Interactive Medical Image Segmentation by Statistical Seed Models

Thiago Vallin Spina, Samuel Botter Martins, and Alexandre Xavier Falcão
Institute of Computing – University of Campinas, Campinas, SP, Brazil

(a) Interactive segmentation with 5 markers.

(b) Interactive segmentation with 7 markers.

(c) Interactive segmentation with 9 markers.

(d) Proposed initial automatic segmentation.

(e) Proposed user correction with 3 markers.

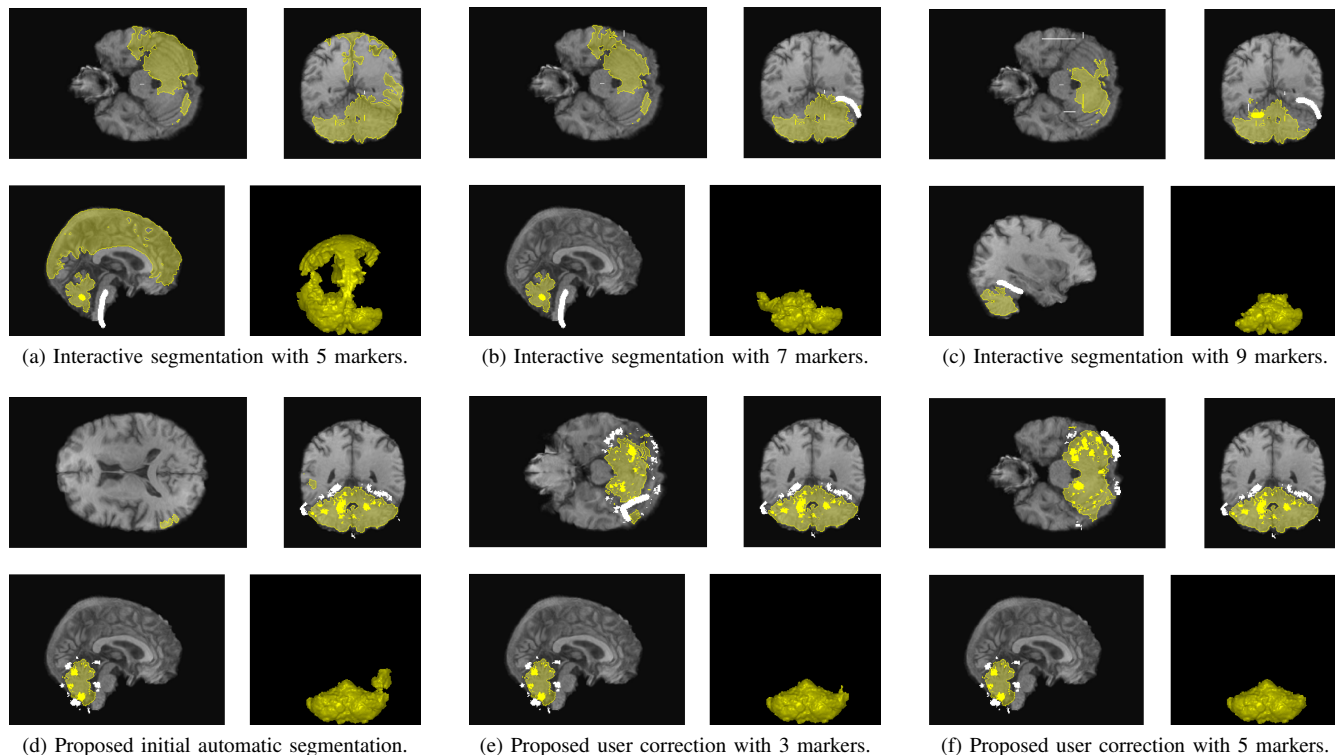(f) Proposed user correction with 5 markers.

Fig. 1. Segmentation results of the cerebellum obtained interactively in 3D without and with the aid of our proposed approach, for user U1 and image 36 in Table I. The 3D label renditions refer to the posterior-anterior coronal view of the organ. **Top row:** the interactive image segmentation of the cerebellum using a standard scribble-based methods [1], after the user added 5 markers to obtain an initial result. We depict next two other iterations when the user attempted to correct the poor result from (a) by adding 4 extra markers in total in different slices/planes. **Bottom row:** the proposed approach for interactive image segmentation assisted by a *Statistical Seed Model*. The result in (d) refers to the automatic segmentation using SSM without user interaction. The next two figures show the addition of only 5 extra markers in total by the user to correct our model's result. For comparison, the user added 15 markers in total in the fully interactive case to achieve a similar, but very crude, result in our experiments (see Figure 5).

