed to be known.

In order to initialize the Kalman filter, we have

$$\hat{\boldsymbol{\alpha}}_0^+ = E(\boldsymbol{\alpha}_0), \tag{2.7}$$

$$P_0^+ = E\left[(\boldsymbol{\alpha}_0 - \hat{\boldsymbol{\alpha}}_0^+)(\boldsymbol{\alpha}_0 - \hat{\boldsymbol{\alpha}}_0^+)^\top\right], \tag{2.8}$$

where $P_0^+$ is the a posteriori estimate for the covariance matrix, representing the uncertainty in our initial estimate of $\boldsymbol{\alpha}_0$. Then, the recursive estimation of $\hat{\boldsymbol{\alpha}}_k^+$ consists of two steps, namely, *prediction* and *update*. Assuming $\hat{\boldsymbol{\alpha}}_{k-1}^+$ is available, the prediction step gives the a priori estimate of $\boldsymbol{\alpha}_k$ and its covariance matrix at time $k$ as

$$\hat{\boldsymbol{\alpha}}_k^- = F_k \hat{\boldsymbol{\alpha}}_{k-1}^+, \tag{2.9}$$

$$P_k^- = F_{k-1} P_{k-1}^+ F_{k-1}^\top + Q_{k-1}. \tag{2.10}$$

At time $k$ after the observation $\boldsymbol{\beta}_k$ is available, it can then be used to update the estimate in order to obtain the a posteriori estimate:

$$K_k = P_k^- H_k^\top \left( H_k P_k^- H_k^\top + R_k \right)^{-1}, \tag{2.11}$$

$$\hat{\boldsymbol{\alpha}}_k^+ = \hat{\boldsymbol{\alpha}}_k^- + K_k \left( \boldsymbol{\beta}_k - H_k \hat{\boldsymbol{\alpha}}_k^- \right), \tag{2.12}$$

$$P_k^+ = (Id - K_k H_k) P_k^-, \tag{2.13}$$

where $K_k$ the Kalman gain and $Id$ is the identity matrix.

The above two steps are iterated, and the a posteriori estimate for $\boldsymbol{\alpha}_k$ is obtained recursively. In the subsequent sections, we detail how the multiatlas segmentation may be cast under the Kalman filtering framework.

**2.3. Kalman multiatlas segmentation.** The computation of the transformations from $I_i, i = 0, \ldots, N-1$, to $I$ can be viewed as a dynamical system and solved via Kalman filtering. More specifically, the Kalman filter is able to provide the optimal recursive least square solution for the estimates of the $\Psi_i, i = 0, \ldots, N-1$.

However, directly applying the Kalman filtering to a nonlinear transformation group is not a feasible solution due to the following reasons. First, the composition of two registrations, $\Psi_{i-1} \circ \Psi_i^{i-1}$, is not a linear operation. This violates the linear requirement of using the Kalman filter. Moreover, while other techniques, such as the extended Kalman filter [52], unscented Kalman filter [32], or particle filter [50], may handle such nonlinear situations, the dimensionality of $\Psi_i$ is huge due to its representing the entire nonlinear deformation field. This makes the computation intractable.

In practice, the accuracy of the overall nonlinear registration depends heavily on that of the affine component. Indeed, when the affine registration does not produce satisfying global matching, it is very difficult for the subsequent deformable step to obtain high accuracy. Therefore, improving the affine registration is crucial for the overall registration results.

For $\boldsymbol{x} = (x_1, x_2, x_3)^\top \in \mathbb{R}^3$, writing $\boldsymbol{x}$ in the homogeneous coordinates, we have $\tilde{\boldsymbol{x}} = (x_1, x_2, x_3, 1)^\top$ and the transformation

$$A_i \boldsymbol{x} + \boldsymbol{t}_i = \begin{pmatrix} A_i & \boldsymbol{t}_i \\ \mathbf{0} & 1 \end{pmatrix} \tilde{\boldsymbol{x}} =: \tilde{A}_i \tilde{\boldsymbol{x}}, \tag{2.14}$$

where $\tilde{A}_i \in \mathbb{R}^{4\times4}$ is the homogeneous representation of the affine transformation pair $(A_i, \boldsymbol{t}_i)$. Hence, we have the transition concatenation equation:

$$\tilde{A}_{i+1} = \tilde{A}_i \cdot \tilde{A}_{i+1}^i. \tag{2.15}$$

Furthermore, we define the system state variable $\boldsymbol{\alpha}_i \in \mathbb{R}^{16}$ as

$$\tilde{A}_i(\mu, \nu) =: \boldsymbol{\alpha}_i(4 \cdot (\mu - 1) + \nu), \quad \mu, \nu \in \{1, 2, 3, 4\}, \tag{2.16}$$

which is the rowwise vectorization of $\tilde{A}_i$, and the system transition equation can be written as

$$\boldsymbol{\alpha}_i = F_{i-1} \boldsymbol{\alpha}_{i-1} + u_{i-1}, \tag{2.17}$$

where $u_{i-1} \in \mathbb{R}^{16}$ is the normally distributed noise and $F_{i-1} \in \mathbb{R}^{16\times 16}$ is a block-diagonal matrix with $(\tilde{A}_i^{i-1})^\top$ on its diagonal blocks:

$$F_{i-1} = \begin{pmatrix} \tilde{A}_i^{i-1} & 0 & 0 & 0 \\ 0 & \tilde{A}_i^{i-1} & 0 & 0 \\ 0 & 0 & \tilde{A}_i^{i-1} & 0 \\ 0 & 0 & 0 & \tilde{A}_i^{i-1} \end{pmatrix}^\top . \tag{2.18}$$

In (2.17), the noise term $u_{i-1}$ models the nondeterministic factors in the registration process, which may include factors such as local minima and/or the discrepancy between the registration metric and the visual appearance.

The observation matrix is an identity matrix, that is, $H_i = Id_{16\times 16}$ $\forall i$. This indicates the fact that the registration parameters can be directly observed by the usual optimization process.

Having cast the multiatlas registration problem as a dynamical system, and having set up the system and observation models, the Kalman filter is applied to solve this estimation problem, which is detailed in Algorithm 1.

**Algorithm 1.** Kalman multiatlas segmentation.

1: Affine register $J_i$ to $J_{i-1}$ for $i = 1, \dots, N-1$ by minimizing the MSE, and record the transformations $\tilde{A}_i^{i-1}$. This can be done offline before $I$ is available
2: Initialize $P_0^+$, $Q$, and $R$
3: When $I$ is available, register $I_0$ to $I$ and record the transformations $A_0$ $(\boldsymbol{\alpha}_0^+)$
4: **for** $i = 1, 2, \dots, N-1$ **do**
5: Predict transformation: $\boldsymbol{\alpha}_i^- = F_i \boldsymbol{\alpha}_{i-1}^+$
6: Covariance matrix: $P_i^- = F_{i-1} P_{i-1}^+ F_{i-1}^\top + Q$
7: Get observation: affine register $I_i$ to $I$ by minimizing the dissimilarity measure and record the transformation $\boldsymbol{\beta}_i$
8: Kalman gain: $K = P_i^- \left(P_i^- + R_i\right)^{-1}$
9: Update state: $\boldsymbol{\alpha}_i^+ = \boldsymbol{\alpha}_i^- + K\left(\boldsymbol{\beta}_i - \boldsymbol{\alpha}_i^-\right)$
10: Update covariance matrix: $P_i^+ = (Id - K)P_i^-$
11: **end for**

This algorithm is a general-purpose scheme and can be combined with any advanced registration method. Furthermore, it adds very little extra (online) computation (the additional matrix computations) as compared to the independently registered multiatlas algorithms, and the extra running time is less 1%.

