Published in final edited form as:
Proc Int Conf Image Proc. 2011 ; : 2177–2180. doi:10.1109/ICIP.2011.6116064.

TEMPORAL REGISTRATION OF PARTIAL DATA USING PARTICLE FILTERING

Guy Nir[1] and **Allen Tannenbaum**[2]
[1]University of British Columbia, Vancouver, BC, Canada
[2]Georgia Institute of Technology, Atlanta, GA, USA

Abstract

We propose a particle filtering framework for rigid registration of a model image to a time-series of partially observed images. The method incorporates a model-based segmentation technique in order to track the pose dynamics of an underlying observed object with time. An applicable algorithm is derived by employing the proposed framework for registration of a 3D model of an anatomical structure, which was segmented from preoperative images, to consecutive axial 2D slices of a magnetic resonance imaging (MRI) scan, which are acquired intraoperatively over time. The process is fast and robust with respect to image noise and clutter, variations of illumination, and different imaging modalities.

Keywords

Index Terms; Image registration; image segmentation; biomedical image processing; active contours; particle filters

1. INTRODUCTION AND RELATED WORK

Image registration is the process of establishing a common geometric reference frame between two or more image data sets, i.e., bringing the images into spatial alignment. The images are typically of the same, or similar, object or scene possibly acquired at different times, from different perspectives or by different imaging techniques. Registration of time-series of images is known as temporal registration.

Registration is a well-known problem in image processing and used in numerous applications where data sets of images are compared or integrated, requiring the mapping of the data into one coordinate system. The motivation for this work arose from the work in medical applications, in which consecutive axial 2D slices of a magnetic resonance imaging (MRI) scan, acquired intraoperatively over time, need to be registered to a 3D model of the anatomical structure, segmented from preoperative images. Such slice-to-volume temporal registration algorithm may be implemented to assist in image guided surgery and therapy.

Since 2D slices (with an infinitesimal thickness) can be viewed as partial data of 3D images, we turn to discuss the generalized problem of registering images with partial or missing data. Then, we may utilize the solution in order to solve the slice-to-volume problem as a special case.

While the model image in our case is assumed to be known in advance, thus can be preprocessed and segmented thoroughly by some algorithm (or hand segmented), the partially observed images are typically acquired in real-time, computationally allows for only partial or no segmentation process. Therefore, we are interested in a model-based registration [1], where a model is transformed to fit the observed images.

We use the model image to provide a shape prior for the underlying object in the partially observed images. In [2], a shape prior is employed to segment medical images by minimizing region-based functionals. We employ this approach to solve the registration of partial data problem by applying appropriate modifications.

Temporal filtering techniques are widely used for tracking objects from video or series of images. In [3] the authors proposed a particle filtering framework for tracking highly deformable objects in the presence of noise and clutter by incorporating dynamic shape priors and image statistics. In this paper, we follow a similar approach in order to solve the temporal registration problem.

2. PROPOSED METHOD

2.1. Chan-Vese on Partial Data

An extensively used region-based active contour model for image segmentation was proposed by Chan and Vese in [4]. In this model, an energy functional is minimized with respect to (w.r.t.) a segmenting hypersurface Γ (a curve in 2D or a surface in 3D) and a pair of parameters c_-, c_+. The Chan-Vese model was generalized to vector-valued image (multichannel) segmentation and to multiple regions (multiphase) segmentation [5]. We propose a model that takes into account only values inside a subdomain on which the image is given. As in [4], we use level sets to represent hypersurfaces implicitly.

Suppose we observe grayscale intensities of a partial image $I: \Omega \to \mathbf{R}_+$ on some observed subdomain $\Omega \subseteq \mathcal{D} \subseteq \mathbf{R}^d$. Let the boundary $\partial\Omega$ of the observed subdomain be embedded as the zero level set of a level set function $\Psi: \mathbf{R}^d \to \mathbf{R}$, with negative values assigned to the region inside and positive values assigned to the region outside. Also, let $\Phi: \mathbf{R}^d \to \mathbf{R}$ be a level set function embedding the segmenting hypersurface Γ, with negative values inside and positive values outside. We define the partial data Chan-Vese energy functional as

