In IEEE International Symposium on BioImaging, ISBI'04, Crystal City, VA, April 2004.

SUBPIXEL REGISTRATION IN RENAL PERFUSION MR IMAGE SEQUENCE

Ying Sun[1], *José M. F. Moura*[1], *and Chien Ho*[2]

[1]Department of Electrical and Computer Engineering
[2]Department of Biological Sciences, Carnegie Mellon University, PA 15213
yings, moura@ece.cmu.edu

ABSTRACT

We propose a subpixel registration algorithm to deal with the motion induced by breathing for renal perfusion MR image sequences. Our approach minimizes an energy functional that integrates a motion model and temporal smoothness constraints. We then present a multistage level set method that segments different kidney structures in the registered MR image sequence. Tests with a rat kidney transplantation study show that the algorithm successfully compensates for the breathing motion and extracts the kidneys, cortex, medulla, and pelvis.

1. INTRODUCTION

We develop image and signal processing methodologies for contrast-enhanced renal perfusion MRI to detect early rejection in kidney transplantation. Our approach is to track the time variation patterns of the concentration of dextran-coated ultra-small superparamagnetic iron oxide (USPIO) particles after a bolus injection in the recipients [1]. We use Brown Norwegian (BN) and D Agouti (DA) rats to study kidney transplantation with focus on isograft rats (BN→BN) and allograft rats (DA→BN) [1]. We have developed a methodology that detects rejection by recognizing when the renal perfusion signals on transplanted kidneys are significantly distinct from the perfusion signals on native kidneys [1, 2], and an energy-based image segmentation algorithm that uses the spatial correlation among pixels in the same image as well as the temporal correlation across the image sequence [3]. Our segmentation method has successfully extracted the cortex under the assumption that the kidneys are motion-free during the perfusion process. To monitor reliably the function across the entire kidney, we propose to first compensate for the breathing motion through registration, then segment the kidneys from the surrounding tissues, and ultimately segment within the kidneys their structures including the cortex, the medulla, and the pelvis.

In dynamic MRI achieving higher temporal resolution and spatial resolution are in general mutually conflicting desirable requirements; because they are perfusion studies, the temporal resolution should not be compromised, which results in image sequences with relatively low spatial resolution. Figure 1 displays 4 MR images from a perfusion sequence for an isograft rat. Experimentally, we observe that the motion caused by breathing is usually head-to-feet with subpixel displacement. Therefore, motion compensation is akin to a local registration problem across the image sequence. This problem is made difficult by the rapidly changing signal intensity and image contrast. Automatic segmentation of the kidney from its surrounding soft tissues is a challenging task. It is even more difficult to segment kidney structures. As shown in Figure 1, it is hard to distinguish from each individual image the different anatomical structures due to their lack of contrast. In our proposed image segmentation scheme, we use the entire MRI sequence in the segmentation to overcome this difficulty.

The paper is organized as follows. In section 2, we present our algorithm for subpixel registration. Section 3 describes the segmentation method. Section 4 presents our experimental results, and section 5 concludes the paper.

2. SUBPIXEL REGISTRATION

The major challenge in registration is that the signal intensity changes rapidly with time. We pose the registration problem as an energy minimization problem through the integration of a subpixel motion model and temporal smoothness constraints.

Problem Formulation. Let $g(i,j,t)$ be the intensity at pixel (i,j) in MRI frame t. Given the observed image sequence g, we seek to find a motion-free and noiseless image sequence f such that through a motion transform H, Hf is close to g, and f is temporally smooth. We use the vec operator that stacks the columns of each image in the image sequence, i.e.,

$$\text{vec}\,(g) = [\, g(1,1,1) \,\ldots\, g(N_i,1,1) \,\ldots\, g(1,N_j,1) \,\ldots\, g(N_i,N_j,1) \,\ldots\, g(1,N_j,N_t) \,\ldots\, g(N_i,N_j,N_t)\,]^{\mathrm{T}}.$$

Let $\mathbf{f} = \text{vec}(f)$ and $\mathbf{g} = \text{vec}(g)$. Given $\mathbf{g}$, we solve for $\mathbf{f}$

This work was supported by grants from NIH grant (P41EB001977 and R01EB/AI-00318).

