

Automatic Positioning of Hippocampus Deformable Mesh Models in Brain MR Images Using a Weighted 3D-SIFT Technique

Matheus Müller Korb, Ricardo José Ferrari(✉), and for the Alzheimer's Disease Neuroimaging

Department of Computing, Federal University of São Carlos, São Carlos, SP 13565-905, Brazil
rferrari@ufscar.br
https://www.bipgroup.dc.ufscar.br

Abstract. Automatic hippocampus segmentationin Magnetic Resonance (MR) images is an essential step in systems for early diagnostic and monitoring treatment of Alzheimer's disease (AD). It allows quantification of the hippocampi volume and assessment of their progressive shrinkage, considered as the hallmark symptom of AD. Among several methods published in the literature for hippocampus segmentation, those using anatomical atlases and deformable mesh models are the most promising ones. Although these techniques are convenient ways to embed the shape of the models in the segmentation process, their success greatly depend on the initial positioning of the models. In this work, we propose a new keypoint deformable registration technique that uses a modification of the 3D Scale-Invariant Feature Transform (3D-SIFT) and a keypoint weighting strategy for automatic positioning of hippocampus deformable meshes in brain MR images. Using the Mann-Whitney U test to assess the results statistically, our method showed an average improvement of 11% over the exclusive use of Affine transformation, 30% over the original 3D-SIFT and 7% over the non-weighted point procedure.

Keywords: 3D-SIFT · Keypoint registration · Deformable mesh positioning · Hippocampus segmentation · Alzheimer's Disease · MRI

1 Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disease associated with age, affecting approximately 10% of the world population over 60 years old [5]. In Brazil, the average increase in mortality with the underlying cause of AD is 11.7% in men and 13.2% in women [27]. The natural evolution of

Alzheimer's Disease Neuroimaging—Disease Neuroimaging Initiative (ADNI) database.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12250, pp. 75–90, 2020.
https://doi.org/10.1007/978-3-030-58802-1_6

this disease manifests cognitive deficiencies that can lead to extreme incapacitation. To date, there is no cure for the AD; however, early diagnosis associated with appropriate treatment helps delay progression of symptoms, improving the patient's quality of life.

In medical centers and research laboratories, the manual hippocampi contouring procedure in magnetic resonance (MR) images used for measuring their volumes is usually conducted by a specialist following some predefined or established protocol [11]. However, the time spent and the visual fatigue of the specialists are limiting factors for many executions of this procedure [2]. As a result, the medical imaging community has become interested in developing fully automatic techniques for hippocampus segmentation in MR images.

In the literature, two automatic approaches have achieved the most promising results for the segmentation of hippocampus in MR images [1]. The first creates and uses multiple atlases, exploring label fusion or aggregation of clinical information [13,15,23], while the second combines atlases and deformable models [14,26,30]. In both cases, image registration techniques play a fundamental role, and an improvement in these algorithms would lead to increase precision of hippocampus segmentation.

In various medical image segmentation scenarios, deformable image registration techniques present the best results because they can successfully deal with anatomical inter-patient brain variation [4]. However, intensity-based registration that rely on optimization of the entire image space are prone to get stuck in local optima, possibly resulting in a large mismatch. By constraining the search space according to anatomical landmarks [6,9], such mismatches are unlikely to occur. Moreover, landmarks-based registration can take advantage of point matches in micro-regions of the brain, which are stable between patients and less susceptible to deformations due to age and neurodegenerative diseases [10].

In this work we present a new technique that uses 3D keypoints to estimate a deformable geometric transformation between a clinical and a template atlas images. The keypoints are automatically detected and matched using a modified version of Scale Invariant Feature Transform (SIFT) algorithm proposed by Rister [24,25], introduced to overcome numerical instabilities presented by the algorithm in the detection of tip-like landmarks. The geometric transformation are used to better positioning deformable hippocampus models in clinical MR images. Besides a modification of the SIFT algorithm, our technique also assigns to each matched pair of points a weighting value, which takes into account its distance to the centroid of hippocampus mesh model positioned via affine registration, to refine the mesh position adaptation.

2 Image Datasets

Two image datasets were used in this study; the first one is composed of synthetic structures and the second contains clinical brain images. In addition, it is presented a description of the topological atlas used for positioning the hippocampus meshes.

2.1 Synthetic Structures

A synthetic image dataset containing 864 tip-like structures was constructed using the parametric model proposed by Wörz et al. [29] with variations of translation, rotation, intensity, tapering, bending, and addition of Gaussian noise and bias field. These structures, which represent small three-dimensional high curvatures, were chosen because their similarity to 3D structures found in specific regions of MR brain images. All images were generated with isotropic resolution of 1 mm, size of $33 \times 33 \times 33$ voxels, and location of the point with maximum curvature mathematically determined.

In addition, a synthetic solid cube of size of $33 \times 33 \times 33$ voxels with constant intensity value 1, and centered in an empty three-dimensional array of size $128 \times 128 \times 128$ voxels, was created and used to investigate the behavior of the 3D-SIFT point detector proposed by Rister et al. [24].

