Feature Based Nonrigid Brain MR Image Registration With Symmetric Alpha Stable Filters

Shu Liao* and Albert C. S. Chung

***Abstract*—A new feature based nonrigid image registration method for magnetic resonance (MR) brain images is presented in this paper. Each image voxel is represented by a rotation invariant feature vector, which is computed by passing the input image volumes through a new bank of symmetric alpha stable ($S\alpha S$) filters. There are three main contributions presented in this paper. First, this work is motivated by the fact that the frequency spectrums of the brain MR images often exhibit non-Gaussian heavy-tail behavior which cannot be satisfactorily modeled by the conventional Gabor filters. To this end, we propose the use of $S\alpha S$ filters to model such behavior and show that the Gabor filter is a special case of the $S\alpha S$ filter. Second, the maximum response orientation (MRO) selection criterion is designed to extract rotation invariant features for registration tasks. The MRO selection criterion also significantly reduces the number of dimensions of feature vectors and therefore lowers the computation time. Third, in case the segmentations of the input image volumes are available, the Fisher's separation criterion (FSC) is introduced such that the discriminating power of different feature types can be directly compared with each other before performing the registration process. Using FSC, weights can also be assigned automatically to different voxels in the brain MR images. The weight of each voxel determined by FSC reflects how distinctive and salient the voxel is. Using the most distinctive and salient voxels at the initial stage to drive the registration can reduce the risk of being trapped in the local optimum during image registration process. The larger the weight, the more important the voxel. With the extracted feature vectors and the associated weights, the proposed method registers the source and the target images in a hierarchical multiresolution manner. The proposed method has been intensively evaluated on both simulated and real 3-D datasets obtained from BrainWeb and Internet Brain Segmentation Repository (IBSR), respectively, and compared with HAMMER, an extended version of HAMMER based on local histograms (LHF), FFD, Demons, and the Gabor filter based registration method. It is shown that the proposed method achieves the highest registration accuracy among the five widely used image registration methods.**

***Index Terms*—Fisher's separation criterion (FSC), heavy-tail behavior, maximum response orientation selection criterion, nonrigid registration, rotation invariance, symmetric alpha stable ($S\alpha S$) filter.**

Manuscript received April 16, 2009; revised June 26, 2009. First published August 07, 2009; current version published January 04, 2010. This work was supported by the K. S. Lo Foundation, Hong Kong. *Asterisk indicates corresponding author.*

*S. Liao is with the Lo Kwee-Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: liaoshu@ust.hk).

A. C. S. Chung is with the Lo Kwee-Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: achung@cse.ust.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2009.2028078

I. Introduction

As NONRIGID image registration plays an important role in medical image analysis, many novel approaches have been developed for this application. They can be broadly classified into three categories: landmark-based, intensity-based, and feature-based registration methods.

Landmark-based registration methods exploit prior knowledge to establish anatomical correspondences. Anatomical features are extracted from manually placed landmark points [1]–[6]. Transformations are then estimated from those anatomical features. The landmark-based registration methods are usually computationally efficient. However, to give more accurate registration results, it requires additional work to manually place sufficient number of landmark points. Intensity-based registration methods guide the registration with the goal of optimizing the image intensity similarity between the source and the target images [7]–[12]. But intensity similarity does not necessary lead to anatomical similarity. Therefore, in some cases the registration process may be trapped at local minima. Feature-based registration methods use feature vectors as signatures to characterize each voxel in the image volume. Image registration is considered as a feature matching and optimization problem, and registration accuracy therefore depends on the extracted features.

HAMMER is a hierarchical feature-based registration algorithm proposed by Shen and Davatzikos [13]. At the initial stage of registration, it selects a small number of the most salient voxels as active points to drive the registration process. The saliency is measured based on the distinctness of the extracted feature vectors of each voxel. Since voxels with the most distinct feature vectors are usually located at the salient regions of brain images such as sulcal roots and gyral crowns, such voxels are more reliable to drive the registration process in the initial stage. As the registration procedure proceeds, more and more voxels are added as active points to refine the registration results. Finally, all voxels are used as active points. The subvolume deformation model is used in the HAMMER framework [13]. More precisely, the displacement vectors of other voxels are interpolated from the displacement vectors of the current active points by using the Gaussian windows which prescribe the influence of the current active points on the displacement of other voxels and fade away with distance. The energy function in HAMMER is implicitly optimized by finding suitable candidate voxels to be deformed to in the template image through comparing the feature vectors of two voxels in the template and subject images. The geometric moment invariants (GMIs) of white matter, gray matter, and cerebrospinal fluid are adopted as features in

the HAMMER algorithm. Therefore, it is obvious that one limitation of the HAMMER algorithm is that images to be registered must be segmented before registration proceeds. Images for some modalities and some organs are difficult, or impossible to give accurate segmentation. To overcome such shortage, Shen [14] applied the regular geometric moments of local histograms as features based on the HAMMER framework. Unlike the GMIs features adopted in the original HAMMER algorithm, the regular geometric moments of local histograms do not need to be calculated from the segmented images.

In this paper, we propose a new feature extraction method by using the $S\alpha S$ filters which are derived based on the $S\alpha S$ distribution [15] and apply the new method with the HAMMER framework. In our proposed method, each voxel is characterized by the $S\alpha S$ filter responses as its signature. The main contributions of this paper are the following. First, it is observed that the energy spectrums of brain MR images often exhibit non-Gaussian heavy-tail behavior which cannot be satisfactorily modeled by the conventional Gabor filters. As will be shown in this paper, such behavior is essential for image registration because voxels, which are located at the tail of the energy spectrums, are more likely located at the salient regions of the brain MR image volumes. To this end, the $S\alpha S$ filter is designed to model this non-Gaussian heavy-tail behavior. The Gabor filter is shown to be a special case of the $S\alpha S$ filter. This feature extraction procedure does not need the segmentation of the input images. Second, as pointed out in [13], rotation invariance is an important property for the feature-based registration methods. To ensure that the features extracted by the $S\alpha S$ filters are rotation invariant, we propose the maximum response orientation (MRO) selection criterion. The MRO criterion not only extracts rotation invariant features, but also significantly reduces the computational burdens during the registration process. Finally, in case quality segmentations of the input image volumes are available, we propose an evaluation protocol which can *directly compare* the discriminating power of different features prior to perform the actual registration. More specifically, the registration accuracy can be affected by many factors in the registration process such as the extracted features, transformation models, and the similarity measures. In this paper, given the segmentation is available for an input image, we design an evaluation protocol for directly comparing the discriminating power of different features based on the Fisher's separation criterion (FSC). This protocol can avoid the interferences from other factors such as transformation models but solely focus on the features during the feature evaluation process. Therefore it can be a useful tool to reliably analyze the performance of different features used in the registration tasks. Moreover, based on FSC and the given segmentations, we can automatically assign weights to each voxel in the image volumes. The larger the weight, the more important the voxel contributes in the registration tasks.

The paper is organized as follows. Section II introduces the $S\alpha S$ filters and demonstrates the non-Gaussian heavy-tail behavior existing in the energy spectrum of the brain MR images. The MRO selection criterion is proposed in this section for extracting rotation invariant features. This section also briefly describes the energy function and deformation model we use in this paper. Section III gives a protocol for *directly comparing* the discriminating power of different features in the registration task based on the FSC in case the segmentations of input images are available. Moreover, a training framework is provided to automatically assign weights to different voxels in the image based on the FSC. Section IV analyzes the experimental results and Section V concludes the paper.

II. Feature Extraction With the $S\alpha S$ Filters

In this section, we introduce the $S\alpha S$ filters which are derived based on the symmetric alpha stable distribution [15]. We show the existence of the non-Gaussian heavy-tail behavior in the energy spectrums of the brain MR images, and highlight the limitations of the Gabor filters in modeling such behavior (Sections II-A and II-B). With the $S\alpha S$ filters, in Section II-C, we propose the maximum response orientation selection criterion for choosing rotation invariant features, which are subsequently exploited in the hierarchical multi-resolution nonrigid registration process (Section II-D).

A. Conventional Gabor Filter

The Gabor filter, first proposed by Gabor [16], is a useful tool for feature extraction and signal decomposition in image processing, computer vision, and medical image analysis. It has been applied to medical image registration [17]. The 3-D Gabor filter is defined as

$$G(x,y,z) = S \cdot \exp\left(-\frac{(x^2+y^2+z^2)}{\sigma^2}\right) \times \exp\left(j2\pi(xu+yv+zw)\right) \quad (1)$$

where S is a normalization factor, $f = \sqrt{u^2+v^2+w^2}$ is the center frequency in the 3-D Fourier domain, $u = f\sin\phi\cos\theta$, $v = f\sin\phi\sin\theta$, and $w = f\cos\phi$. $0 \le \phi \le \pi$ and $0 \le \theta \le \pi$ define the orientation in the 3-D frequency domain. σ characterizes the shape of the Gaussian envelope. The Gabor filter defined in (1) assumes that the Gaussian window is symmetric and isotropic. In general case, the Gaussian window needs not be isotropic. However, in this paper, we follow the Gabor filter definition in [17] as defined in (1), which assumes that the Gaussian kernel is isotropic.

From (1), it is observed that the Gabor filter can be viewed as a Gaussian kernel modulated by a complex sinusoid in the frequency domain. The center of the Gaussian kernel is located at the center frequency f, with a specific orientation defined by ϕ and θ. Therefore, to satisfactorily describe the input image by using the Gabor filters, the basic assumption is that the image Fourier spectrums of various frequency bands follow the Gaussian distributions. Otherwise, important information may be lost which is contained in the frequency bands lying outside the coverage of the Gaussian kernel bandwidth. However, there are many signals in practical applications do not follow the Gaussian distributions [15]. Using the Gaussian distribution to model non-Gaussian behaviors can lead to significant performance degradation [15]. In the next subsection, it will be demonstrated that the non-Gaussian behavior exists in the energy spectrums of the brain MR images and we will show how the $S\alpha S$ filters can be used to model such behavior.