Furthermore, the choice of filtering the affine transformation, instead of the entire nonlinear transformation, can further be justified from a prediction accuracy perspective. Evidently, if the prediction is of poor accuracy, incorporating such information may even impair the overall performance. To solve such a problem and achieve better accuracy in the prediction step, as can be observed in Algorithm 1, the computation of $\tilde{A}_i^{i-1}$'s should be based on the registration between $J_i$ and $J_{i-1}$, rather than $I_i$ and $I_{i-1}$. This is due to two reasons. First, it is commonly recognized that the registration of binary images can often be solved with

Algorithm 2. Deformable Kalman multiatlas segmentation.

```
for $i = 0, 2, ..., N-1$ do
    Compute the affine registration $\boldsymbol{\alpha}_i^+$ $(A_i)$ using Algorithm 1
    Register $I_i \circ A_i$ to $I$ by minimizing NNC with respect to B-spline deformation field
    Record the final deformation field $\psi_i$ and registration cost $c_i$
end for
Sort $c_i$'s so that $c_{k_i} \leq c_{k_{i+1}} \forall i = 0, \dots, N-2$
Define $w_{k_i} = \exp(c_{k_0} - c_{k_i}) \forall i = 0, \dots, N-1$ and normalize $w_i$'s s.t. $\sum_{i=0}^{N-1} w_i = 1$
$P(\boldsymbol{x}) = \sum_{i=0}^{N-1} w_{k_i} J_{k_i} (\psi_{k_i} (A_{k_i} (\boldsymbol{x})))$
```

higher accuracy than the registration between two grayscale images, which may even be of different modality with different fields of view. Second, for the multiatlas segmentation task, the ultimate concern is the segmentation of the target. Hence the registration transformation should provide good matching on and around the target region. This is naturally achieved by registering the binary images but is not emphasized in the registration between $I_i$ and $I_{i-1}$. Due to these reasons, filtering the affine transform is a good choice not only because of its low extra computational load, but also it is a natural choice from the point of view of accuracy.

In order to address the nonlinear deformation between the novel image and the training images, in addition to the affine registrations, a further nonlinear registration is performed after the aforementioned affine registration. More explicitly, B-splines are used to model the free-form deformation and the image dissimilarity measure (e.g., negative normalized correlation (NNC)) is minimized. Indeed, we usually have available several training images with the same modality as the novel image. Therefore, NNC is found sufficient in practice to balance the computation load and the capability of handling the intramodality variation. Nevertheless, the proposed framework does not preclude the use of other registration metrics. After the nonlinear registration, the transformed label images are combined, weighted by the final accuracy of the registration ($w_i$ as in Algorithm 2), to form a prior probability map $P(\boldsymbol{x})$. The detailed algorithm is given in Algorithm 2. Finally, the prior probability map $P(\boldsymbol{x})$ represents the relative likelihood of a certain point $\boldsymbol{x}$ being inside the target object. In what follows, such prior information is combined with the likelihood, computed from the intensity distribution, to form a posterior map and the final segmentation.

**2.4. Bayesian framework to incorporate novel image information and multiatlas.** Since $P(\boldsymbol{x})$ represents the likelihood of a point $\boldsymbol{x}$ being inside the target object, we defined $P_B(\boldsymbol{x}) = 1 - P(\boldsymbol{x})$ to represent the probability of the point $\boldsymbol{x}$ being in the background.

The likelihood map is computed from the intensity information. Specifically, the intensity distribution function $l(x)$ (defined in step 2 of Algorithm 3) within the target regions of all the training images may be learned by the nonparametric kernel density estimation technique with the kernel density automatically estimated as in [8]. Before the density estimation, the histograms of all the grayscale images are matched with that of the first training image $I_0$. This is because although the images are of the same modality, different imaging settings may cause discrepancy among them.

While the original Bayesian decision theory involves the comparison of the posterior between two (or more) categories, many previous methods only learn and utilize the likelihood within only one group (the target region) to guide segmentation [11, 24, 41]. This may be partially due to the fact that the information outside the target can be very heterogeneous, since it contains everything other than the target. As a result, learning is severely unbalanced and the learned "background characteristics" may become so varied that everything may look like background, which does not provide useful inference. To treat such a difficulty, and to take full advantage of Bayesian decision theory, in this work only the information in the local background region is utilized. Moreover, because the approximate location of the target has already been represented by $P(\boldsymbol{x})$, we therefore only have to distinguish the target from its surroundings. To that end, for each $J_i$, the distance map is constructed from the target region utilizing the method in [12]. Then, a threshold is chosen so that the narrow banded surrounding region has the same volume as the target region in $J_i$. The threshold is applied and those intensity values in $I_i$ in this banded region are collected, from which the distribution function of the local background, $l_B(x)$ (defined in step 3 of Algorithm 3), is computed.

With all the prior and likelihood information, the final segmentation decision is drawn by the Bayesian decision rule:

$$D(\boldsymbol{x}) = \begin{cases} 1 & \text{if } P(\boldsymbol{x}) \cdot l(\boldsymbol{x}) > P_B(\boldsymbol{x}) \cdot l_B(\boldsymbol{x}); \\ 0 & \text{otherwise.} \end{cases} \tag{2.19}$$

The overall segmentation algorithm is summarized in Algorithm 3.

Algorithm 3. Overall segmentation algorithm.

1: Match the histogram of $I$ and $I_i$ to $I_0$ for $i = 1, \dots, N-2$
2: Compute the distribution $l(x)$ of the set $\{I_i(\boldsymbol{x}) : J_i(\boldsymbol{x}) = 1\ \forall \boldsymbol{x} \in \mathbb{R}^3; i = 0, \dots, N-1\}$
3: Compute the distribution $l_B(x)$ of the set $\{I_i(\boldsymbol{x}) : \boldsymbol{x}$ in local background $\forall \boldsymbol{x} \in \mathbb{R}^3; i = 0, \dots, N-1\}$
4: Compute target prior $P(\boldsymbol{x})$ using Algorithm 2
5: Compute background prior $P_B(\boldsymbol{x}) = 1 - P(\boldsymbol{x})$
6: Compute segmentation using (2.19)

**2.5. Estimation of covariance matrices.** In the Kalman filtering framework, one important issue is the determination of the covariance matrices, which is in general a very difficult task. In our specific situation, in Algorithm 1, three covariance matrices $P_0^+$, $Q$, and $R$ have to be determined.

Among them, $P_0^+$ and $R$ characterize the performance of the grayscale image registration. In order to estimate them, we assume the registration of the binary images gives the "ground truth" and then measure the uncertainty of the grayscale image registration. Specifically, we perform $M$ affine registrations between $I_i$ and $I_j$ with $i, j \in \{1, \dots, N\}$ and $i \neq j$ and record the resulting parameters $\{\hat{\boldsymbol{\alpha}}_k \subset \mathbb{R}^{16}; k = 1, \dots, M\}$. Given $N$ training images, there are $N(N-1)/2$ distinct pairs of images. However, registering all those pairs is tedious work, even if this is carried out offline. As a result, $M$ pairs are randomly chosen (with uniform distribution) from the $N(N-1)/2$ pairs. Similarly, $M$ registrations between $J_i$ and $J_j$ are also performed and the results are recorded as $\{\check{\boldsymbol{\alpha}}_k \subset \mathbb{R}^{16}; k = 1, \ldots, M\}$. Therefore, we estimate $P_0^+$ and $R$ as

$$\boldsymbol{\xi} := \frac{1}{M} \sum_{k=1}^{M} (\hat{\boldsymbol{\alpha}}_k - \check{\boldsymbol{\alpha}}_k), \tag{2.20}$$

$$P_0^+ = R := \frac{1}{M} \sum_{k=1}^{M} (\hat{\boldsymbol{\alpha}}_k - \check{\boldsymbol{\alpha}}_k - \boldsymbol{\xi}) (\hat{\boldsymbol{\alpha}}_k - \check{\boldsymbol{\alpha}}_k - \boldsymbol{\xi})^\top . \tag{2.21}$$

The covariance matrix $Q$, however, represents the certainty of the prediction for the registration between $I_{i+1}$ with $I_i$, based on the registration between $I_i$ with $I$ and the registration between $J_i$ with $J_{i+1}$. Since the novel image $I$ is involved, one may not be able to estimate $Q$ offline when $I$ is not available. To solve this issue, $I_j, j \neq i$, is used to take the place of $I$. By doing this, the estimate of $Q$ is the same as that for $P_0^+$ and $R$.