These synthetic images were specifically constructed to evaluate the 3D-SIFT point detector [24] and its modification, making possible to determine the most sensible parameters of the method.

2.2 EADC-ADNI Clinical Dataset

The Alzheimer's Disease Neuroimaging Initiative (ADNI[1]) is part of a project created in 2003 by three research institutions from the United States (the National Institute of Aging, the Institute of Biomedical Imaging and Bioengineering and the Food and Drug Administration), with the aim of promoting the study of AD by making available standardized biomarker data.

The European Alzheimer's Disease Consortium (EADC-ADNI[2]) established a harmonized protocol for the manual segmentation of hippocampus in MR images, allowing more effective comparison between automatic and manual methods. The EADC-ADNI provides 135 T1-weighted (T1-w) RM images with binary masks of manually segmented hippocampi using a standardized protocol. The database contains images of patients between 60 and 90 years old and provides information about the cognitive health condition of each individual: Normal, Mild Cognitive Impairment (MCI), Late MCI (LMCI) and AD.

2.3 Topological Atlas

The topological atlas employed in this study is from the Neuroimage Analysis Center[3] (NAC). The atlas provides manual markings of brain structures of a healthy patient made by specialists in a T1-w MR image, with resolution of $1 \times 1 \times 1$ mm^3 and size of $256 \times 256 \times 256$ voxels. For each demarcated structure, a binary mask was extracted and a triangular mesh was constructed from it. In addition to a T1-w image used as a reference, this atlas also provides 149 meshes of various brain structures, including the left and right hippocampus.

[1] http://adni.loni.usc.edu/.
[2] http://www.hippocampal-protocol.net.
[3] http://www.spl.harvard.edu/publications/item/view/2037.

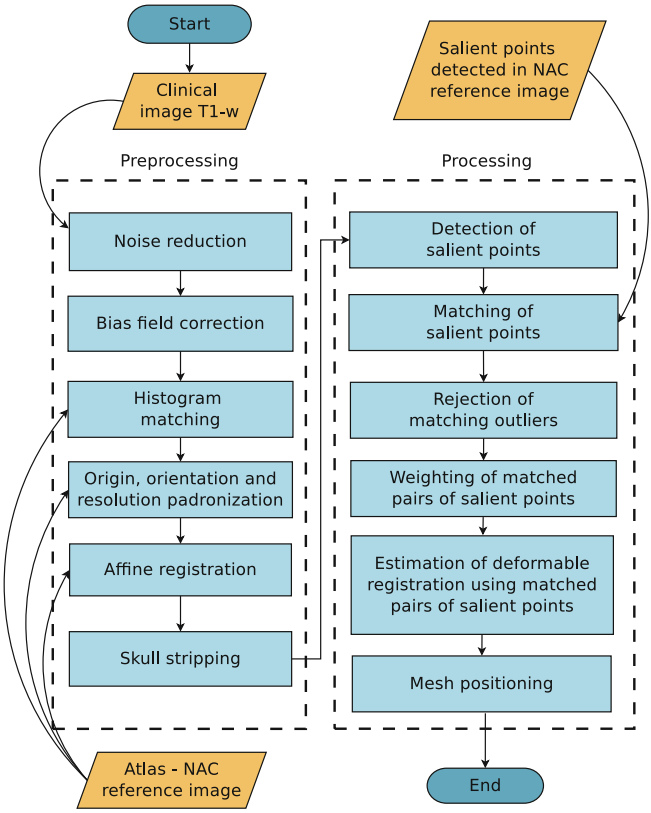


Fig. 1. Flowchart of the proposed method.

3 Methodology

This work proposes a technique for automatic hippocampus mesh positioning in MR images guided by automatically detected 3D salient points. Figure 1 depicts the sequence of steps used in the work.

3.1 Preprocessing

The preprocessing steps include image noise reduction using the Non-Local Means [3], bias field correction using the N4-ITK [17], intensity standardization using the technique of Nyul et al. [21], affine registration using the Nifty-Reg tool [22], skull stripping using the technique of Iglesias et al. [16] and, finally, the transformation of binary masks into three-dimensional meshes using an ITK-Mesh[4] image function.

[4] https://itk.org/Doxygen/html/classitk_1_1BinaryMask3DMeshSource.html.

3.2 Keypoint Extraction

Detection. The SIFT algorithm [25] uses a Gaussian scale-space (GSS) representation to detect keypoints. Formally, each level of the GSS representation is defined as a family of derived signals $L(\boldsymbol{x}, \sigma)$ computed as the convolution of the input image $I(\boldsymbol{x})$ with a Gaussian function, $G(\boldsymbol{x}, \sigma)$, as

$$L(\boldsymbol{x}, \sigma) = G(\boldsymbol{x}, \sigma) * I(\boldsymbol{x}), \tag{1}$$

where σ is the standard deviation, $\boldsymbol{x} = (x, y, z)$ corresponding to a spatial position, and $*$ is the convolution operation.

