Midsagittal plane extraction from brain images based on 3D SIFT

Huisi Wu\textsuperscript{1,6}, Defeng Wang\textsuperscript{2,3,4,5}, Lin Shi\textsuperscript{2,3}, Zhenkun Wen\textsuperscript{1,6} and Zhong Ming\textsuperscript{1}

\textsuperscript{1} College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong, People’s Republic of China
\textsuperscript{2} Department of Imaging and Interventional Radiology, The Chinese University of Hong Kong, Shatin, NT, Hong Kong
\textsuperscript{3} Research Center for Medical Image Computing, The Chinese University of Hong Kong, Shatin, NT, Hong Kong
\textsuperscript{4} The Chinese University of Hong Kong Shenzhen Research Institute, Shenzhen, People’s Republic of China
\textsuperscript{5} Department of Biomedical Engineering and Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong, Shatin, NT, Hong Kong

E-mail: hswu@szu.edu.cn and wenzk@szu.edu.cn

Received 2 June 2013, revised 6 December 2013
Accepted for publication 24 December 2013
Published 28 February 2014

Abstract

Midsagittal plane (MSP) extraction from 3D brain images is considered as a promising technique for human brain symmetry analysis. In this paper, we present a fast and robust MSP extraction method based on 3D scale-invariant feature transform (SIFT). Unlike the existing brain MSP extraction methods, which mainly rely on the gray similarity, 3D edge registration or parameterized surface matching to determine the fissure plane, our proposed method is based on distinctive 3D SIFT features, in which the fissure plane is determined by parallel 3D SIFT matching and iterative least-median of squares plane regression. By considering the relative scales, orientations and flipped descriptors between two 3D SIFT features, we propose a novel metric to measure the symmetry magnitude for 3D SIFT features. By clustering and indexing the extracted SIFT features using a $k$-dimensional tree (KD-tree) implemented on graphics processing units, we can match multiple pairs of 3D SIFT features in parallel and solve the optimal MSP on-the-fly. The proposed method is evaluated by synthetic and in vivo datasets, of normal and pathological cases, and validated by comparisons with the state-of-the-art methods. Experimental results demonstrated that our method has achieved a
real-time performance with better accuracy yielding an average yaw angle error below 0.91° and an average roll angle error no more than 0.89°.

Keywords: symmetry analysis, MSP extraction, 3D SIFT, GPU KD-tree

(Some figures may appear in colour only in the online journal)

1. Introduction

The left and right hemispheres of human brains are closely symmetrical because most tissues in the left hemisphere have a mirrored counterpart on the right with apparently similar scale and appearance. Research studies demonstrated that the symmetrical structure of human brain degenerates in traumatic brain injury (Gilles et al. 2003, Kumar et al. 2005), brain infections (Corti et al. 2008, Hermes et al. 2008), brain tumors (Joshi et al. 2003, Yu et al. 2012), metabolic disorders (Herbert et al. 2005, Cullen et al. 2006, Takao et al. 2010, Oertel-Knöchel and Linden 2011), and perinatal brain lesions (Tommasi et al. 2009, Tilman et al. 2010, Roussigne et al. 2012, Saenger et al. 2012), which caused abnormal tissue intensity or texture alterations to break the symmetrical structure of human brains. To enhance the use of computer-aided technology in disease diagnosis of human brain, researchers have to understand the relationship between asymmetrical brain structure and specific brain pathologies. By measuring structural and radiometric asymmetries based on an efficient multi-scale registration algorithm, Lorenzen et al. (2001) proposed a framework to identify brain tumors. They registered the brain tissues with their counterpart reflections across the midsagittal plane (MSP) using adaptive image deformation and intensity warping. By analyzing the texture asymmetry between two brain hemispheres based on heuristic minimization search, Bergo et al. (2008a) successfully segmented the focal cortical dysplasia, the most common malformation region in patients with intractable epilepsy. By quantifying brain asymmetry across the MSP in multi-scale spaces based on a non-supervised manifold learning algorithm, Liu et al. (2007) and Teverovskiy et al. (2008) discovered the biomarkers for Alzheimer disease (AD) from the deformation and tensor fields in MR brain images, which played an important role in early diagnosis of AD. The key issue to perform the above brain symmetry analysis is the MSP extraction from 3D brain images. A fast and robust method to automatically extract the MSP from 3D brain images could improve the accuracy and efficiency.

