

# Midbrain Segmentation in Transcranial 3D Ultrasound for Parkinson Diagnosis

Seyed-Ahmad Ahmadi<sup>1</sup>, Maximilian Baust<sup>1</sup>, Athanasios Karamalis<sup>1</sup>,  
Annika Plate<sup>2</sup>, Kai Boetzel<sup>2</sup>, Tassilo Klein<sup>1</sup>, and Nassir Navab<sup>1</sup>

<sup>1</sup> Computer Aided Medical Procedures, Technische Universität München, Germany  
ahmadi@cs.tum.edu

<sup>2</sup> Department of Neurology, Ludwig-Maximilians-University of Munich, Germany

**Abstract.** Ultrasound examination of the human brain through the temporal bone window, also called transcranial ultrasound (TC-US), is a completely non-invasive and cost-efficient technique, which has established itself for differential diagnosis of Parkinson's Disease (PD) in the past decade. The method requires spatial analysis of ultrasound hyperechogenicities produced by pathological changes within the Substantia Nigra (SN), which belongs to the basal ganglia within the midbrain. Related work on computer aided PD diagnosis shows the urgent need for an accurate and robust segmentation of the midbrain from 3D TC-US, which is an extremely difficult task due to poor image quality of TC-US. In contrast to 2D segmentations within earlier approaches, we develop the first method for semi-automatic midbrain segmentation from 3D TC-US and demonstrate its potential benefit on a database of 11 diagnosed Parkinson patients and 11 healthy controls.

## 1 Introduction and Medical Motivation

The midbrain, also called mesencephalon, is an approximately 2x2x1cm sized region of the brain, which forms the upper part of the brainstem. It contains several conglomerates of nerve cells, so called nuclei, which are involved in visual, auditory, and motoric functions of the human brain. It is well known that a degeneration of the *Substantia nigra* (SN), cf. Fig. 1, which is located in the midbrain, is the cause of Parkinson's Disease (PD) [16]. This degeneration, which is accompanied by an increased concentration of ferrite deposits, depletes other parts of the brain from the neurotransmitter dopamine, thereby causing the cardinal motor symptoms of Parkinson's disease, i.e. slowed movement, tremor, and muscular rigidity. These ferrite deposits can be visualized by transcranial B-mode ultrasound (TC-US) in form of hyperechogenic areas within the SN [1].

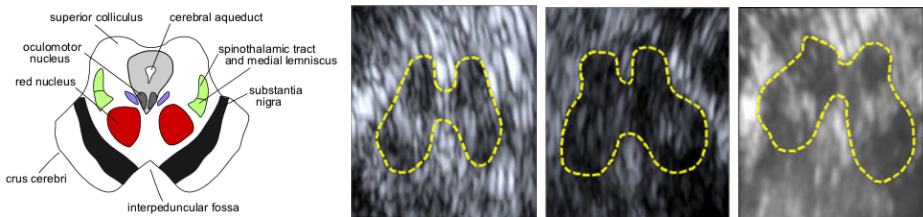
Several studies using this completely non-invasive method [1,16,2] found that it is possible to distinguish PD patients and PD types with sensitivities and specificities of above 90%, based on the size of the hyperechogenic area within the SN. There is also the indication that SN deterioration is visible in US *before* onset of motor symptoms, which is of high medical relevance, because an early diagnosis of PD may lead to better treatment and deceleration of disease

progression. Consequently, the hope of medical researchers is to be able to use TC-US as a screening method for PD in future, especially since the degeneration of SN cells *cannot* be visualized by Magnetic Resonance Imaging (MRI) [12].