Using the GSS, stable feature point locations can be efficiently detected as extrema out of the convolution of the difference-of-Gaussian (DoG) function with the image, $D(\boldsymbol{x}, \sigma)$, which is computed from the difference of two nearby scales in the GSS separated by a constant multiplicative factor k as

$$\begin{aligned} D(\boldsymbol{x}, \sigma) &= L(\boldsymbol{x}, k\sigma) - L(\boldsymbol{x}, \sigma) \\ &= (G(\boldsymbol{x}, k\sigma) - G(\boldsymbol{x}, \sigma)) * I(\boldsymbol{x}). \end{aligned} \tag{2}$$

The DoG function has a close approximation to the scale-normalized Laplacian-of-Gaussian (LoG) [12], $\sigma^2 \nabla^2 G$, with the advantage of being computationally more efficient. As demonstrated by Lowe [19], this LoG-DoG relation can be expressed as

$$\left((k-1)\sigma^2\right) \nabla^2 G \approx G(\boldsymbol{x}, k\sigma) - G(\boldsymbol{x}, \sigma). \tag{3}$$

Equation 3 incorporates the scaling normalization factor σ^2, and guides the construction of the GSS, i.e, a DoG function pyramid, which is dependent on three factors: the number of Gaussian filter applications (scales) at each downsampling (octave), the number of octaves and, the standard deviation used.

From the DoG pyramid, the minimum and maximum values of the GSS are identified by comparing each point with the six nearest neighbors (up, down, front, back, left, and right) and then with the seven voxels of the next scale and so on. If the point remains as a minimum or a maximum, then it is considered a keypoint candidate. After that, a thresholding based on the relationship between the keypoint magnitude and the highest magnitude found among all keypoint candidates is applied [24], as described by

$$|D(\boldsymbol{x}, \sigma)| < \alpha \max_{\boldsymbol{x}, \sigma} |D(\boldsymbol{x}, \sigma)|, \tag{4}$$

where α is a parameter that is set to a value in the range $[0, 1]$. Equation 4 helps to adjust the threshold according to the contrast of the image, instead of using absolute values.

Before the GSS construction, the image is upsampled by a factor 2 followed by an initial smoothing and downsampling, also by a factor of 2. According to [19] this process assists in the detection of keypoints from information of the high frequencies of the image.

Rotation Invariance. To assign invariance to rotation, each keypoint and its neighborhood in the descriptor vector are rotated until its main orientation aligns with the direction given by the peak in a gradient orientation histogram [19]. In three or more dimensions this repositioning is not so simple, and an applicable alternative involves the analysis of the eigenvectors of a structure tensor (ST) [24], which is computed as

$$ST(\boldsymbol{x}) = \sum_{\boldsymbol{x} \in W} w(\boldsymbol{x}) \nabla I(\boldsymbol{x}) \nabla I^T(\boldsymbol{x}), \tag{5}$$

where $\nabla I(\boldsymbol{x})$ is the gradient of the image I at the spatial position $\boldsymbol{x}$, and $w(\boldsymbol{x})$ is a Gaussian weighted window centered on point $\boldsymbol{x}$. An advantage of this analysis is that it is less sensitive to noise than the partial derivatives [24]. In the matrix form, a structure tensor is represented by

$$ST(\boldsymbol{x}) = \begin{bmatrix} I_x^2(\boldsymbol{x}) & I_x(\boldsymbol{x}) I_y(\boldsymbol{x}) & I_x(\boldsymbol{x}) I_z(x) \\ I_y(\boldsymbol{x}) I_x(\boldsymbol{x}) & I_y^2(\boldsymbol{x}) & I_y(\boldsymbol{x}) I_z(\boldsymbol{x}) \\ I_z(\boldsymbol{x}) I_x(\boldsymbol{x}) & I_z(\boldsymbol{x}) I_y(\boldsymbol{x}) & I_z^2(\boldsymbol{x}) \end{bmatrix}. \tag{6}$$

The analysis of eigenvalues (λ_i) and eigenvectors ($\boldsymbol{q}_i$) of the ST matrix helps to determine the predominant edge orientations and the neighborhood isotropy of a point. Each eigenvector has an associated eigenvalue that implies a direction certainty. The local geometry of a point can thus be assessed by using a function of the eigenvalues of the ST matrix to distinguish edges and corners. This analysis also allows discarding salient points whose eigenvalues magnitudes are close to zero, which is a condition that produces instabilities [20,24]. Another possibility is to use the relation between the three eigenvalues to discard tubular structures, which do not vary in at least one direction.

To represent the gradient information, the SIFT algorithm [19] uses a vector whose magnitude reflects the maximum change in intensity values and the orientation of the vector corresponds to the direction of the intensity change. In the case of isotropic structures, this representation is problematic because there is no preferred gradient direction. To exemplify this, consider a keypoint located at the tip of a tip-like structure on a less detailed (smoothed) scale. In this case, the side edges are equal and opposite, generating a gradient magnitude value close to zero in this direction. When the relation of the eigenvalues of the tensor structure is used, this type of instability is minimized.