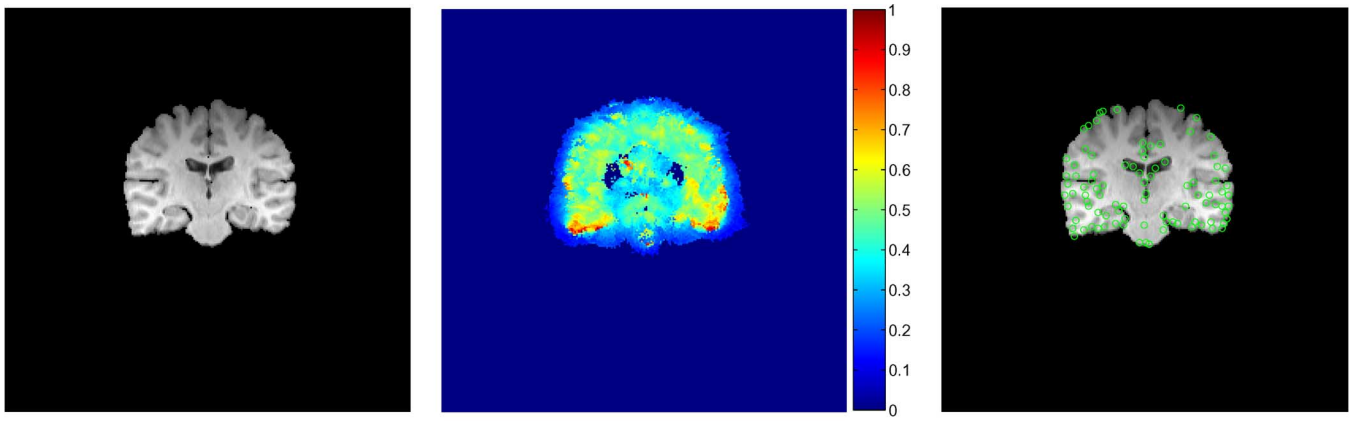

Fig. 1. (a) T1 MR image slice obtained from the IBSR website; (b) color map of the image obtained by transforming the original image to the Fourier domain, preserving only the smallest 10% and largest 10% of the energy magnitudes in various frequency bands (lie on the tail), then transform back to the spatial domain. Here, red color denotes large response values while blue color denotes low response values. (c) Voxels highlighted with green circle are the top 10% (most dominant) voxels with the largest response values in (b).

B. $S\alpha S$ Filters and the Heavy-Tail Behavior

In this subsection, we analyze the Fourier spectrums of the brain MR images, demonstrate the existence of the heavy-tail behavior, and introduce the $S\alpha S$ filters.

Fig. 1(a) is a T1-MR image slice with parameters: TR $=$ 40 ms, TE $=$ 8 ms, flip angle $=$ 50 degrees, field of view $=$ 30 cm obtained from the Internet Brain Segmentation Repository (IBSR) project [18]. Its normalized energy magnitude distributions of various frequency bands are plotted in blue bars in Fig. 2 from column 1 to column 3. The histograms plotted in Fig. 2 are computed by using ideal band-pass filters whose passing bands are determined by the frequency ranges specified in Fig. 2 which shows the frequency spectrum distributions of different frequency bands. It is observed that the energy magnitude distributions of various frequency bands exhibit non-Gaussian heavy-tail behavior. Such behavior is more obvious in the mid- and high-frequency bands. The red curves on the first row of Fig. 2 are the best fitted Gaussian models estimated by using the maximum likelihood estimation for various frequency bands. It is obvious that Gaussian distributions cannot satisfactorily model the heavy-tail behavior.

To further investigate the property of the heavy-tail behavior, we first transformed the input image, as shown in Fig. 1(a), to the Fourier representation. We only maintained the smallest 10% and the largest 10% (i.e., the tails) of the energy magnitudes in various frequency bands, and filtered out other frequency information. Then we transformed the image back to the spatial domain, the corresponding color map is shown in Fig. 1(b). Therefore, the image shown in Fig. 1(b) only contains the frequency information which lies on the tails of the various frequency bands. The voxels with small response values can be eliminated as they are not significant during the process of registration. However, those with large response values play an important role during registration as they contain rich information. We highlight the top 10% voxels which have the largest response values in Fig. 1(b) with green circles in Fig. 1(c). It is found that most of these voxels are located at the salient regions of the input image such as sulcal roots and gyral crowns which are important for registration process as stated in [13].

The $(S\alpha S)$ distribution is a useful tool to model the heavy-tail behavior [15]. Its characteristic function is given as

$$\varphi(t) = \exp\left(jat - \gamma \cdot |t|^{\alpha}\right) \tag{2}$$

where a is the location parameter and becomes the mean value when $1 < \alpha \leq 2$ and the median value when $0 < \alpha < 1$. γ is the scale parameter, also called the dispersion, which is similar to the variance of Gaussian distributions. α is the characteristic exponent, which measures the "thickness" of the tails of the distribution. The smaller the value of α is, the heavier the tails are. The Gaussian distribution is a special case of the $S\alpha S$ distribution (i.e., when $\alpha = 2$). Generally, there is no closed form expression for the probability density function of the $S\alpha S$ distribution. But it can be satisfactorily approximated by the power series expansions [15].

The red curves on the second row of Fig. 2 are the best fitted $S\alpha S$ distributions obtained via maximum likelihood estimation. It is demonstrated that the heavy-tail behaviors are well modeled as compared with the Gaussian distributions. Table I lists the average sums of squared errors modeled by the Gaussian and the $S\alpha S$ distributions for various frequency bands for the 20 normal subject images obtained from the IBSR website.

In order to further illustrate the general existence of the non-Gaussian heavy-tail behavior in the brain MRI images, Table II lists the characteristic exponents α computed via maximum likelihood estimated together with their corresponding 95% confidence intervals, for a set of 20 image volumes of different subjects obtained from the IBSR website of the low, mid and high frequency bands.

It is observed from Table II that the real 20 image volumes obtained from the IBSR website exhibit different degrees of non-Gaussian heavy-tail behavior in various frequency bands, ranging from $\alpha = 1.004$ (approximate to the Cauchy distribution) to $\alpha = 2.000$ (the Gaussian distribution). Therefore, it can be seen that such non-Gaussian heavy-tail behavior generally

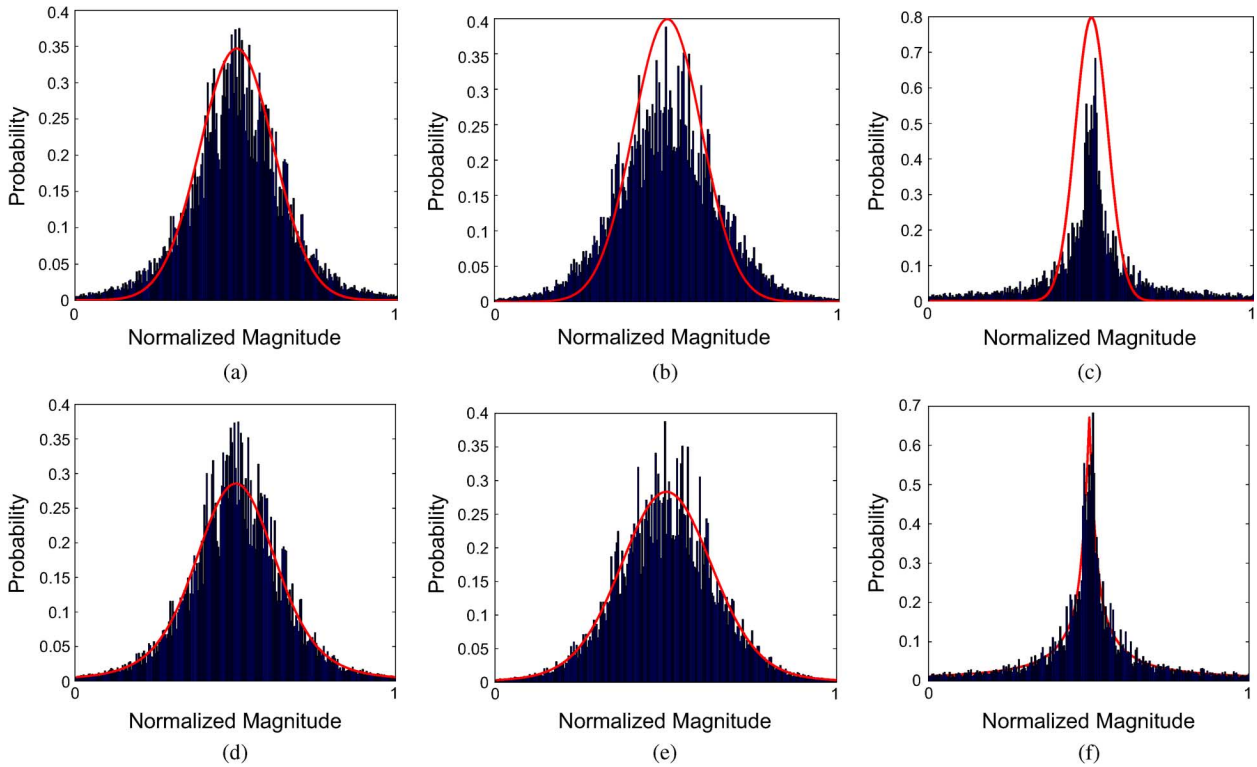

Fig. 2. Blue bars from column 1 to column 3 are the histograms of the normalized frequency magnitude distribution of low frequency band: 2.00 to 3.17, mid frequency band: 3.17 to 5.04, and high frequency band: 5.04 to 9.00 (cycles/image) refer to the input image in Fig. 1(a). (a)–(c): The red curves are the best fitted Gaussian distributions to model the normalized frequency magnitude distribution of different frequency bands obtained via maximum likelihood estimation. (d)–(f): The red curves are the best fitted $S\alpha S$ distributions to model the normalized frequency magnitude distribution of different frequency bands obtained via maximum likelihood estimation. (a) Low frequency; (b) mid frequency; (c) high frequency; (d) low frequency; (e) mid frequency; and (f) high frequency.