In MSP extraction methods (Prima et al. 2002, Teverovskiy and Liu 2006, Volkau et al. 2006, Song et al. 2007, Prima and Daesslé 2007, Zhang and Hu 2008, Liu et al. 2011, Roy and Bandyopadhyay 2012), the head volume is usually considered as a whole to maximize bilateral symmetry. They are more likely to capture the global symmetry of the human brain compared to symmetry axis detection for each slice. Currently, existing MSP extraction methods can be divided into two categories, e.g., shape-based and content-based methods (Liu 2009). Content-based methods determine the symmetrical plane by using registration between two hemispheres based on some kind of internal signal feature, such as voxel intensity, local edge and gray distribution (Ardekani et al. 1997, Stegmann et al. 2005, Grigaitis and Meilunas 2007), while shape-based methods identify MSP mainly relying on the extracted geometric landmarks and topological features of the human head, such as the moments of inertia and inter-hemispheric fissure (Tuzikov et al. 2003, Bergo et al. 2008b, Liu et al. 2011, Roy and Bandyopadhyay 2012). Since MSP provides an intuitive 3D symmetrical division, quantitative asymmetrical evaluation can be performed for the global human brain. Junck et al. (1990) first relied on
cross-correlation analysis to extract fissure planes from PET and SPECT volumes, which the optimal MSP was determined by finding the translation and rotation that yielded the highest correlation between left and right hemispheres. By establishing a scan-independent coordinate system for the human brain, Brummer (1991) formulated the MSP extraction as an image registration problem and applied the Hough transformation to identify MSP. Liu et al. (2001) further employed an edge-based and cross-correlation approach to decompose the MSP fitting problem into discovery of two-dimensional (2D) symmetry axes on each slice, and estimated the MSP parameters using linear regression. By exhaustively searching fissure line segments using histogram-based outlier removal, Hu and Nowinski (2003) also successfully extracted MSP by plane fitting based on the dominant orientation and selected fissure segments. Song et al. (2007) determined the MSP based on a group of assistant parallel lines and correlation of moment of gravitational forces. By representing the image with a polar form, Costantini and Casali (2007) successfully exploited the nonlinear dynamic behavior of cellular neural networks to identify MSP. Liu et al. (2008) also successfully solved the MSP by minimizing the statistical dissimilarity between paired regions in opposing hemispheres and formulating the MSP extraction as an optimization problem. More recently, Ruppert et al. (2011) relied on image features detected by the 3D Sobel edge operator and multi-scale correlation to extract the optimal MSP, which is sensitive to image noises and deformations. Although encouraging improvements have been made in both methodology and performance, existing methods are still not reliable enough or too time consuming to be used in conventional scanning. Most of them relied on gray level, skull shape or edges to infer the symmetry plane, which may easily suffer from image noise, local deformation or the deficiencies of edge detectors. In addition, due to their optimization schemes and iterative natures, none of them can achieve real-time performance.

In this paper, we present a fast and robust MSP extraction method based on parallel 3D scale-invariant feature transform (SIFT) matching and voting. Our framework is shown in figure 1. To collect a set of distinguished features from the brain volume, we have extended the conventional SIFT algorithm (Lowe et al. 2004) from 2D to 3D and implemented it using graphics processing units (GPU). Since 3D SIFT descriptor is invariant to scaling, orientation, affine distortion and intensity changes, our method has higher tolerance to image noise, bias inhomogeneity and local deformation. Also, we have built a GPU KD-tree to index the clustered 3D SIFT features to achieve parallel matching and voting multiple pairs of 3D SIFT features. To effectively measure symmetrical magnitude between two 3D SIFT features, we have designed a novel symmetric similarity metric by combining their relative scales,
orientations and flipped descriptors. Finally, an iterative least-median of squares (LMS) plane regression is performed to solve the optimal MSP. Our method has been evaluated on both synthetic mirrored images and in vivo datasets from normal and pathological cases. Results from experiment comparisons have demonstrated the effectiveness of our method.

2. Methodology

2.1. 3D sift detection

Although 3D extension of the SIFT algorithm has been recently used for generating correspondences between two volume images (Scovanner et al 2007, Allaire et al 2008, Flitton et al 2010), there are rare reports of searching inner correspondences by matching pairs of 3D SIFT features within a volume image. Since the human brain exhibits an approximately symmetrical structure that contains a set of 3D mirrored features, extraction of MSP from a brain volume by matching pairs of inner 3D SIFT features is feasible. Considering that 3D extension of the SIFT algorithm is straightforward and has been implemented in Scovanner et al (2007), Allaire et al (2008) and Flitton et al (2010), we can implement the 3D SIFT based on the above approaches by selecting a best treatment and a best parameter setting in each step.

To detect 3D SIFT keypoints from a brain volume, we first calculate the difference of Gaussian (DoG) volumes. Suppose the given volume image is \( I(x, y, z) \), and 3D Gaussian filter is \( G(x, y, z, k\sigma) \), then the DoG volumes can be written as,

\[
\text{DoG}(x, y, z, k') = G(x, y, z, k'\sigma) \ast I(x, y, z) - G(x, y, z, k\sigma) \ast I(x, y, z)
\]

where \( \ast \) is the convolution operation in \( x, y \) and \( z \); \( k'\sigma \) is the scale of the volume. Similar to Scovanner et al (2007), we set \( k = \sqrt{2}, \sigma = 1.6 \) and \( i \in \{1, 2, \ldots, 5\} \) in our experiments.

Once the DoG volumes are obtained, a set of candidate keypoints can be identified as the local extrema (minima or maxima) in the DoG volumes across the scale spaces. Specifically, we have compared each voxel in the DoG volumes with its 26 neighbors at the same scale and 27 corresponding neighboring voxels in each of the two neighboring scales. Then, we mark it as a candidate keypoint if this voxel is the maximum or minimum in terms of intensity among all compared voxels. To eliminate the false candidates that produce unstable descriptors, two kinds of erroneous candidates (e.g. poorly localized along an edge or with poor contrast) have been removed. To discard the keypoints with poor contrast, we have applied the treatment from Allaire et al (2008), in which the candidate keypoint is discarded if its density is below a threshold \( (\tau_c = 0.05) \). On the other hand, we have also used a \( 3 \times 3 \) Hessian matrix \( H = (D_{ij})_{3\times3} \) to describe the local curvature and measure the edge responses for the candidate keypoints, where \( D_{ij} \) is the second derivative in the DoG volume. Similar to Flitton et al (2010), we can use the following equation to identify the false candidates that are poorly localized along an edge,

\[
\frac{\text{Trace}^3(H)}{\text{Det}(H)} < \frac{(2\tau_e + 1)^3}{(\tau_e)^2}
\]

where \( \tau_e \) is a threshold. According to Flitton et al (2010), we can also set \( \tau_e = 30 \) to remove the false candidates. After removing the two kinds of false candidates, localizations of the true keypoints are determined in the DoG volumes. Figure 2 shows typical 3D SIFT keypoints detected in a volume brain image.