Proposed Modifications to the SIFT Technique. To better assess the performance of the SIFT, in this study we initially tested the algorithm on synthetic images, represented by a cube and tip-like structures, as described in Sect. 2.1. However, to our surprise, the algorithm completely failed to detect both the cube corners and the tips of the tip-like structures. This happened even by testing the algorithm with a large variation of its parameter set. After a thorough investigation, we found that the problem was in the strategy used by Rister et al. [24] to discard unstable keypoints. In their work, unstable keypoints are discarded

by thresholding the angle between the image gradient $\boldsymbol{d}$ and eigenvectors $\boldsymbol{q}_i$ of the structure tensor ST, computed as

$$\cos(\theta_i) = \frac{\boldsymbol{q}_i^T \boldsymbol{d}}{\|\boldsymbol{q}_i\| \, \|\boldsymbol{d}\|}, \tag{7}$$

and analyzing the relation between the eigenvalues as

$$\max_i \left(\left| \frac{\lambda_i}{\lambda_{i+1}} \right| \right) > \beta, \tag{8}$$

where the eigenvalues λ_i are organized in ascending order, i.e., $(\lambda_1 < \lambda_2 < \lambda_3)$ and β is set equal to 0.9 in their study.

Because the cube corners and the tip-like structures present approximate isotropy in two of the three dimensions [29], the strategy proposed by Rister et al. (Eqs. 7 and 8) failed to detect them. As a solution, we have eliminated the use Eq. 7 to reduce outliers and focused on the assessment of the local geometry in the neighbor of the point. To this end, we replaced Eq. 8 by

$$T_1 < \left| \frac{\lambda_i}{\lambda_{i+1}} \right| < T_2, \tag{9}$$

where $T_1 = 0.1$ is used to ensure the existence of variance in all 3 directions, even if minimal, and $T_2 = 1.0$ acts as an anisotropic filter for discarding isotropic curvature corners. Therefore, if the absolute eigenvalue ratios of a point are in the interval given by Eq. 9, then it is considered for further analysis, otherwise the point is discarded. Absolute values of the eigenvalue ratios close to 0 indicate a geometric line in at least one of the three dimensions or a numerical instability. By applying the above procedure, points belonging to tubular or plates structures are discarded.

Descriptor. The keypoint descriptor proposed by Rister et al. [25] and used in this research is defined by distinct histograms in a matrix of $4 \times 4 \times 4$ cubic sub-regions, with 12 vertices per histogram, resulting in 768 dimensions that is stored in a one-dimensional vector.

Matching. The matching between two SIFT salient points is performed by comparing the Euclidean distances between the descriptors, represented by uni-dimensional vectors. The analysis of the nearest neighbor was also incorporated to reject a matching if the nearest neighbor is very close to the second nearest neighbor. Such a process has shown to result in the elimination of 90% false positive matches.

Removal of Matching Points Outliers. Because curvatures of the human brain have similar characteristics, during preliminary tests we found matchings between very distant salient points, which turns out to be located in different

hemispheres. Such behavior, generated by the rotation invariance of the SIFT [24], occurs because the structural characteristics of the left hemisphere of a brain image may, eventually, be similar to the right hemisphere of another brain image. To circumvent such a problem, Euclidean distance thresholding is used to discard matched points that are more than 50 voxels of distance apart, which are defined in this paper as matching outliers.

3.3 Local Image Transformation

The set of m pairs of 3D salient points, $\{(P_{1,m}\,,\, P_{2,m})\}$, is used to estimate a B-spline transformation [18] that maps the spatial locations of the points $P_{1,m}$ to $P_{2,m}$. The resulting transformation, represented by a deformation vector field, is applied to the vertices of the reference mesh, configuring its position and shape to the clinical space of the image.

3.4 Weight Function

In this study, we proposed a weight function based on the Mahalanobis distance [8] between each pair of matched salient points and the centroid of the reference model (NAC hippocampal mask). This function permits a more adequate weight distribution to the cylindrical, elongated and slightly curved shape of the hippocampus, which is the structure of interest in this research.

To fit the ellipsoid in the hippocampus binary mask, we used the 3D ellipsoid fitting plugin from the ImageJ software[5], which returns the centroid coordinates, $\boldsymbol{c}_{x,y,z}$, and a covariance matrix, Σ, representing the shape and orientation of the ellipsoid. By using these parameters, the Mahalanobis distance from any point $\boldsymbol{p}_{x,y,z}$ in the image to the centroid $\boldsymbol{c}_{x,y,z}$ of the ellipsoid can be calculated as

$$D_{\mathrm{M}}(\boldsymbol{p}) = \sqrt{(\boldsymbol{p}-\boldsymbol{c})^T \Sigma^{-1} (\boldsymbol{p}-\boldsymbol{c})}. \tag{10}$$

Considering the B-spline function requires a weight in the $]0, 1]$ range for each point, in this study we defined the weight function

$$\rho(\boldsymbol{p}) = e^{-\gamma D_{\mathrm{M}}(\boldsymbol{p})} \tag{11}$$

to change the influence of each pair of matched salient point in the mesh deformation process. Parameter γ, called herein as Mahalanobis weight, is a constant that controls the exponential decay as the Mahalanobis distance (D_{M}) increases.