TABLE I
AVERAGE SUMS OF SQUARED ERRORS OF MODELING THE ENERGY MAGNITUDE DISTRIBUTION USING THE GAUSSIAN AND THE $S\alpha S$ DISTRIBUTIONS FOR DIFFERENT FREQUENCY BANDS FOR THE 20 NORMAL SUBJECT IMAGES OBTAINED FROM THE IBSR WEBSITE

	Low Frequency	Mid Frequency	High Frequency
Gaussian	0.3628	0.5293	0.4831
$S\alpha S$	0.0673	0.0586	0.0714

TABLE II
ESTIMATED CHARACTERISTIC EXPONENTS VIA THE MAXIMUM LIKELIHOOD ESTIMATION AND THEIR CORRESPONDING 95% CONFIDENCE INTERVALS OF THE 20 IMAGE VOLUMES OBTAINED FROM THE IBSR WEBSITE OF THE LOW-, MID-, AND HIGH-FREQUENCY BANDS

	Low Frequency	Mid Frequency	High Frequency
Subject 1	1.782 ± 0.058	1.413 ± 0.079	1.163 ± 0.026
Subject 2	1.822 ± 0.047	1.215 ± 0.028	1.267 ± 0.014
Subject 3	1.718 ± 0.072	1.782 ± 0.033	1.535 ± 0.042
Subject 4	1.306 ± 0.016	1.216 ± 0.074	1.285 ± 0.007
Subject 5	2.000 ± 0.022	1.226 ± 0.029	1.386 ± 0.085
Subject 6	1.492 ± 0.013	1.606 ± 0.082	1.335 ± 0.049
Subject 7	1.984 ± 0.058	1.775 ± 0.048	1.052 ± 0.052
Subject 8	1.597 ± 0.024	1.466 ± 0.028	1.208 ± 0.011
Subject 9	1.885 ± 0.072	1.617 ± 0.094	1.371 ± 0.037
Subject 10	1.906 ± 0.037	1.841 ± 0.023	1.263 ± 0.073
Subject 11	1.835 ± 0.026	1.602 ± 0.016	1.538 ± 0.044
Subject 12	1.817 ± 0.063	1.257 ± 0.072	1.289 ± 0.003
Subject 13	1.693 ± 0.052	1.393 ± 0.037	1.004 ± 0.017
Subject 14	1.927 ± 0.073	1.905 ± 0.029	1.781 ± 0.055
Subject 15	1.526 ± 0.017	1.511 ± 0.092	1.500 ± 0.038
Subject 16	1.931 ± 0.060	1.836 ± 0.058	1.672 ± 0.021
Subject 17	1.812 ± 0.092	1.883 ± 0.042	1.717 ± 0.005
Subject 18	1.996 ± 0.048	1.937 ± 0.015	1.752 ± 0.081
Subject 19	1.625 ± 0.014	1.331 ± 0.082	1.163 ± 0.063
Subject 20	1.806 ± 0.042	1.526 ± 0.033	1.392 ± 0.029

exists in various frequency bands of brain MR images. Therefore, we are motivated to introduce the $S\alpha S$ filter to extract features from brain MR images for registration task. The 3-D $S\alpha S$ filter is defined as

$$\psi_{\alpha,\gamma,f,\theta,\phi}(x,y,z) = f_{\alpha,\gamma}(x,y,z) \cdot \exp\left(j2\pi(xu+yv+zw)\right) \tag{3}$$

where $f_{\alpha,\gamma}(x,y,z)$ is the zero-mean $S\alpha S$ kernel defined as

$$f_{\alpha,\gamma}(x,y,z) = \left(\frac{1}{\sqrt{2\pi}}\right)^3 \int_u \int_v \int_w A \cdot B \, dw dv du \tag{4}$$

where A and B are defined by (5) and (6), respectively

$$A = \exp\left\{-\gamma(x^2+y^2+z^2)^{\frac{\alpha}{2}}\right\} \tag{5}$$

$$B = \exp\left(j(xu+yv+zw)\right). \tag{6}$$

In (3) and (6), u, v, and w are 3-D frequency components which are the same with the (1). $u = f\sin\phi\cos\theta$, $v = f\sin\phi\sin\theta$, and $w = f\cos\phi$. $0 \le \phi \le \pi$ and $0 \le \theta \le \pi$ define the orientation in the 3-D frequency domain. $f = \sqrt{u^2+v^2+w^2}$ is the center frequency. Gabor filter is a special case of the $S\alpha S$

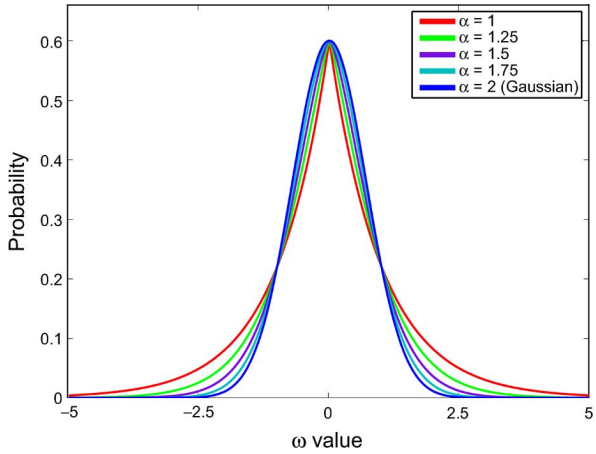

Fig. 3. Illustration of the 1-D zero mean $S\alpha S$ kernels with unit dispersion parameter and different α coefficients in the frequency domain. It is shown that the heavy-tail behavior has the dual property across the spatial and frequency domains.

filters, when $\alpha = 2$. It should be noted that based on the definitions stated in (5) and (6), we assume that the $S\alpha S$ kernel is isotropic, following the similar assumption in [17] for the conventional Gabor filter case. The location parameter of the $S\alpha S$ kernel is zero and the same as in the conventional Gabor filter case [17] because the center of the kernel is determined by the frequency of the modulated sinusoid to extract features from different frequency bands in the Fourier domain.

It is worth pointing out that the Gaussian kernel used in the conventional Gabor filter has the dual property across the spatial and Fourier domains. More specifically, after transforming the Gaussian kernel in the spatial domain to the Fourier domain, it remains its Gaussian property. Similarly, the heavy-tail property of the $S\alpha S$ kernel also has the dual property across the spatial and Fourier domains. For illustration, Fig. 3 shows the zero mean $S\alpha S$ kernels in the frequency domain with different α coefficients, which are obtained by applying the Fourier transform to the 1-D $S\alpha S$ kernel described in

$$F_{\alpha,\gamma}(\omega) = \int_R f_{\alpha,\gamma}(x) \cdot \exp(-j\omega x)dx \tag{7}$$

where $f_{\alpha,\gamma}(x)$ is the 1-D zero mean $S\alpha S$ kernel in the spatial domain, and the dispersion parameter γ is set to 1 for illustration. It is observed in Fig. 3 that the heavy-tail behavior of the $S\alpha S$ kernels has the dual property across the spatial and frequency domains. This dual property is similar to the dual property of the Gaussian kernels across the spatial and frequency domains.

The input image is represented by the $S\alpha S$ filter responses. More specifically, the input image is convolved with a set of the $S\alpha S$ filters with various characteristic exponents, center frequencies, and orientations. The magnitudes of the filtered responses are normalized to the range [0, 1] and adopted as features. In this paper, the center frequencies and orientations are defined as $f_i = f_{\max}/(\sqrt{2})^i$ $(i = 0, 1, \ldots, N_f)$, $\theta_j = j\pi/N_\theta$ $(j = 0, 1, \ldots, N_\theta)$, $\phi_k = k\pi/N_\phi$ $(k = 0, 1, \ldots, N_\phi)$. N_f, N_θ, and N_ϕ are the numbers of center frequencies and orientations to be decomposed in the 3-D frequency domain, $f_{\max}$ is the highest center frequency to be analyzed. In this paper, $f_{\max} = 16.00$ (cycles/image), $N_f = 5$, $N_\theta = 6$, and $N_\phi = 6$. The frequency unit cycles/image is a widely used unit to characterize different frequency bands, which can be found in [19]–[21]. Six different characteristic exponents are used to capture the heavy-tailed behavior: $\alpha_1 = 1.0$, $\alpha_2 = 1.2$, $\alpha_3 = 1.4$, $\alpha_4 = 1.6$, $\alpha_5 = 1.8$, and $\alpha_6 = 2.0$. The scale parameter γ is set for each filter so that each filter has a half peak radial bandwidth of one octave.

Moreover, to demonstrate the fitness of using the $S\alpha S$ kernel to capture the anatomical information of voxels which lie on the tails (i.e., the smallest 10% and largest 10% of the energy magnitudes) in various frequency bands, Table III lists the average proportions of the voxels which lie on the tails of the normalized magnitude distribution in various frequency bands covered by the Gaussian kernel, the proposed $S\alpha S$ kernel and two possible alternative kernels: log-normal kernel and t-distribution kernel in the Fourier domain within the half peak radial passing band for the 20 image volumes of IBSR. Each f_i in Table III is calculated in the same way as the one in the previous paragraph (i.e., $f_i = f_{\max}/(\sqrt{2})^i$ $(i = 0, 1, \ldots, N_f)$, $f_{\max} = 16.00$, (cycles/image), $N_f = 5$). It is observed in Table III that for each frequency band, using the log-normal kernel, the t-distribution kernel, and the $S\alpha S$ kernel all have larger coverage proportions of the voxels which lie on the tails of the normalized magnitude distribution of the corresponding frequency band than the Gaussian kernel within their passing bands. This observation reflects that a significant amount of voxels which lie on the tails of the normalized magnitude distribution of each frequency band indeed locates outside the passing band of the Gaussian kernel. As illustrated in Figs. 1 and 2, failure to model the voxels lying on the tails on each frequency band can lead to significant performance degradation. Therefore, kernels with heavy-tail properties should be adopted.

It is also reflected by Table III that, comparing to other two possible alternative kernels with heavy-tail properties: log-normal kernel and t-distribution kernel, the $S\alpha S$ kernel consistently has larger coverage proportions than both the log-normal kernel and the t-distribution kernel. The reported coverage proportions of the log-normal kernel and the t-distribution kernel are calculated as follows: six values are adopted for the standard deviation parameter σ which controls the shape of the tail of the log-normal kernel, where $\sigma_1 = 3/2$, $\sigma_2 = 1$, $\sigma_3 = 1/2$, $\sigma_4 = 1/4$, $\sigma_5 = 1/8$, and $\sigma_6 = 1/16$, then the log-normal kernel with the highest coverage proportion of the voxels lying on the tail of the normalized magnitude distribution of each frequency band is selected as the reported coverage proportion value. Similar for the t-distribution kernel, where six values are used for the number of degrees of freedom parameter v which controls the shape of the tail of the t-distribution kernel, where $v_1 = 1$, $v_2 = 2$, $v_3 = 3$, $v_4 = 4$, $v_5 = 5$, and $v_6 = 6$ and the highest coverage proportions among those six t-distribution kernels are reported. It is also worth pointing out that both the t-distribution kernel and the $S\alpha S$ kernel have larger coverage proportions of the voxels lying on the tails than the log-normal kernel almost in every frequency band. This illustrates the importance of the symmetric property of the kernels.