$$E_{CV}^{\Psi}[\Phi, c_-, c_+] = \int_{\mathcal{D}} (I(x)-c_-)^2 \mathscr{H}(-\Phi(x))\mathscr{H}(-\Psi(x))\,dx + \int_{\mathcal{D}} (I(x)-c_+)^2 \mathscr{H}(\Phi(x))\mathscr{H}(-\Psi(x))\,dx, \quad (1)$$

where $\mathcal{H}$ denotes the Heaviside function. We have used the fact that $\mathcal{H}(-\Phi)\,\mathcal{H}(-\Psi)$ and $\mathcal{H}(\Phi)\,\mathcal{H}(-\Psi)$ are the characteristic functions of the regions {inside(Γ) ∩ Ω} and {outside(Γ) ∩ Ω}, respectively.

Note that it does not matter how I is extended onto the entire domain $\mathcal{D}$, since values outside Ω are not considered. Also, neither the subdomain Ω nor the hypersurface Γ needs to be simple, keeping in mind that in the level set formulation it is inherently dealt with.

Similar to the classical Chan-Vese model, for a fixed Γ (thus a fixed Φ), the parameters c_-, c_+ that minimize the functional (1) are the mean intensities of the observed image on the intersection of the subdomain Ω with the regions inside and outside Γ, respectively.

We can express measures of Γ in terms of Ω, such as

$$\begin{aligned} &\text{volume inside } (\mathit{outside}): A_{\mp}^{\Psi}[\Phi] = \int_{\mathcal{D}} \mathscr{H}(\mp\Phi)\mathscr{H}(-\Psi)\,dx, \\ &\text{sum inside (outside)}: S_{\mp}^{\Psi}[\Phi] = \int_{\mathcal{D}} I\mathscr{H}(\mp\Phi)\mathscr{H}(-\Psi)\,dx, \end{aligned} \quad (2)$$

NIH-PA Author Manuscript
NIH-PA Author Manuscript

where the notations $\cdot_{-}^{\Psi}$ and $\cdot_{+}^{\Psi}$ refer to the regions {inside(Γ) ∩ Ω} and {outside(Γ) ∩ Ω}, respectively. We can then write the mean intensities as

$$c_{\mp}^{\Psi}[\Phi]=\frac{S_{\mp}^{\Psi}[\Phi]}{A_{\mp}^{\Psi}[\Phi]}. \quad (3)$$

Therefore, Γ has the role of partitioning the partially observed image to minimize the intensity variances on the regions inside and outside it. Assuming that the object and the background have different intensity ranges (but not necessarily clear edges), the optimal Γ is expected to approximate the boundary of the underlying object on the given subdomain Ω. However, this does not ensure Γ will fit the full boundary of the object, since on the rest of the domain $\mathcal{D}\backslash\Omega$ it can take any form (subject to continuity). In other words, the segmentation of partial data problem is ill-posed since there may be more than one global minimizer, and, as long as we are not given any other information, any one of them is as good as the other.

In the next section, we will use the shape of the object in the model image as a prior. In addition, we may be given observations on several subdomains, and the greater is the region they cover, the better are the odds the hypersurface that fits all of them together is the sought-after one.

2.2. Model-Based Registration of Partial Data

Suppose we are given, *a priori*, a model image I_0: $\mathcal{D} \to \mathbf{R}_+$ on the domain $\mathcal{D} \subseteq \mathbf{R}^d$. As discussed in Section 1, we assume that I_0 is segmented, and the hypersurface Γ_0 describing the shape of the model object is extracted and embedded as the zero level set of Φ_0.

Let $x \in \mathbf{R}^d$ be a d-dimensional spatial coordinates vector. Then, a rigid transformed model hypersurface $\tilde{\Gamma}_0$ is embedded as the zero level set of the following transformed function

$$\tilde{\Phi}_0(x;s)=\Phi_0(T_s^{-1}x), \quad (4)$$

where $T_s \in \mathbf{R}^{(d+1)\times(d+1)}$ is the rigid transformation matrix w.r.t. the pose vector $s \in \mathbf{R}^{(\frac{d(d+1)}{2}+1)}$ comprising the pose parameters: d translation values, a (non-zero) scale factor and $\frac{d(d-1)}{2}$ rotation angles. Note that T_s operates on homogeneous coordinates $(x^T\ 1)^T$, and we denote this linear mapping as $T_s^{-1}x$ for short. Details and explicit expressions of T_s in 2D and 3D can be found, e.g., in [6].