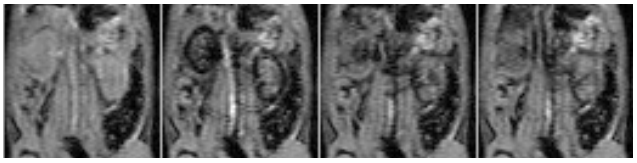


Fig. 1. USPIO-enhanced dynamic MR image sequence

that minimizes

$$E = (\mathbf{g} - \mathbf{H}\mathbf{f})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{g} - \mathbf{H}\mathbf{f}) + \alpha(\mathbf{D}_1\mathbf{f})^{\mathrm{T}}\mathbf{W}(\mathbf{D}_1\mathbf{f}) + \beta(\mathbf{D}_2\mathbf{f})^{\mathrm{T}}\mathbf{W}(\mathbf{D}_2\mathbf{f}), \quad (1)$$

where $\mathbf{H}$ is the motion transform matrix, and $\mathbf{\Sigma}$ is the covariance matrix of the noise. Matrices $\mathbf{D}_1$ and $\mathbf{D}_2$ are the first order and second order time derivative operators. The diagonal weight matrix $\mathbf{W}$ is derived such that only the intensity oscillations due to the motion and the noise are penalized, while the intensity oscillations induced by the bolus injection are preserved. The dimensions of these matrices are $(N_iN_jN_t) \times (N_iN_jN_t)$. In this formulation, the first term is a data fidelity constraint, and the last two terms impose temporal smoothness constraints. Here, α and β are positive scalars that control the tradeoffs between the data fidelity and the temporal smoothness constraints.

Modeling the motion: We adopt a first order motion model under the assumption that the movement is head-to-feet (vertical) within one pixel, and that all the pixels along the same horizontal line i experience the same motion λ_{it}. We illustrate how to model the movement at sub-pixel level using Figure 2. When the motion is from head to feet by half a pixel, the intensity of one pixel in the observed image can be viewed as an equal weighted linear combination of the intensity of the same pixel and the intensity of its upper nearest neighbor in the motion-free image.

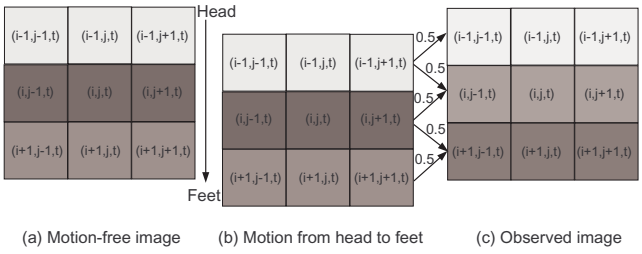


Fig. 2. Motion model

Let $d_t = +1$ represent the direction of the motion from head to feet with respect to the first (reference) image, and $d_t = -1$ represent the opposite direction, then, $\forall j$,

$$g(i,j,t) = \frac{1+d_t}{2}\lambda_{it} f(i-1,j,t) + (1-\lambda_{it})\, f(i,j,t) + \frac{1-d_t}{2}\lambda_{it} f(i+1,j,t) + n(i,j,t), \quad (2)$$

where $n(i,j,t)$ represents the white noise at pixel (i,j) at time t, and λ_{it} stands for the amount of the motion of the ith horizontal line at time t as explained before. Moreover, we observe that the motion is not uniform across the image. The closer to the diaphragm a pixel is, the larger the motion it experiences. To capture this, we model the non-uniformity of the motion by

$$\lambda_{it} = r^{(i-1)}\lambda_t, \quad \text{for } 1 \le i \le N_i \quad (3)$$

where $0 \le \lambda_t \le 1$ is the maximum amount of motion at time t, and $0 \le r \le 1$ is the decay rate of the vertical motion, i.e., $r = 0.95$. The subpixel motion is modeled by the following $N_i \times N_i$ matrix $\mathbf{M}_t$:

$$\begin{bmatrix} 1-\lambda_{1t} & \frac{1-d_t}{2}\lambda_{1t} & & \mathbf{0} \\ \frac{1+d_t}{2}\lambda_{2t} & \ddots & \ddots & \\ & \ddots & 1-\lambda_{(N_i-1)t} & \frac{1-d_t}{2}\lambda_{(N_i-1)t} \\ \mathbf{0} & & \frac{1+d_t}{2}\lambda_{N_it} & 1-\lambda_{N_it} \end{bmatrix}$$