Once the location and the scale of a keypoint are determined, we can calculate its dominant orientation based on local image gradients to achieve orientation invariance. Unlike the 2D SIFT keypoint that only requires one angle to determine the dominant orientation, a 3D case
requires three angles. As shown in figure 3, they are azimuth angle $\phi_a \in [-\pi, \pi]$, elevation angle $\phi_e \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, and tilt angle $\phi_t \in [-\pi, \pi]$. Therefore, we can denote each keypoint as a seven-dimensional vector $k_i = (x_i, y_i, z_i, s_i, \phi_a, \phi_e, \phi_t)$, where $(x_i, y_i, z_i)$ and $(\phi_a, \phi_e, \phi_t)$ describe its location, scale and orientation, respectively. To encode the local appearance of the keypoint $k_i$, we have also generated a 3D SIFT descriptor $p_i$ and normalize its scale and orientation. As shown in figure 5(a), we choose $N_s \times N_s \times N_s$ sub-volumes surrounding each keypoint to build the 3D SIFT descriptor $p_i$. Each sub-volume includes $N_s \times N_s \times N_s$ voxels, which are the basic elements to be described using gradient histograms. For each voxel, we split it into 45° bins along its azimuth and elevation directions to calculate a Gaussian weighted gradient histogram, so we obtain $8 \times 4$ bins in each gradient histogram. Finally, each 3D SIFT
descriptor $p_i$ includes $N_s^3 \times N_v^3 \times 8 \times 4$ components. We set $N_s = N_v = 2$, and our 3D SIFT descriptor can be written as $p_i = (v_1, v_2, \ldots, v_{2048})$. Our experiment demonstrated that increasing the value of $N_s$ or $N_v$ does not improve the accuracy much but slows down the performance. Note that, in our experiments, the orientation of $p_i$ is normalized by always setting the dominant orientation $(\phi_a, \phi_e)$ of the keypoint $k_i$ as the starting point for each gradient histogram. In addition, the scale of each component in $p_i$ is also normalized to unity. So far, since no additional location, scale or orientation information is recorded in the 3D SIFT descriptor $p_i$, it satisfies the properties of spatial invariance, scale invariance and orientation invariance.

2.2. Clustering and indexing

Given a set of 3D SIFT keypoints within a brain volume, we can perform the MSP extraction by matching pairs of symmetric 3D SIFT keypoints. Multiple pairs of 3D SIFT features of symmetric voxels within the approximately symmetrical structure should have similar scale $s_i$ and tilt angle $\phi_t$. Clustering on the 3D SIFT keypoints can accelerate the voxel matching. Thus, we have employed the adaptive mean-shift clustering algorithm (Comaniciu et al. 2002) to group the extracted 3D SIFT keypoints. The major advantage of adaptive mean-shift clustering is to automatically determine the number of clusters; unlike other techniques such as $k$-means, which require to measure the number of clusters in advance. In our implementation, the scales $s_i$ and the tilt angle $\phi_t$ of the 3D SIFT keypoints are fed for adaptive mean-shift clustering in order to group features with similar shapes and tilt orientations. Specifically, a 2D feature space $(s_i, \phi_t)$ is regarded as the probability density function (pdf) of the represented 3D SIFT keypoints. Dense regions in the feature space correspond to local maxima of the pdf, which are also named the modes of the density. Therefore, based on the local structure of the feature space $(s_i, \phi_t)$, we can obtain the number and the shapes of all clusters by identifying the locations of the modes. Since mean-shift clustering is a nonparametric clustering technique, which does not require prior measure of the number of clusters, and does not constrain the shape of the clusters, we start the mode finding procedure on the extracted 3D SIFT keypoints. Then we run mean-shift procedure to find the stationary points of the density function, and retain those of local maxima. By repeating the above procedures, we have finally obtained the clustering results when the set of all locations converge to the same modes.

To facilitate the following $k$-nearest neighbor (KNN) searching in symmetric voxel matching, we further index the scales $s_i$ and tilt angles $\phi_t$ of the keypoints in each cluster using a $k$-dimensional tree (KD-tree), which is a well-known acceleration structure to partition and organize points in $k$-dimensional space. For fast retrieval, we have implemented the KD-tree on GPU to support the parallel KNN searching. Specifically, we have applied compute unified device architecture to build all tree nodes in breadth-first search to fully exploit the fine-grained parallelism of modern GPUs at all stages of KD-tree construction (Zhou et al. 2008). In terms of speed, such an implementation is much faster than well-optimized single-core CPU algorithms and competitive with multi-core CPU algorithms.

2.3. Symmetric similarity metric

Symmetric pairs of 3D SIFT keypoints in each cluster form by matching a voxel with its nearest neighbors. However, if we directly form pair of voxels by searching nearest neighbors in the GPU KD-tree, the resulting pairs of voxels may potentially contain different things, as the mean-shift clustering and following indexing are only based on a crude measurement of the scale $s_i$ and tilt angle $\phi_t$ without considering the volume content. Hence, a more sophisticated metric is required to validate and select neighbors during the selection of a symmetric pair of
Figure 4. Orientation symmetric similarity of two 3D SIFT features.

voxels. In the following, we will define a novel metric to quantify the symmetric similarity of two neighboring voxels in terms of scale, orientation and appearance.