The model-based rigid registration of partial data problem is finding the pose vector $s^{\star}$, which yields the rigid transformed model $\tilde{\Phi}_0$ that fits optimally, in some sense, the observed image I on the observed subdomain Ω. The problem can thus be formulated as $\min_s E[\tilde{\Phi}_0(\cdot;s)]$, where E is a functional that measures similarity, e.g., the sum of squared differences (SSD), likelihood measurement, correlation ratio, normalized correlation, or mutual information [1]. We use the Chan-Vese energy functional above to measure similarity. This choice was proved to be useful for the purposes of our work. However, other functionals can be easily incorporated into the proposed framework.

We take $\Phi = \tilde{\Phi}_0$ in the partial data Chan-Vese energy (1), so as to use the model as the shape prior for the segmenting hypersurface, allowing only rigid transformations. By opening the squared terms in (1) and using (2), we have that the Chan-Vese functional to be minimized is equivalent (up to a term which does not depend upon Φ) to the following function of the pose vector

$$E_{CV}^{\Psi}(s)=-\left(\frac{(S_{-}^{\Psi}(s))^2}{A_{-}^{\Psi}(s)}+\frac{(S_{+}^{\Psi}(s))^2}{A_{+}^{\Psi}(s)}\right), \quad (5)$$

where we have used abbreviated notations to describe the dependence of the functionals on s (e.g., by $S_{-}^{\Psi}(s)$ we mean $S_{-}^{\Psi}[\tilde{\Phi}_0(\cdot;s)]$).

The rigid pose parameters that minimize the energy function (5) lead to the optimal transformed model hypersurface, which approximates the boundary of the object on the observation image. Therefore, the optimal pose is the solution of the model-based rigid registration of partial data problem.

Starting in a pose $s^{(0)}$, we minimize (5) using gradient descent, i.e., in iteration k the pose is $s^{(k)}=s^{(k-1)}-\alpha\nabla_s E_{CV}^{\Psi}(s^{(k-1)})$, with a small $\alpha > 0$. The gradient of E_{CV}^{Ψ} using (3) is

$$\nabla_s E_{CV}^{\Psi}(s)=-2\{c_{-}^{\Psi}(s)\nabla_s S_{-}^{\Psi}(s)-c_{+}^{\Psi}(s)\nabla_s S_{+}^{\Psi}(s)\}+\{(c_{-}^{\Psi}(s))^2\nabla_s A_{-}^{\Psi}(s)+(c_{+}^{\Psi}(s))^2\nabla_s A_{+}^{\Psi}(s)\}, \quad (6)$$

Taking the gradients of (2), we have that

$$\begin{aligned}\nabla_s A_{\mp}^{\Psi} &= \mp\int_{\mathscr{D}}\delta(\mp\tilde{\Phi}_0)\mathscr{H}(-\Psi)\nabla_s\tilde{\Phi}_0 dx,\\ \nabla_s S_{\mp}^{\Psi} &= \mp\int_{\mathscr{D}} I\delta(\mp\tilde{\Phi}_0)\mathscr{H}(-\Psi)\nabla_s\tilde{\Phi}_0 dx,\end{aligned} \quad (7)$$

where δ denotes the Dirac delta function. Using (4) and the chain rule, the lth component of the gradient of the transformed level set function $\nabla_s\tilde{\Phi}_0(\cdot;\, s)$ is given by

$$\frac{\partial}{\partial s_l}\tilde{\Phi}_0(x;s)=\frac{\partial}{\partial s_l}\Phi_0(T_s^{-1}x)=\nabla\Phi_0(T_s^{-1}x)^T\frac{\partial}{\partial s_l}T_s^{-1}x, \quad (8)$$

which can be computed as shown in [2].

Figure 1 shows the result of employing the proposed algorithm on an image of a Formula One car. Note that the resulted transformed model, overlaid on the (unknown) complete image in Figure 1(c), does not fully match the image on the unobserved subdomain.

2.3. Temporal Registration

Particle filters (PFs) [7] are sequential Monte Carlo methods based on point mass (or “particle”) representations of probability density functions (pdf’s), which can be applied to any state-space model (i.e., nonlinear and non-Gaussian). As the number of particles becomes very large, this characterization becomes an equivalent representation to the usual functional description of pdf, and PF approaches the optimal Bayesian estimate.