**2.6. Effect of different orders.** In the original independent registration scheme, there is no difference in performing the registration in different orders. However, with the Kalman filter, we impose some sequential order for the training images. Usually all the training images are available during the segmentation and an arbitrary order could be used. Therefore, a natural question is whether the order would affect the final results.

Unfortunately, we did not find a way to theoretically characterize the influence of different orderings. However, in section 3, we perform some experimental tests and found that though order affects the final outcome, they all improve the results quite significantly over the scheme without Kalman filtering. We will return to the point in subsection 3.5.

**3. Experiments and results.** We used the Kalman filter multiatlas method as a core component in the proposed segmentation framework and tested the segmentation on three sets of images: striatum, prostate, and left atrium. Furthermore, by removing the Kalman filtering and leaving everything else the same, the altered method is also tested on the same data sets and statistically significant improvements are observed. By doing this, we show not only the value of the proposed Kalman multiatlas framework but also the applicability of such a method as a component in other multiatlas techniques, hence its potential of being included in more application-specific methods.

**3.1. Striatum.** The striatum is a subcortical structure in the forebrain which contains the caudate nucleus, putamen, and fundus. It has been shown that the striatum is associated with several neural psychiatrical diseases such as Parkinson's disease [33], schizophrenia [37], and Huntington's disease [45]. A key step in evaluating the patient involves the segmentation of the striatum from magnetic resonance (MR) imagery. While manual delineation is a time-consuming process, automatic segmentation from the MR imagery is a very challenging task due to the complex convolved shape of the striatum and the low contrast with the surrounding tissue. In this section, we report the results of 27 spoiled gradient-recalled volumetric T1-weighted images acquired at Brigham and Women's Hospital (BWH), Harvard Medical School. The image resolution is $1 \times 1 \times 1$ mm$^3$. Fortunately, we have all the images segmented by a radiologist at BWH. In order to make full use of all the data, we choose the

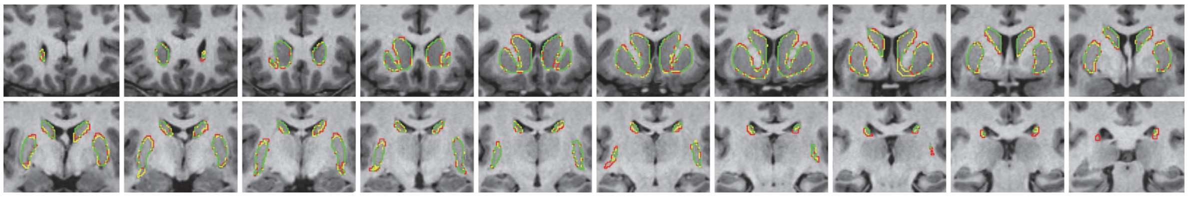

**Figure 4.** *Coronal slices (every three image slices) of the segmentation result whose DC is the median among all. The yellow contours are generated by the proposed automatic segmentation algorithm, while the red contours are drawn manually by an expert radiologist at BWH, and the green contour is where the automatic ones coincide with the manual ones.*

leave-one-out strategy to perform the shape training and segmentation. The segmentation is fully automatic, so that no human interaction is needed.

In Figure 4, we show the segmentation result of one data set in coronal slice views. The coronal view is used because it better shows the complexity of the target shape. In this result, the striatum boundary generated by the proposed algorithm is colored yellow, whereas the manual boundary, drawn by an expert radiologist, is colored red and used as the reference. We use green when the two results meet. As can be observed from the comparison, the two sets of contours match very well in all the slices. Due to space constraints, only one subject from which every third coronal slice is shown here. However, note that all the 27 experiments work consistently, and the detailed statistical analysis is given below.

Moreover, in order to quantify the differences between the two sets of results, the distance maps are computed. Given any point on the manually segmented surface, the distance map value is the closest distance to the automatically extracted surface. Altogether, they form a scalar function defined on the manually generated surface. The distance maps are illustrated as the colors on the manual surfaces. Given any location on the surface, red (respectively, blue) means that there exists a relatively large (respectively, small) difference between the surface generated by the expert and the algorithm. As can be seen in Figure 5, except at a few isolated regions, the automatic results are very consistent with the expert manual results. It is also noted that due to space limits, we only show the surfaces of five uniformly randomly chosen results.

In order to further analyze the results quantitatively, the Dice coefficients (DC) [14] and the 95% Hausdorff distances [25, 4] with respect to the expert segmentation results are provided in Table 1. The accuracy is higher than the previously reported results on the same structure [18].

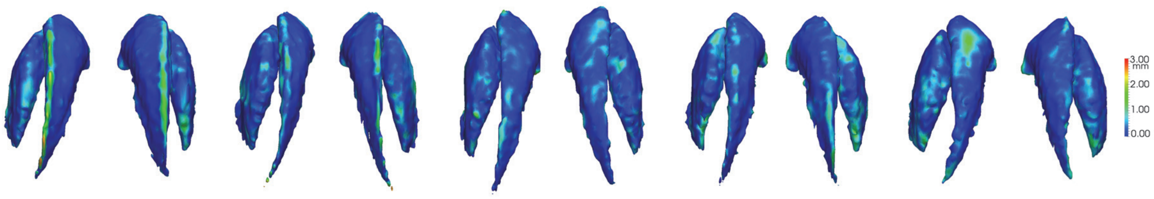


**Figure 5.** *The surfaces of five randomly selected striatums. On the surfaces, the color shows the difference between the automatic and manual surfaces.*

**Table 1**

*DCs and the 95% Hausdorff distances (HD) between the expert segmentation and the automatic result generated by the proposed method for the striatums. The unit for the Hausdorff distance is mm.*

| Subject ID | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 |
|---|---|---|---|---|---|---|---|---|---|---|
| DC$\times 10^2$ | 91.5 | 91.0 | 92.0 | 89.5 | 89.1 | 90.1 | 90.0 | 90.6 | 91.5 | 91.6 |
| 95% HD | 3.78 | 1.07 | 2.85 | 1.88 | 2.10 | 1.66 | 3.38 | 1.49 | 3.90 | 3.82 |
| Subject ID | #11 | #12 | #13 | #14 | #15 | #16 | #17 | #18 | #19 | #20 |
| DC$\times 10^2$ | 91.5 | 89.9 | 91.5 | 90.4 | 91.7 | 91.9 | 91.8 | 90.8 | 91.5 | 91.5 |
| 95% HD | 1.11 | 3.09 | 2.86 | 3.79 | 3.22 | 1.54 | 1.50 | 1.85 | 2.41 | 1.59 |
| Subject ID | #21 | #22 | #23 | #24 | #25 | #26 | #27 | **Mean** | **STD** | |
| DC$\times 10^2$ | 91.1 | 91.4 | 90.8 | 91.0 | 91.5 | 91.4 | 91.5 | **91.1** | **0.91** | |
| 95% HD | 1.17 | 1.23 | 2.40 | 3.52 | 2.99 | 3.49 | 1.21 | **2.40** | **0.99** | |