Abstract—**Interactive 3D object segmentation is an important and challenging activity in medical imaging, although it is tedious and error-prone to be done. Automatic segmentation methods aim to replace the user altogether, but require user interaction to produce training data sets of segmented masks and to make error corrections. We propose a complete framework for interactive medical image segmentation, which reduces user effort by automatically providing an initial segmentation result. We develop a Statistical Seed Model (SSM) to this end, that improves from seed sets selected by robot users when reconstructing masks of previously segmented images. The SSM outputs a seed set that may be used to automatically delineate a new test image. The seeds provide both an implicit object shape constraint and a flexible way of interactively correcting segmentation. We demonstrate that our framework decreases the amount of user interaction by a factor of three, when segmenting MR-images of the cerebellum.**

Keywords-**Interactive Medical Image Segmentation; Object Models; Robot Users; Segmentation Resuming**

I. INTRODUCTION

Object segmentation is one of the most crucial and challenging activities in medical imaging. The goal is to assign a label to every element (voxel) in the image that indicates whether it belongs to one out of multiple objects of interest (body anatomic structures) or the background. As such, image segmentation may be divided into object *recognition* and object *delineation*. Object recognition involves approximately locating the objects in the image, as well as verifying the result of segmentation to point out errors for correction. Delineation is the process that actually performs label assignment. In general, object recognition is a simpler task for humans, while

the computers are often more precise in delineation.

Interactive image segmentation techniques usually exploit the synergism between the human operator for object recognition and the machine for object delineation, in order to increase effectiveness. The user can provide input using a few mouse clicks, often in the form of scribbles (*seed* voxels, also called *markers*) drawn over the foreground and the background, while the computer considers this input to delineate the objects. Many approaches follow this intuitive form of user interaction [2], [3], [4], [5], [1], [6], but the key problem of minimizing the amount of user intervention and the time necessary for obtaining an accurate segmentation [7], while maintaining the user's control over the process, still remains.

Image segmentation is not a single shot task. It requires the user to draw a set of scribbles to locate the object for a first delineation to be automatically performed (Figure 1a). Then, the user must verify if the result is correct and add more scribbles when errors occur (Figure 1b). This is an iterative process that may heavily burden the user (Figures 1b- 1c), who may even have to delete markers that were inadvertently misplaced. This is particularly troublesome in medical image segmentation, since objects such as the cerebellum are three-dimensional and usually connected to others with similar texture (e.g., the spinal cord, Figure 1), thereby requiring the user to carefully analyze multiple axes of orientation in search for errors.

To mitigate the above problem, several approaches have been proposed for automatic object segmentation, such as statistical atlases [8], [9], [10], [11], fuzzy object models [12], [13], active shape models [14], and template matching [15]. They differ from interactive methods in that they desire to replace the user altogether, by computing a model from prior object information (shape and/or texture) to treat object recognition. However, interactive segmentation is still required by automatic approaches since they consider a training data set containing segmentation masks of the object of interest for object modeling, which can only be accurately created under an expert user's supervision. Moreover, those approaches aim to solve object recognition, but they still demand interactive correction when errors in automatic segmentation occur.

In spite of several advances in object modeling, the persistent refusal in acknowledging the user's contribution to the development and application of automatic approaches has left two major open research questions. First, automatic methods have not been designed to aid the user in the very process of constructing the required data sets for training them, in order to reduce the spent amount of user effort. A related problem is that the update of the models is usually disregarded in case more unlabeled images become available for training, which must be once again segmented by the user from scratch. Second, since their goal is accuracy, the output segmentation produced by object models cannot be easily corrected by the user when the models fail.