3.5 Description of the Model Parameters

In our analysis, we first used the synthetic cube image because of its simplicity and symmetry characteristics. However, to our surprise, the original SIFT algorithm fails badly to detect the its corners, even using a large variation of

[5] https://imagejdocu.tudor.lu/tutorial/plugins/3d_ellipsoid.

its parameters. This fact led us to investigate the reasons for the algorithm's failure and to propose the modifications described in Sect. 3.2 to overcome this limitation.

Since the total number of parameters in the method is relatively large (22 in total), we assessed the performance of the SIFT_orig with different variations of its parameters for the detection salient points using synthetic images and selected a subset of the most sensitive ones. These parameters were further optimized to best operate on T1-weighted MR images. A brief description of the selected parameters is given as follows.

The peak threshold parameter (T_{peak}) of the SIFT technique filters small peaks in the DoG scale space and, in this study, it was set to a small value because of the low contrast of the brain tissues in the region of interest (ROI), defined by a hippocampal binary mask. The initial standard deviation (σ_{init}) parameter is related to the image resolution and, in general, a small value should be used to avoid a large discard of high-frequency information. The standard deviation of the DoG pyramid (σ_{DOG}) is a parameter responsible for the construction of the scale-space and consequently the frequency bands of the filters. The number of octaves (N_{oct}), excluding the initial upsampling operation, was fixed to 2 in this work to avoid an excessive number of downsamplings and processing, as images greatly degraded by such operations and, therefore, will not contain relevant information. The number of levels per octave ($N_{\text{levels_oct}}$) is intrinsically linked to the standard deviation of the DoG pyramid; it adjusts the number of scales that will be in each octave. A small number generates few frequency bands for analysis, while a large number generates too many. The matching threshold (T_{match}) controls the importance of the point descriptor. Low values of the matching threshold, for instance, overwhelm the role of the descriptors, making difficult for matchings and assigning a greater relevance to them in discarding possible outliers. High values, on the other hand, reduce the importance of the descriptors, which results in accepting a greater number of matchings. The maximum matching distance ($T_{\text{max_dist}}$) parameter is used to filter the matching points that are too far apart. In this study, this was useful to eliminate matching points located in different hemispheres. The Mahalanobis weight (γ) parameter in Eq. 11 allows us to adjust which points will be the most relevant to the B-spline grid distortions based on the positioning of the meshes in the reference image. With this, greater importance can be attributed to the points close to the hippocampus to estimate the deformable transformation.

3.6 SIFT Parameter Optimization

The Simple Genetic Algorithm (SGA) method, which is available in the PyGMO scientific library[6], was used for fine-tuning the parameters. To this end, only one island, with a population of 50 individuals and five generations, was used. To define the objective function, images from six patients (two from each group Normal, MCI/LMCI, and AD) were selected from the EADC-ADNI dataset using

[6] https://esa.github.io/pagmo2/index.html.

random stratified sampling based on the presence of neurodegenerative disease and gender. These images represent approximately 15% of the each stratified population. As for the objective function, we used the mean Dice similarity coefficient of 12 hippocampi, two per patient. This optimization task resulted in the following parameter values: $T_{\text{peak}} = 0.1$, $\sigma_{\text{init}} = 1.0$, $\sigma_{\text{DoG}} = 1.85$, $N_{\text{levels_oct}} = 2$, $T_{\text{match}} = 0.85$, $T_{\text{max_dist}} = 50$, and $\gamma = 0.01$.

3.7 Statistical Comparative Analysis

The comparative analysis between different techniques (SIFT [24] and phase congruency (PC) [28] variants, and affine transformation) was conducted using the Mann-Whitney U test [7], which is a non-parametric test of the null hypothesis (H_0) that two samples come from the same population against an alternative hypothesis (H_1), comparing the mean values of the two samples. The test allows us to check whether the difference between the mean Dice values from two different techniques is statistically significant.

4 Results and Discussion

The experiments in this paper are organized into two distinct groups; the first uses synthetic images to evaluate the modifications introduced in the original SIFT technique, and the second uses all clinical MR images from the EADC-ADNI dataset to assess our method. For the following sections, we will refer to the assessed techniques as follows:

- AFFINE: application of the affine transformation only;
- SIFT_orig: original SIFT without the weight function;
- SIFT_orig_w: original SIFT with the weight function;
- SIFT_modif: modified SIFT without the weight function;
- SIFT_modif_w: SIFT modified with the weight function;
- PC: phase congruency without the weight function;
- PC_w: phase congruency with the weight function.

4.1 Evaluation of the Proposed Method on the EADC-ADNI Dataset

In these tests, we assessed the behavior of three methodological variations (AFFINE, SIFT_orig_w, and SIFT_modif_w) for the mesh positioning. The graphs in Fig. 2, which were used to illustrate the stratified results by diagnosis and gender groups, are similar to the population pyramid charts, where two bar graphs, one for each hippocampus, are arranged in a mirrored form with horizontal bars. Each bar represents the mean Dice of a subgroup with the corresponding standard deviation indicated at the bar end by a small horizontal line. In each graph, the two dashed vertical lines represent the mean Dice for the left and right hippocampus of the entire population.