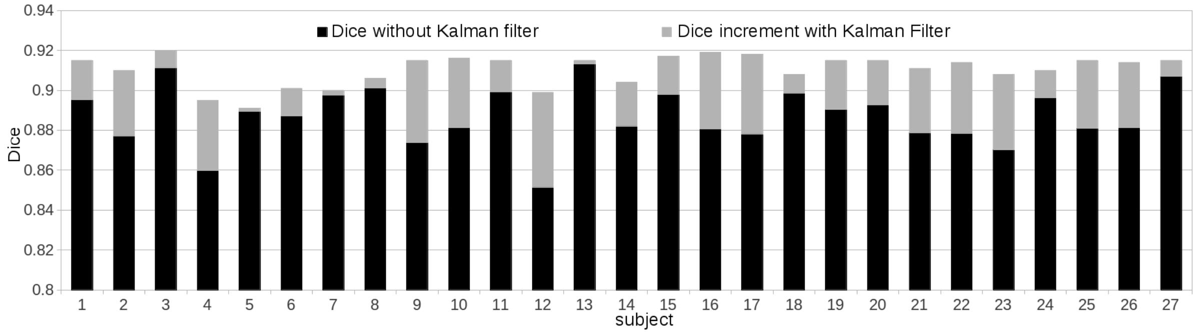


**Figure 6.** *DCs of the striatum segmentation using the proposed Algorithm 3 without the Kalman filtering (black bars). The increment of DC with the Kalman filtering is in gray. Therefore the total bar heights correspond to the values in Table 1. Note that the vertical axis starts from* 0.8 *to clarify the increment so the gray/black bar height ratios do not represent relative increments.*

Since the main theme of the present work is to propose a new approach to multiatlas segmentation, keeping the overall scheme relatively simple is useful in isolating the effect of adopting such a perspective with respect to the "normal control." As a result, the fusion step here does not involve state-of-the-art techniques such as those introduced in [49, 55, 56], nor did we further combine the atlas segmentation with shape-based active contour evolution to fine tune the results [19, 23, 16].

The influence of adopting the dynamical system perspective and solving it using the Kalman filter is further quantified in Figure 6. In the figure, we show the DCs of the segmentations generated with a single change in the entire algorithm: we replace the Kalman filter scheme with conventional independent registration. The improvement of the Kalman filtering over the normal control can be clearly seen. In addition to visual inspection, we perform the single side $t$-test on the DCs of the segmentation methods with and without the Kalman filtering. The null hypothesis that the difference between the two sets of DCs are due to random perturbation is rejected at a significance level of 0.01.

**3.2. Prostate.** Prostate cancer ranks among one of the most widespread of all cancers for the U.S. male population [29]. MR imaging is an attractive method for guiding and monitoring interventions since it provides superior visualization of the prostate as well as its substructures. Extracting the prostate from MR imagery is a challenging and important task for prostate cancer assessment and surgical planning [22].

**Table 2**

*DCs and 95% HD (in mm) using the proposed Kalman multiatlas method for the prostate MR images. It is noted that the patient IDs are taken from the clinical study and they are not contiguous.*

| Patient ID | 29 | 35 | 36 | 41 | 45 | 49 | 54 | 55 | 56 |
|---|---|---|---|---|---|---|---|---|---|
| DC$\times 10^2$ | 78.8 | 78.2 | 71.1 | 72.5 | 70.3 | 75.8 | 81.9 | 77.5 | 71.7 |
| 95% HD | 9.75 | 9.07 | 12.2 | 13.2 | 13.9 | 6.44 | 12.7 | 12.3 | 4.68 |
| Patient ID | 60 | 64 | 69 | 73 | 85 | 88 | **Mean** | **STD** | |
| DC$\times 10^2$ | 74.5 | 76.2 | 81.0 | 80.5 | 73.9 | 79.1 | **76.2** | **3.8** | |
| 95% HD | 4.01 | 9.89 | 17.0 | 7.43 | 11.2 | 5.63 | **9.96** | **3.7** | |

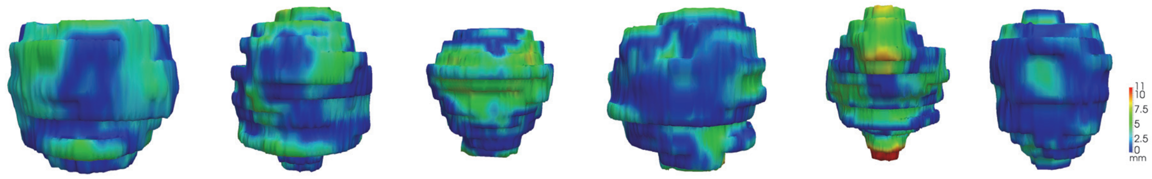

**Figure 7.** *The surfaces of six randomly picked prostates. On the surfaces, the color shows the difference between the automatic and manual surfaces.*

In this work, 30 data sets from 15 patients were downloaded from the online publicly available repository [42]. Both T1 and T2 MR prostate images were performed on each patient from a 1.5T MR scanner. However, given that only the expert segmentations for T2 images are provided, in order to compare with the expert segmentation, we opted to segment the 15 T2 images. The ages of the patients range from 50 to 80 with a mean age of 63.7. After applying the proposed method (leave-one-out scheme), the DCs with respect to the expert segmentation results given on the website are provided in Table 2. The distance map on six randomly chosen manual surfaces is shown in Figure 7.

Comparing with [22, 23, 53], the performance is not necessarily higher. However, in the present study, we ignore those aspects of other methods that are not relevant to the main theme of Kalman filtering to elucidate our approach to multiatlas segmentation. By doing this, the impact of adopting the new Kalman filtering perspective in the multiatlas segmentation can be justified more clearly, as shown in the gray increment bars of the plots in Figure 8. The single side $t$-test on the hypothesis that the difference between the two sets of DCs is due to random perturbation is rejected at a significance level of 0.01. Consequently, it is believed that

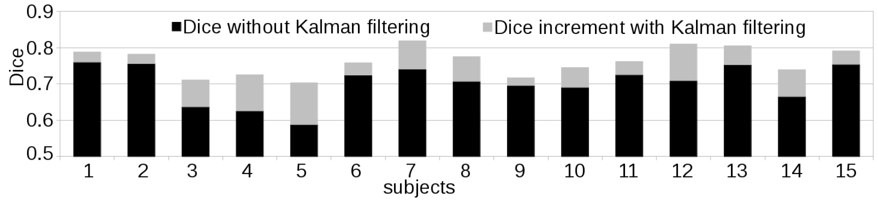


**Figure 8.** *DCs of the prostate segmentation using the proposed Algorithm 3 without the Kalman filtering (black bars). The increment of DC with the Kalman filtering is in gray. Therefore the total bar heights corresponds to the values in Table 2.*

most of the current multiatlas segmentation schemes can benefit from adopting the dynamical system perspective which utilizes the intercorrelation among the training images, though a thorough testing with all the many methods is beyond the scope of this work.