Though bearing a lot of potential, PD diagnosis based on transcranial US is not easy and suffers from a few issues. Due to the scanning at low frequencies (2-4 MHz), transcranial US is unfortunately poor in resolution and characterized by high levels of noise and large speckle patterns (see Fig. 1). Also, the US acquisition through the temporal bone window, a layer of bone thin enough to be penetrated by low-frequency US (typically 2-4MHz), introduces additional phase aberrations and interferences [9] and higher absorption than in regular tissue, causing low contrast. Moreover, a high amount of experience is necessary to find an optimal scanplane in 2D and to interpret TC-US images correctly [15]. Thus, PD diagnosis based on TC-US suffers from high intra- and inter-rater variability. We strongly believe that this procedure can be improved significantly by the usage of 3D-US, making the selection of a cross-section superfluous, as well as an automated processing of the volumetric data, making the segmentation results reproducible.

The crucial part of any computer aided diagnosis system is an accurate segmentation of the whole midbrain. In fact, it is advisable to segment the midbrain first and the SN afterwards, because SN echogenicities appear to be almost identical to speckle and intensity patterns outside the midbrain area (cf. US images in Fig. 1). This was also proven by related work on computer aided PD diagnosis, where either a manual [10,4] or an automatic [6] midbrain segmentation is used.

As a consequence, any method which provides a robust and reliable midbrain segmentation from transcranial US greatly facilitates the subsequent steps of SN segmentation and classification. Please note that the only related approach proving an automated midbrain segmentation is the one of [6]. This method is, however, restricted to 2D and based on a complex and computationally intensive finite-element model. We argue that no further attempts at computer-assisted segmentation of transcranial B-Mode US have been performed yet, because a careful choice of prior knowledge, data term and optimization method is necessary (see Results section). Indeed, even recent reviews, e.g. [8], show that the application of shape models to 3D-US segmentation has not been addressed extensively, so far. To the best of our knowledge, this is the first work on *volumetric* midbrain segmentation from transcranial 3D-US.



**Fig. 1.** From left to right: Midbrain and basal ganglia such as substantia nigra, along with typical 2D TC-US images (middle), and a slice through one of our 3D volumes

## 2 Materials and Methods

In the following, we will show that it is possible to construct a method which is able to provide a reliable segmentation of the midbrain in the area around the SN. The proposed approach is based on three components: Firstly, the generation of a statistical shape model. Secondly, the combination of this shape model with an active surface framework. Thirdly, the active polyhedron framework of [14] to implement the discrete surface evolution. Before we describe these three steps, we briefly describe the acquired data and the necessary pre-processing.

**Data Acquisition and Pre-processing.** For our study, we collected data from 23 subjects, 11 previously diagnosed PD patients and 11 healthy controls (mean age 60.7 yrs., 60% female, 40% male). One subject was excluded due to insufficient bone window (4.3% exclusion rate). A medical expert scanned the 22 subjects with a medical ultrasound machine<sup>1</sup> at 3 MHz. Scans were performed bi-laterally, i.e. through the left and right temporal bone window, taking care that the entire midbrain area was contained in the scanned area. Bi-lateral 2D US planes were reconstructed into a 3D Freehand US volume using a backward-warping compounding technique [17] at an isotropic resolution of 0.45mm. Afterwards, the medical expert segmented all 22 volumes in a blinded, anonymized fashion into 3 regions, midbrain, SN left and SN right. Segmentation was performed only for those slices in which the SN was clearly visible and considered relevant for a diagnostic decision. Manual midbrain segmentations were then transformed into 3D meshes using the MATLAB iso2mesh toolbox [7] and default parameters (i.e. head surface element size bound of 2, keep ratio of 0.05).

**Statistical Shape Model Creation.** For SSM creation, we utilize the comprehensive SPHARM-MAT toolbox of [13]. The toolbox was slightly modified to take a set of training meshes as input and generate the SSM model in one function. The only modified parameter was the SPHARM degree for expansion and it was set to 15 for all meshes, otherwise default parameters as downloaded with the toolbox were used. A 5-fold cross-validation was performed, i.e. the 22 meshes were split into five groups and five SSM models were generated from four groups (training set) to perform segmentation and evaluation on the fifth group (test set). The output of the SPHARM toolbox trained on  $M + 1$  training shapes is a SSM with  $M$  modes of variation  $\mathcal{S}_i$ ,  $i = 1 \dots M$ :

$$\mathcal{S}_\alpha = \mathcal{S}_\mu + \sum_{i=1}^M \alpha_i \mathcal{S}_i, \quad (1)$$

where  $\mathcal{S}_\mu$  denotes the mean shape and  $\alpha_1, \dots, \alpha_M \in \mathbb{R}$ . The shape model  $\mathcal{S}_\alpha$  can be considered as a parametrized family of shapes, where a specific configuration is completely determined by the shape vector  $\alpha = (\alpha_1, \dots, \alpha_M)^T$ .