In our temporal registration of partial data problem, we want to track the pose of the observed object. Therefore, we define the state vector to be the pose vector s_t of the transformed segmenting hypersurface at a discrete time t. At each time $t \geq 1$, we observe a noisy partial image I_t on an observed subdomain Ω_t, whose boundary is embedded in Ψ_t. Note that the underlying object in I_t may have rigid dynamics over time. We use PF to

NIH-PA Author Manuscript

NIH-PA Author Manuscript

recursively obtain $p(s_t|\Psi_{1:t})$, the posterior pdf of the current pose parameters given the past and current observed subdomains.

The state process is assumed to be Markov, i.e., $p(s_t|s_{0:t-1}) = p(s_t|s_{t-1})$, and the observations are assumed to be conditionally independent given the current state, i.e., $p(\Psi_t|s_{0:t}) = p(\Psi_t|s_t)$. Since it is difficult to sample from the posterior pdf, we use sampling importance re-sampling (SIR) [8], which is based upon a scheme proposed in [3], and described briefly here.

The algorithm starts with sampling N particles (states) from the initial state distribution $p(s_0)$, in order to approximate it by $p(s_0) \approx \frac{1}{N}\sum_{i=1}^{N}\delta(s_0 - s_0^i)$, where s_0^i is the ith particle. Also, an initial importance weight $\omega_0^i = \frac{1}{N}$ is set to each particle. Then, at each time step, the particles propagate according to the stages below.

Prediction Stage—We assume the variation in the pose of the observed object over time to have noisy characteristic, and generate

$$\hat{s}_t^i = s_{t-1}^i + n_t, \qquad 1 \le i \le N, \qquad (9)$$

where $\hat{s}_t^i$ is the predicted ith particle at time t, and $n_t \in \mathbf{R}^{(\frac{d(d+1)}{2}+1)}$ is an identically and independently distributed (i.i.d.) process noise sequence with known pdf.

In our experiments, we used $n_t \sim \mathcal{N}(0, \Sigma_n)$, namely, a zero mean Gaussian noise vector with a diagonal covariance matrix Σ_n. The variances σ_n^2 in the diagonal express the magnitude of the assumed dynamic, with larger values correspond to rapid dynamics. Although PF allows more complex dynamic model than this “random velocity”, it proved to provide plausible results in the absence of information about the dynamics.

Update Stage—When an observation Ψ_t becomes available, we employ the model-based registration of partial data, described above in Section 2.2, over all past and current observations $\Psi_{1:t}$. The process is applied to each particle i, with its predicted state $\hat{s}_t^i$ taken as the initial pose vector. The resulted pose parameters are put together as the new state s_t^i of the ith particle. This stage is notated as

$$s_t^i = f_{reg}^L(\hat{s}_t^i), \qquad (10)$$

where f_{reg}^L stands for performing L iterations of gradient descent using (6)–(8) to minimize $\sum_{\tau=1}^{t}\gamma^{(t-\tau)}E_{CV}^{\Psi_\tau}$, with $0 < \gamma \le 1$ as a discount factor. This type of update process can be seen as an importance sampling, as explained in [3].

We perform only L iterations, rather than running the process until convergence, since a local minimizer hypersurface would be highly dependent upon the observations, leading the state to lose its dependence on the previous one, and causing loss of track in case the observations are bad. Thus, the choice of L expresses the trust in the observations, with large values correspond to high trust in the observations, while small values correspond to high trust in the system model (9).

The choice of the discount factor γ also provides a measure of trust in the observations. A value of $\gamma \approx 1$ results in approximately uniform weights assigned to the current and past observations, while a value of $\gamma \approx 0$ results in very low weighting of past observations. Therefore, the choice of this parameter expresses the trust in old observations and can be used to discard them, e.g., in case of large variations in pose, where old observations are irrelevant.

Weighting and Resampling—After all particles were sampled from (10), we are to assign an importance weight to each one. We define the likelihood pdf, i.e., the probability of the current observation given the state vector, as

$$p(\Psi_t|s_t) \propto \exp\left\{-E_{CV}^{\Psi_t}(s_t)/\sigma_v^2\right\}, \quad (11)$$

where σ_v^2 is another parameter that decides how much one trusts the system model versus the observation model, with larger values in favor of the system model. The parameter depends on the observed image statistics. If the image measurement noise is modeled as a Gaussian noise v_t, then we can take σ_v^2 to be its variance. Note that, from the discussion on the Chan-Vese model in Section 2.1, if the pose s_t provides a good segmentation of the partial data, the energy takes small values, which by (11) results in high likelihood.