**3.3. Left atrium.** Atrial fibrillation is the most common cardiac arrhythmia and is characterized by unsynchronized electrical activity in the left atrium (LA) of the heart. One of the first-line treatments for atrial fibrillation is radiofrequency (RF) ablation therapy. RF ablation attempts to isolate spurious sources of electrical activity in the LA by selectively scarring the LA wall tissue using RF energy that is delivered with an intracardiac catheter [28]. Successful ablation can eliminate the problematic sources of electrical activity and, often, completely cure atrial fibrillation. Typically, RF ablations are performed under guidance by a combination of real-time fluoroscopy and a preoperatively acquired LA wall segmentation, obtained from either CT or MRI. The latter provides the ablationist with a three-dimensional context for catheter navigation and, potentially, the location of diseased heart tissue and previous LA scarring, which may be important in planning repeat RF ablation procedures. Late gadolinium enhancement MR (LGE-MR), for example, has recently been shown to be effective for identification of fibrotic and scarred LA wall tissue, which appear as brighter regions in the wall, due to differential diffusion rates of gadolinium through these tissues [54]. Quantitative assessment of fibrotic regions has also been show to be useful in planning treatment for atrial fibrillation patients and assessing the likelihood of successful ablation outcomes [1, 40].

While the LGE-MR sequence is highly useful for planning and guidance of RF ablations, the noise level and bias field effect in these images are higher than typical cardiac MR sequences, which is not ideal for structural tissue segmentation and makes the segmentation/registration of LGE-MR images an especially challenging task [20]. Indeed, the registration may be misled by artifacts in the images that can cause misalignment of the training images with the novel image to be segmented. Similar to the situation where the satellite signal is corrupted by the clutter of buildings and becomes less reliable, with Kalman filtering, the combination of the prediction and observation effectively improves the registration and hence the segmentation.

In this work, we tested the algorithm on the LGE-MR images available from [9]. In total, there are 32 subjects, each of which has the preablation and postablation LGE-MR images. We tested our algorithm on the both the pre- and postablation sets separately. That is, only preablation (respectively, postablation) images are used as training images for segmenting preablation (respectively, postablation) images. The results for preablation images are shown in Figure 9 and Table 3 and those for postablation images are shown in Figure 10 and Table 4.

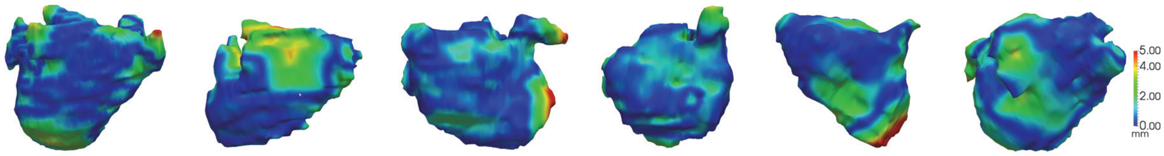


**Figure 9.** *The surfaces of six randomly chosen preablation LAs. On the surfaces, the color shows the difference between the automatic and manual surfaces.*

**Table 3**

*DCs and the 95% Hausdorff distances between the expert segmentation and the automatic result generated by the proposed method for the preablation LA image. The unit for the Hausdorff distance is mm.*

| Patient ID | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 |
|---|---|---|---|---|---|---|---|---|---|---|
| DC$\times 10^2$ | 68.8 | 61.4 | 85.2 | 80.9 | 78.3 | 68.3 | 70.8 | 80.3 | 81.4 | 81.4 |
| 95% HD | 4.47 | 3.29 | 2.01 | 2.62 | 2.36 | 5.12 | 2.72 | 3.82 | 3.06 | 3.94 |
| Patient ID | #11 | #12 | #13 | #14 | #15 | #16 | #17 | #18 | #19 | #20 |
| DC$\times 10^2$ | 83.9 | 83.0 | 67.4 | 67.6 | 68.9 | 81.6 | 63.5 | 70.5 | 70.3 | 81.5 |
| 95% HD | 2.60 | 2.08 | 2.07 | 3.62 | 3.08 | 3.06 | 4.37 | 3.46 | 3.38 | 4.51 |
| Patient ID | #21 | #22 | #23 | #24 | #25 | #26 | #27 | #28 | #29 | #30 |
| DC$\times 10^2$ | 83.3 | 73.8 | 73.8 | 80.3 | 75.2 | 76.5 | 82.1 | 64.8 | 78.1 | 71.1 |
| 95% HD | 2.57 | 3.78 | 3.89 | 3.02 | 3.39 | 2.27 | 2.29 | 3.30 | 3.80 | 5.31 |
| Patient ID | #31 | #32 | **Mean** | | | | **STD** | | | |
| DC$\times 10^2$ | 64.8 | 60.7 | **74.4** | | | | **7.35** | | | |
| 95% HD | 2.67 | 3.37 | **3.29** | | | | **0.86** | | | |

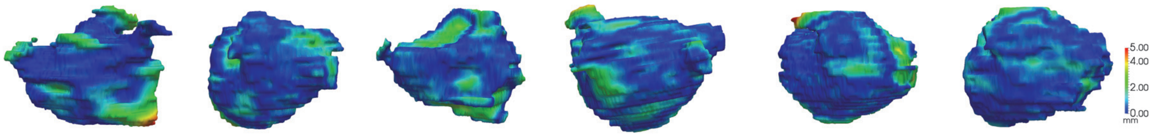


**Figure 10.** *The surfaces of six randomly chosen postablation LAs. On the surfaces, the color shows the difference between the automatic and manual surfaces.*

**Table 4**

*DCs and the 95% Hausdorff distances between the expert segmentation and the automatic result generated by the proposed method for the postablation LA images. The unit for the Hausdorff distance is mm.*

| Patient ID | #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 | #9 | #10 |
|---|---|---|---|---|---|---|---|---|---|---|
| DC$\times 10^2$ | 69.6 | 61.6 | 81.8 | 76.9 | 83.6 | 64.2 | 81.6 | 80.7 | 79.2 | 71.4 |
| 95% HD | 3.15 | 1.65 | 2.86 | 1.76 | 3.90 | 2.50 | 3.30 | 2.77 | 3.46 | 2.73 |
| Patient ID | #11 | #12 | #13 | #14 | #15 | #16 | #17 | #18 | #19 | #20 |
| DC$\times 10^2$ | 84.9 | 83.6 | 63.5 | 69.7 | 63.7 | 71.2 | 66.8 | 52.8 | 62.8 | 86.0 |
| 95% HD | 3.61 | 3.81 | 4.60 | 3.06 | 1.48 | 2.53 | 4.31 | 2.35 | 3.99 | 3.32 |
| Patient ID | #21 | #22 | #23 | #24 | #25 | #26 | #27 | #28 | #29 | #30 |
| DC$\times 10^2$ | 76.6 | 79.0 | 78.2 | 79.6 | 77.5 | 72.4 | 72.0 | 73.8 | 84.9 | 69.7 |
| 95% HD | 4.38 | 1.63 | 3.05 | 2.24 | 5.57 | 3.88 | 4.33 | 2.36 | 2.84 | 2.99 |
| Patient ID | #31 | #32 | **Mean** | | | | **STD** | | | |
| DC$\times 10^2$ | 76.6 | 73.6 | **74.0** | | | | **8.11** | | | |
| 95% HD | 4.10 | 2.99 | **3.17** | | | | **0.95** | | | |

The improvement from adopting the new view for the multiatlas segmentation is evident from the gray bars in Figures 11 and 12. The single side $t$-test on the hypothesis that the difference between the two sets of DCs are due to random perturbation, for either pre- or postablation, is rejected at a significance level of 0.01.

