RELIABLE EXTRACTION OF THE MID-SAGITTAL PLANE IN 3D BRAIN MRI VIA HIERARCHICAL LANDMARK DETECTION

*Alexander G. Schwing**

ETH Zurich
Department of Computer Science
Zurich, Switzerland

Yefeng Zheng

Imaging and Computer Vision
Siemens Corporate Technology
Princeton, NJ, USA

ABSTRACT

Reliable detection of the mid-sagittal plane is the key for brain image registration, asymmetry analysis, and group studies. Although the brain presents most of the time a regular structure, outliers in the data consisting of brain tumors or various deformations pose challenges to the existing approaches. We propose in this paper a robust approach for mid-sagittal plane extraction based on hierarchical landmark detection. Cross-validated results demonstrate comparable accuracy ($1.08°$ plane normal error) to those of human experts on a volumetric data set that contains pediatric patients as well as elderly with different diseases.

Index Terms— Mid-Sagittal Plane Detection, Brain Landmark Detection, Magnetic Resonance Imaging

1. INTRODUCTION

Spatial normalization, i.e., placing volumetric magnetic resonance imaging (MRI) data into a common coordinate system, is a critical step in most neuro-science multi-subject studies. A significant effort has been invested recently, to obtain automatic, fast, and accurate mid-sagittal plane (MSP) alignment [1]. Not only scientists, but most importantly radiologists working with clinical data on a daily basis require precise and fast localization of standard planes. As the patient's pose during different acquisitions changes with at least six degrees of freedom (translation and rotation), it is time consuming to manually find the same cross-section for comparison to previously obtained volumetric datasets. A reliable diagnosis and tracking of the course of disease is however essentially based on precise alignment of longitudinal studies. To ensure comparability and reliability, robust intra- as well as inter-patient alignment is required even for anomalous cases. Another application that relies on the MSP is scout imaging. Hereby, we obtain a low resolution volume and estimate the MSP to guide the following high resolution scan.

It is well known that the cerebral hemispheres are approximately symmetric and separated by the inter-hemispheric
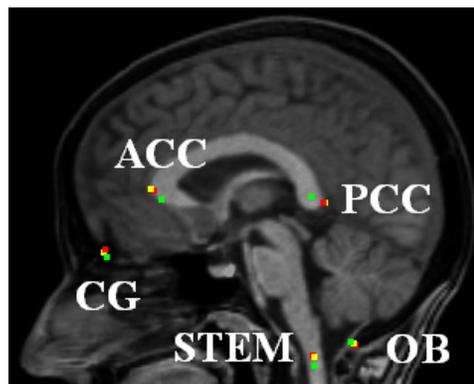

Fig. 1. The five landmarks on the mid-sagittal plane with red for the ground truth, green for initial estimate after global landmark object detection, and yellow for refined detection.

fissure. Thus, methods for extracting the MSP are usually divided into two main groups [1, 2]: (i) approaches maximizing a symmetry measure and (ii) algorithms that detect the inter-hemispheric fissure. Methods in the first group [3, 4] usually define a symmetry measure and search for a plane maximizing this score. Challenges occur due to accentuated asymmetries found in patients with schizophrenia, epilepsy and Alzheimer's diseases as well as cerebral malformations caused eventually by surgical procedures. Approaches in the second group try to detect the fissure mainly by applying snakes [5], the Hough transform [6], or Kullback-Leibler divergence [7]. The fact that the fissure is often not planar even in normal brains limits robustness of these approaches [3].