2.3.1. Scale similarity. Even when the symmetric volumes are slightly deformed, two symmetric voxels remain close in scale or size. Hence, the scale similarity $S_{ij}$ of two 3D SIFT features $k_i$ and $k_j$ is a function of their scale

$$S_{ij} = \exp \left( \frac{-\|s_i - s_j\|^2}{\sigma_s} \right)$$

(3)

where $\sigma_s$ controls the weighting of the scale variation. We set $\sigma_s = 2.0$ in all our experiments.

2.3.2. Orientation symmetric similarity. If two keypoints $k_i$ and $k_j$ exhibit reflected symmetry, their orientations should also satisfy a mirror relationship with each other. This suggests that we have to consider their orientation symmetric similarity in designing the global symmetric similarity metric. Hence, we have defined an orientation symmetric similarity $\Phi_{ij}$ to measure the mirror relationship between $k_i$ and $k_j$ in terms of their directions. As shown in figure 4, suppose the median plane between $(x_i, y_i, z_i)$ and $(x_j, y_j, z_j)$ is $P$, we denote the orientation symmetric similarity $\Phi_{ij}$ as

$$\Phi_{ij} = \exp \left( \frac{-d_{ij}}{2\pi} \right) \cdot \frac{1 - \cos(\theta_i - \theta_j) \cos(\psi_i - \psi_j)}{2}$$

(4)

where $\theta_i/\theta_j$ is the angle between $k_i/k_j$ and $P$; $A_i/A_j$ is the intersection point between $k_i/k_j$ and $P$; $d_{ij}$ is the distance from $A_i$ to $A_j$; $\psi_i/\psi_j$ is the counterclockwise/clockwise tilt angle of $k_i/k_j$. From the formulation, we can determine that $\Phi_{ij} \in [0, 1]$, with larger value when the orientations of $k_i$ and $k_j$ are more symmetric with each other.

2.3.3. Appearance symmetric similarity. Since neither scale similarity nor orientation symmetric similarity between keypoints $k_i$ and $k_j$ have considered the volume content, we
Figure 5. (a) The relationship between two symmetric 3D SIFT descriptors. Note that flipping does not change the elevation, but only inverts the azimuth of the 3D orientation. (b) Formation of the mirrored 3D SIFT descriptor $q_i$ by reordering sub-volumes, voxels and gradient histograms of $p_i$, where $N_x = N_y = 2$ and the dimension of $p_i$ is 2048.

Further measure their appearance symmetric similarity by comparing their descriptors $p_i$ and $p_j$, which encode the local volumes associated with $k_i$ and $k_j$, as shown in figure 3. The detailed mathematical definition of the 3D SIFT descriptor and its implementation can be found in Flitton et al (2010). Although the 3D SIFT descriptor satisfies scale-independent and orientation-independent after scale and orientation normalization, we still cannot define the appearance symmetric similarity by directly calculating the Euclidean distance between $p_i$ and $p_j$, because the 3D SIFT descriptor does not satisfy flipping-invariance. The relationship between two symmetric 3D SIFT descriptors is as shown in figure 5. Two symmetric 3D SIFT descriptors can be converted into each other by reordering the 2048 elements and warping.
their orientations. Hence, we need to form a mirrored 3D SIFT descriptor $q_i$ by reordering the 2048 elements of $p_i$. Recall that we have set $N_s = N_v = 2$ in section 2.1, as shown in figure 5(b), our sub-volumes are sorted as eight octants in a standard Cartesian coordinate system, e.g., $p_i = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8)$, where each $a_i$ denotes a sub-volume and includes $N_v^3 \times 8 \times 4$ of 2048 elements. For sub-volume level, we can reorder $p_i$ and obtain $q_i = (a_2, a_1, a_4, a_3, a_5, a_6, a_8, a_7)$ according to the flipping relationship of the eight octants. Similarly, for voxel level, we sort the eight voxels as eight octants and reorder them as in sub-volume level. Finally, for the gradient histogram bin level, we observe that the flipping does not change the order of elevation bins, but only inverts the order of azimuth bins. As shown on the right of figure 5(b), suppose the 32 bins of gradient histogram are sorted first according to azimuth and then elevated, then we can obtain the new gradient histogram of $q_i$ only by inverting the order of the eight azimuth bins in the original gradient histogram.

According to the above three-level reordering, we obtain a mirrored version 3D SIFT descriptor $q_i$, and denote the appearance symmetric similarity $\Omega_{ij}$ as follows

$$\Omega_{ij} = \exp\left(\frac{-||q_i - p_j||^2}{\sigma_a}\right), \tag{5}$$

where $\sigma_a$ controls the weighting of the appearance variation, and in our experiment, $\sigma_a = 0.8$. After the scale similarity, orientation symmetric similarity and appearance symmetric similarity have been defined, we can combine them to form an overall symmetric similarity metric $W_{ij}$ for each pair of keypoints $k_i$ and $k_j$

$$W_{ij} = S_{ij} \times \Phi_{ij} \times \Omega_{ij}. \tag{6}$$

Therefore, we can use $W_{ij} \in [0, 1]$ to quantify the symmetric magnitude for each pair of 3D SIFT keypoints. Keypoints are identified as a symmetric pair if their symmetric magnitude $W_{ij} > \omega$. In our experiment, we set $\omega = 0.6$ to select a number of symmetric pairs to fit the best MSP, where most outliers will be filtered out.