Statistical atlases [8], [9], for instance, produce a *prior object probability map* (Figure 2a) that imposes a hard shape constraint about where the object's boundary may exist in

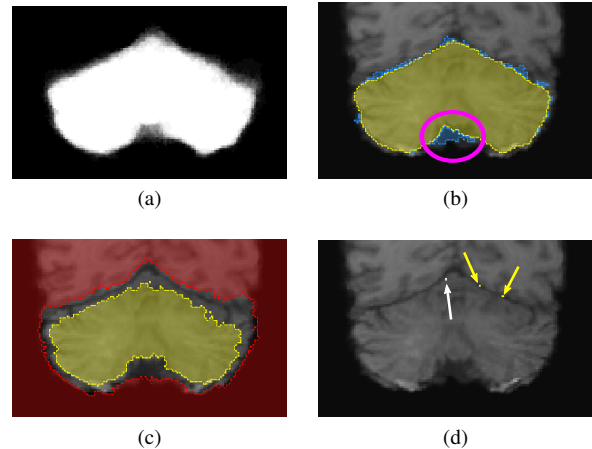

(a)   (b)

(c)   (d)

Fig. 2.   (a) Coronal view of the prior probability map for a statistical atlas of the cerebellum. (b) Error in atlas-based segmentation (yellow) circled in magenta (ground truth in blue). (c) The voxels shaded in yellow and red indicate the internal and external regions of the probability map, respectively, and may be considered as seeds. (d) The seed set generate by the method in [16], [17] to reproduce the atlas-based segmentation for correction (yellow and white arrows indicate the foreground and background seed voxels, respectively).

the test image. When errors occur (Figure 2b), one has to convert the internal and external portions of the map into seed voxels for interactive segmentation. However, those seeds are too dense to allow effective corrections (Figure 2c), even if a dilation/erosion is applied. To circumvent this problem, ad-hoc resuming techniques such as those in [16], [17] allow the conversion of *any* previous segmentation mask into a seed set with minimum cardinality that perfectly reproduces the input mask with a delineation algorithm. This time, the generated seeds are too sparse and make the user lose control over the process when she/he adds or removes seeds (Figure 2d).

As a first step towards answering the aforementioned open research questions, we propose a complete framework for *interactive image segmentation* that considers an object model for automatic segmentation to produce an initial delineation result (Figure 1d), which may then be evaluated and naturally corrected by the user with reduced effort (Figures 1e-1f. To this end, we design our framework by considering the very interactive image delineation algorithm that will be used for corrections in the training of our model, such that the output of our model *is* a suitable seed voxel set. This set strikes a balance between the strict object shape constraint of statistical atlases and the sparseness of the seeds produced by the resuming techniques in [16], [17].

We develop our framework by understanding how the user segments an image interactively, under the following premise: for a given object of interest (e.g., a body organ such as the cerebellum) and a specific interactive image segmentation method, an expert human operator places seeds roughly at similar (critical) locations to segment a training set of images (e.g., close to the boundary between the cerebellum and the spinal cord). Hence, if the training images are in a same
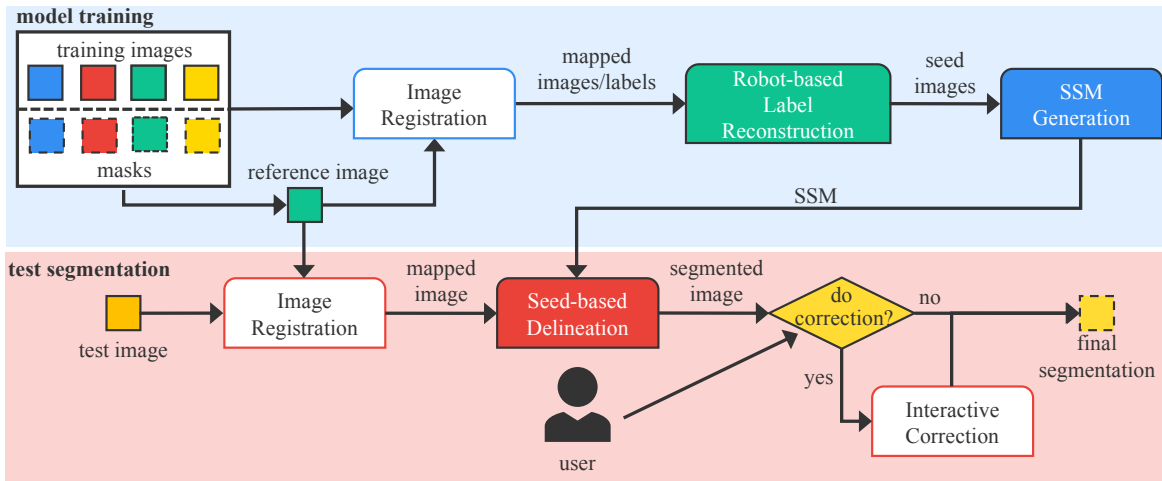
Fig. 3. Pipeline of the Statistical Seed Model construction and test image segmentation.

coordinate space, it should be possible to learn the statistical distribution of those seed locations and to build a *Statistical Seed Model* (SSM), rather than computing a prior probability map as statistical atlases. Once the distribution is estimated, our method selects a probabilistic seed set to automatically segment new test images (Figure 4c), which implicitly models the object's shape in a more flexible way. Interactive corrections may be directly done by the addition of new seeds and the removal of old ones.