TABLE III
AVERAGE PROPORTIONS AND STANDARD DEVIATION (IN PERCENT) OF THE VOXELS WHICH LIE ON THE TAILS OF THE NORMALIZED MAGNITUDE DISTRIBUTION OF DIFFERENT FREQUENCY BANDS COVERED WITHIN THE PASSING BAND OF VARIOUS KERNELS: GAUSSIAN KERNEL, LOG-NORMAL KERNEL, T-DISTRIBUTION KERNEL AND THE PROPOSED $S\alpha S$ KERNEL FOR THE 20 IMAGE VOLUMES IN IBSR. f_i IS CALCULATED BY: $f_i = f_{\max}/(\sqrt{2})^i$ $(i = 0, 1, \ldots, 5)$, WHERE $f_{\max} = 16.00$, (CYCLES/IMAGE)

	$0 \sim f_1$	$f_1 \sim f_2$	$f_2 \sim f_3$	$f_3 \sim f_4$	$f_4 \sim f_5$
Gaussian	76.28 ± 4.02	71.63 ± 3.47	73.17 ± 5.83	62.44 ± 6.37	53.72 ± 6.81
Log-normal	82.61 ± 3.85	80.26 ± 5.56	76.74 ± 6.37	78.80 ± 5.46	71.69 ± 4.18
t-distribution	79.35 ± 4.53	83.90 ± 5.22	87.64 ± 3.96	85.47 ± 3.14	78.06 ± 4.52
$S\alpha S$	84.93 ± 2.73	90.21 ± 4.28	89.35 ± 3.08	88.17 ± 4.79	82.63 ± 3.18

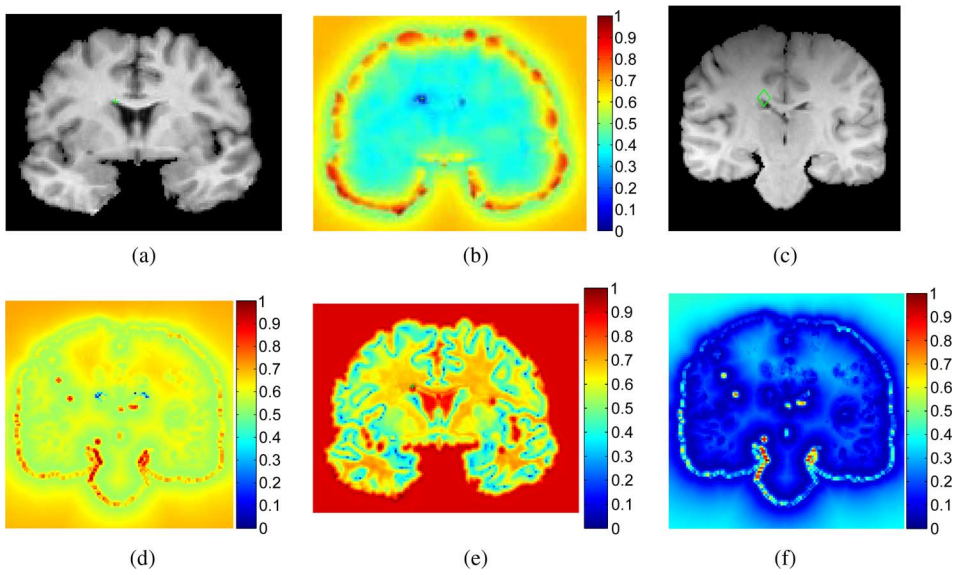

Fig. 4. (a) Image slice obtained from the IBSR website. The green cross point located at the corner of ventricle is the point whose feature vector is compared with the feature vectors of other points. (b) Resulting difference map. Blue color represents high similarity and red color denotes low similarity. (c). Image slice obtained from the IBSR website, which is of different subject as compared with the image slice in (a). The diamond point is the voxel whose feature vector is the most similar to the feature vector of the green cross point in (a). (d) Resulting color-coded difference map obtained by comparing the difference between the feature vector of the green cross point in (a) and the feature vectors of all points in (c). Blue color denotes high similar and red color denotes low similarity. For comparison purpose, the corresponding color-coded difference maps generated by using the Gabor + MRO features are also included in (e) and (f), respectively.

C. Maximum Response Orientation Selection Criteria

The definition of $S\alpha s$ filter and the way to use the filter to extract features from brain MR images have been described in Section II-B. However, the extracted features up to the current stage are not rotation invariant (i.e., when the input image rotates at a specific angle, the elements of the feature vectors representing each voxel change). As pointed out by [13], rotation invariance is a desired property for feature based registration methods. Therefore, we propose the maximum response orientation selection criterion to choose the rotation invariant features. Suppose for a given characteristic exponent α and center frequency f, according to the formulation described in Section II-B, there are $N_\theta \cdot N_\phi$ responses for each voxel with respect to different orientations. Then the maximum response for each voxel is defined as

$$R(x, y, z) = \max_{j,k} \left| I * \psi_{\alpha,\gamma,f,\theta_j,\phi_k}(x, y, z) \right| \tag{8}$$

where I denotes the input image and the $*$ symbol means convolution. Therefore, the $R(x, y, z)$ value for each voxel is rotation invariant because no matter how the images are rotated, the maximum response value still remains the same for a particular direction. Another advantage of the MRO selection criterion is that the feature dimension is significantly reduced. Without the MRO selection criterion, each voxel is represented by a $6 * 5 * 6 * 6 = 1080$ dimension feature vector (six characteristic exponents, five center frequencies, six θ directions, and six ϕ directions). After applying the MRO selection criterion, the feature dimension is reduced to $6 * 5 * 1 * 1 = 30$. Therefore, the computational burden is significantly reduced.

To demonstrate the discriminating ability of the proposed feature, Fig. 4(a) is a brain MR image slice obtained from the IBSR website. Fig. 4(b) shows the color-coded difference map of input image, as shown in Fig. 4(a). The green cross point in Fig. 4(a) is the point whose extracted feature vector is to be compared with feature vectors of other points. We adopt the similarity measure function between two feature vectors $\vec{v_1}$ and $\vec{v_2}$ proposed in HAMMER [13], which is defined as

$$S(\vec{v_1}, \vec{v_2}) = \prod_{i=1}^{m} (1 - |\vec{v_1}(i) - \vec{v_2}(i)|) \tag{9}$$

where $\vec{v_1}(i)$ and $\vec{v_2}(i)$ denote the ith element of feature vectors $\vec{v_1}$ and $\vec{v_2}$, respectively. m denotes the dimensions of the feature vectors $\vec{v_1}$ and $\vec{v_2}$. It is worth reminding that each dimension of feature vectors has been normalized to the range [0, 1] as mentioned in Section II-B.

It is observed from the difference map [Fig. 4(b)] that the voxel highlighted with green cross in Fig. 4(a) is only similar to the voxels within a small neighborhood and voxels with symmetric locations. All these voxels are with similar anatomical property (i.e., around the corners of ventricle). Also, for the feature based image registration methods, it is essential that the corresponding points across different subjects have similar feature vectors. Between different subjects, in order to investigate the differences in feature vectors of the proposed method, we compare the feature vector of the green cross point in Fig. 4(a) with the feature vectors of all the points in Fig. 4(c). Fig. 4(a) and Fig. 4(c) are image slices obtained from the IBSR website of different subjects. The corresponding difference map is shown in Fig. 4(d).

For comparison purpose, the difference maps generated by using the Gabor + MRO features are also included in Fig. 4(e) and (f) which are corresponding to the intraperson difference map and the inter-person difference map respectively. It is observed from Fig. 4(e) that using the Gabor + MRO features results in many false similarity matches. Most of the false similarity matches occur at the region of sulcal roots and gyral crowns, which are the salient regions corresponding to the voxels lying on the tail of different frequency bands as illustrated in Figs. 1 and 2. Therefore, the advantage of using the $S\alpha S$ kernel is strongly implied. Furthermore, for the intersubject difference map obtained by Gabor + MRO shown in Fig. 4(f), it is observed that the Gabor + MRO features cannot effectively distinguish the differences between the referencing voxel and other voxels. It partially reflects the stronger discriminating power of the $S\alpha S$ + MRO features than the Gabor + MRO features.

Based on the resulting difference map [Fig. 4(d)], the feature vector of the green cross point in Fig. 4(a) is only similar to a small group of points which have similar anatomical structures in Fig. 4(c) (i.e., at or around the ventricle corners). The voxel with the most similar feature vector as compared with the green cross point in Fig. 4(a) is highlighted with a diamond in Fig. 4(c). It is shown that the most similar voxel is located at the same relative position (i.e., the upper left corner of ventricle). Therefore, the proposed feature not only can discriminate voxels with different anatomical structures from the same subject, but also can help establish voxel correspondences between different subjects.

D. Energy Function and Deformation Model

In this paper, we adopt the energy function used in [14], which can be expressed as

$$E = \sum_{u\in \mathrm{ACT}(T)} W_T(u)\cdot \Gamma_t + \sum_{v\in \mathrm{ACT}(S)} W_S(v)\cdot \Gamma_s + \beta \sum_{u\in \mathrm{ACT}(T)} \left\|\nabla^2 d(u)\right\| \quad (10)$$

where

$$\Gamma_t = \left(\frac{\sum_{z\in n(u)} \varepsilon(z)\left(1 - S\left(\vec{v_T}(z), \vec{v_S}\left(h(z)\right)\right)\right)}{\sum_{z\in n(u)} \varepsilon(z)}\right)$$

$$\Gamma_s = \left(\frac{\sum_{z\in n(v)} \varepsilon(z)\left(1 - S\left(\vec{v_T}\left(h^{-1}(z)\right), \vec{v_S}(z)\right)\right)}{\sum_{z\in n(v)} \varepsilon(z)}\right)$$

where $\mathrm{ACT}(T)$ and $\mathrm{ACT}(S)$ denote the current sets of active points in the template image and the subject image respectively. $\vec{v_T}(z)$ and $\vec{v_S}(z)$ represent the feature vectors of voxel z in the template and the subject images, respectively. $d(u)$ is the displacement field from which the template image deforms to the subject image. Therefore, we obtain the forward transformation $h(u) = u + d(u)$. $h^{-1}(\cdot)$ denotes the inverse transformation. $W_T(u)$ is the weight for voxel u in the template image. It implies the importance of voxel u during the registration process. Similar to [14], $W_T(u)$ is determined by the edge response value by using the Canny edge detector. $n(u)$ is the neighborhood system of voxel u, which is a sphere centered at voxel u with a radius r. In this paper, r is set by following the same criterion in [14] (i.e., $r = 8$ voxels in the initial stage of registration, and decreases to 1 voxel at the end of the registration process). $\varepsilon(z)$ is a weight which can be determined similar to $W_T(u)$.