2.4. MSP fitting

Given a number of symmetric pairs of keypoints, we first calculate a set of midpoints $M = \{M_1, M_2, \ldots, M_n\}$ for the symmetric pairs to fit the best MSP within a brain volume, as shown in figure 6. The formulation of MSP is written as,

$$aX + bY + cZ + d = 0 \tag{7}$$

where $a$, $b$, $c$ and $d$ are the four parameters to uniquely determine the plane. The MSP detection problem has now changed to fitting a least-squares plane to a set of points $M = \{M_1, M_2, \ldots, M_n\}$. To perform a robust linear regression, we have adopted the LMS estimator (Mount et al 2007) to solve $a$, $b$, $c$ and $d$ because it can handle large fractions of outliers, unlike M-estimator because it tolerates much lower percentages of outliers and requires a careful initialization to avoid local optimization. We have further developed a two-stage filtering process to improve the efficiency of the LMS estimator. Specifically, after solving each tentative MSP $P[a_i, b_i, c_i, d_i]$, we use two regularizations to filter out two clusters of outliers. The first kinds of outliers are the points $O_1 = \{M_i\}$, for which the distance of $M_i$ to $P[a_i, b_i, c_i, d_i]$ is larger than 2 mm, as shown in figure 6(b). Also, we filter out another kind of outliers by measuring if $\theta_f$ satisfies $|\theta_f - \frac{\pi}{2}| < \frac{\pi}{2}$, where $\theta_f$ is the angle between the line passing through the original pair of keypoints (e.g., $k_3, k_4$) and $P[a_i, b_i, c_i, d_i]$. An interactive LMS plane fitting is performed and the best MSP is solved until all outliers are filtered out using the two conditions.
3. Results

We have implemented our method in C language on a PC with 4 Intel(R) Xeon(TM) CPUs 3.20 GHz, 16 GB RAM, and nVidia GeForce GTX 690 GPU with 4 GB video memory. To evaluate the effectiveness of our proposed method, we have tested it on both synthetic and in vivo MRI datasets.

3.1. Evaluation on synthetic datasets

To evaluate the accuracy of our method in extracting MSP, a set of reflectional symmetrical volumes is created. Then a fissure plane is figured out followed by flipping one side of the brain volume to form another mirrored hemisphere. The left and right symmetrical volume can be perfectly stitched and the ground truth MSP is known. By synthesizing perfectly symmetrical brain volumes, conventional subjective factors in human visual evaluations also can be avoided. In our experiments, 40 perfectly symmetrical brain volumes were synthesized. For different types of evaluations, a certain amount of Gaussian noise, blur, deformation or asymmetric structures have been added to the datasets.

After the synthetic data are completed, MSP is extracted using the proposed method. To clearly observe the accuracy of the proposed method, several representative slices perpendicular to the extracted MSP have been clipped. As shown in figure 7, the green lines are the intersecting lines between the clipped slices and the MSP. Compared with the top row, we can observe that our method can accurately extract the MSP, where the intersecting lines conform to symmetry axes presented in the brain slices. To evaluate the ability of our method in tolerance of noise, blur, distortions and asymmetric structures, we have also added Gaussian noise, Gaussian blur,
deformation and artificial tumors to the synthetic volumes, and applied our method to identify the MSP. The results from the second to fifth rows of figure 7 demonstrate the stability of our method. In our experiments, we synthesized an artificial tumor by drawing a sphere with a uniform gray level, where the size and the location of the tumor is randomly generated within the synthesized brain, so its scale and position can be changed, as shown in the bottom row of figure 7. Although the artificial tumors or the distortion regions cover about one third of the human brain, our method successfully identified the MSP.

Figure 7. MSP extracted from synthetic volumes with known ground truth. Gaussian noise, blur, local distortions and artificial tumors are added in the second, third, fourth and fifth rows, respectively. All synthetic volumes have a resolution of $256 \times 256 \times 256$. 
Figure 8. Visual comparison of different methods in extracting MSP from synthetic datasets. (a) Synthetic volumes. MSP extracted from (a) using Zhang and Hu (2008), Ruppert et al (2011), Liu et al (2011) and our method are shown in (b), (c), (d) and (e), respectively.

Also, we have compared the proposed method with three state-of-the-art methods, including Zhang and Hu (2008), Ruppert et al (2011) and Liu et al (2011). We have randomly selected 30 perfectly symmetrical brains and resynthesized into four types of datasets, e.g. data with different degree of rotation, Gaussian noise, Gaussian blur and artificial tumors, which combined a total of 120 synthesized brains for comparison. The MSP extraction results for different methods are as shown in figures 8(b)–(e), respectively. As Zhang and Hu (2008) only relied on the principal component analysis for gray similarity and used a linear combination to approximately construct the MSP, it could not always obtain a correct MSP with high accuracy, especially when the brain volumes suffered from rotation, noise, blur or local asymmetry structures (figure 8(b)). By detecting principal edges using a 3D Sobel operator and fitting the MSP based on distinctive edge matching, Ruppert et al (2011) achieved a better performance than Zhang’s method, as shown in figure 8(c). By representing the head volume as a reparameterized surface and searching the best MSP using surface matching, Liu et al
obtained comparable results versus Ruppert’s method, as shown in figure 8(d). But without relying on a better distinctive 3D image feature like 3D SIFT to extract MSP, all of the above methods have low tolerance to image rotation, noise, blur and asymmetry. Instead, by extending the well validated SIFT feature from 2D to 3D, and using it to detect MSP, our method outperforms the three competitors in terms of accuracy of symmetry detection.