Let $\mathbf{I}_N$ stand for the identity matrix of size $N \times N$, then the motion transform matrix

$$\mathbf{H} = \mathrm{diag}\begin{pmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \dots & \mathbf{H}_{N_t} \end{pmatrix} \quad (4)$$

where $\mathbf{H}_t = \mathbf{I}_{N_j} \otimes \mathbf{M}_t$ and $\mathbf{H}_1 = \mathbf{I}_{N_j} \otimes \mathbf{I}_{N_i}$ ($\otimes$ stands for Kronecker product). The covariance matrix $\mathbf{\Sigma}$ is assumed to be diagonal and is given by

$$\mathbf{\Sigma} = \mathrm{diag}\left(\sigma_s^2(1), \sigma_s^2(2), \dots, \sigma_s^2(N_t)\right) \otimes \mathbf{I}_{N_j} \otimes \mathbf{I}_{N_i}. \quad (5)$$

Since the muscle tissue does not take up iron, we use pixels in the muscle to estimate the variances $\sigma_s^2(t)$.

Temporal smoothness constraints: The first order time derivative operator $\mathbf{D}_1$ and the second order time derivative operator $\mathbf{D}_2$ are

$$\mathbf{D}_1 = \mathbf{D}_t \otimes \mathbf{I}_{N_j} \otimes \mathbf{I}_{N_i} \quad \text{and} \quad \mathbf{D}_2 = \mathbf{D}_{tt} \otimes \mathbf{I}_{N_j} \otimes \mathbf{I}_{N_i},$$

where

$$D_t = \frac{1}{2}\begin{bmatrix} -1 & 1 & & & \mathbf{0} \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ \mathbf{0} & & & -1 & 1 \end{bmatrix},$$

$$D_{tt} = \begin{bmatrix} -2 & 2 & & & \mathbf{0} \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ \mathbf{0} & & & 2 & -2 \end{bmatrix}.$$

The weight matrix $\mathbf{W} = \mathrm{diag}(\mathrm{vec}(w))$, whose diagonal element $w(i,j,t)$ is computed by

$$w(i,j,t) = \frac{1}{\sigma_t^2(t)} \exp(1/2) \exp\left(-\frac{p^2(i,j,t)}{2\sigma_t^2(t)}\right)$$

$$p^2(i,j,t) = \frac{1}{2m+1}\sum_{k=t-m}^{t+m} \left(g(i,j,k) - \bar{g}_m(i,j,t)\right)^2$$

where $\bar{g}_m(i,j,t) = \frac{1}{2m+1}\sum_{k=t-m}^{t+m} g(i,j,k)$ and m is a user-defined parameter. The variance $\sigma_t^2(t)$ is estimated from pixels in the muscle. Larger $p^2(i,j,t)$ results in smaller $w(i,j,t)$ and hence less penalty on the temporal intensity discontinuities. Thus we can selectively smooth the temporal signal.

Energy Minimization. Keeping $\mathbf{H}$ fixed, we find the solution $\mathbf{f}^*$ to $\frac{\partial E}{\partial \mathbf{f}} = 0$ as

$$\mathbf{f}^* = \left[\mathbf{H}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{H} + \alpha\left(\mathbf{D}_1^{\mathrm{T}}\mathbf{W}\mathbf{D}_1\right) + \beta\left(\mathbf{D}_2^{\mathrm{T}}\mathbf{W}\mathbf{D}_2\right)\right]^{-1}\mathbf{H}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{g}$$