---

<sup>1</sup> Siemens Accuson Antares, PX4-1 transducer.

**Evolving the Shape Model.** In order to evolve the shape model towards the desired configuration, we iteratively minimize an active surface energy of the form

$$E(\mathcal{S}) = \int_{int\mathcal{S}} f_i dx + \int_{ext\mathcal{S}} f_e dx, \quad (2)$$

where  $\mathcal{S}$  denotes the surface,  $int\mathcal{S}$  the region inside  $\mathcal{S}$ , and  $ext\mathcal{S}$  the region outside  $\mathcal{S}$ . As our shape model provides enough regularity itself we employ no additional regularizer, such as the surface area. Due to the highly inhomogeneous nature of US images, foreground and background regions cannot be described by global statistics. Thus, we use a localized version of the Chan-Vese model [3] proposed by [11]

$$\begin{cases} f_i = (I - c_i(x))^2, & \text{where } c_i(x) = \frac{\int_{int\mathcal{S}} B_\epsilon(x) I(x) dx}{\int_{int\mathcal{S}} B_\epsilon(x) dx}, \\ f_e = (I - c_e(x))^2, & \text{where } c_e(x) = \frac{\int_{ext\mathcal{S}} B_\epsilon(x) I(x) dx}{\int_{ext\mathcal{S}} B_\epsilon(x) dx}, \end{cases} \quad (3)$$

where  $B_\epsilon(x)$  denotes a ball of radius  $\epsilon$  centered at  $x$ . In order to derive the evolution equation for the shape model, i.e. the gradient descent for the shape vector  $\alpha$ , we plug the shape model (1) into (2) and compute the partial derivatives with respect to  $\alpha_j$ :

$$\frac{\partial}{\partial \alpha_j} E(\mathcal{S}_\alpha) = \frac{\partial E}{\partial \mathcal{S}} \frac{\partial \mathcal{S}}{\partial \alpha_j} = \int_{\mathcal{S}} (f_i - f_e) N \cdot \mathcal{S}_j ds, \quad (4)$$

where  $N$  denotes the surface normal. Put together, these partial derivatives yield the gradient of  $E$  with respect to the shape vector  $\alpha$ , which we denote by  $\nabla_\alpha E$ . In order to implement a gradient descent for  $\alpha$ , we have to discretize the expression in (4).

**Active Polyhedron Framework.** We decided to use an explicit surface representation based on a triangular mesh as it is given by the active polyhedron method described in [14]. However, as the created shape model provides enough regularity we do not use the regularization described in [14] and evolve the model directly according to  $\nabla_\alpha E$ . For the evaluation of the cost function, we need to determine which voxels are inside or outside a given mesh. This is trivial from an algorithmic perspective, however, it is the computational bottleneck of the optimization process as this task is performed for each iteration. We addressed this issue by integrating a simple but efficient GPU accelerated voxelization algorithm introduced by Crane et al. [5] into our framework. Once all voxels inside and outside the shape are determined, we can compute the local mean values  $c_i$  and  $c_e$ , which eventually allows us to approximate the expression in (4) as follows:

$$\int_{\mathcal{S}} (f_i - f_e) N \cdot \mathcal{S}_j ds \approx \sum_{k=1}^N [f_i - f_e]_k [N]_k \cdot [\mathcal{S}_j]_k, \quad (5)$$

where  $k$  denotes the vertex number and  $[\cdot]_k$  denotes the evaluation at vertex  $k$ .