To estimate the seed locations, the ideal procedure would be to have expert users segment a training set of images while recording those seeds for posterior learning, or even to construct our model while labeling the data set to apply it to aid the user in this process. For simplicity, we instead assume that some training images have already been segmented, and propose a solution based on robot users [18], [19], [20], [21]. These robots are algorithms developed for evaluating interactive delineation methods by simulating the user's behaviour in segmentation. Given an input segmentation mask, they try to approximately reconstruct it with the chosen delineation method by interacting with the method as the user would (e.g., by iteratively placing seeds at the geodesic center of segmentation error components [21]). Hence, the robots may select consistent seed sets if they are carefully elaborated.

A recent method proposed in [22] overlaps with our vision of a complete system for medical image segmentation. The authors have developed a template matching-based model that accepts user-drawn scribbles as hard constraints to force the numerical optimization in their framework to fix automatic segmentation results. However, according to their evaluation, their object shape regularization term makes it difficult to perform the interactive correction of large errors. We argue that ultimately their model was not designed by understanding how the user interacts with interactive segmentation methods. Unfortunately, their code is not available for comparison.

Our first contribution is an interactive medical image segmentation framework that facilitates user intervention, by

automatically providing an initial delineation result that can be easily and intuitively corrected. The second is a label reconstruction procedure-based on robot users, which may approximately reconstruct the result of *any* input segmentation mask. Our final contribution involves making seed estimation in existing robot users [19] more robust for label reconstruction. SSMs work for multiple objects simultaneously, if supported by the delineation method, although we leave such an improvement and its evaluation as future work. We compare our method against standalone interactive image segmentation conducted by expert users, which is the natural baseline of our approach, on a data set with MR images of the cerebellum. Section II describes the proposed framework, followed by our experimental evaluation (Section III), discussions about the method (Section IV), and conclusions (Section V).

II. PROPOSED FRAMEWORK

The next sections describe each step of our framework, whose pipeline we illustrate in Figure 3. As in statistical atlases, the input to our framework is a training set of segmented images, which must be placed in a same coordinate system. Hence, we first elect a reference image from the training set and perform deformable registration of the remainder with it, as commonly done for statistical atlases. Then, we map the segmentation masks according to the deformation fields of their corresponding original images. Instead of building a prior probability map by combining the mapped images as in statistical atlases [8], [9], we apply a label reconstruction procedure using robot users on each training image to reconstruct the mapped labels. Our method evaluates the resulting seed set locations across the training images and defines a common probabilistic seed set as the output for SSM. To delineate a test image, we place it in the reference image's coordinate space, and apply the SSM's seed set with a certain threshold for automatic segmentation. The user evaluates the result and corrects it interactively if necessary.

A. Image registration

We find the reference image I^* by selecting the one whose segmentation mask has the highest average Dice similarity score with respect to all other training images in the data set, after centralizing them according to the masks' geometric centers. Intuitively, this criterion selects as reference the image whose object is most similar to all others, thereby decreasing the amount of required deformation for registration.

Deformable image registration and mapping rely on the *Elastix* algorithm [23] to place the remaining training images and masks in the reference coordinate space. The Elastix algorithm computes a transformation $T = T_A \circ T_B$ that first rigidly maps the image I on to the reference image I^* with an affine transform T_A, and then deforms it with a B-spline transformation T_B to place it in the coordinate space of I^*. Transformation T then allows to map the segmentation mask L of I on to I^*. We apply the same algorithm to register new unlabeled test images with the reference image I^*.

B. Robot-based label reconstruction

Label reconstruction is an important step in our framework. As aforementioned, we intend to determine key locations for seed placement when segmenting a given object of interest, which depend on characteristics of the delineation algorithm A that we rely on.