This needs inverting a $(N_iN_jN_t)\times(N_iN_jN_t)$ matrix. In our study, $Ni = Nj = 64, N_t = 128$, resulting in a very large matrix of $2^{19}\times 2^{19}$. Although this is an L-block banded matrix ($L = 2$) and highly sparse, its inversion is a full matrix. To alleviate the computational cost, we seek to decouple the spatial and temporal correlations in Equation (1) by introducing an auxiliary vector $\tilde{\mathbf{g}} = \mathrm{vec}(\tilde{g})$, where $\tilde{g}$ is a motion-free version of g. Now minimizing the energy function in Equation (1) is equivalent to minimizing the following two energy functions

$$\begin{aligned} E_1 &= (\tilde{\mathbf{g}}-\mathbf{f})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\tilde{\mathbf{g}}-\mathbf{f}) + \alpha(\mathbf{D}_1\mathbf{f})^{\mathrm{T}}\mathbf{W}(\mathbf{D}_1\mathbf{f}) \\ &\quad +\beta(\mathbf{D}_2\mathbf{f})^{\mathrm{T}}\mathbf{W}(\mathbf{D}_2\mathbf{f}) \\ E_2 &= (\mathbf{g}-\mathbf{H}\mathbf{f})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{g}-\mathbf{H}\mathbf{f}), \quad \tilde{\mathbf{g}} = \mathbf{H}^{\mathrm{T}}\mathbf{g} \end{aligned} \tag{6}$$

Since $\boldsymbol{\Sigma}$ is diagonal, the energy E_1 can be decomposed into $N_i \times N_j$ separate pixel-based optimization problems. On the other hand, the energy E_2 can be decomposed into N_t separate frame-based optimization problems. The above two equations are solved iteratively by first assuming that $\tilde{\mathbf{g}}$ is known, find the best estimation for $\mathbf{f}$ by minimizing E_1, then assuming that $\mathbf{f}$ is known and finding the best estimate for $\mathbf{H}$ by minimizing E_2, hence $\tilde{\mathbf{g}}$. At each iteration, the covariance matrix $\boldsymbol{\Sigma}$ and the weight matrix $\mathbf{W}$ are re-estimated using the updated $\tilde{\mathbf{g}}$. The initial condition is $\tilde{\mathbf{g}} = \mathbf{g}$, i.e., the original image sequence. The iteration continues until $\boldsymbol{\Sigma}$ can not be reduced further.

3. KIDNEY SEGMENTATION

We extend our work on kidney segmentation in [3] to segment the kidney into three regions: cortex, medulla, and pelvis. We propose to distinguish between these kidney structures using intensity-time information through a multistage level set approach. At each stage, we assume that there are two regions in the image sequence whose temporal signals follow closely two different unknown dynamic profiles, similar to the piecewise-constant intensity case in image segmentation [4].

At the first stage, we segment the cortex from the entire kidney. As we are interested in the temporal dynamics, we remove the temporal mean from the perfusion signal and then collect this mean removed time signal at each pixel (x,y) in the vector $\mathbf{f}(x,y)$. Let C_1 be the curve that is the boundary between the cortex and the medulla. We denote by Ω_1^{i} and Ω_1^{o} the inside and the outside of the curve C_1, respectively. We introduce our energy functional as

$$\begin{aligned} E(C_1) &= \mu\cdot\mathrm{Length}(C_1) + \lambda_1\int_{\Omega_1^{\mathrm{i}}}\mathrm{dis}^2(\mathbf{f}(x,y),\bar{\mathbf{f}}_1^{\mathrm{i}})dxdy \\ &\quad +\lambda_2\int_{\Omega_1^{\mathrm{o}}}\mathrm{dis}^2(\mathbf{f}(x,y),\bar{\mathbf{f}}_1^{\mathrm{o}})dxdy, \end{aligned} \tag{7}$$

where $\bar{\mathbf{f}}_1^{\mathrm{i}}$ is the average perfusion signal vector inside the curve C_1, while $\bar{\mathbf{f}}_1^{\mathrm{o}}$ is the average perfusion signal vector outside the curve C_1. The distance metric between two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is defined as