The prior pdf, i.e., the probability of the current state given the previous state vector, is given by (9) with $n_t \sim \mathcal{N}(0, \Sigma_n)$, as

$$p(s_t|s_{t-1}) \propto \exp\left\{-\frac{1}{2}(s_t-s_{t-1})^T \sum\nolimits_n^{-1}(s_t-s_{t-1})\right\}. \quad (12)$$

We use (11), (12) according to [8] and get the following recursion for importance weights

$$\omega_t^i \propto \omega_{t-1}^i \exp\left\{-E_{CV}^{\Psi_t}(s_t)/\sigma_v^2 - \frac{1}{2}(s_t-s_{t-1})^T \sum\nolimits_n^{-1}(s_t-s_{t-1})\right\}, \quad (13)$$

normalized such that $\sum_{i=1}^{N}\omega_t^i=1$.

The posterior pdf of the state given past and current observations is then approximated by $p(s_t|\Psi_{1:t}) \approx \sum_{i=1}^{N}\omega_t^i\delta(s_t-s_t^i)$, from which an optimal estimate of the state, w.r.t. any criterion, may be obtained. In our experiments we used the maximum *a posteriori* (MAP) state, which corresponds to the state of the highest weighted particle, i.e.,

$$s_t^{MAP}=s_t^j, \quad \text{with} \quad j=\arg\max_{1\le i\le N}\omega_t^i. \quad (14)$$

Finally, resampling is performed. We resample N times according to the approximated posterior distribution above, to have new particles $\{s_t^i\}_{i=1}^N$ that replace the current. This step eliminates particles that have very low weights and concentrates on higher weighted particles. The importance weights for the resampled particles are reset to $\omega_t^i=\frac{1}{N}$ (thus, ω_{t-1}^i can be omitted from (13)).

3. EXPERIMENTS

We demonstrate the proposed method on two scenarios inspired by the application of a slice-to-volume temporal registration. The image domain is the 3D unit box, i.e., $\mathcal{D} = [0, 1]^3$ ($d = 3$). The observations are consecutive 2D slices of the observed image, sampled equally spaced along the z axis, i.e., at each time t an observed image I_t is sampled on a corresponding plane $z = z_t$. We use zero-order hold to interpolate the intermediate slices between the sampled ones, therefore, we can define the observed subdomain Ω_t as the 3D volume $\Omega_t = \{(x, y, z): (x, y, z) \in [0, 1] \times [0, 1] \times [z^t, z^{t+1}]$. Initial states are uniformly distributed in $\pm\frac{1}{4}$ and $\pm\frac{\pi}{4}$ intervals around the translation and rotation parameters of the model, respectively.

Brain MRI—In this example we have constructed a head model from the boundaries of 27 slices of a brain MRI scan. The observed slices are translated and rotated compared to the model. Also, a zero mean Gaussian noise with a variance of $\sigma_v^2 = 10^{-2}$ was added to each 64 × 64 slice. A variance of value $\sigma_n^2 = 10^{-2}$ was used for the pose prediction. Since the observation is static, only $N = 10$ particles were used, with $L = 25$ iterations, and a discount factor of $\gamma = 1.0$ to take into account all of the available data.

The results of the registration process at selected times are illustrated in Figures 2(a)–2(d). Top row shows the MAP model (blue mesh) and the underlying observation (red surface). The observed slice at each time (green plane) is also shown. Bottom row shows the intersection (cyan contour) of the MAP model surface with the corresponding observed slice.

Dynamic Monkey—In this example we test our method on a toy monkey model. Here, the observation is being translated and rotated along some trajectory in 3D over 49 time steps, during which 64×64 slices are observed. A variance of value $\sigma_n^2 = 10^{-2}$ was used for the pose prediction. Due to the dynamics, a higher number of $N = 25$ particles were used, with $L = 25$ iterations, and a discount factor of only $\gamma = 0.5$ to discard previous observed slices, since they are no longer represent the true state. Similar to Figure 2, the results of this registration process are illustrated in Figures 3(a)–3(d). Notice the topological changes in the observed images.