For the assessment of the methods regarding normal and abnormal (neurodegenerative conditions) individuals, the population was divided into four groups, normal cognitive aging (referred herein simply as Nornal) patients, patients with MCI and LMCI, and patients with AD. These corresponding groups contain 42, 27, 16 and 43 individuals, respectively.

As can be observed in Fig. 2(a), the presence of AD worsens the mesh positioning when using the AFFINE method. By analyzing the Normal and AD groups, the mean Dice decreases 16.6% and 24.4% for the male right and left hippocampus, respectively, from Normal to AD. Correspondingly, the results for the female group are 25.5% and 21.4%.

This behavior also occurs when using the SIFT_orig_w, as can be seen in Fig. 2(b). However, the large number of outliers worsen the average positioning of the SIFT_orig_w in relation to the AFFINE method. In this case, the mean Dice results for the male right and left hippocampus decreased 14% and 23.8%, respectively while for the female right and left hippocampus, the corresponding decreases were 13.3% and 36.1%.

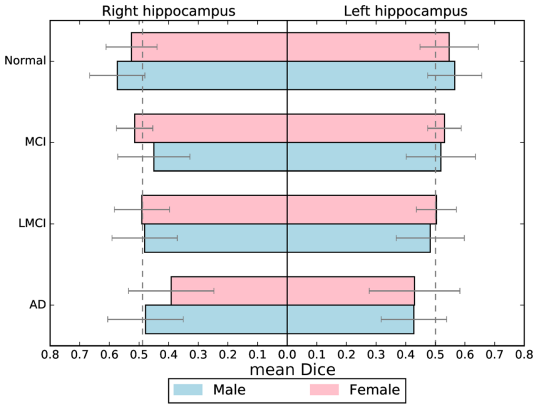
The results obtained with the proposed SIFT_modif_w method, shown in Fig. 2(c), indicate an improvement in the mean mesh positioning in relation to age. They also indicate the success obtained by such modifications due to the lower variance of the mean Dice results across the neurodegenerative conditions when compared to the other methods.

Because of the low number of individuals in the MCI (18 males and 11 females) and LMCI (9 males and 8 females) subgroups, the results presented for them precludes any meaningful statistical analysis. These subgroups, together or alone, present themselves as a structural transition between a healthy condition and Alzheimer disease.

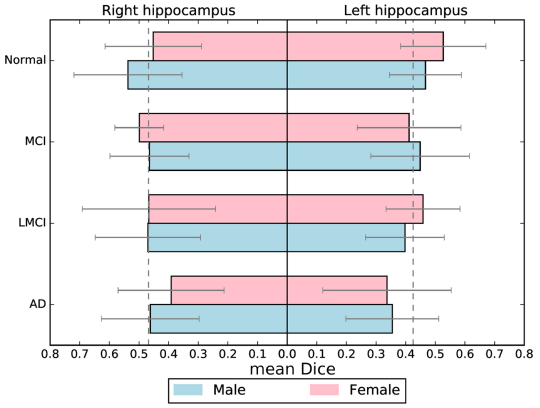
Finally, although the resulting Dice values for all experiments seem to be low, it is important to notice that for the best case scenario, i.e., complete overlapping between the ground truth and the positioned meshes, the Dice value will be lower than 1 because the size differences between the meshes. The goal is provide a better mesh positioning to improve the success of mesh adaptation.

4.2 Comparative Analysis Between AFFINE, SIFT Variants, and PC

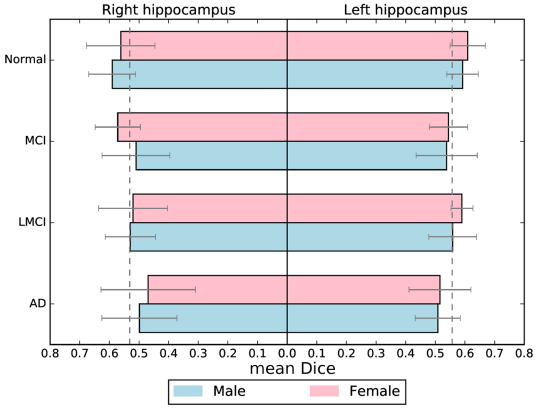
The results of mesh positioning using the AFFINE, SIFT variants, and PC methods are shown in Table 1. To facilitate the comparison with the other works, results are presented using the Dice metric, Jaccard, Hausdorff distance, and Hausdorff average distances. Although, the preprocessing steps were applied to all 135 images of the EADC-ADNI dataset, the results only refer to 129 individuals, since 6 images were excluded for previous adjustment of the method parameters. The results of the proposed SIFT_modif_w achieved an improvement of approximately 11% in relation to the exclusive use of affine transformation, 30% in relation to the SIFT without modifications, and 7% in relation to the positioning without using the weight function approach.