### 3 Experiments and Results

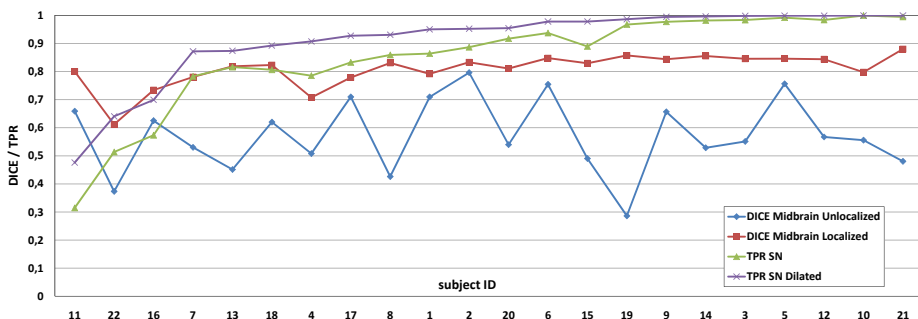
We performed the automated segmentation for all 22 subjects, given five folds of SSM models. The only user interaction necessary was the rigid placement of the mean shape  $S_\mu$  into the midbrain region. For all 22 experiments we kept the same parameter settings: the gradient descent step size was  $\tau = 0.05$ , the radius for the localization sphere  $B_\epsilon(x)$  in (3) was set to 15 voxels ( $\hat{=} 6.75\text{mm}$ ), and the maximum iteration number was set to 100.

In order to demonstrate that the localization of the data term (cf. (3)) is necessary, we performed another series of experiments with a standard Chan-Vese model [3]. Further, we investigated whether preprocessing the image data with a few anisotropic diffusion steps yields better results, but we could not observe a significant improvement. In all quantitative evaluations, only slices with manual ground truth were considered, since those were the only slices in which the SN was clearly visible and manually segmentable by the medical expert, i.e. we restrict our evaluation to those slices with medical relevance to the diagnostic problem.

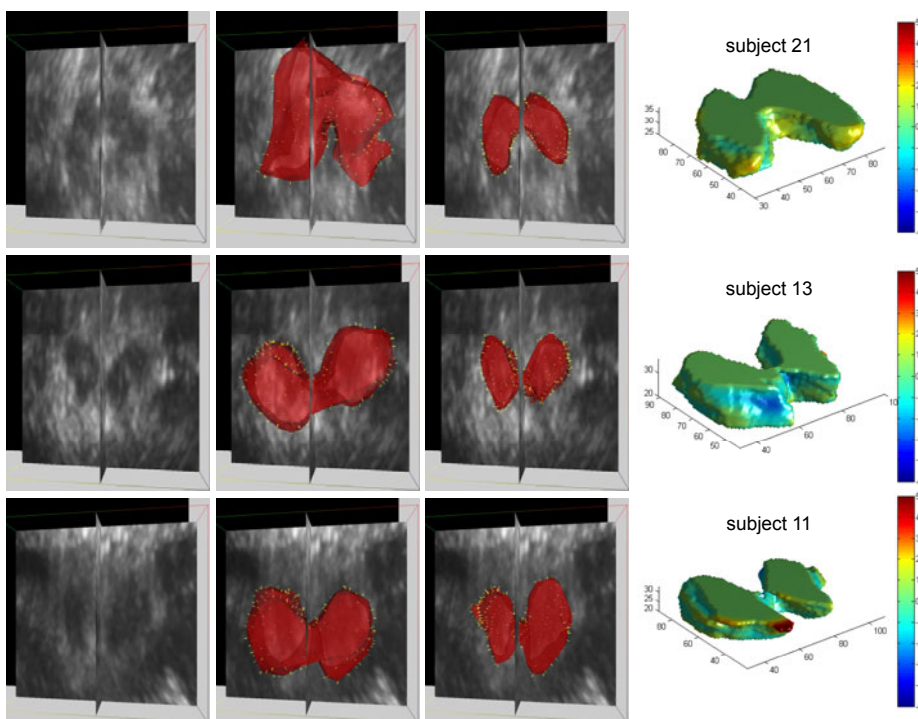
Regarding the midbrain segmentation, we calculated the DICE coefficient of overlap between ground truth segmentations and the segmentations obtained by our method. We can observe in Fig. 2 that using the localized data term improves the segmentation results significantly. Despite the poor image quality, the median of the DICE overlap of midbrain voxels across 22 subjects is 0.83, which means that in 50% of all cases the DICE coefficient is *at least* 0.83. In contrast to this, when the un-localized data term is used, we only achieve a median DICE of only 0.55, again supporting the assumption that global foreground and background statistics are not valid in the modality of ultrasound.