The first energy term focuses on the data term similarity in the template image domain, and the second energy term focuses on the data term similarity in the subject image domain. The third energy term focuses on the smoothness constraint on the deformation field, where $\nabla^2 d(u)$ denotes the Laplacian operator of the displacement field $d(u)$. The coefficient β controls the importance of the smoothness constraint, in this paper, $\beta = 0.5$. Adding the first and second energy terms can guarantee the symmetric property with respect to the data term. However, the smoothness term is not inverse consistent. To reduce the computational time and simplify the formulation, we enforce inverse consistence for the data term only to the active points during registration. The detail procedure for enforcing such inverse consistent is described in [13].

We use the subvolume deformation model adopted in [13] and [14]. The multiresolution scheme is used. More specifically, the registration is performed in three resolution levels of the original image. The highest resolution level is the original image, the medium resolution level is obtained by downsampling the original image by a factor of two, and the lowest resolution level is obtained by downsampling the medium resolution level image by a factor of two. The same registration process is performed at each resolution level, from the lowest resolution level to the highest resolution level. The registration result at the end of each resolution level is used as the initial stage to the next resolution level.

III. Feature Discriminating Power Comparison and Optional Training Step With Fisher's Separation Criteria

Section II describes the way to use $S\alpha S$ filters for extracting features from the input images and use the MRO selection criterion to choose rotation invariant features. It is reminded that up to the current stage, the proposed framework is an unsupervised process. This means that there is no training step and the framework does not require the segmentations of the input images. In this section, we introduce the FSC and propose a protocol for *direct comparison* of the discriminating ability between different kinds of features if the segmentations of the input images are available. Furthermore, based on the FSC, if the segmentations are available, weights can also be assigned to each voxel to denote its importance during registration.

TABLE IV
MEAN VALUES OF THE FISHER'S SEPARATION CRITERIA FUNCTION IN EQUATION (13) FOR DIFFERENT KINDS OF TISSUE PAIRS: GRAY MATTER AND WHITE MATTER, GRAY MATTER, AND THE CEREBROSPINAL FLUID, WHITE MATTER AND THE CEREBROSPINAL FLUID OF VARIOUS FEATURES OF THE 20 NORMAL SUBJECT IMAGES OBTAINED FROM THE IBSR WEBSITE

Features	Intensity	GMIs	LHFs	Gabor + MRO	Log-normal + MRO	t-distribution + MRO	$S\alpha S$ + MRO
Gray & White	4.31	13.82	8.16	11.38	13.94	14.15	15.25
Gray & CSF	3.79	15.27	11.72	10.41	10.76	13.62	18.94
White & CSF	3.06	16.96	10.80	10.16	11.27	13.83	17.96

A. Feature Discriminating Power Comparison Based on FSC

It is noted that the proposed method is a feature based registration method and therefore the performance of the feature based registration methods are directly related to the types of features adopted. A problem naturally arises: how different types of features are compared. One way is to observe the final registration accuracies by using different kinds of features. However, comparing features in this way has limitations because the final registration accuracy is also affected by many other factors in the registration process such as the similarity measures and deformation models adopted. Some features may fit for a specific energy function but some are not. Therefore, judging the performance of different kinds of features by merely observing the final registration result may have bias. In this paper, we propose a protocol to *directly compare* the discriminating power of different kinds of features prior to perform the actual registration based on the FSC if segmentation of an input image is available. FSC is also named as the Fisher's linear discriminant analysis in the literature.

Let G_s denote the source image, and it is segmented into c classes of tissues. N_i $(i = 1, 2, \ldots, c)$ represents the number of voxels belonging to the ith tissue. Let $\vec{f_{i,j}}$ denote the feature vector of jth voxel belonging to the ith tissue. Then the mean feature vector $\vec{m_i}$ $(i = 1, 2, \ldots, c)$ of each class of tissue can be calculated by

$$\vec{m_i} = \frac{\sum_{j=1}^{N_i} \vec{f_{i,j}}}{N_i}. \tag{11}$$

The principle of FSC [22] is that it aims to project the original feature space onto a 1-D space. This 1-D space is determined such that it maximizes the separability between two clusters of different classes. It measures the relationship between two clusters and evaluates their interclass distance. Based on FSC [22], such projection can be formulated as a linear transformation obtained by

$$k_z = (\vec{m_1} - \vec{m_2})^T S^{-1} \vec{f_z} \tag{12}$$

where $\vec{m_1}$ and $\vec{m_2}$ are the mean feature vectors from two different classes of tissues obtained from (11), S^{-1} is the inverse of the pooled covariance matrix. $\vec{f_z}$ is the feature vector of voxel z belonging to either tissue classes 1 or 2. k_z is the projected value of $\vec{f_z}$ in the 1-D space. The projection of (12) maximizes the following Fisher's criterion [23]

$$f = \frac{|C_1 - C_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{13}$$

where σ_1 and σ_2 are the variances of the projected feature vectors belonging to tissue classes 1 and 2, respectively. C_1 and C_2 are the projected values of the mean feature vectors $\vec{m_1}$ and $\vec{m_2}$ onto the 1-D space obtained by

$$C_1 = (\vec{m_1} - \vec{m_2})^T S^{-1} \vec{m_1} \tag{14}$$

$$C_2 = (\vec{m_1} - \vec{m_2})^T S^{-1} \vec{m_2}. \tag{15}$$

Equation (13) aims to maximize the distance between the mean feature vector of two classes of tissues (i.e., the numerator) while minimizing the intraclass variance of two classes (i.e., the denominator). The larger the value of f in (13), the better the feature discrimination between two classes of tissues.

For different kinds of features, we can use this protocol to directly compare the discriminating power of different kinds of features among different kinds of tissues. The most simple feature is the voxel intensity alone. Therefore, the value of (13) obtained by using voxel intensity value alone can be served as a baseline for feature comparison. In this case, the feature vector $\vec{f_z}$ of each voxel z reduces to a single value (i.e., the intensity value of z).

We compare seven different kinds of features with each other by using the proposed protocol: i) the intensity value alone, ii) the geometric moment invariant (GMIs) features adopted in HAMMER [13], iii) the local histogram based features used in (LHF) [14], iv) the Gabor filtered response with the MRO selection criteria (Gabor + MRO), v) the log-normal kernel based filtered response with MRO (Log-normal + MRO), vi) the t-distribution kernel base filtered response with MRO (T-distribution + MRO), and vii) the proposed features ($S\alpha S$ + MRO). The average FSC values of (13) for different kinds of tissue pairs for those seven types of features of the 20 normal subject images obtained from the IBSR website are listed in Table IV.

For clarity purpose, the details of the GMIs features adopted in HAMMER [13] and the LHF features used in [14] are given as follows.

The GMI features of each tissue class is defined based on the 3-D regular moments in (17)

$$M_{p,q,r} = \iiint_{(x_1)^2+(x_2)^2+(x_3)^2<R^2} x_1^p x_2^q x_3^r \Upsilon dx_1 dx_2 dx_3 \tag{16}$$

where $\Upsilon = f_{\text{tissue}}(x_1, x_2, x_3)$ where the order of the 3-D regular moment is determined by $(p + q + r)$, (x_1, x_2, x_3) are the coordinates of the voxel under consideration. In (17), it is also assumed that the origin of the coordinate system has already been shifted to (x_1, x_2, x_3) [13]. Therefore, the integration is defined within a sphere with radius R, where R determines the scale of interest. $f_{\text{tissue}}(x_1, x_2, x_3)$ is the membership function with each tissue (i.e., WM, GM, and CSF).

The GMI features are defined based on the 3-D regular moments. For example, the GMI features of zero-order and second-order are defined as

$$\begin{aligned}
I_1 &= M_{0,0,0} \\
I_2 &= M_{2,0,0} + M_{0,2,0} + M_{0,0,2} \\
I_3 &= M_{2,0,0}M_{0,2,0} + M_{2,0,0}M_{0,0,2} + M_{0,2,0}M_{0,0,2} \\
&\quad - M_{1,0,1}^2 - M_{1,1,0}^2 - M_{0,1,1}^2, \\
I_4 &= M_{2,0,0}M_{0,2,0}M_{0,0,2} - M_{0,0,2}M_{1,1,0}^2 \\
&\quad + 2M_{1,1,0}M_{1,0,1}M_{0,1,1} - M_{0,2,0}M_{1,0,1}^2 - M_{2,0,0}M_{0,1,1}^2.
\end{aligned}$$

HAMMER [13] uses the GMI features with zero order, second order, and third order, which results in a 13-D GMI feature for each tissue class and each scale of an individual voxel.

The reader should be reminded that HAMMER [13] requires the segmentation results of input images in order to calculate the GMI features. In some image modalities, accurate segmentation results are often difficult, or impossible to obtain. Therefore, Shen [14] extended HAMMER by using a new type of features, called the local histogram features (LHF).

LHF [14] adopts the local histogram features which do not require the segmentation results of input images. For each voxel x of the input image, LHF [14] first defines a local spherical neighborhood system centered at x, with radius R. Therefore R defines the scale of interest of the feature vectors. Then, the intensity histogram within the spherical neighborhood system is calculated, denoted as $h(x)$. The regular geometric moments calculated based on $h(x)$ is expressed by

$$m(x,p) = \sum_i i^p h(x,i) \tag{17}$$

where p denotes the order of the regular geometric moment and $h(x,i)$ denotes the frequency of intensity i in histogram $h(x)$. In LHF [14], p takes the value from 0 to 2. Therefore, a 3 $\times$ 1-D feature vector can be calculated from the local histogram for each voxel x. We denote this 3 $\times$ 1-D feature vector signature for voxel x as $a(x)$. It is obvious that $a^{\mathrm{Hist}}(x)$ is rotation invariant because the local histogram $h(x)$ is calculated from the spherical neighborhood system centered at x. In LHF [14], in order to extract multiscale features from each voxel x, the regular moment features are extracted in a multiresolution manner. More precisely, the regular moment features are extracted in three different resolutions of the input images, with the first resolution level the original image, the middle resolution level by downsampling the original image with a factor of two, the coarsest resolution level by downsampling the middle resolution level image with a factor of two. We denote the regular moment features extracted from theses three resolution levels as $a_1(x)$, $a_2(x)$ and $a_3(x)$, respectively.