Since our 3D SIFT features are invariant to scaling, orientation, affine distortion and intensity changes, our method is more stable in tolerance to rotation, noise, blur or asymmetry, and obtains better results than the other three state-of-the-art methods. As shown in figure 8(e), although the datasets have quite heavy noise or blur disturbance, our method can still robustly extract the MSP even for the cases with more than one tumor. Note that, since our 3D SIFT features are identified as the local extrema (minima or maxima) in the DoG volumes across the scale spaces, our method has an inherent tolerance to Gaussian blur, as shown in the third row of figures 7 and 8, and statistics results shown in figure 11 demonstrate its effectiveness. Even for the cases with more than one tumor, our method can still accurately identify the MSP as long as their total volume is not larger than one third of the whole brain (as shown in the bottom row of figure 8).

Besides visual comparison, we also calculated the yaw angle error (in degrees) and roll angle error (in degrees) between detected MSP $P_i\{a_i, b_i, c_i, d_i\}$ and the ground truth MSP $P_0\{a_0, b_0, c_0, d_0\}$ to quantitatively compare our method with the state-of-the-art methods. In our experiments, polar angle $\theta_p$ and roll angle $\theta_r$ are defined as

$$\theta_p = \arctan \frac{b}{a}$$

(8)

$$\theta_r = \arctan \frac{-c}{\sqrt{a^2 + b^2}}$$

(9)

where $\theta_p$ and $\theta_r$ can uniquely describe a MSP. So we can measure the accuracy of different methods in MSP extraction based on $\theta_p$ and $\theta_r$. After applying our method and the competitors over the synthetic datasets with different ratio of rotation, noise, blur and asymmetry, we calculated the mean squared error for polar angle $\theta_p$ and roll angle $\theta_r$ between detected $P_i$ and the ground truth $P_0$, as shown in figures 9–12. From the comparison, the proposed method generally outperforms the other three in terms of accuracy (both yaw angle and roll angle),
with more significant improvement when there is larger rotation angle (figure 9), heavier noise (figure 10), blur (figure 11) and larger asymmetry ratio (figure 12). This points out the effectiveness of using 3D SIFT feature to detect MSP.

Table 1 shows the statistics of yaw angle error and roll angle error of three synthetic datasets. From the statistics results, our method has improved the MSP extraction in terms of accuracy in both yaw angle and roll angle. To evaluate if our improvement is statistically significant, we conducted a multivariate analysis of variance (MANOVA) to compare our method with the three other state-of-the-art methods. In our experiments, yaw angles and roll angles obtained with the four competitors are fed for MANOVA. As shown in the bottom of table 1, the MANOVA yields a result of $F$-value = 13.04 and $P$-value = 0.0012, which indicated that there are significant differences among the methods. Besides, we also conducted a simple planned comparison between our method and each competitor to evaluate if there is a significant difference. The results are as shown in the two $P$-value columns of table 1.
Figure 12. Comparison of different methods in tolerance to asymmetry.

Table 1. Accuracy statistics of different methods over all simulated datasets. Each P-value in the columns is obtained using a simple planned comparison performed between our method and each competitor.

<table>
<thead>
<tr>
<th>Method</th>
<th>Yaw angle error (deg)</th>
<th>Roll angle error (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Zhang and Hu (2008)</td>
<td>2.563</td>
<td>0.356</td>
</tr>
<tr>
<td>Ruppert et al (2011)</td>
<td>2.382</td>
<td>0.301</td>
</tr>
<tr>
<td>Liu et al (2011)</td>
<td>1.857</td>
<td>0.225</td>
</tr>
<tr>
<td>Our method</td>
<td>0.926</td>
<td>0.198</td>
</tr>
</tbody>
</table>

MANOVA: F-value = 13.04, P-value = 0.0012.

Table 2. Running time statistics of different methods over all simulated datasets.

<table>
<thead>
<tr>
<th>Method</th>
<th>Average running time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>128 × 128 × 128</td>
</tr>
<tr>
<td>Zhang and Hu (2008)</td>
<td>9.856</td>
</tr>
<tr>
<td>Liu et al (2011)</td>
<td>7.231</td>
</tr>
<tr>
<td>Our method</td>
<td>0.639</td>
</tr>
</tbody>
</table>

From the planned comparison results in table 1, we can see that our method has a significant improvement in terms of accuracy at the 5% level (P-value is less than 0.05). Table 2 shows the statistics of running time for different methods over all three simulated datasets. To evaluate the speed of each method on the datasets with different resolutions, we resynthesized each brain volume from 256 × 256 × 256 to 128 × 128 × 128/512 × 512 by interpolation, and collected the running time of each method. The results indicated that the proposed method outperforms the other three in running time by 20 times. As shown in table 2, even when the volume size is 512 × 512 × 512, it still achieves a real-time performance.
3.2. Evaluation on in vivo datasets