The overall idea of robot-based reconstruction is to select an optimum seed set S by considering the input label as the ground truth that must be achieved via segmentation. A generic robot user works as follows, for a given image I and input label L:

1) Select an initial seed set $S \in I$ with voxels on the object of interest in L and the background according to a criterion C;
2) Run delineation algorithm A on I with S as input;
3) Measure the delineation accuracy of the resulting label L' by considering the input image L as the ground truth;
4) Select a new seed set S' with wrongly labeled voxels in L', also according to criterion C;
5) Update $S \leftarrow S \cup S'$ and iterate from 2 until a certain accuracy is reached or $S' = \emptyset$.

Steps 1-5 mimic how a human user segments images interactively. Criterion C is the core routine to be defined since it selects which voxels must be used as seeds. In its simplest form, C could select spherical *brush strokes* (markers) at the geodesic center of the largest segmentation error components as seed voxels at each iteration, as in [19] and [21]. This is how a non-expert human user traditionally uses interactive methods. However, C can be optimized for the chosen delineation algorithm A, given that A may perform better with fewer seeds if they are placed on strategic image locations.

In our framework, we have chosen algorithm A as the Image Foresting Transform by *optimum seed competition* [1] (IFT-SC). The IFT is a tool for the development of image processing and pattern recognition operators based on optimum connectivity, which generalizes Dijkstra's algorithm for multiple sources

and *smooth* path cost functions. The seeds in IFT-SC compete among themselves to conquer all of the remaining voxels in the image. Although IFT achieves state-of-the-art results, leakages in segmentation may occur around regions of the real object's boundary with lower image gradient. Those are the best locations where seed voxels may be selected to achieve accurate segmentation with minimum user intervention [2], [24], [1].

The robot user proposed in [19] was designed with a criterion C that prioritizes regions of the object's boundary in L with low gradient magnitudes. It ranks the voxels on a certain distance to the boundary by the gradient magnitude value of the closest voxels to them on the object's border in L as candidate seed centers, and then selects at each iteration a voxel p^* with lowest score as the center of a spherical brush stroke, if p^* was wrongly segmented by A at L'. Such criterion selects all voxels within a given radius from p^* as seeds, as long as they belong to the same label as p^* in L.

We argue that by selecting seeds as spherical brush strokes the criterion above does not follow the same principle as expert users do when drawing scribbles close to the object's boundary. The user-drawn scribbles tend to follow the object's contour and therefore protect better those regions, if they are selected around locations with low image gradient. Therefore, we update criterion C with the following modifications that compute scribble-like seed voxels in 3D:

C1. We select as candidate seed "centers" the voxels on the object's border in L, sorted by increasing order of image gradient magnitude value;[1]
C2. The currently selected "center" voxel p^* must not only be wrongly segmented in L', but also belong to the largest segmentation error component at that iteration;
C3. We mark a voxel q as a seed for S' if q is wrongly segmented in L', within a distance range $[d_{\min}, d_{\max}]$ to the object's border in L, and its closest voxel q' on the contour is within a given radius ρ of p^*.

Conditions C1 and C2 sort the selected scribbles not only by gradient value, but also by the size of the segmentation error components. Hence, it mimics expert users who add scribbles around weak parts of the object's boundary, prioritizing largest errors first to complete segmentation faster. Condition C3 further makes the seed voxels be selected on a geodesic band around the border that follows the object's contour, aiming to simulate scribbles instead of drawing spheric markers. Figure 4b depicts examples of seeds selected over the foreground and background by the robot user, until reaching an acceptable reconstruction accuracy. In general, we stop segmentation when the Dice similarity between the segmented image L' and the input label L is above a threshold α. In our experiments, we have found that the generation of our Statistical Seed Model achieves better results with our robot than with the one proposed in [19].