Besides the features extracted from the local histogram of each voxel x, LHF [14] also incorporates the edge information of each voxel x as the second layer feature to characterize the anatomical structures. The Canny edge detector is applied to the input image. The output of the Canny edge detector is a point-wise boundary map, whose values represent the strength of boundary. For each voxel x, let $b(x)$ denote its corresponding Canny edge detector response value. Similar to the regular moment features extracted from the local histogram of x, the edge features are also extracted by referring to three different image resolution levels. We denote them as $b_1(x)$, $b_2(x)$ and $b_3(x)$, respectively. Therefore, the final feature signature for each voxel x is expressed by

$$f(x) = \left[\left[a_1(x)\ \ b_1(x)\right], \left[a_2(x)\ \ b_2(x)\right], \left[a_3(x)\ \ b_3(x)\right]\right]. \tag{18}$$

It is observed that all the other six kinds of features outperform the baseline (i.e., using intensity only as feature) as all the values obtained from (13) by using different features have significant larger values between different kinds of tissues compared with the baseline. Moreover, the GMIs features used in HAMMER [13] can better discriminate between the classes of white matter, gray matter and cerebrospinal fluid as compared with the local histogram features used in LHF [14]. However, we should bear in mind that the GMIs features require the segmentation results of input images. Features extracted by using Log-normal + MRO, t-distribution + MRO, and $S\alpha S$ + MRO are all have better discriminating power than features extracted by using Gabor + MRO, which strongly implies the importance of modeling the heavy-tail behavior. Also, comparing to the other two alternative (i.e., log-normal and t-distribution) kernels which also have heavy-tail properties and the GMIs features, the $S\alpha S$ kernel still has the best discriminating power, which reflects its fitness for capturing the information from the voxels lying on the tails of different frequency bands for the brain MR images.

B. Optional Training Step With Segmentation

In Section III-A, a protocol is presented for evaluating the discriminating power of different kinds of features directly. In this section, based on FSC [22], [25], we also provide a training framework which can automatically assign weights to different voxels. The larger the weight, the more important the voxel is.

We denote the input image as G, which is segmented into c classes of tissues, and define a spherical neighborhood for each reference voxel p with radius R. All voxels on or inside this spherical neighborhood will be taken into consideration. For all the voxels under consideration, let Y_i $(i = 1, 2, \ldots, c)$ denote the number of voxels belonging to the ith tissue class. $\vec{f_{i,j}}$ denotes the feature vector of jth voxel belonging to the ith tissue obtained after passing G through the $S\alpha S$ filters with the MRO selection criterion. Then, we first calculate the intraclass mean and variance by (19) and (20), respectively

$$m_{\mathrm{intra}} = \frac{1}{c}\sum_{i=1}^{c} \frac{2}{Y_i(Y_i-1)} \sum_{k=2}^{Y_i}\sum_{j=1}^{k-1} Eu(\vec{f_{i,j}}, \vec{f_{i,k}}) \tag{19}$$

$$\sigma_{\mathrm{intra}} = \left(C \cdot \sum_{i=1}^{c}\sum_{k=2}^{Y_i}\sum_{j=1}^{k-1} \left(Eu(\vec{f_{i,j}}, \vec{f_{i,k}}) - m_{\mathrm{intra}} \right)^2 \right)^{\frac{1}{2}} \tag{20}$$

where $C = 2/((\sum_{i=1}^{c} Y_i(Y_i-1)) - 2)$.

Then the interclass mean and variance can be calculated by

$$m_{\text{inter}} = \frac{2}{c(c-1)} \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} \frac{1}{Y_i Y_j} \sum_{k=1}^{Y_i} \sum_{l=1}^{Y_j} Eu(\vec{f_{i,k}}, \vec{f_{j,l}}) \quad (21)$$

$$\sigma_{\text{inter}} = \left(D \cdot \sum_{i=1}^{c-1} \sum_{j=i+1}^{c} \sum_{k=1}^{Y_i} \sum_{l=1}^{Y_j} \left(Eu(\vec{f_{i,k}}, \vec{f_{j,l}}) - m_{\text{inter}} \right)^2 \right)^{\frac{1}{2}} \quad (22)$$

where $D = 1/((\sum_{i=1}^{c-1} \sum_{j=i+1}^{c} Y_i Y_j) - 1)$ and $Eu(\cdot)$ denotes the Euclidean distance. Finally, the weighting factor for voxel p is obtained by

$$w_p = \frac{(m_{\text{intra}} - m_{\text{inter}})^2}{(\sigma_{\text{intra}})^2 + (\sigma_{\text{inter}})^2}. \quad (23)$$

It should be noted that if this training step is adopted, the energy function in (11) needs to be slightly modified in order to make use of the weight of each voxel calculated from (23). More precisely, the weighting factors $W_T(u)$ in the first energy term and $W_S(v)$ in the second energy term now become

$$W_T(u) = \lambda \cdot \text{Canny}_T(u) + (1-\lambda) \cdot \text{FSC}_T(u) \quad (24)$$

$$W_S(v) = \lambda \cdot \text{Canny}_S(v) + (1-\lambda) \cdot \text{FSC}_S(v) \quad (25)$$

where $\text{Canny}_T(u)$ and $\text{Canny}_S(v)$ denote the weights obtained from the edge response value by using the Canny edge detector, $\text{FSC}_T(u)$ and $\text{FSC}_S(v)$ denote the weights obtained by using FSC from (23). λ is a coefficient range from 0 to 1 which reflects whether we focus more on the edge weights or the FSC weights, in this paper, $\lambda = 0.5$. Similar modifications are also applied to $\varepsilon(z)$ in the first and second energy terms in (11).

IV. Experimental Results

In this section, we evaluate the registration performance of the proposed method on both simulated and real 3-D datasets obtained from the Simulated Brain Database (BrainWeb)[1] and the Internet Brain Segmentation Repository (IBSR), respectively. The proposed method is also compared with other five widely used nonrigid image registration methods, including HAMMER [13], an extended version of HAMMER based on local histograms as features (LHF) [14], FFD [12], Demons [11], and Gabor based registration method (Gabor) [17]. FFD [12] is an intensity based registration method which first superimposes a control point grid with uniform spacing onto the original image grid. In [12], the normalized mutual information (NMI) is adopted as the similarity metric and it is optimized via gradient descent. FFD first estimates the displacement vector of all the control points by optimizing the NMI metric. Then, the displacement vectors of other voxels are interpolated from the displacement vectors of the control points by using the tri-cubic *B*-spline basis function. The Demons algorithm [11] is also an intensity based registration method. Different from FFD [12], in Demons [11] each voxel can be freely displaced off the original position. The displacement vector of each voxel is estimated by the optical flow equation in Demons [11], which is solely based on the voxel intensity information of the template and subject images. The Gabor base registration method [17] adopts the Gabor filtered responses of the original image as features. In [17], only the filtered responses with the maximum output magnitude for each voxel are adopted as features to represent the input images. In all the experiments, the control point spacing of FFD was set to 2.5 mm, as suggested in [26]. The maximum number of iterations for Demons was set to 100. The variance of the Gaussian kernel used to smooth the displacement field of Demons in each iteration was 1. The elementary demon forces were computed according to the optical flow as defined in [11]. The implementations of FFD and Demons were obtained from the Insight Segmentation and Registration Toolkit (ITK).[2] For the Gabor filter based method, the center frequencies and number of orientations were the same with the $S\alpha S$ filters. For the HAMMER algorithm, the maximum number of iterations was set to 50. Three resolution levels were used to drive the registration [13]. The radii of the spherical subvolumes used to calculate the GMIs features in the low level resolution, the middle level resolution, and the high level resolution were set to 3, 3, and 7, respectively, by following the same line as described in [13]. For the LHF based method, the registration process was performed on three resolution levels [14], the radii of the spherical neighborhood used to calculate the local histogram features in each resolution level were decreased from 8 to 1 voxels [14]. The LHF [14] method is implemented by ourselves.

A. Experiments With Simulated Images

We have evaluated the proposed method on simulated 3-D images obtained from BrainWeb. Twenty image volumes from different subjects were used. One of the image volumes was served as the template image, and the others were used as the subject images. The size of each image volume is $256 \times 256 \times 181$ pixels. From the BrainWeb, all voxels have already been segmented and labeled with one of the three tissue classes which are white matter, gray matter, and cerebrospinal fluid. Fig. 5 shows sample slices with the same cross section of the 20 image volumes.

It is observed in Fig. 5 that brains in these 20 image volumes have significant structural differences, especially in the regions of ventricle and cortex. In order to compare the proposed method with HAMMER [13] and LHF [14], each image volume was preprocessed to remove skull because skull stripping is a required step for HAMMER [13], [14]. We used the software Brain Suite version 2 obtained from USC[3] to accomplish the skull removing process. The proposed method (without training) and other five approaches (FFD, Demons, Gabor, HAMMER, and LHF) were used to independently warp the subject images to the template image. To visualize the registration accuracy of the proposed method (without training), the average brain images were reconstructed from the warped subject images as shown in Fig. 6 of various methods. For comparison, the template image is also included on the first and second rows, first column of Fig. 6.

[1]http://www.bic.mni.mcgill.ca/brainweb/.

[2]http://www.itk.org/.

[3]http://www.brainsuite.usc.edu/.

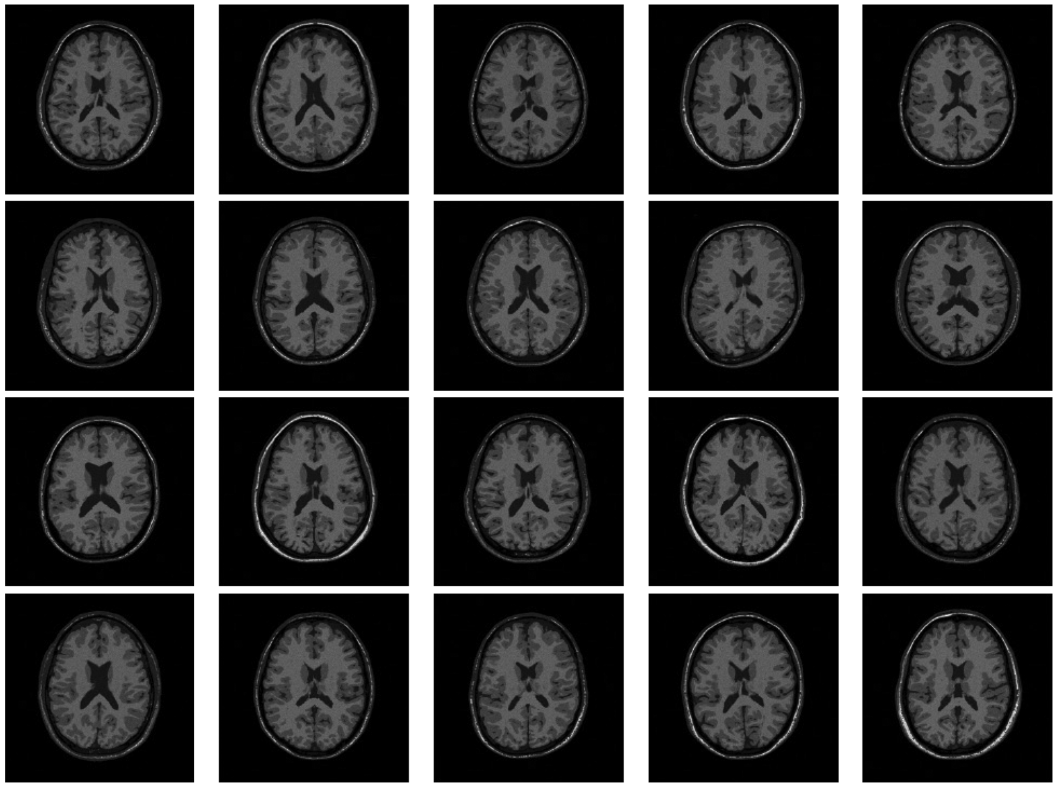

Fig. 5. Sample slices with the same cross section of the 20 image volumes obtained from BrainWeb.