Indeed, the MAP state tracks the true state. To quantify the performance, we compute the error between the MAP state and the ground truth state at each time. Figure 4 shows the (normalized) error for both simulations, and illustrates the fast rate of convergence.

4. CONCLUSION AND FUTURE WORK

We presented a model-based approach to registration of partial data. A model shape is rigidly transformed to fit the partially observed image by minimizing the Chan-Vese energy functional. The method was then incorporated into a particle filtering framework, allowing temporal registration of partial data.

Advantages of the method are robustness to noise and different imaging modalities. It can be adapted to employ other region-based models, such as binary-mean or binary-variance [2]. It took about 2sec per particle to process one slice with an un-optimized MATLAB code on a 2.93GHz Quad Core CPU machine. However, optimized implementation on a GPU may achieve real-time performance. Disadvantages are usual problems of a gradient descent algorithm, which can be overcome using standard techniques such as multiresolution optimization or line search.

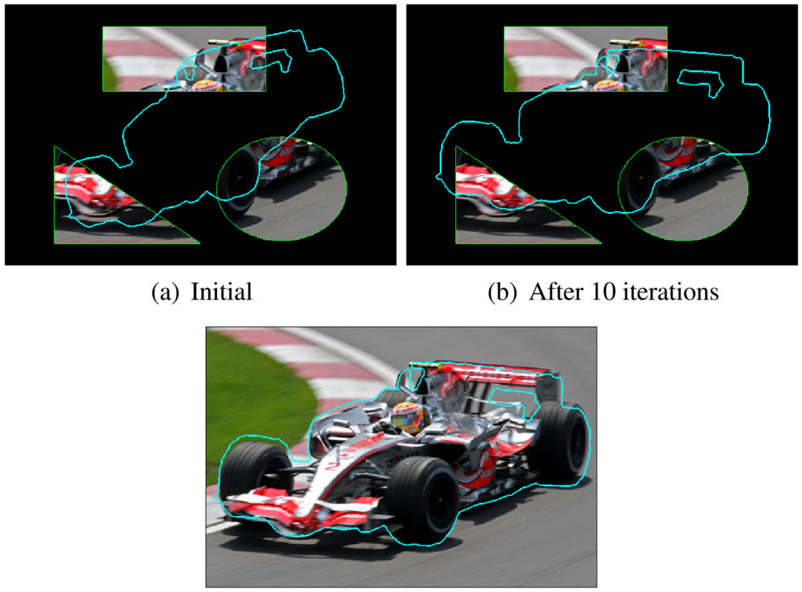
The framework can be generalized to capture deformations of the observation, by extracting shape variations from a set of model images using, e.g., PCA as in [2]. Since a particle filter allows for nonlinearity and non-Gaussianity, a complex dynamic model may be develop in order to simulate more accurately the movements of the anatomical structure *in vivo*.

Acknowledgments

This work was supported in part by grants from NSF, AFOSR, ARO, MURI, MRI-HEL as well as by a grant from NIH (NAC P41 RR-13218) through Brigham and Women's Hospital. This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from http://nihroadmap.nih.gov/bioinformatics.