**3.4. Comparison with other registration methods.** In this work, the main idea is to view the multiatlas segmentation as a dynamical system and use the Kalman filtering to stabilize the estimation. Adopting such a perspective, the Kalman filtering can be inserted as a component to most of the existing segmentation-by-registration schemes. Because of

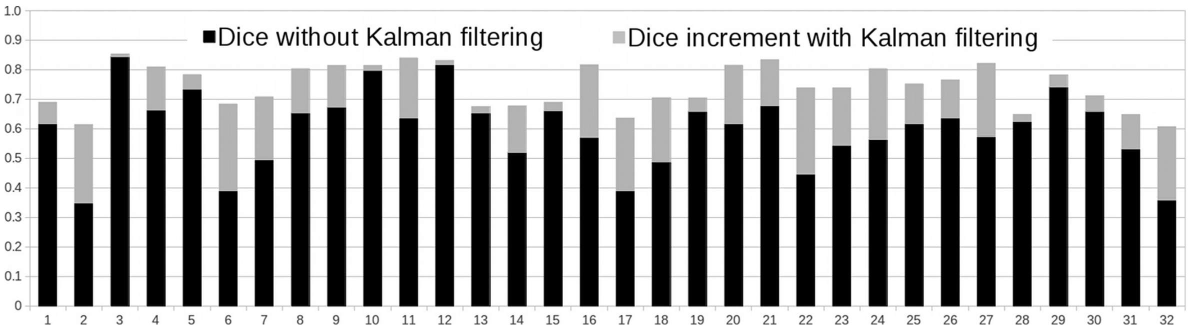


**Figure 11.** *DCs of the preablation LA image segmentation using the proposed Algorithm 3 without the Kalman filtering (black bars). The increment of DC using the Kalman filter is in gray. Therefore the total bar heights corresponds to the values in Table 3.*

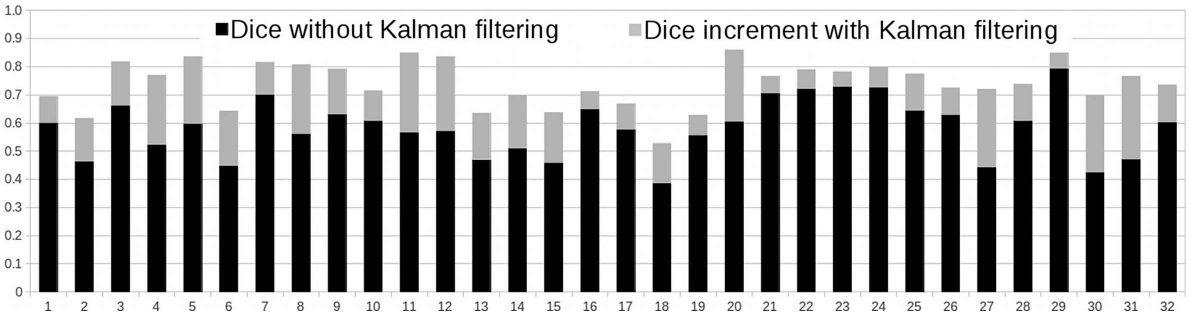


**Figure 12.** *DCs of the postablation LA image segmentation using the proposed Algorithm 3 without the Kalman filtering (black bars). The increment of DC using the Kalman filter is in gray. Therefore the total bar heights corresponds to the values in Table 4.*

this, we compare a scheme without such filtering and with such filtering, in order to have a controlled evaluation of the added value of such a component. For the same reason, we did not treat it as yet another standalone multiatlas algorithm. That said, the proposed algorithm is also compared with different state-of-the-art registration packages: ANTs [6], DRAMMS [44], FSL [30], and Elastix [35]. For the groupwise registration, we used GLIRT (Groupwise and Longitudinal Image Registration Toolbox) [57]. Both affine and deformable registrations are performed using the above packages. Then Bayesian fusion is computed with our own program. Since we have the ground truth of manual segmentation (striatum, LA, and prostate), a leave-one-out validation is used. The results of the compared packages are shown in Table 5. From the comparison, several aspects of the proposed framework can be observed.

First, while the Kalman multiatlas achieves good overall accuracy, the standard deviations of the DCs remain the smallest in all tests. Indeed, modeling the entire registration process with a dynamical system whose state is estimated using the Kalman filter is the key component that brings such robustness.

Second, in the striatum segmentation, the performances of different algorithms do not differ much. This may be caused by the fact that brain images have less variability both in their content as well as in the imaging conditions. Once the entire brain is registered, the

**Table 5**

*DCs between the expert segmentation and the automatic result generated by all the packages. The* largest *mean DCs and* smallest *standard deviations are in bold face.*

| | Striatum | | Prostate | | Pre-op LA | | Post-op LA | |
|---|---|---|---|---|---|---|---|---|
| | Mean | STD | Mean | STD | Mean | STD | Mean | STD |
| ANTs | **92.0** | 1.37 | 71.7 | 3.58 | 71.3 | 8.04 | 71.10 | 9.34 |
| Elastix | 90.7 | 1.55 | 76.1 | 6.15 | 72.7 | 7.69 | 72.5 | 9.15 |
| DRAMMS | 91.1 | 1.19 | 71.1 | 4.66 | 72.6 | 7.96 | 72.5 | 8.32 |
| FSL | 88.1 | 2.24 | 70.6 | 4.53 | 71.1 | 8.19 | 72.1 | 8.28 |
| GLIRT | 90.9 | 1.17 | 75.9 | 5.72 | **74.5** | 7.95 | 73.8 | 8.57 |
| Proposed without Kalman filtering | 89.8 | 2.17 | 72.7 | 6.71 | 71.0 | 8.39 | 71.13 | 9.18 |
| Proposed with Kalman filtering | 91.1 | **0.91** | **76.2** | **3.8** | 74.4 | **7.35** | **74.0** | **8.11** |

internal structures are relatively well aligned. In such cases, the statistical estimation does not provide much extra benefit. This is the same as other optimal estimation tasks, such as object tracking. When the observation step alone already achieves good accuracy, the dynamic modeling and prediction step does not add too much to the overall accuracy and robustness.

However, for the more challenging cases of the pelvic and cardiac registrations where much stronger variances exist in image contents, regions of interest, intensity differences, etc., the robustness becomes a major issue. In such cases, the Kalman filtering successfully reduces the variation and finally provides a better segmentation.

This is consistent with the performance of the groupwise registration technique. For the relatively less challenging striatum cases, the improvement of groupwise registration is not significant. However, for the prostate and the atrium volumes, by unifying the costs of all the training images together, the groupwise registration also achieves better performance, though at much higher extra computational load when compared to the recursive least square approach in the Kalman filter, where no additional computation cost is observed.

Last, as has been alluded to, the proposed framework stems from the perspective of treating the multiatlas construction process as a dynamical system, whose optimal estimation is performed by the Kalman filter. Such a system perspective and filtering scheme is generic enough and may be combined with almost any of the existing registration and fusion techniques. Because of this, we believe other registration methods may also benefit from adopting such a system perspective into their pipeline. That is, by going deep into the source code of various packages, one may augment, for example, DRAMMS with Kalman filtering for even better performance.

**3.5. Order matters.** We conducted experiments to access whether the order of the training image registration under the Kalman filtering framework would affect the final segmentation performance. Given $N$ training images, there are $N!$ different orders. Unfortunately, this is too many for an exhaustive test for realistic values of $N$. Therefore, we randomly chose 100 different permutations of the array $0, \dots, N$ and followed that sequence to perform the proposed segmentation. The resulting mean DCs are recorded and plotted in Figure 13. In the figure, each of the four bars records the 100 trails in different registration orders. The arrows to the right of the barplots indicate the mean DCs found in Tables 1, 2, 3, and 4.