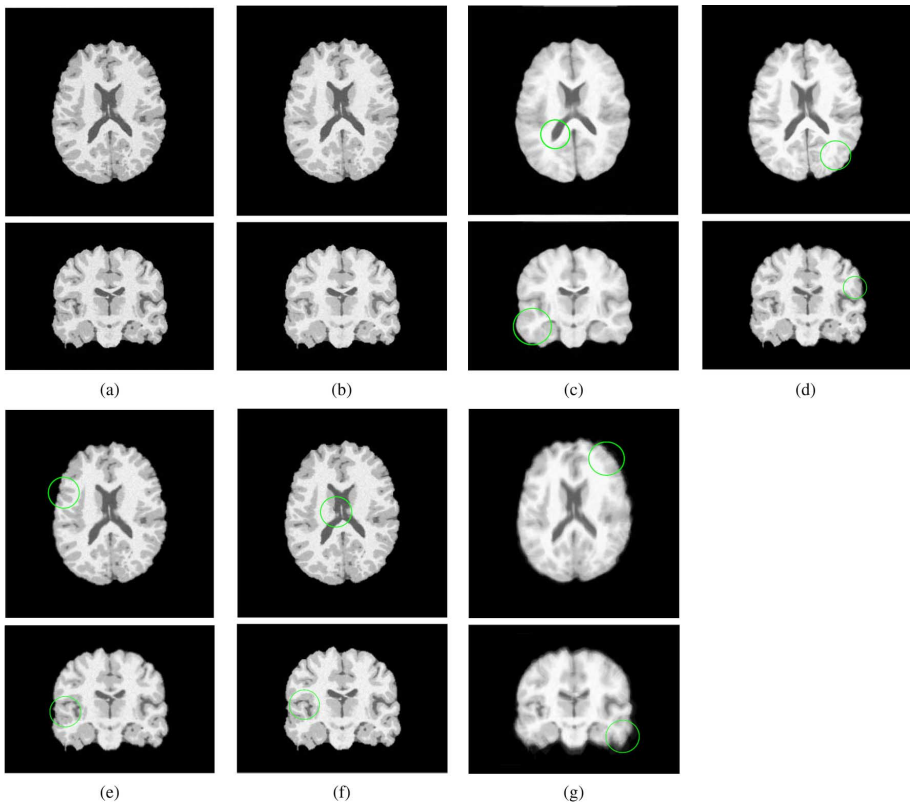

Fig. 6. Reconstructed average brain images of the proposed method, LHF [14], HAMMER [13], Gabor filter based method [17], Demons [11], and FFD [12] for the simulated datasets. The template images are also provided for reference. Regions with most significant differences are highlighted with green circles. Note that the skulls have been removed in the preprocessing step. (a) Template. (b) Our method. (c) LHF. (d) HAMMER. (e) Gabor. (f) Demons. (g) FFD.

It is revealed from Fig. 6 that, as compared to LHF, HAMMER, FFD, Gabor, and Demons, the average brain images reconstructed from the warped subject images using the proposed method are more similar to the template image shown

TABLE V
MEAN VALUES OF P AND SDS OF THE TISSUES OF GM, WM, AND CSF WITH DIFFERENT METHODS ON THE BRAINWEB DATABASE. BR DENOTES BEFORE REGISTRATION, $S\alpha S$ (WOT) DENOTES USING $S\alpha S$ FILTERS WITHOUT TRAINING, $S\alpha S$ (WT) DENOTES USING $S\alpha S$ FILTERS WITH TRAINING. THE HIGHEST VALUES OF P ARE BOLDED FOR EACH TISSUE

Tissue	BR	FFD	Demons	Gabor	Gabor + MRO	HAMMER	LHF	$S\alpha S$ (WOT)	$S\alpha S$ (WT)
Gray	0.461±0.07	0.721±0.06	0.782±0.07	0.718±0.04	0.802±0.05	0.836±0.03	0.790±0.06	0.850±0.05	$\mathbf{0.874 \pm 0.06}$
White	0.444±0.05	0.767±0.03	0.760±0.03	0.744±0.06	0.790±0.04	0.841±0.04	0.804±0.03	0.863±0.02	$\mathbf{0.882 \pm 0.04}$
CSF	0.379±0.04	0.733±0.05	0.758±0.07	0.728±0.05	0.784±0.05	0.827±0.02	0.801±0.03	0.842±0.06	$\mathbf{0.875 \pm 0.06}$

on the first column, first and second rows of Fig. 6, which can be evident from the highly similar ventricular and cortical regions. The sharpness of the average brain images obtained by using the method in [14] (extension of HAMMER, using local histogram based features) is comparable with those obtained by HAMMER [13]. The average brain images obtained by using the proposed method are sharper than those obtained by HAMMER, as more structural details are preserved. The average brains obtained via the Gabor filter based methods [17], Demons [11] and FFD [12] are also shown in Fig. 6. It is shown from the figure that the average brains reconstructed by using the proposed method are sharper and can preserve more structural details than those methods. Therefore, Fig. 6 partially implies the registration accuracy of the proposed method. It echoes the results measured by the proposed feature evaluation protocol described in Section III-A.

To precisely analyze the registration accuracy of the proposed method, we have adopted the registration evaluation criterion proposed in [26]. This evaluation criterion is based on the proportions of tissue overlaps of gray matter (GM), white matter (WM), and the cerebrospinal fluid (CSF). The more the tissues overlapping between the template the warped subject images, the better the registration accuracy. The evaluation measure function is defined as

$$P = \frac{N(A \cap B)}{N(A \cup B)} \tag{26}$$

where A and B denote the regions of a specific type of tissues in template and warped subject image, respectively. Besides the HAMMER algorithm [13] and its extension [14] based on LHF, we have also compared the proposed method with three widely used nonrigid registration methods such as FFD [12], Demons [11] and Gabor filter based method [17]. In order to test the relative importance of the $S\alpha S$ kernel, the Gabor filter based method with the MRO criteria is also included for comparison.

Table V lists the registration accuracies of seven approaches based on the evaluation function defined in (26).

It is observed that the proposed method has the highest registration accuracy among all the other five methods. Also, the registration accuracy of Gabor + MRO is improved as compared with the conventional Gabor filter based approach [17], which illustrates the importance of the rotation invariant property in feature based nonrigid image registration methods and the contributions of MRO. The untrained $S\alpha S$ kernel + MRO outperforms the Gabor kernel + MRO approach. It implies the contribution and superiority of the $S\alpha S$ kernel. Such results and the results provided by the FSC evaluation protocol in Table IV are matched. If the training framework proposed in Section III-B is adopted, the registration accuracy can be further improved. Moreover, the results listed in Table V are consistent with the findings obtained from the feature evaluation protocol, as proposed in Section III-A.

B. Experiments With Real Images

As shown in Section IV-A, the proposed method achieves the highest registration accuracy among the five compared methods on the simulated datasets obtained from BrainWeb. In this section, we evaluate the proposed method on real 3-D images obtained from the IBSR website. Twenty skull stripped real T1-weighted normal subjects were selected from the IBSR website in order to fairly compare the proposed method with HAMMER [13] and LHF [14], as skull stripping is a required preprocessing step for HAMMER [13] and LHF [14]. The 20 selected image volumes have been manually segmented into three classes of tissues: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF) as provided by the IBSR website. Each image volume has resolution of $256 \times 256 \times 60$ approximately. Fig. 7 shows the image slices of the 20 image volumes with the same cross section.

Similar to Section IV-A, one of the image volumes was used as the template, others were served as the subject images. The proposed method then independently warped the subject images to the template image. In order to visualize the registration accuracy of the proposed method, the average brain images were reconstructed from the warped subject images, as shown in Fig. 8. The template image is also provided on the first column of Fig. 8 for reference. It is observed that the proposed method visually outperforms all the other five methods as mirrored by the more salient sharpness of the reconstructed average brain images, which matches the findings obtained from the proposed feature evaluation protocol given in Section III-A. It is also shown that the reconstructed images are very similar to the template image, especially reflected by the ventricular and cortical regions.

The measure function, as stated in (26), was used to statistically measure the registration accuracy. Table VI lists the results obtained via different approaches. The control point spacing for the FFD method was set to 2.5 mm. The parameter settings for Demons [11], Gabor filter based method were the same with the parameter settings used in Section IV-A. From the table, it is shown that the proposed method maintains the highest registration accuracy among the other five methods and the Gabor + MRO approach. The registration accuracy and robustness of the proposed method are strongly implied.