Recently, we proposed marginal space learning (MSL) [8] to estimate an object pose (including position, orientation, and scale) robustly and efficiently, which can be applied to detect the MSP. Clinically, an error of around $1°$ is desired for MSP plane orientation estimation, which is quite challenging. In MSL, the object detection problem is formulated as a binary classification problem: whether a given pose hypothesis is a good estimate or not. The classifiers in MSL have a certain limit: If the good and bad hypotheses have a distance less than this limit, the classifiers cannot distinguish them any
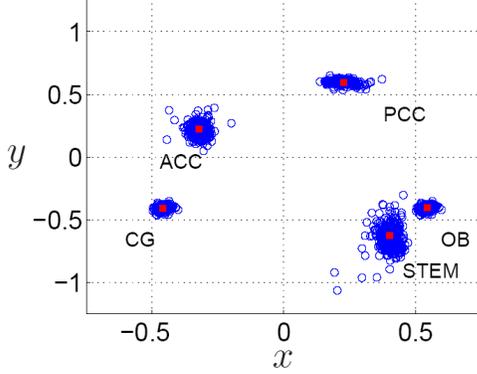
Fig. 2. Distribution of five mid-sagittal plane landmarks of a training set after normalizing the global translation, rotation, and scaling. Red rectangles indicate the inferred positions after global landmark object detection.

more. To obtain an accurate yet efficient estimate, we propose to detect five distinguishable landmarks on the MSP. By fitting a plane to those landmarks, the MSP can be detected consistently even if the inter-hemispheric fissure moderately deviates from a plane. Some of the most discriminative points on the MSP, also referred to as landmarks within this work, are the crista galli (CG), the tip of the occipital bone (OB), the anterior of the corpus callosum (ACC), the posterior of the corpus callosum (PCC), and a landmark in the brain stem (STEM) as highlighted with red markers in Fig. 1. Independent detection of each landmark is however time consuming and subject to detection outliers. In this work, we therefore exploit a hierarchical landmark detection approach. We first detect a global object composed of all landmarks. After estimating the position, orientation, and size of the global object using MSL, the rough estimate of the landmark positions can be inferred, which are then further refined within a local region.

2. MSP EXTRACTION BASED ON HIERARCHICAL LANDMARK DETECTION

The proposed hierarchical approach first detects a global object which permits to infer initial estimates of the five aforementioned landmarks. The estimates are illustrated with green markers in Fig. 1 and are subsequently refined with particularly trained classifiers to result in the locations highlighted with yellow color in Fig. 1. To this end, we employ a total of $N = 5$ landmark locations $z_i \in \mathbb{R}^3$, $i \in \{1, \ldots, N\}$ for every training volume. To facilitate detection of a global object, we need to define its position, orientation, and scale. The position is given as the arithmetic mean of the five landmarks within the volume, i.e., $p = \sum_{i=1}^{N} z_i$. Using the arithmetic mean p together with two chosen landmarks (CG and PCC in this work), we obtain a right-handed coordinate

system after employing the Gram-Schmidt procedure. Consequently, we define the three orientation parameters θ for a rotation with respect to the volume axis and the three scaling parameters s of the coordinate axis.

Given the chosen global object, we obtain a local representation $\tilde{z}_i$ of the position of the landmarks via

$$\tilde{z}_i = S^{-1} R^{\mathrm{T}} (z_i - p) \quad \forall i \in \{1, \ldots, N\}, \tag{1}$$

with an axis-aligned scaling matrix $S = \mathrm{diag}(s) \in \mathbb{R}^{3 \times 3}$ and the rotation matrix $R \in \mathrm{SO}(3)$ obtained from the rotation angles θ. The local coordinates of this model obtained by applying Eq. (1) are illustrated for different training volumes by the blue circles in Fig. 2. To obtain the mean representation, we compute the arithmetic mean over the training volumes. The mean shape coordinates $\bar{z}_i$ are illustrated via red rectangles in Fig. 2.

We parse a volume using the MSL classifiers trained with an adaptive Random Forest [9] to detect the global object. We thus obtain an estimated location $\hat{p}$, orientation $\hat{\theta}$, and scale $\hat{s}$. Having access to the mean model coordinates $\bar{z}_i$ computed with Eq. (1), we use our location, orientation and scale estimates to reverse Eq. (1), i.e., we apply

$$\check{z}_i = \hat{R} \hat{S} \bar{z}_i + \hat{p} \quad \forall i \in \{1, \ldots, N\} \tag{2}$$

to obtain an estimate $\check{z}_i$ for the position of all the landmarks. For clarity we use the hat $(\hat{\cdot})$ to denote the estimate of a classifier whereas we denote the estimate obtained when applying the geometric model with a check $(\check{\cdot})$. We illustrate this estimate with green markers in Fig. 1.