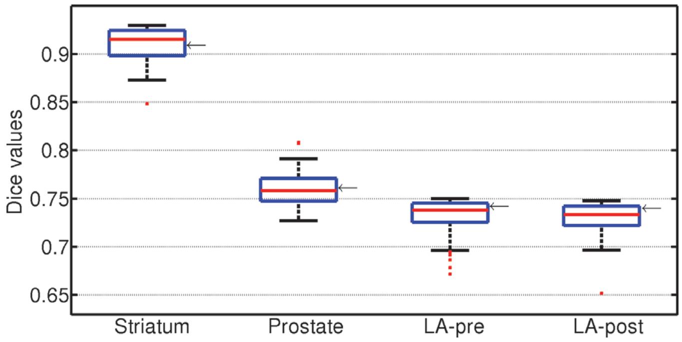


**Figure 13.** *Bar plots of the mean DCs for striatum, prostate, LA preablation, and LA postablation. Each bar records* 100 *trials using different orders. The arrow to the right of each bar is the mean DC reported in each of Tables* 1, 2, 3, *and* 4.

Admittedly, since $N!$ is a very large number, the 100 trials may be too few to reveal the overall influence of order. Nevertheless, it can be observed for most of the orders tested that the resulting mean DCs are greater than the DC without Kalman filtering scheme. Indeed, finding an optimal order, even local optimal, is a topic for future research. For example, given a novel image to be segmented, it may be possible to perform independent pairwise registrations to compute each final cost. Then, based on the similarity measured by the final costs, one may determine an order to perform the Kalman filtering multiatlas again. For example, it may be preferred that the most similar training image be registered first, so that its accurate transformation can be propagated along to improve the subsequent, possibly less accurate registration. In addition, after certain number of registrations the algorithm may be able to automatically determine it is sufficient to stop and discard the subsequent, presumably not very similar training images. We, however, have not been able to prove the optimality of such an approach and we are still working in such a direction.

**4. Discussion, conclusions, and future work.** In this work, we proposed a new dynamical system perspective to the commonly used multiatlas segmentation framework. By adopting such a point of view, the previously independently performed atlas registrations are now carried out under a filtering scheme. As a result, the accuracy and the robustness of the registrations and the overall segmentation have been significantly statistically improved. Furthermore, this new dynamical system framework may be employed in most of the existing multiatlas segmentation schemes and is expected to improve the overall segmentation accuracy by stabilizing the affine registration. However, we still need further tests to confirm whether the performance would indeed increase with other state-of-the-art registration and label fusion algorithms. Further, such a dynamical system point of view can be regarded as "fusion in the registration step." This may provide the impetus for new ideas in label fusion as well.

Regarding the type of transformation in the nonlinear registration step, vector field based deformation may also be employed. In this regard, we tested the diffeomorphic demon technique. However, due to the significant intensity variances among the images and within the same image, we found optical flow based deformable registration (demons and their variants) not to be applicable, in particular for the heart and prostate images. As a result, the final segmentation results have Dice values of about 40% for the LA and 50% for the prostate. On the other hand, B-spline with normalized correlation based registration cost handles the intensity variances robustly. In fact, regardless of the specific registration method being used, using the dynamical system point of view and filtering may help to stabilize the registration process. Incorporating vector field based deformable registration with various metrics is part of our ongoing research.

It is also noted that there are many state-of-the-art fusion techniques. On the other hand, in this work the main theme is the use of statistical filtering in order to exploit the correlation between training images. As a result, we mentioned but did not implement more sophisticated fusion techniques, in order to keep the overall algorithm better controlled and to avoid implementation variations. Moreover, it has been indicated by several researchers that the computationally more intensive fusion techniques do not necessarily give better results than even majority voting; see, for example, Figure 4 of [15] and Figure 5 of [36]. Indeed, it would be out of the scope of the current work to evaluate which fusion algorithm fits the problem the best. As a result, we performed the controlled experiment with the focus on the main filtering theme. In fact, we believe that using Kalman filtering together with different fusion techniques, such as STAPLE, is an ongoing and future direction.

Future work includes incorporating more sophisticated registration techniques and label fusion algorithms into the Kalman multiatlas framework. By doing so, it is expected that the overall segmentation performance would be further improved. Moreover, as mentioned above, finding an optimal or better order for performing the filtering is also an ongoing research problem.

## REFERENCES


[1] N. Akoum, M. Daccarett, C. McGann, N. Segerson, G. Vergara, S. Kuppahally, T. Badger, N. Burgon, T. Haslam, and E. Kholmovski, et al., *Atrial fibrosis helps select the appropriate patient and strategy in catheter ablation of atrial fibrillation: A DE-MRI guided approach*, J. Cardiovascular Electrophysiology, 22 (2011), pp. 16–22.

[2] P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, *Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy*, Neuroimage, 46 (2009), pp. 726–738.

[3] S. Allassonnière, Y. Amit, and A. Trouvé, *Towards a coherent statistical framework for dense deformable template estimation*, J. R. Stat. Soc. Ser. B Stat. Methodol., 69 (2007), pp. 3–29.

[4] N. Archip, O. Clatz, S. Whalen, D. Kacher, A. Fedorov, A. Kot, N. Chrisochoides, F. Jolesz, A. Golby, P. M. Black, et al., *Non-rigid alignment of pre-operative MRI, fMRI, and DT-MRI with intra-operative MRI for enhanced visualization and navigation in image-guided neurosurgery*, Neuroimage, 35 (2007), pp. 609–624.

[5] X. Artaechevarria, A. Muñoz-Barrutia, and C. Ortiz-de Solorzano, *Combination strategies in multi-atlas image segmentation: Application to brain MR data*, IEEE Trans. Medical Imaging, 28 (2009), pp. 1266–1277.

[6] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, *A reproducible evaluation of ANTs similarity metric performance in brain image registration*, Neuroimage, 54 (2011), pp. 2033–44.

[7] D. J. Blezek and J. V. Miller, *Atlas stratification*, Medical Image Anal., 11 (2007), pp. 443–457.

[8] Z. I. Botev, J. F. Grotowski, and D. P. Kroese, *Kernel density estimation via diffusion*, Ann. Statist., 38 (2010), pp. 2916–2957.

[9] CARMA *Left Atrial MRI*, http://www.insight-journal.org/midas/collection/view/197.

[10] D. L. Collins and J. C. Pruessner, *Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting animal with a template library and label fusion*, NeuroImage, 52 (2010), pp. 1355–1366.

[11] T. F. Cootes, G. J. Edwards, and C. J. Taylor, *Active appearance models*, IEEE Trans. Pattern Analysis Machine Intelligence, 23 (2001), pp. 681–685.

[12] P. E. Danielsson, *Euclidean distance mapping*, Comput. Graphics Image Processing, 14 (1980), pp. 227–248.

[13] M. Depa, G. Holmvang, E. J. Schmidt, P. Golland, and M. R. Sabuncu, *Towards efficient label fusion by pre-alignment of training data*, in Proceedings of MICCAI, 2011, pp. 38–46.

[14] L. R. Dice, *Measures of the amount of ecologic association between species*, Ecology, 26 (1945), pp. 297–302.

[15] J. Doshi, G. Erus, Y. Ou, B. Gaonkar, and C. Davatzikos, *Multi-atlas skull-stripping*, Academic Radiology, 20 (2013), pp. 1566–76.

[16] M. Erdt, S. Steger, and G. Sakas, *Regmentation: A new view of image segmentation and registration*, J. Radiation Oncology Informatics, 4 (2012), pp. 1–23.

[17] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. Dale, *Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain*, Neuron, 33 (2002), pp. 341–355.

[18] E. Frenoux, V. Barra, and J. Y. Boire, *Segmentation of the striatum using data fusion*, in Proceedings of EMBS, Vol. 3, IEEE, 2001, pp. 2630–2633.

[19] Y. Gao, B. Corn, D. Schifter, and A. Tannenbaum, *Multiscale 3D shape representation and segmentation with applications to hippocampal/caudate extraction from brain MRI*, Medical Image Anal., 16 (2012), pp. 374–385.