$$\mathrm{dis}\,(\mathbf{v}_1,\mathbf{v}_2) = \left|\sin\frac{\theta}{2}\right| = \sqrt{\frac{1-\cos\theta}{2}} \tag{8}$$

where $\cos\theta$ is given by the correlation coefficient $c(\mathbf{v}_1,\mathbf{v}_2)$:

$$\cos\theta = c(\mathbf{v}_1,\mathbf{v}_2) = \frac{\langle\mathbf{v}_1,\mathbf{v}_2\rangle}{\|\mathbf{v}_1\|_2\|\mathbf{v}_2\|_2}. \tag{9}$$

$\langle\cdot\rangle$ denotes the inner product and $\|\cdot\|_2$ is the Euclidean L_2 norm. The parameters λ_1, λ_2, and μ are positive scalars. The integrals in (7) sum over the pixels, while the distances in the integrands sum over the frames in the sequence.

In our energy functional, the first term penalizes the total length of the boundary, while the last two terms control the data fidelity. We restrict the pixels in Ω_1^{i} to follow the temporal profile of one single vector $\bar{\mathbf{f}}_1^{\mathrm{i}}$ and the pixels in Ω_1^{o} to follow another vector $\bar{\mathbf{f}}_1^{\mathrm{o}}$. Our energy functional exploits the spatially or intra-image, as well as the temporally or across the images, correlations.

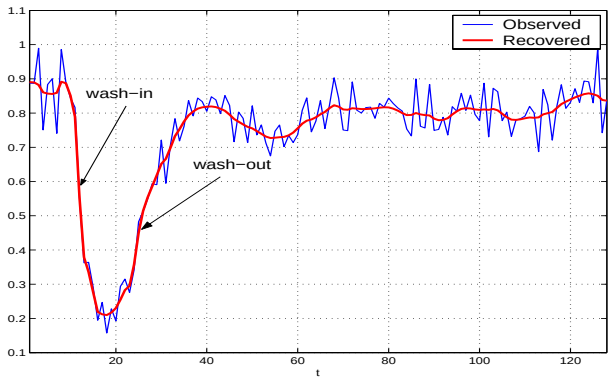
At the second stage, to distinguish between the medulla and the pelvis, we replace Ω_1 by $\Omega_2 = \Omega_1^{\mathrm{i}}$ and curve C_1 by curve C_2, and so on. Segmentation of the kidney structures is now equivalent to minimizing the energy functional (7) stage by stage. At each stage k, we minimize the functional (7) iteratively using a level set based method and update the estimation of $\bar{\mathbf{f}}_k^{\mathrm{i}}$ and of $\bar{\mathbf{f}}_k^{\mathrm{o}}$, $k = 1, 2$.

4. EXPERIMENTAL RESULTS

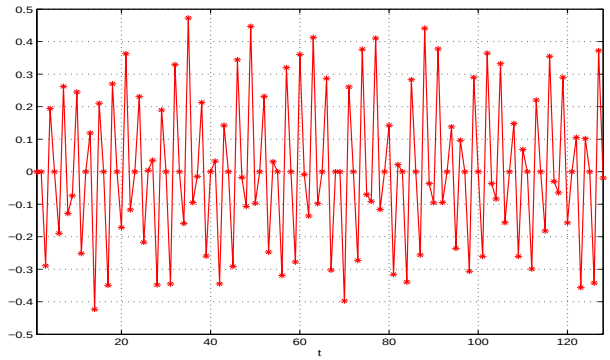
Our data are collected with four groups of rats: 5 BN rats, 5 DA rats, 6 allograft rats, and 4 isograft rats. The rats were bolus injected with dextran-coated USPIO particles at a dose of 6 mg Fe/kg body weight to evaluate first-pass renal perfusion [1]. All rats underwent 128 consecutive snapshot fast low angle shot coronal dynamic studies in 43 s, on a 4.7-T, 40-cm horizontal bore Bruker AVANCE DRX MR instrument using a 7-cm diameter Bruker volume transceiver coil. The image matrix was 64×64 pixels.