In order to evaluate the quality of the segmentation with respect to the diagnostic problem, we calculated the True Positive Rate (TPR) for SN segmentations, i.e. how many SN voxels were retained within the mesh after segmentation. We found that a median of 89% of SN voxels is retained by the midbrain ROI. We also post-processed all obtained segmentations with a dilation by one voxel ( $\hat{=} 0.45\text{mm}$ ), because we observed that the expert segmentation was systematically similar in shape to our segmentation, but segmented slightly further outwards, i.e. more into the hyper-echogenic regions surrounding the midbrain. In contrast, our segmentation converges slightly before these hyper-echogenic regions are reached. As we perform a ROI segmentation, a dilation of one voxel does not contradict the overall purpose. However, one should note that we only use dilation as one possible selection for a ROI, while our proposed method does not rely on dilation. This post-processing raises the median to 95%, i.e. half of our segmentations are able to retain more than 95% of voxels with diagnostic relevance, and five volumes achieve a perfect preservation of SN voxels after segmentation. Moreover, a dilation also improves the automatic midbrain segmentation, yielding a median DICE value of 0.86, which again shows that our segmentation is very similar in shape to the expert opinion.

In addition to the quantitative evaluation presented in Fig. 2, we want to compare one of the poorly segmented cases (case 11) to two other cases with



**Fig. 2.** DICE coefficients for overlap of our automatic segmentation with the midbrain (left) and True Positive Rate (TPR) with the substantia nigra (right). Subject indices are ordered by TPR SN Dilated, the percentage of voxels with diagnostic value retained.



**Fig. 3.** Exemplary Segmentation Results: Rows: data from three subjects. Columns: sample slice through volume with midbrain visible (left), segmentation result without data term localization (middle left), segmentation result with localization (middle right), mesh surface distance map between result and ground truth (colorbar in mm).

excellent (case 21) and medium (case 13) segmentation results (cf. Fig. 2), which serve as good examples for the performance of the algorithm on the remaining data sets. The upper, middle and bottom row in Fig. 3 show cases 21, 13, and 11, respectively. The last column of the figure shows mesh surface distance maps between the ground truth segmentation and the final automatic segmentation with localized data term, showing that the automatically segmented shape corresponds well with the ground truth. The reason for the relatively poor performance in case 11 can be explained by the comparatively bad image quality and the rather unusual shape of the midbrain, making also the manual segmentation very difficult - even for a medical expert. The same holds true for cases 16 and 22.

## 4 Discussion and Conclusion

We have presented a robust and largely automated method for segmentation of the midbrain in 3D TC-US, which is, to the best of our knowledge, the first approach for *volumetric* segmentation of the midbrain from TC-US. The performed experiments clearly demonstrate that the segmentation performance is consistently high across 19 out of 22 subjects, although the quality of US volumes differs highly due to different thicknesses of the temporal bone windows, which proves the robustness of our method. As the image quality also depends on the used US settings and hardware, the localized region-based data term provides the advantage that it does not depend on these factors, in contrast to an Active Appearance Model for instance. Thus, the presented method can be readily applied to any form of 3D B-mode volume generation, such as wobbler probes or 2D matrix arrays.