The result obtained by applying the mean model does not satisfy our expectation of accuracy. But as mentioned before, we are now able to constrain the detection range of classifiers designed for each particular landmark around the geometric model estimates $\check{z}_i$, $\forall i \in \{1, \ldots, N\}$. To be specific, we train a classifier using 3D Haar-like features [8] and the adaptive Random Forest [9] to precisely estimate which voxel corresponds to the targeted landmark position. The size of the constrained region around the initial estimate $\check{z}_i$ is dependent on the deviation observed during training of the mean model. Comparing our proposed hierarchical approach to standard landmark detection, we only parse the entire volume once rather than N times. We refer to the final refined estimates by $\hat{z}_i$ and illustrate the result with yellow markers in Fig. 1. We hence observe a significant visual improvement of $\hat{z}_i$ compared to the geometric inference results $\check{z}_i$ highlighted with green markers in Fig. 1.

To obtain the MSP parameters (c for plane center and n for plane normal), we solve the least squares minimization problem

$$\min_{n, c} \sum_{i=1}^{N} \left[(\hat{z}_i - c)^T n \right]^2 \quad s.t. \ \|n\| = 1. \tag{3}$$

Table 1. The four-fold cross-validated detection results using independent landmark (Lmrk.) detection.

Lmrk.	Mean Error	Med. Error	Std. Dev.
CG [mm]	3.42	2.32	10.50
OB [mm]	3.70	1.80	14.84
ACC [mm]	1.97	1.47	2.41
PCC [mm]	2.28	1.54	4.72
STEM [mm]	3.26	1.76	12.93
PNE [deg]	1.27	0.86	2.65
IPE [deg]	1.66	0.68	7.87

Table 2. The four-fold cross-validated detection results using hierarchical landmark (Lmrk.) detection.

Lmrk.	Mean Error	Med. Error	Std. Dev.
CG [mm]	2.79	2.37	1.98
OB [mm]	2.83	1.78	3.43
ACC [mm]	1.87	1.50	1.79
PCC [mm]	2.06	1.57	1.91
STEM [mm]	2.30	1.77	1.87
PNE [deg]	1.08	0.92	0.76
IPE [deg]	1.07	0.70	1.31

3. EXPERIMENTS

To validate our approach, the experiments are conducted on 509 volumes including male and female patients of all ages with various diseases. Due to this heterogeneity, we cover a huge set of variations. All the reported results are obtained after four-fold cross-validation, i.e., dividing the available volumes into four equally sized sets whereas three are used for training and the fourth set is used for testing. We rotate the configuration until each set is used for testing once. The detection accuracy of the landmarks (Lmrk.) is evaluated as well as the plane normal error (PNE) and the in-plane error (IPE).

First, we use the global object as defined in Section 2 to represent the pose of the MSP and use MSL to automatically estimate its location, orientation and scale. To accurately detect the orientation of MSP, we need to use a small searching step size for orientation. However, we found there was a certain limit beyond which the trained binary classifier could not distinguish a good orientation hypothesis from a bad one. The smallest error we can achieve is $1.65°$ of PNE and $2.09°$ of IPE using a searching step size of $3.26°$.

In the next experiment, we detect each of the five landmarks independently and then fit a plane to the detected landmarks as the MSP. The result is given in Table 1 for the detection accuracy of the landmarks (Lmrk.) as well as PNE and IPE. The PNE of $1.27°$ is much smaller than $1.65°$ achieved in the global approach. However, the high standard deviation (Std. Dev.) of landmark detection errors clearly indicates that some outliers cause the discrepancy between mean and median error (Med. Error). Note that due to treating the detectors independently, the detection time is approximately 80 seconds for a volume with $192 \times 192 \times 149$ voxels.