(a) AFFINE



(b) SIFT_orig_w



(c) SIFT_modif_w

Fig. 2. Results of hippocampus mesh positioning for the stratified data.

Table 1. Results of the hippocampus mesh positioning for the different methods applied to 129 images of the EADC-ADNI dataset. The p-values were obtained by testing each method against the AFFINE using the mean Dice. A $p-$value smaller than 0.05 indicates a statistical significance between the mean Dices from a given method and the AFFINE.

Metric		AFFINE	SIFT_orig	SIFT_orig_w	SIFT_modif	SIFT_modif_w	PC	PC_w
Dice	left hip.	0.49 ± 0.12	0.40 ± 0.17 $U = 12270$ $(p = 0.0000)$	0.42 ± 0.17 $U = 11286$ $(p = 0.0007)$	0.50 ± 0.10 $U = 8832$ $(p = 0.6625)$	**0.55 ± 0.09** $U = 6526$ $(p = 0.0001)$	0.52 ± 0.09 $U = 8322$ $(p = 0.2182)$	0.52 ± 0.08 $U = 8194$ $(p = 0.1522)$
	right hip.	0.48 ± 0.12	0.43 ± 0.18 $U = 10425$ $(p = 0.0408)$	0.46 ± 0.17 $U = 9418$ $(p = 0.6345)$	0.51 ± 0.13 $U = 7895$ $(p = 0.0578)$	**0.53 ± 0.12** $U = 7118$ $(p = 0.0019)$	0.50 ± 0.11 $U = 8328$ $(p = 0.2214)$	0.51 ± 0.11 $U = 8241$ $(p = 0.1746)$
Jaccard	left hip.	0.34 ± 0.10	0.26 ± 0.13	0.28 ± 0.14	0.34 ± 0.09	**0.38 ± 0.08**	0.35 ± 0.08	0.36 ± 0.09
	right hip.	0.33 ± 0.10	0.29 ± 0.14	0.32 ± 0.14	0.35 ± 0.12	**0.37 ± 0.11**	0.34 ± 0.09	0.35 ± 0.10
Hausdorff dist.	left hip.	7.29 ± 1.38	9.09 ± 2.96	8.45 ± 3.08	7.72 ± 2.41	**6.68 ± 1.38**	6.79 ± 1.27	6.73 ± 1.22
	right hip.	7.33 ± 1.55	8.72 ± 2.99	8.38 ± 3.01	7.36 ± 1.85	**6.84 ± 1.51**	7.16 ± 1.45	7.12 ± 1.44
Hausdorff avg. dist.	left hip.	1.02 ± 0.49	1.75 ± 1.34	1.66 ± 1.45	1.03 ± 0.51	**0.79 ± 0.28**	0.91 ± 0.31	0.91 ± 0.32
	right hip.	1.12 ± 0.53	1.64 ± 1.45	1.05 ± 1.83	1.07 ± 0.57	**0.95 ± 0.50**	1.04 ± 0.43	1.03 ± 0.43

5 Conclusions and Future Works

In this study we presented an automatic approach for hippocampus deformable mesh positioning using a modified version of the 3D-SIFT salient point detector. The proposed detector was assessed using synthetic volumetric and real MR images. By using the synthetic images, we could identify some limitations of the original 3D-SIFT algorithm and also determine its most sensitive parameters. These parameters were further optimized by a genetic algorithm to best perform on brain MR images.

A weight function was introduced to deliberately change the influence of the detected salient points in the hippocampus mesh positioning. Points closer to the hippocampus received a higher weight and, as a consequence, acted more effectively on the mesh positioning. The weight function has shown to produce better results for the mesh positioning.

Our method was evaluated using clinical MR images stratified on gender and diagnosis conditions. Results showed that the modifications made in the original SIFT significantly decrease the sensitivity of the method to the presence of the Alzheimer's disease.

As future work we intend to explore a more elaborate technique to reject matching outliers and use the result of our mesh initialization technique as input to a deformable simplex mesh model. External energy based on image gradient and texture information will be considered for the model adaptation.

Acknowledgments. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding Statement. This study was financed in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant number 2018/08826-9) and the

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Finance Code 001.