In terms of usability, the proposed semi-automatic segmentation method reduces the overall segmentation time from approximately 20 minutes per patient for manual segmentation to around 1.5 minutes. As an accurate, robust, and user-friendly midbrain segmentation from 3D TC-US is of high importance for a subsequent segmentation *and* classification of the SN, we strongly believe that an important step has been made towards a computer aided diagnosis of Parkinson's disease, possibly improving the chance for early detection of the disease and early onset of therapy for affected patients in future.

**Acknowledgments.** This work was partly sponsored by EU grant FP7-ICT-2009-6-270460 ACTIVE.

## References

1. Becker, G., Seufert, J., Bogdahn, U., Reichmann, H., Reiners, K.: Degeneration of substantia nigra in chronic parkinson's disease visualized by transcranial color-coded real-time sonography. *Neurology* 45(1), 182–184 (1995)
2. Berg, D., Godau, J., Walter, U.: Transcranial sonography in movement disorders. *Lancet Neurology* 7, 1044–1055 (2008)
3. Chan, T., Vese, L.: Active contours without edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001)

4. Chen, L., Seidel, G., Mertins, A.: Multiple feature extraction for early parkinson risk assessment based on transcranial sonography image. In: 2010 17th IEEE International Conference on Image Processing (ICIP), pp. 2277–2280 (2010)
5. Crane, K., Llamas, I., Tariq, S.: Real-time simulation and rendering of 3d fluids. In: Nguyen, H. (ed.) GPU Gems 3, ch. 30. Addison Wesley Professional, Reading (2007)
6. Engel, K., Toennies, K.D.: Segmentation of the midbrain in transcranial sonographies using a two-component deformable model. *Annals of the BMVA* (4), 1–12 (2009)
7. Fang, Q., Boas, D.: Tetrahedral mesh generation from volumetric binary and grayscale images. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009, pp.1142–1145 (2009)
8. Heimann, T., Meinzer, H.P.: Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis* 13(4), 543–563 (2009)
9. Ivancevich, N., Dahl, J., Trahey, G., Smith, S.: Phase-aberration correction with a 3-d ultrasound scanner: feasibility study. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 53(8), 1432–1439 (2006)
10. Kier, C., Cyrus, C., Seidel, G., Hofmann, U.G., Aach, T.: Segmenting the substantia nigra in ultrasound images for early diagnosis of parkinson’s disease. *International Journal of Computer Assisted Radiology and Surgery* 2(S1), 83–85 (2007)
11. Lankton, S., Tannenbaum, A.: Localizing region-based active contours. *IEEE Transactions on Image Processing* 17(11), 2029–2039 (2008)
12. Michaeli, S., Oz, G., Sorce, D., Garwood, M., Ugurbil, K., Majestic, S.: Assessment of brain iron and neuronal integrity in patients with Parkinson’s disease using novel MRI contrasts. *Movement Disorders: Official Journal of the Movement Disorder Society* 22(3), 334–340 (2007)
13. Shen, L., Farid, H., McPeck, M.A.: Modeling three-dimensional morphological structures using spherical harmonics. *Evolution: International Journal of Organic Evolution* 63(4), 1003–1016 (2009)
14. Slabaugh, G., Unal, G.: Active polyhedron: surface evolution theory applied to deformable meshes. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 2, pp. 84–91 (2005)
15. Vlaar, A., de Nijs, T., van Kroonenburgh, M., Mess, W., Winogrodzka, A., Tromp, S., Weber, W.: The predictive value of transcranial duplex sonography for the clinical diagnosis in undiagnosed parkinsonian syndromes: comparison with spect scans. *BMC Neurology* 8(1), 42 (2008)
16. Walter, U., Dressler, D., Probst, T., Wolters, A., Abu-Mugheisib, M., Wittstock, M., Benecke, R.: Transcranial brain sonography findings in discriminating between parkinsonism and idiopathic parkinson disease. *Archives of Neurology* 64(11), 1635–1640 (2007)
17. Wein, W., Pache, F., Röper, B., Navab, N.: Backward-warping ultrasound reconstruction for improving diagnostic value and registration. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 750–757. Springer, Heidelberg (2006)