For comparison, the results using hierarchically constrained detection are given in Table 2. Note that, we achieve decreased mean errors compared to those of independent detection (see Table 1) while obtaining dramatically smaller standard deviations. We successfully prevent outliers and achieve a PNE of $1.08°$ and an IPE of $1.07°$.

We also compare our errors to the inter-observer variability obtained by expert annotations. We have 40 volumes annotated by four different physicians for four out of five land-

Table 3. The inter-observer variability measured on 40 expert-annotated datasets. In brackets we indicate the difference between our proposed approach (Table 2) and the inter-observer variability.

Lmrk.	Mean Error	Med. Error
CG [mm]	3.15 (+0.36)	2.76 (+0.39)
OB [mm]	2.29 (-0.54)	2.10 (+0.32)
ACC [mm]	1.62 (-0.25)	1.50 (+0.00)
PCC [mm]	2.13 (0.07)	1.87 (+0.30)
PNE [deg]	1.09 (+0.01)	0.98 (+0.06)
IPE [deg]	0.94 (-0.13)	0.73 (+0.03)

marks. The expert variability for those four landmarks are given in Table 3 together with the deviations from our result presented in Table 2. Hereby, a '+' indicates that the proposed approach has better accuracy. The proposed approach yields a PNE similar to the one achieved by physicians.

As shown in Fig. 3, our algorithm produces reliable results even in difficult cases. In the illustrations, the annotated ground truth is highlighted with red markers and the detected result is indicated with yellow dots. To illustrate the robustness of the proposed approach we draw the readers attention to the tumor covering almost the entire right cerebral hemisphere of the cross-section illustrated in Fig. 3 (a) or the over-exposure observed in (b). Despite these differences to standard volumes we detect the desired landmarks and therefore the MSP without problems. For acquisition of the volume shown in Fig. 3 (c), the head was rotated about 45 degrees as observable by comparing its pose to the orientation of the shoulders. The fourth volume originates from a pediatric clinic. Hence anatomical structures have different scales compared to adult's. Nonetheless we achieve reasonable accuracy after applying the same mean model. To indicate the difference of anomalous cases to usual volumes, we show detection results of two "standard" subjects in Fig. 4.

The accuracy we achieve is comparable to the state-of-the-art performance. It performs similarly to the inter-observer variability on our data set and mostly meets the clinical requirement of a one degree error on largely varying
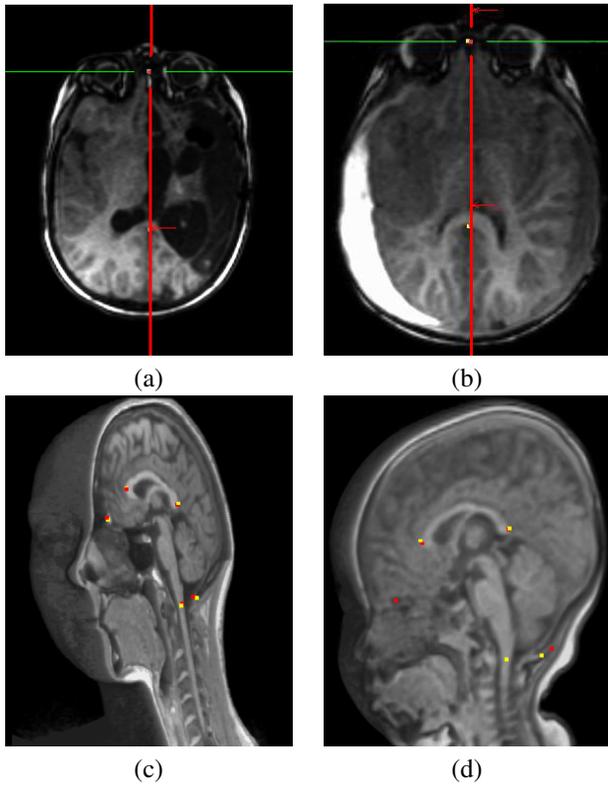
Fig. 3. Mid-sagittal plane detection results on some anomalous cases. (a) A tumor in the right cerebral hemisphere. (b) A bright spot in the left cerebral hemisphere. (c) A head rotated about 45 degrees during acquisition. (d) The head of a baby having a different shape to adult's. The results are indicated by either red vertical lines in (a) and (b), or detected landmarks (yellow dots) in (c) and (d).
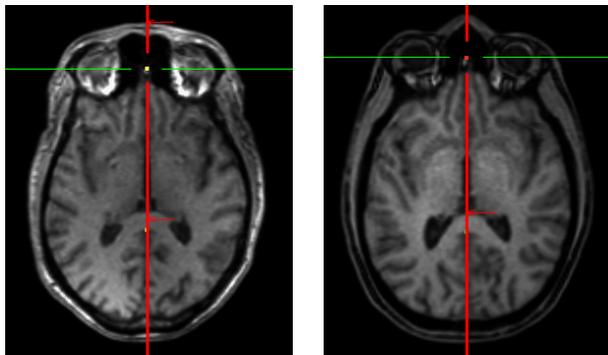