V. CONCLUSION

First, it is demonstrated in this paper that the energy spectrums of brain MR images often exhibit non-Gaussian heavy-tail behavior. This behavior cannot be satisfactorily modeled by the conventional Gabor filters. Such behavior

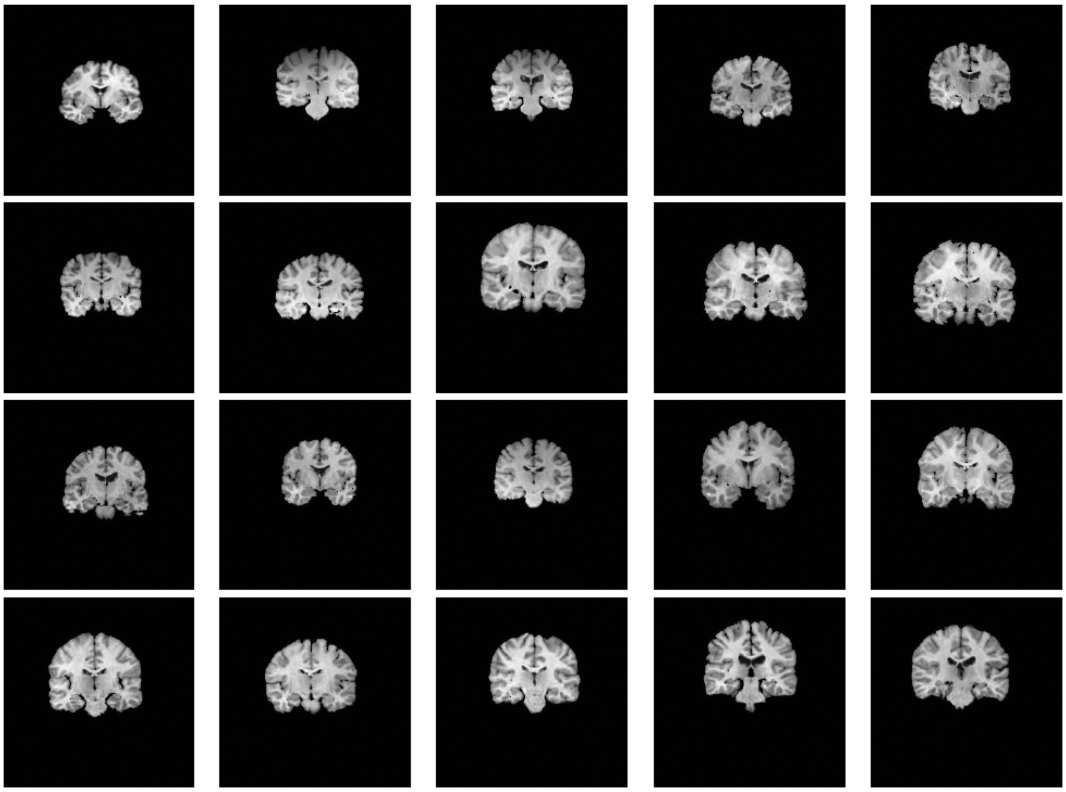

Fig. 7. Sample slices with the same cross section of the 20 image volumes obtained from IBSR.

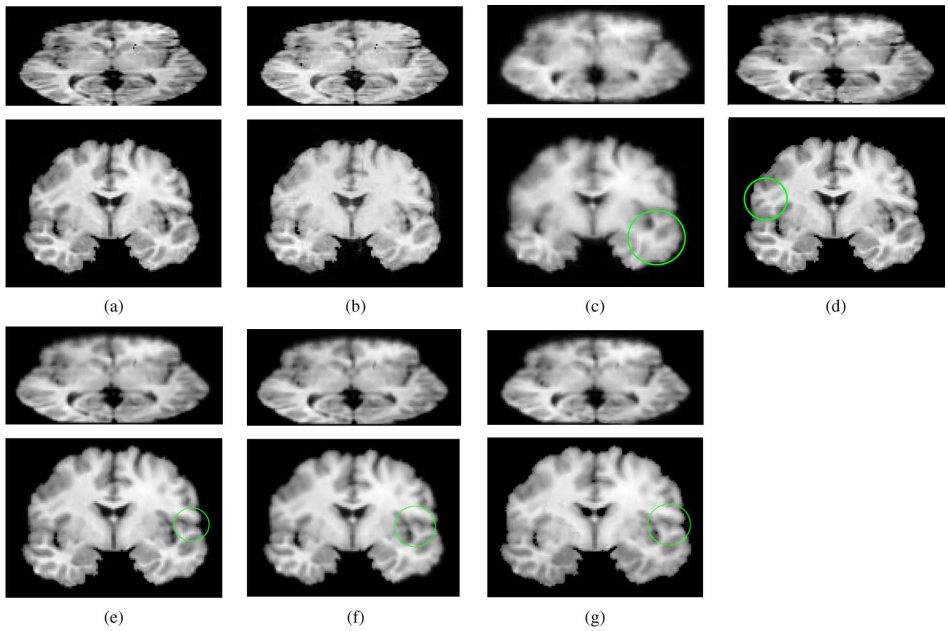

Fig. 8. Reconstructed average brain images of the proposed method, LHF [14], HAMMER [13], Gabor filter based method [17], Demons [11], and FFD [12] for the IBSR real datasets. The template images are also provided for reference. Regions with most significant differences are highlighted with green circles. (a) Template. (b) Our method. (c) LHF. (d) HAMMER. (e) Gabor. (f) Demons. (g) FFD.

TABLE VI
MEAN VALUES OF P AND SDS OF THE GM, WM, AND CSF WITH DIFFERENT METHODS ON THE IBSR DATABASE. BR DENOTES BEFORE REGISTRATION, $S\alpha S$ (WOT) DENOTES USING $S\alpha S$ FILTERS WITHOUT TRAINING, $S\alpha S$ (WT) DENOTES USING $S\alpha S$ FILTERS WITH TRAINING. HIGHEST VALUES OF P ARE BOLDED FOR EACH TISSUE

Tissue	BR	FFD	Demons	Gabor	Gabor + MRO	HAMMER	LHF	$S\alpha S$ (WOT)	$S\alpha S$ (WT)
Gray	0.525±0.05	0.751±0.04	0.774±0.02	0.742±0.08	0.803±0.04	0.817±0.02	0.783±0.05	0.835±0.05	$\mathbf{0.862 \pm 0.05}$
White	0.552±0.03	0.769±0.06	0.781±0.03	0.730±0.06	0.796±0.03	0.821±0.03	0.790±0.04	0.849±0.04	$\mathbf{0.875 \pm 0.03}$
CSF	0.319±0.04	0.750±0.03	0.772±0.06	0.721±0.02	0.761±0.04	0.804±0.05	0.776±0.03	0.838±0.03	$\mathbf{0.867 \pm 0.05}$

should be modeled because most of the voxels which lie on the tails of the energy spectrums are located at the salient regions of the brain MR images such as sulcal roots and gyral crowns. Therefore, the $S\alpha S$ filter is introduced to effectively model the image energy spectrums and subsequently extract features from the brain MR images. Second, to extract rotation invariant features, the MRO selection criterion has been designed in this work. Given a specific center frequency, the MRO selection criterion only takes the maximum response value of a particular orientation as features. The selected features are orientation invariant because no matter how the images are rotated, the maximum response value remains unchanged in a particular orientation. Third, based on the FSC, a new feature evaluation protocol is proposed if the input image segmentations are available. This protocol can help directly compare different types of features without biasing toward other factors in the registration process such as transformation models and energy functions. The proposed protocol provides a mean to evaluate different types of features prior to perform the actual registration process. It is experimentally shown that the proposed feature gives the best discriminating power as compared with the GMIs features used in HAMMER [13] as well as the local histogram based features used in the HAMMER's extension (LHF) [14]. We have also proposed a new training framework to assign suitable weights to different voxels. The larger the weight, the more important the voxel during the registration process. Finally, with the extracted features and the estimated weights, the proposed method conducts registration in a hierarchical and multiresolution manner similar to HAMMER [13]. The proposed method has been evaluated by performing nonrigid registration experiments on both simulated and real 3-D datasets obtained from BrainWeb and IBSR respectively. The experimental results reveal that the proposed method achieves the highest registration accuracies on both datasets among the five methods including HAMMER, FFD, Demons, LHF, and the Gabor filter based registration method.

REFERENCES

[1] K. Rohr, "Image registration based on thin plate splines and local estimates of anisotropic landmark localization uncertainties," in *Proc. Conf. Med. Image Computing Computer-Assisted Intervention (MICCAI'98)*, 1998, pp. 1174–1183.

[2] P. Thompson and A. W. Toga, "A surface-based technique for warping three-dimensional images of the brain," *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 402–417, Aug. 1996.

[3] H. Chui *et al.*, "A unified feature registration method for brain mapping," *Inf. Process. Med. Imag.*, pp. 300–314, 2001.

[4] M. Breijl and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples," *IEEE Trans. Med. Imag.*, vol. 19, no. 10, pp. 973–985, Oct. 2000.

[5] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.

[6] C. Davatzikos and R. N. Bryan, "Using a deformable surface model to obtain a shape representation of the cortex," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 785–795, Dec. 1996.

[7] R. Gan and A. C. S. Chung, "Multi-dimensional mutual information based robust image registration using maximum distance gradient magnitude," *Inf. Process. Med. Imag.*, pp. 210–221, 2005.

[8] Y. Wang and L. H. Staib, "Boundary finding with prior shape and smoothness models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 7, pp. 738–743, Jul. 2000.

[9] P. E. Roland *et al.*, "Human brain atlas: For high-resolution functional and anatomical mapping," *Hum. Brain Mapp.*, vol. 1, pp. 173–184, 1994.

[10] S. Liao and A. C. S. Chung, "Multi-modal image registration using the generalized survival exponential entropy," in *Proc. Conf. Med. Image Comput. Computer-Assisted Intervention (MICCAI'2006)*, 2006, pp. 964–971.

[11] J. P. Thirion, "Image matching as a diffusion process: An analogy with maxwell's demons," *Med. Imag. Anal.*, vol. 3, pp. 243–260, 1998.

[12] D. Rueckert *et al.*, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.

[13] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1421–1439, Nov. 2002.

[14] D. Shen, "Image registration by local histogram matching," *Pattern Recognit.*, vol. 40, pp. 1161–1172, 2007.

[15] C. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*. New York: Wiley, 1995.

[16] D. Gabor, "Theory of communication," *IEEE J. Inst. Elect. Eng.*, vol. 93, pp. 429–457, 1946.

[17] J. Liu, B. Vemuri, and J. Marroquin, "Local frequency representations for robust multimodal image registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 462–469, May 2002.

[18] Internet Brain Segmentation Repository (IBSR) [Online]. Available: http://www.cma.mgh.harvard.edu/ibsr/

[19] R. Porter and N. Canagarajah, "Robust rotation-invariant texture classification: Wavelet, Gabor filter and GMRF based schemes," *Proc. Inst. Elect. Eng. Vis. Image Signal Process.*, pp. 180–188, 1997.

[20] A. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 55–73, Jan. 1990.

[21] A. Teuner, O. Pichler, and B. J. Hosticka, "Unsupervised texture segmentation of images using tuned matched Gabor filters," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 863–870, Jun. 1995.

[22] A. Fisher, *The Mathematical Theory of Probabilities*. New York: MacMillan, 1923.

[23] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1990.

[24] R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: Wiley, 1992.

[25] R. Parry and I. Essa, "Feature weighting for segmentation," in *Proc. ICMIR*, 2004, pp. 116–119.

[26] W. Crum and D. Rueckert *et al.*, "A framework for detailed objective comparison of non-rigid registration algorithms in neuroimaging," in *Proc. Conf. Med. Image Computing Computer-Assist. Intervention (MICCAI'2004)*, 2004, pp. 679–686.