References

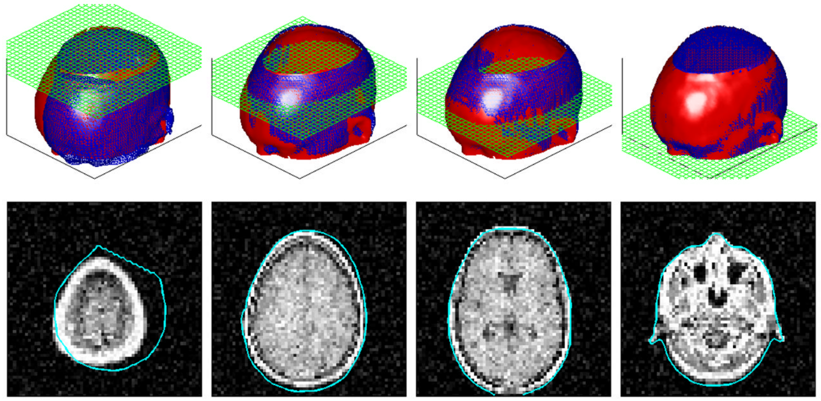
1. Maintz J, Viergever M. A survey of medical image registration. Med Im Anal. 1998; 2(1):1–36.
2. Tsai A, Yezzi A, et al. A shape-based approach to the segmentation of medical imagery using level sets. IEEE Trans on Medical Imaging. Feb; 2003 22(2):137–154.
3. Rathi Y, Vaswani N, Tannenbaum A. A generic framework for tracking using particle filter with dynamic shape prior. IEEE Trans on Im Proc. 2007; 16(5):1370–1382.
4. Chan T, Vese L. Active contours without edges. IEEE Trans on Im Proc. Feb.2001 10:266–277.
5. Vese L, Chan T. A multiphase level set framework for image segmentation using the Mumford and Shah model. Int J Comp Vis. 2002; 50(3):271–293.
6. Hartley, R.; Zisserman, A. Multiple View Geometry in Computer Vision. 2. Cambridge University Press; 2004.
7. Gordon NJ, Salmond DJ, Smith AFM. Novel approach to nonlinear/non-gaussian bayesian state estimation. IEE Proc F. 1993; 140(2):107–113.
8. Arulampalam MS, et al. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. IEEE Trans on Signal Proc. 2002; 50(2):174–188.



(a) Initial (b) After 10 iterations

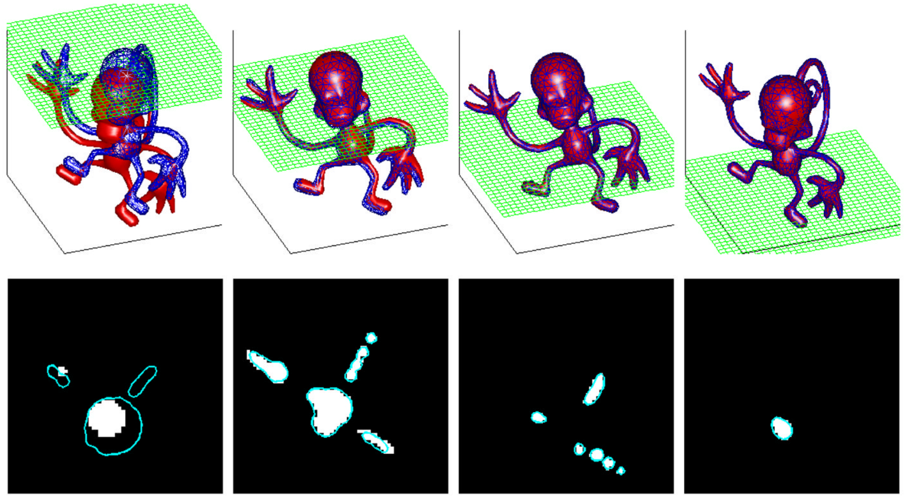
(c) Final (24 iterations)

Fig. 1.
Model-based registration of partial data set. The model curve is transformed to match the image on the observed subdomains.



(a) $t = 1$ (b) $t = 9$ (c) $t = 17$ (d) $t = 27$

Fig. 2.
Temporal registration of brain MRI slices (see text).



(a) $t = 1$ (b) $t = 17$ (c) $t = 33$ (d) $t = 49$

Fig. 3.

Temporal registration of “dynamic monkey” slices (see text).

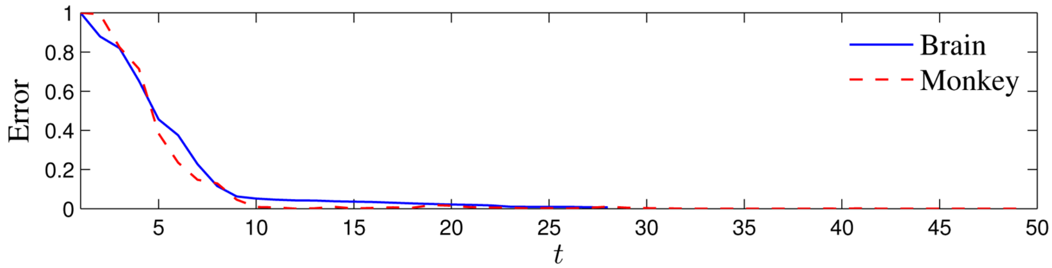


Fig. 4.
Error norm between MAP state and the ground truth state.