Fig. 4. Mid-sagittal plane detection results on two normal cases (indicated by red vertical lines).

scans. A lot of methods proposed in literature are validated on less than 100 volumes. Some of them use healthy subjects or phantoms for testing and others just allow limited variation of a "ground truth" position and orientation. We test on 509 volumes with variations in age, pathology, and pose. The time for the algorithms varies a lot and so does the accuracy. With an accuracy of $1.08°$ in 12 seconds we achieve state-of-the-art results. To give an example, the approach presented in [1] was tested on 64 volumes and achieved an error of $1.64°$ in about 60 seconds.

4. CONCLUSION

We proposed a reliable method for extracting the mid-sagittal plane based on hierarchical detection of anatomical landmarks. The method has been cross-validated on a heterogeneous data set of 509 3D volumes and achieves the state-of-the-art accuracy.

5. REFERENCES

[1] F. Bergo, A. Falcao, C. Yasuda, and G. Ruppert, "Fast, Accurate and Precise Mid-Sagittal Plane Location in 3D MR Images of the Brain," in *Proc. Biomedical Engineering Systems and Technologies*, 2008, pp. 278 – 290.

[2] P. Anbazhagan, A. Carass, P.-L. Bazin, and J.-L. Prince, "Automatic estimation of midsagittal plane and AC-PC alignment on nonrigid registration," in *Proc. IEEE Int'l Sym. Biomedical Imaging*, 2006, pp. 828–831.

[3] L. Teverovskiy and Y. Liu, "Truly 3D Midsagittal Plane Extraction for Robust Neuroimage Registration," in *Proc. IEEE Int'l Sym. Biomedical Imaging*, 2006, pp. 860 – 863.

[4] A. V. Tuzikov, O. Colliot, and I. Bloch, "Evaluation of the symmetry plane in 3D MR brain images," *Pattern Recogn. Lett.*, vol. 24, no. 14, pp. 2219–2233, 2003.

[5] R. Guillemaud, P. Marais, A. Zisserman, B. McDonald, T.J. Crow, and M. Brady, "A three dimensional mid sagittal plane for brain asymmetry measurement," *Schizophrenia Research*, vol. 18, no. 2, pp. 183–184, 1996.

[6] M. E. Brummer, "Hough transform detection of the longitudinal fissure in tomographic head images," *IEEE Trans. Medical Imaging*, vol. 10, no. 1, pp. 74–81, 1991.

[7] I. Volkau, K.N.B. Prakash, A. Ananthasubramaniam, A. Aziz, and W.L. Nowinski, "Extraction of the midsagittal plane from morphological neuroimages using the Kullback-Leibler's measure," *Medical Image Analysis*, vol. 10, pp. 863 – 874, 2006.

[8] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu, "Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features," *IEEE Trans. Medical Imaging*, vol. 27, no. 11, pp. 1668–1681, 2008.

[9] A.G. Schwing, C. Zach, Y. Zheng, and M. Pollefeys, "Adaptive random forest – how many "experts" to ask before making a decision?," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011, pp. 1377–1384.