[20] Y. Gao, B. Gholami, R. S. MacLeod, J. Blauer, W. M. Haddad, and A. R. Tannenbaum, *Segmentation of the endocardial wall of the left atrium using local region-based active contours and statistical shape learning*, in Proc. SPIE, 7623 (2010), 76234Z1–8.

[21] Y. Gao, Y. Rathi, S. Bouix, and A. Tannenbaum, *Filtering in the diffeomorphism group and the registration of point sets*, IEEE Trans. Image Processing, 21 (2012), pp. 4383–4396.

[22] Y. Gao, R. Sandhu, G. Fichtinger, and A. Tannenbaum, *A coupled global registration and segmentation framework with application to magnetic resonance prostate imagery*, IEEE Trans. Medical Imaging, 29 (2010), pp. 1781–1794.

[23] Y. Gao and A. Tannenbaum, *Combining atlas and active contour for automatic 3D medical image segmentation*, in Proceedings of ISBI, IEEE, 2011, pp. 1401–1404.

[24] S. Haker, G. Sapiro, and A. Tannenbaum, *Knowledge-based segmentation of SAR data with learned priors*, IEEE Trans. Image Processing, 9 (2000), pp. 299–301.

[25] F. Hausdorff, *Set Theory*, Chelsea, New York, 1957.

[26] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, *Automatic anatomical brain MRI segmentation combining label propagation and decision fusion*, NeuroImage, 33 (2006), pp. 115–126.

[27] I. Isgum, M. Staring, A. Rutten, M. Prokop, M. A. Viergever, and B. van Ginneken, *Multi-atlas-based segmentation with local decision fusion: Application to cardiac and aortic segmentation in CT scans*, IEEE Trans. Medical Imaging, 28 (2009), pp. 1000–1010.

[28] P. Jaïs, R. Weerasooriya, D. C. Shah, M. Hocini, L. Macle, K. J. Choi, C. Scavee, M. Haïssaguerre, and J. Clémenty, *Ablation therapy for atrial fibrillation (AF)*, Cardiovascular Research, 54 (2002), pp. 337–346.

[29] A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, and M. J. Thun, *Cancer statistics*, CA Cancer J. Clin., 57 (2007), pp. 43–66.

[30] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, *FSL*, NeuroImage, 62 (2012), pp. 782–90.

[31] S. Joshi, B. Davis, M. Jomier, and G. Gerig, *Unbiased diffeomorphic atlas construction for computational anatomy*, NeuroImage, 23 (2004), pp. S151–S160.

[32] S. J. Julier and J. K. Uhlmann, *A new extension of the Kalman filter to nonlinear systems*, in Proceedings of the International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Vol. 3, SPIE, Bellingham, WA, 1997.

[33] S. J. Kish, K. Shannak, and O. Hornykiewicz, *Uneven pattern of dopamine loss in the striatum of patients with idiopathic Parkinson's disease*, New England J. Med., 318 (1988), pp. 876–880.

[34] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, *On combining classifiers*, IEEE Trans. Pattern Analysis Machine Intelligence, 20 (1998), pp. 226–239.

[35] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, *ELASTIX: A toolbox for intensity based medical image registration*, IEEE Trans. Medical Imaging, 29 (2010), pp. 196–205.

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

[37] J. J. Levitt, L. K. Rosow, U. Khan, R. Kikinis, F. A. Jolesz, R. W. McCarley, and M. E. Shenton, *The ventral and dorsal striatum in schizophrenia: An MRI study*, Schizophrenia Bull., 33 (2007), p. 342.

[38] J. M. P. Lotjonen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, D. Rueckert, et al., *Fast and robust multi-atlas segmentation of brain magnetic resonance images*, Neuroimage, 49 (2010), pp. 2352–2365.

[39] B. Ma and R. E. Ellis, *Surface-based registration with a particle filter*, in Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Comput. Science, 3216, 2004, pp. 566–573.

[40] N. F Marrouche, D. Wilber, G. Hindricks, P. Jais, N. Akoum, F. Marchlinski, E. Kholmovski, N. Burgon, N. Hu, et al., *Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: The DECAAF study*, J. AMA, 311 (2014), pp. 498–506.

[41] J. Melonakos, Y. Gao, and A. Tannenbaum, *Tissue Tracking: Applications for Brain MRI Classification*, in Proc. SPIE (2007).

[42] *Prostate Segmentation Challenge*, http://wiki.na-mic.org/Wiki/index.php/Training_Data_Prostate_Segmentation_Challenge_MICCAI09.

[43] M. H. Moghari and P. Abolmaesumi, *Point-based rigid-body registration using an unscented Kalman filter*, IEEE Trans. Medical Imaging, 26 (2007), pp. 1708–1728.

[44] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, *DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting*, Medical Image Anal., 15 (2011), pp. 622–39.

[45] J. S. Paulsen, P. C. Nopoulos, E. Aylward, C. A. Ross, H. Johnson, V. A. Magnotta, A. Juhl, R. K. Pierson, J. Mills, D. Langbehn, et al., *Striatal and white matter predictors of estimated diagnosis for huntington disease*, Brain Research Bull., 82 (2010), pp. 201–207.

[46] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, *Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains*, NeuroImage, 21 (2004), pp. 1428–1442.

[47] T. Rohlfing, D. B. Russakoff, and C. R. Maurer, *Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation*, IEEE Trans. Medical Imaging, 23 (2004), pp. 983–994.

[48] M. R. Sabuncu, S. K. Balci, M. E. Shenton, and P. Golland, *Image-driven population analysis through mixture modeling*, IEEE Trans. Medical Imaging, 28 (2009), pp. 1473–1487.

[49] M. R. Sabuncu, B. T. T. Yeo, K. V. Leemput, B. Fischl, and P. Golland, *A generative model for image segmentation based on label fusion*, IEEE Trans. Medical Imaging, 29 (2010), pp. 1714–1729.

[50] D. Salmond, N. Gordon, and A. Smith, *Novel approach to nonlinear/non-Gaussian Bayesian state estimation*, in IEE Proc. F, 140 (1993), pp. 107–113.

[51] R. Sandhu, S. Dambreville, and A. Tannenbaum, *Particle filtering for registration of 2D and 3D point sets with stochastic dynamics*, in Proceedings of IEEE CVPR, 2008, pp. 1–8.

[52] D. Simon, *Optimal State Estimation: Kalman, $H_\infty$ and Nonlinear Approaches*, LibreDigital, 2006.

[53] R. Toth, P. Tiwari, M. Rosen, G. Reed, J. Kurhanewicz, A. Kalyanpur, S. Pungavkar, and A. Madabhushi, *A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation*, Medical Image Anal., 15 (2011), pp. 214–225.

[54] J. Vogel-Claussen, C. E. Rochitte, K. C. Wu, I. R. Kamel, T. K. Foo, J. A. C. Lima, and D. A. Bluemke, *Delayed enhancement mr imaging: Utility in myocardial assessment*, Radiographics, 26 (2006), pp. 795–810.

[55] H. Wang, J. W. Suh, S. Das, J. Pluta, M. Altinay, and P. Yushkevich, *Regression-based label fusion for multi-atlas segmentation*, in Proceedings of CVPR, IEEE, 2011, pp. 1113–1120.

[56] H. Wang, J. W. Suh, S. Das, J. Pluta, C. Craige, and P. Yushkevich, *Multi-atlas segmentation with joint label fusion*, IEEE Trans. Pattern Anal. Machine Intelligence, 35 (2012), pp. 611–623.

[57] G. Wu, H. Jia, Q. Wang, and D. Shen, *SharpMean: Groupwise registration guilded by sharp mean image and tree-based registration*, NeuroImage, 54 (2011), pp. 1968–1981.