---

[1]We compute our image gradient by taking into account object information to enhance differences between the foreground and the background as in [8]. We linearly combined this enhanced gradient with the image's using a weight of 0.8 in its favor.
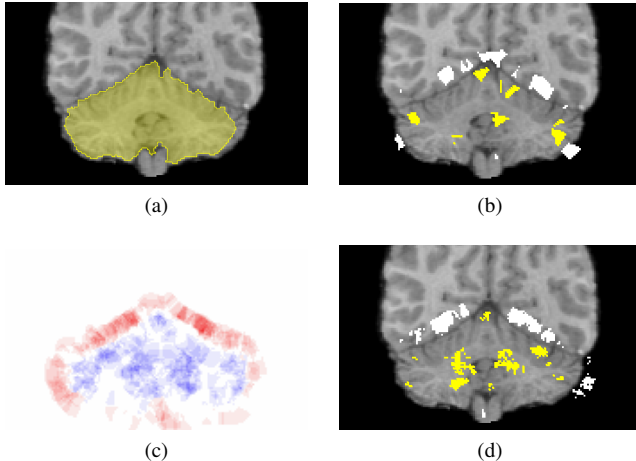
Fig. 4. (a) Input label for robot reconstruction. (b) Reconstruction seeds selected by the robot inside (yellow) and outside (white) the cerebellum. (c) Probabilistic seed set image in which blue-shaded voxels were more frequently labeled as foreground seeds, and red-shaded voxels were more frequently labeled as background seeds. (d) The final seed set $\mathcal{S}^*$ after thresholding the SSM from (c).

C. Statistical Seed Model generation

Given input seed sets $\mathcal{S}^1, \mathcal{S}^2, \ldots, \mathcal{S}^k$ for each of the k training images in a data set, the generation of a Statistical Seed Model is straightforward. Similarly to statistical atlases, those seed sets indicate whenever a voxel that belongs to a same location in the body anatomy has been selected as the background, as the foreground, or when it simply wasn't labeled (recall that the training images were registered). In statistical atlases, the training masks are averaged to output an estimated prior probability that the voxel has of belonging to the object of interest, thereby imposing a strict shape constraint.

The great advantage of SSMs over statistical atlases is that we select only a subset of voxels from the training masks that must be labeled as seeds for delineation, strictly where they are required to obtain accurate results. Hence, SSMs provide implicit shape constraints that might be nearly as effective as the ones imposed by atlases, with the added bonus of facilitating user corrections due to sparseness.

To compute an SSM, we place the seeds back in image coordinates and average them according to their label. We count the frequency that a voxel p has of belonging to the foreground $M_f(p)$ or to the background $M_b(p)$, if it was ever selected as seed. Let $M(p) = \max\{M_f(p), M_b(p)\}$ denote the maximum label frequency between both labels, representing the probabilistic seed set as depicted in Figure 4c, we select voxel p as part of the final seed set $\mathcal{S}^*$ if $M(p) \geq \tau$. We then assign the corresponding seed label as foreground or background depending on whether $M_f(p) > M_b(p)$. Hence, if a voxel has been frequently labeled as seed across all training images and mostly as foreground, it will be selected as a foreground seed. Figure 4d depicts the seed set $\mathcal{S}^*$.

D. Model-aided interactive segmentation of a test image

To segment a new test image I^t, we first register it with the reference image I^*. Then, automatic segmentation applies algorithm A with seed set $\mathcal{S}^*$ directly on I^t to obtain the initial segmentation for the user to visualize. If the image registration is perfect and the seed set $\mathcal{S}^*$ is properly positioned, we expect that the resulting segmentation be highly accurate with no need for user intervention. When that assumption fails, the seed set $\mathcal{S}^*$ is sparse and informative enough for the user to make corrections by removing misplaced seeds in $\mathcal{S}^*$ or by drawing new scribbles. Hence, with little effort the user may conduct hard segmentation given the model's implicit shape constraint.

III. EXPERIMENTAL RESULTS

It is very important to note that our end goal is to perform interactive image segmentation. Hence, our natural baseline is to use the same delineation algorithm applied during training interactively, without the aid of our model. We aim to show that Statistical Shape Modeling may reduce user intervention while achieving high accuracy in this setting. Moreover, SSM is a framework general enough in that we can change both the robot user selected for label reconstruction and the scribble-based delineation algorithm with any other existing in the literature. Hence, the advantages we want to demonstrate should extend to those cases as well, although we leave such an evaluation as future work.

We have conducted both automatic and interactive experiments to validate our approach, using a data set containing 37 MR T1-weighted images of the brain with voxel size $0.98 \times 0.98 \times 0.98$ mm^3 and their corresponding binary masks for the cerebellum. The cerebellum is a particularly challenging organ for interactive segmentation, since it presents intense connection with the spinal cord that makes it difficult to distinguish a clear-cut boundary between them. Ground truths were generated for all images by experts using manual and interactive segmentation tools.