The proposed method has been tested on 136 in vivo datasets, in which 30 cases are from normal subjects and 106 cases from patients with stroke, infarct, Alzheimer and tumor, and compared with the state-of-the-art methods. Extracted MSP using Zhang and Hu (2008), Ruppert et al (2011), Liu et al (2011) and the proposed method are shown in figures 13(c)–(f), respectively. Since there are no ground truths for the in vivo datasets, a specialist was invited to manually draw the MSP as shown in figure 13(b) to compare all methods with the human performance. From the visual comparison in figure 13, none of Zhang and Hu (2008), Ruppert et al (2011) and Liu et al (2011) can achieve a stable symmetrical detection because their methods rely on the gray principal component analysis, 3D edges and reparameterized surface, which are sensitive to distorted regions, noise or asymmetry, to determine the MSP. Instead, by using
Table 3. In vivo datasets collected for evaluation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Pathology</th>
<th># of volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stroke</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>Infarct</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>Alzheimer</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>Tumor</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>Normal</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 4. Statistical analysis of MANOVA on each dataset. Yaw angles and roll angles obtained in each dataset are fed for MANOVA.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14.05</td>
<td>0.0015</td>
</tr>
<tr>
<td>2</td>
<td>13.40</td>
<td>0.0025</td>
</tr>
<tr>
<td>3</td>
<td>13.62</td>
<td>0.0029</td>
</tr>
<tr>
<td>4</td>
<td>14.98</td>
<td>0.0010</td>
</tr>
<tr>
<td>5</td>
<td>15.08</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

Table 5. Accuracy statistics of different methods for each in vivo dataset. Each P-value is obtained using a simple planned comparison performed between our method and each indicated state-of-the-art method.

### Yaw angle error (deg)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>P-value</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>2.612</td>
<td>0.385</td>
<td>0.0007</td>
<td>2.438</td>
</tr>
<tr>
<td>2</td>
<td>2.653</td>
<td>0.379</td>
<td>0.0003</td>
<td>2.433</td>
</tr>
<tr>
<td>3</td>
<td>2.621</td>
<td>0.372</td>
<td>0.0009</td>
<td>2.432</td>
</tr>
<tr>
<td>4</td>
<td>2.637</td>
<td>0.376</td>
<td>0.0007</td>
<td>2.441</td>
</tr>
<tr>
<td>5</td>
<td>2.648</td>
<td>0.381</td>
<td>0.0007</td>
<td>2.435</td>
</tr>
</tbody>
</table>

### Roll angle error (deg)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>P-value</td>
<td>Mean</td>
</tr>
<tr>
<td>1</td>
<td>2.739</td>
<td>0.411</td>
<td>0.0008</td>
<td>2.507</td>
</tr>
<tr>
<td>2</td>
<td>2.736</td>
<td>0.421</td>
<td>0.0005</td>
<td>2.512</td>
</tr>
<tr>
<td>3</td>
<td>2.741</td>
<td>0.415</td>
<td>0.0008</td>
<td>2.515</td>
</tr>
<tr>
<td>4</td>
<td>2.737</td>
<td>0.419</td>
<td>0.0009</td>
<td>2.509</td>
</tr>
<tr>
<td>5</td>
<td>2.698</td>
<td>0.423</td>
<td>0.0008</td>
<td>2.513</td>
</tr>
</tbody>
</table>

3D SIFT feature to determine MSP, our method generally outperforms the other methods, especially on the pathological brain volumes that usually present accentuated asymmetry due to the presence of tumors or variation of the brain tissues. As shown in figure 13(f), results obtained from the proposed method are more identical to the expert’s results in figure 13(b), when compared to the other three methods in figures 13(c)–(e).

Besides visual comparison, we have further calculated the yaw angle error (in degrees) and roll angle error (in degrees) between detected MSP \( \{a_i, b_i, c_i, d_i\} \) and the human expert’s result \( P_0\{a_0, b_0, c_0, d_0\} \) to quantitatively compare our method with the other methods. Table 5 shows the statistics of accuracy for different methods on each dataset in table 3. From the statistics comparison in table 5, our method generally outperforms the other three competitors.
in terms of accuracy. To evaluate if our improvement is statistically significant, a MANOVA also has been carried out. Similar to the synthesized dataset, yaw angles and roll angles obtained with the four competitors are fed for MANOVA. Table 4 shows the results of MANOVA conducted in each in vivo dataset, which suggest that there are significant differences among the methods. In addition, we also conducted a simple planned comparison between our method and each competitor to evaluate if there is a significant difference. The results are as shown in table 5, which also indicates that our method has a significant improvement in terms of accuracy at the 5% level (P-value is less than 0.05). We further compared the running time of our method with the other three methods. From the results using the in vivo dataset as shown in table 6, the proposed method outperforms the other three methods by 20 times. Based on the GPU implementation of 3D SIFT matching and MSP fitting, we obtained a real-time performance on the in vivo brain datasets.

4. Discussion and conclusion

We have developed a novel MSP extraction method based on 3D SIFT by extending the conventional SIFT feature from 2D. The metric is based on scale, orientation and flipped descriptor to measure the symmetric similarity between two brain volumes. After clustering and indexing the 3D SIFT features using mean-shift and GPU KD-tree, the optimal MSP presented in a brain volume is solved on-the-fly. To the best of our knowledge, our method is the first to employ 3D SIFT to detect MSP from brain volumes, and it is more stable and available for real-time performance.

Compared with the existing methods for MSP detection from 3D brain images, the proposed method has the following advantages: (1) improved tolerance to image rotation, noise, blur and asymmetry, unlike other methods that use gray similarity, 3D edges or parameterized surface to detect MSP; (2) an intuitive and effective way to measure the symmetric similarity for each pair of 3D SIFT features, especially where 3D SIFT descriptor is proved to be invariant to scaling, orientation, affine distortion and intensity changes; (3) a real-time performance by clustering and indexing all detected 3D SIFT features using mean-shift and GPU KD-tree. Experiments on both synthetic and in vivo datasets have demonstrated the above advantages of our method in term of both accuracy and running time comparisons.