References

1. Achuthan, A., Rajeswari, M.: Prior integrated segmentation for brain structures: a review. Malays. J. Med. Health Sci. **14**(SUPP1), 190–200 (2018)
2. Bartel, F., Vrenken, H., Herk, M.V., Ruiter, M., Belderbos, J., Hulshof, J., Munck, J.C.: FAst Segmentation Through SURface Fairing (FASTSURF): a novel semi-automatic hippocampus segmentation method. PLoS ONE **14**(1), e0210641 (2019)
3. Buades, A., Coll, B., Morel, J.: A non-local algorithm for image denoising. In: IEEE Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, vol. 2, pp. 60–65, 20–25 June 2005
4. Carmichael, O.T., Aizenstein, H.A., Davis, S.W., Becker, J.T., Thompson, P.M., Meltzer, C.C., Liu, Y.: Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment. Neuroimage **27**(4), 979–990 (2005)
5. Chan, M.: Dementia: a public health priority. Technical report, World Health Organization and Alzheimer's Disease International, Geneva, Switzerland (2017)
6. Cheung, W., Hamarneh, G.: N-SIFT: N-dimensional scale invariant feature transform for matching medical images. In: 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 720–723. Arlington, VA, USA, 12–15 June 2007
7. Corder, G.W., Foreman, D.: Nonparametric Statistics: A Step-by-Step Approach, 2nd edn. Wiley, New York (2014)
8. Duda, R., Hart, P., Stork, D.: Pattern Classification, 2nd edn. Wiley, New York (2001)
9. Ferrari, R.J., Allaire, S., Hope, A., Kim, J., Jaffray, D., Pekar, V.: Detection of point landmarks in 3D medical images via phase congruency model. J. Braz. Comput. Soc. **17**(2), 117–132 (2011). https://doi.org/10.1007/s13173-011-0032-8
10. Fjell, A.M., et al.: Volumetric and microstructural regional changes of the hippocampus underlying development of extended delay long-term memory. bioRxiv - (2019). https://doi.org/10.1101/595827
11. Frisoni, G.B., et al.: The EADC-ADNI harmonized protocol for manual hippocampal segmentation on magnetic resonance: evidence of validity. Alzheimers Dementia **11**(2), 111–125 (2015)
12. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Prentice-Hall Inc, Upper Saddle River (2006)
13. Hao, Y., et al.: Local label learning (LLL) for subcortical structure segmentation: application to hippocampus segmentation. Hum. Brain Mapp. **35**(6), 2674–2697 (2014)
14. Hu, S., Coupé, P., Pruessner, J.C., Collins, D.L.: Appearance-based modeling for segmentation of hippocampus and amygdala using multi-contrast MR imaging. Neuroimage **58**(2), 549–559 (2011)
15. Iglesias, J.E., Leemput, K.V., Augustinack, J., Insausti, R., Fischl, B., Reuter, M.: Bayesian longitudinal segmentation of hippocampal substructures in brain MRI using subject-specific atlases. NeuroImage **141**, 542–555 (2016)
16. Iglesias, J.E., Liu, C., Thompson, P.M., Tu, Z.: Robust brain extraction across datasets and comparison with publicly available methods. IEEE Trans. Med. Imaging **30**(9), 1617–1634 (2011)

17. Juntu, J., Sijbers, J., Van Dyck, D., Gielen, J.: Bias Field Correction for MRI Images, vol. 30, 1st edn. Springer, Heidelberg (2005). https://doi.org/10.1007/3-540-32390-2_64
18. Lee, S., Wolberg, G., Shin, S.Y.: Scattered data interpolation with multilevel B-splines. IEEE Trans. Vis. Comput. Graph. **3**(3), 228–244 (1997)
19. Lowe, D.: Distinctive image features from scale invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
20. Ni, D., et al.: Volumetric ultrasound panorama based on 3D SIFT. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008. LNCS, vol. 5242, pp. 52–60. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85990-1_7
21. Nyul, L.G., Udupa, J.K., Zhang, X.: New variants of a method of MRI scale standardization. IEEE Trans. Med. Imaging **19**(2), 143–150 (2000)
22. Ourselin, S., Stefanescu, R., Pennec, X.: Robust registration of multi-modal images: towards real-time clinical applications. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2489, pp. 140–147. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45787-9_18
23. Pipitone, J., et al.: Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage **101**, 494–512 (2014)
24. Rister, B., Horowitz, M.A., Rubin, D.L.: Volumetric image registration from invariant keypoints. IEEE Trans. Image Process. **26**(10), 4900–4910 (2017)
25. Rister, B., et al.: Scale-and orientation-invariant keypoints in higher-dimensional data. In: IEEE International Conference in Image Processing, Quebec City, QC, Canada, pp. 3490–3494, 27–30 September 2015
26. Shao, Y., Kim, J., Gao, Y., Wang, Q., Lin, W., Shen, D.: Hippocampal segmentation from longitudinal infant brain MR images via classification-guided boundary regression. IEEE Access **7**, 33728–33740 (2019)
27. Teixeira, J.B., Junior, P.R.B.S., Higa, J., Filha, M.M.T.: Mortality from Alzheimer's disease in Brazil, 2000–2009. Cad. Saúde Pública **31**(4), 850–860 (2015)
28. Villa-Pinto, C.H., Ferrari, R.J.: Initialization of deformable models in 3D magnetic resonance images guided by automatically detected phase congruency point landmarks. Pattern Recogn. Lett. **79**, 1–7 (2016)
29. Wörz, S., Rohr, K.: Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. Med. Image Anal. **10**(1), 41–58 (2006)
30. Zarpalas, D., Gkontra, P., Daras, P., Maglaveras, N.: Hippocampus segmentation through gradient based reliability maps for local blending of ACM energy terms. In: 10th IEEE International Symposium on Biomedical Imaging, San Francisco, CA, USA, 7–11 April 2013, vol. 10, pp. 53–56 (2013)