We randomly divided the data set 10 times into 70% for training and 30% for testing to evaluate our method *without* user intervention. We then executed the automatic part of our pipeline for all folds, and measured the final accuracy of our method as the *Average Symmetric Surface Distance* (ASSD) between the segmentation result and the ground truth. Our method achieved 1.20 mm of mean error across all images and folds, which is state-of-the-art accuracy when comparing with other model-based automatic methods (e.g., see [8]).

To assess our method interactively, we would have to require the users to segment the images in each of the previously described test folds to measure accuracy and effort, since the resulting SSM was different for each one. Instead, we selected among all folds of the automatic evaluation step the 10 non-repeated test images in which the automatic method presented worse ASSD scores, and had two expert users U1 and U2 segment them with and without the aid of the corresponding SSM.

Given the difficulty in segmenting the cerebellum interactively, to demonstrate reduction in user effort we limited the

TABLE I

Experimental results for the cerebellum data set, comparing the worst results obtained automatically by SSM across all folds for 10 test images with the interactive segmentation of the corresponding images by two different expert users, U1 and U2, with and without its aid. The metrics refer to the ASSD scores in mm obtained for each segmented image. We present the interactive obtained scores using both 5 and 15 markers without SSM. Best scores denoted in bold.

Image Id	Aut. SSM	Interact. U1 (5 markers)	Interact. U1 (15 markers)	Proposed SSM-aided Interact. U1 (5 markers)	Interact. U2 (5 markers)	Interact. U2 (15 markers)	Proposed SSM-aided Interact. U2 (5 markers)
05	1.73	5.02	1.28	**1.06**	3.26	1.18	**1.10**
08	3.27	3.54	1.73	**1.42**	8.23	1.60	**1.34**
15	1.30	1.54	1.35	**1.02**	1.63	1.27	**0.92**
18	1.64	3.27	1.80	**1.15**	3.87	1.92	**1.05**
24	2.56	5.05	**1.73**	2.45	4.79	1.36	**1.07**
25	1.31	8.03	1.68	**1.29**	2.23	1.32	**1.24**
27	1.26	2.36	1.93	**1.16**	3.00	1.42	**1.03**
28	1.84	3.73	2.21	**1.25**	3.12	1.20	**1.13**
36	2.75	23.78	2.00	**1.27**	8.19	1.52	**1.30**
37	1.24	1.96	1.10	**0.98**	2.16	1.07	**0.97**
Avg.	1.89	5.83	1.68	**1.31**	4.05	1.39	**1.12**

amount of user intervention to the maximum addition of 5 scribbles per image in SSM-aided interactive segmentation. For the fully interactive case with IFT-SC, we allowed the users to add up to 15 scribbles.

Table I presents the measures obtained by SSM automatically and for both users interactively with and without model initialization. The ASSD error dramatically decreases by 0.58 mm for U1 and 0.77 mm for U2 when using SSM to initialize interactive segmentation. This result also improves in 0.37 mm for U1 and 0.27 mm for U2 the error measure when compared to performing interactive segmentation from scratch using 15 markers, even though the users added *three times less markers* with SSM. If we had had stopped fully interactive segmentation at 5 markers as well, as also demonstrated in Table I, the results would not even have been comparable since the mean ASSD error would have risen to 5.83 mm for U1 and 4.05 mm to U2. Figure 5 depicts some results from Table I.

Note that, prior to both the interactive and automatic evaluations, we validated SSM parameters via grid search and 10-fold cross validation in a subset with 50% of the data set images, which had no overlap with the test subsets used for evaluation. This validation resulted in the following configuration: $d_{min} = 2$ and $d_{max} = 5$ voxels for the distances that dictate the width of the robot user's scribbles; $\rho = 10$ voxels for the scribble's length; $\alpha = 0.935$ as the Dice similarity threshold for reconstruction, and $\tau = 25\%$ of the maximum frequency value in M to determine $\mathcal{S}^*$.

We performed image registration with the software Elastix [23] with the parameter file par0000.[2] All experiments used a desktop computer with a Core Intel i7-3770k CPU and 32GB of RAM, running Ubuntu Linux 14.04. Our research code was implemented in C language with little optimization. Overall, the most time-consuming part is the registration, which takes between 2 and 3 minutes per image. The model's training and testing stage are very efficient

---

[2]http://elastix.bigr.nl/wiki/index.php/Parameter_file_database

afterwards, since the bulk processing time can be done offline.