A few limitations should be taken into account. The proposed method cannot perform well for data with high levels of noise when the signal-to-noise ratio is less than –30 dB due to the difficulty in extracting the real 3D SIFT keypoints. Also, when the blur radius (sigma) is larger than 12 mm, or when the data are suffering from other kinds of blur or bias, such as unconstrained box blur and regional wrong focus bias, the accuracy of MSP extraction will be affected even though our method is able to tolerant a certain degree of Gaussian blur because our 3D SIFT features are identified as the local extrema (minima or maxima) in the difference

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18.351</td>
<td>16.139</td>
<td>13.653</td>
<td>0.876</td>
</tr>
<tr>
<td>2</td>
<td>18.348</td>
<td>16.202</td>
<td>13.572</td>
<td>0.869</td>
</tr>
<tr>
<td>3</td>
<td>18.401</td>
<td>16.161</td>
<td>13.499</td>
<td>0.873</td>
</tr>
<tr>
<td>4</td>
<td>18.383</td>
<td>16.152</td>
<td>13.586</td>
<td>0.871</td>
</tr>
<tr>
<td>5</td>
<td>18.395</td>
<td>16.144</td>
<td>13.552</td>
<td>0.875</td>
</tr>
</tbody>
</table>
of Gaussian volumes across the scale spaces. Furthermore, our method is limited to a certain ratio of asymmetry. It can correctly identify the MSP when a brain suffers from more than one tumor, but is not able to accurately perform when the total volume of the tumors is larger than one third of the whole brain. In addition, another side effect of our system to obtain a real-time performance is that the parameter setting of our method is somehow complex, because several steps of our implementation strategies are still a little heuristic, including $\tau_c = 0.05$ and $\tau_e = 30$, which are inherited from the original SIFT algorithm (Lowe 2004) to discard false candidates of keypoints that are localized along an edge or with poor contrast, and $\omega = 0.6$, which is also empirically set to quickly filter out outliers and select a number of symmetric pairs to fit the best MSP. Although our current parameter setting is optimized based on our synthesized datasets and collected in vivo datasets, our experiments suggest that such a parameter setting works well for new collected in vivo datasets as long as the noise or blur bias does not exceed upper limits mentioned above. Finally, similar to the previous methods of MSP detection from a brain volume, our method is also under the assumption that an ideal MSP is a plane. This assumption may not reflect the true anatomical structure of a human brain because the fissure plane could be a curved surface even for a normal person. But, since there exist many situations that only require finding a reference MSP in the brain volume, our method still potentially provides a useful tool for many clinical applications.

In conclusion, we have presented a more accurate, efficient and robust MSP detection method based on 3D SIFT features, which are detected, clustered and indexed under a novel parallel framework. The proposed method has been validated on both synthetic and in vivo datasets from normal and pathological cases. Our future work is to extend our method to curved symmetry surface detection for 3D brain images.

Acknowledgments

The authors would like to thank our anonymous reviewers for their valuable comments. This work was supported in part by grants from National Natural Science Foundation of China (no. 61303101, 61170326, 61170077, 81201157), the Natural Science Foundation of Guangdong Province, China (no. S2012040008028, S2013010012555), the Shenzhen Research Foundation for Basic Research, China (no. JCYJ20120613170718514, JCYJ20130326112201234), the Shenzhen Peacock Plan (no. KQCX20130621101205783), the Start-up Research Foundation of Shenzhen University (no. 2012–801, 2013–000009), the Research Grants Council of the Hong Kong Special Administrative Region, China (no. CUHK 475711, 416712, 473012) and the Innovation and Technology Commission (no. ITS/225/12).

References


Costantini G and Casali D 2007 Detection of symmetry axis by a CNN-based algorithm *Proc. 11th WSEAS Int. Conf. on Circuits* pp 46–49


Grigaitis D and Meilunas M 2007 Automatic extraction of symmetry plane from falx cerebri areas in CT slices *Bildverarbeitung fur die Medizin 2007* (Berlin: Springer) pp 267–71

Hermes G et al 2008 Neurological and behavioral abnormalities, ventricular dilatation, altered cellular functions, inflammation, and neuronal injury in brains of mice due to common, parasitic infection *J. Neuroinflammation* 5 1–37

Hu Q and Nowinski W L 2003 A rapid algorithm for robust and automatic extraction of the midsagittal plane of the human cerebrum from neuroimages based on local symmetry and outlier removal *NeuroImage* 20 2153–65


Oertel-Knochel V and Linden D E 2011 Cerebral asymmetry in schizophrenia *Neuroscientist* 17 456–67


Roussigné M, Blader P and Wilson S W 2012 Breaking symmetry: the zebrafish as a model for understanding left-right asymmetry in the developing brain *Neurobiology* 72 269–81


Stegmann M, Skoglund K and Ryberg C 2005 Mid-sagittal plane and mid-sagittal surface optimization in brain MRI using a local symmetry measure Proc. SPIE 5747 568–79
Teverovskiy L A and Liu Y 2006 Truly 3D midsagittal plane extraction for robust neuroimage registration Proc. 3rd IEEE Int. Symp. on Biomedical Imaging pp 860–3