Subpixel registration: We obtain excellent results using the proposed registration algorithm with $\alpha = 1, \beta = 2$, and

$m = 4$ in Equation (1). Figure 3 (a) plots the observed (thin) and recovered (thick) temporal signals, respectively, at a cortex pixel. The recovered signal is locally smooth while preserving the wash-in and the wash-out slopes. Figure 3 (b) shows the estimated maximum movement $d_t \lambda_t$ at each time point, where the sign indicates the direction of the movement and the magnitude represents the amount of the movement. If we consider a jump of $d_t \lambda_t$ from negative to positive as one breathing cycle, the estimated breathing cycle is 0.96 second, consistent with an average actual breathing cycle of 1 second for rats under sedation.



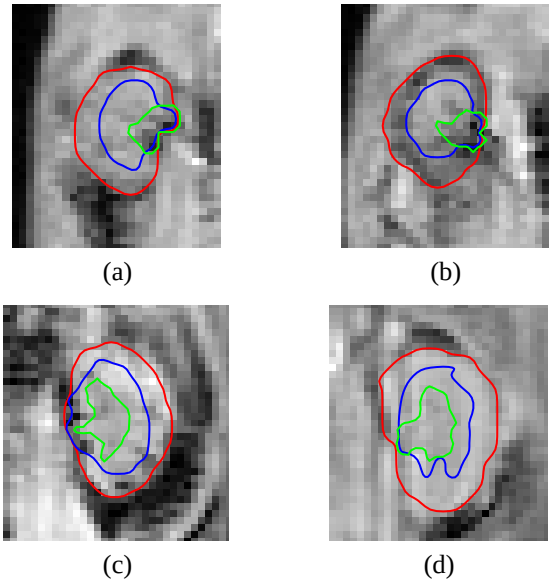
(a) Recovered temporal signal



(b) Estimated movement

Fig. 3. Experimental results of subpixel registration.

Kidney segmentation: In our experiments, we chose the parameters in Equation (7) as follows: $\lambda_1 = \lambda_2 = 2500$, and $\mu = 80$. It is not necessary to perform simulated annealing for λ_1, λ_2, and μ, because the sensitivity of the final segmentation results to the values of these parameters is quite low. Figures 4 demonstrates representative segmentation results for the native and the transplanted kidneys. As shown, the algorithm has successfully located the kidneys and segmented each kidney into three regions: the cortex (outer layer), the medulla (middle layer), and the pelvis (inner layer). The segmentation results are reasonably consistent with the manual segmentations by experts.



(a) (b) (c) (d)

Fig. 4. Segmentation results. (a)-(b): native kidneys; (c)-(d): transplanted kidneys.

5. CONCLUSION

We describe a subpixel registration algorithm that has successfully corrected the breathing motion in renal perfusion MRI. We also present an automatic segmentation algorithm that extracts different kidney structures from the registered MRI sequence. Our experimental results with normal and transplanted rats have shown that the proposed algorithms perform well. These algorithms are very helpful in automatic detection of early rejection in transplanted kidneys.

6. REFERENCES

[1] D. Yang, Q. Ye, M. Williams, Y. Sun, T. Hu, D. Williams, J. M. F. Moura, and C. Ho, "USPIO-enhanced dynamic MRI: Evaluation of normal and transplanted rat kidneys," *Magnetic Resonance in Medicine*, vol. 46, no. 6, pp. 1152–1163, Dec 2001.

[2] Y. Sun, D. Yang, Q. Ye, C. Ho, and J. M. F. Moura, "A novel approach in analyzing dynamic intrarenal signals using USPIO and MRI," in *Proc. 9th ISMRM/ESMRMB Joint Annual Meeting*, Glasgow, Scotland, April 2001.

[3] Y. Sun, J. M. F. Moura, D. Yang, Q. Ye, and C. Ho, "Kidney segmentation in MRI sequences using temporal dynamics," in *2002 IEEE International Symposium on Biomedical Imaging*, Washington, D.C., July 2002.

[4] T. F. Chan and L. A. Vese, "Active contour without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.