IV. Discussion

We have designed Statistical Seed Models for interactive image segmentation, aiming to facilitate manual curation in real applications. As such, we have experimentally demonstrated that the required amount of user effort can be dramatically reduced when using our method while increasing accuracy, after training it on a data set with segmented MR images of the cerebellum obtained from control patients. SSM was designed with some ideas from the traditional statistical atlas model, and may still be limited by similar drawbacks. In particular, the registration procedure might be an issue when dealing with pathological data. If errors in registration occur due to the presence of anomalies such as tumors, the segmentation of the body anatomical structure might be compromised.

We envision three possibilities to circumvent the aforementioned problem. The first one is to have the user correct errors in segmentation when abnormalities occur. This is reasonable to assume since we expect those errors to occur in exceptional locations not captured by the model during training. Moreover, our premise is that SSMs facilitate, but do not fully replace, the user's interaction. The second possibility is to train the model on a data set containing pathological cases to verify its accuracy in this setting. Since tumors and other problems may vary in appearance and position, we believe that this may be difficult to capture with our model. A third and final possibility is to compute the inverse mapping from the reference coordinate space on to the test image, and then apply the mapping to the generated seed set. Segmentation would then take place in the test coordinate system directly. This alternative has the advantage of not deforming the test image, thereby preventing the risk of missing pathologies. Since our seed set provides a sparse, although strict, shape constraint, we believe that corrections of the model's segmentation in the test space may be easier to do, since the pathology will be more evident in this scenario.

V. Conclusions and Future Works

We have presented a novel framework for the interactive segmentation of medical images, which includes an automatic segmentation step as initialization based on Statistical Seed Modeling. Our model trains with the help of robot users and outputs a seed voxel set that can be used to segment test images automatically, and which may be later modified with the addition or removal of scribbles to correct segmentation. We have demonstrated that our framework can decrease the amount of user interaction by threefold, while increasing accuracy in the segmentation of MR images of the cerebellum.

SSM is a general framework with building blocks that may be improved in several ways. Firstly, the selected delineation algorithm may be changed to others based on scribble selection, such as graph cuts by the min-cut/max-flow algorithm [5], [4], random walks [25], and fuzzy connectedness [6]. Secondly, the robot user selected for adding the seeds may be replaced as well by any other in the literature [21], as long as the selected seed set be consistently selected for the delineation algorithm on critical locations of the object of interest. Finally, different strategies for creating and testing the model may be employed, such as replacing deformable registration by simple alignment, as in Fuzzy Object Modeling [12], [13].

We have designed our method aiming to be a first step towards answering the following open questions. 1) Can we develop an object model to aid in the interactive segmentation of images, such that the very training of the model can be done while labeling the input data set? 2) Can this model be tested on new images and allow simple interactive corrections? We have demonstrated that 2) is possible by starting from an initial set of segmented images. In future works, we aim to evaluate if 1) can be resolved by our model, which would be trained directly with the input of expert human users. Lastly, once all images in a data set is segmented, we may use them to train other object model approaches for automatic segmentation.

Future works also involve extending our model to multiple objects, and including search optimizations that may reposition the seeds locally for improved segmentation as in [8]. We further intend to test our model with pathological data to evaluate the required amount of user effort spent in correcting errors, following the three alternatives presented in Section IV.

References

[1] A. X. Falcão, J. Stolfi, and R. A. Lotufo, "The image foresting transform: theory, algorithms, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26(1), pp. 19–29, 2004.

[2] T. V. Spina, P. A. V. Miranda, and A. X. Falcão, "Hybrid approaches for interactive image segmentation using the Live Markers paradigm," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5756–5769, Dec 2014.

[3] P. Miranda and L. Mansilla, "Oriented image foresting transform segmentation by seed competition," *IEEE Trans. Image Process.*, vol. PP, no. 99, pp. 1–1, 2013.

[4] V. Gulshan, C. Rother, A. Criminisi, A. Blake, and A. Zisserman, "Geodesic star convexity for interactive image segmentation," in *CVPR*, San Francisco, USA, 2010, pp. 3129–3136.

[5] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, Nov 2006.

[6] J. K. Udupa, P. K. Saha, and R. A. Lotufo, "Relative fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 1485–1500, 2002.

[7] A. X. Falcão, J. K. Udupa, S. Samarasekera, S. Sharma, B. E. Hirsch, and R. A. Lotufo, "User-steered image segmentation paradigms: Livewire and live-lane," *Graph. Model. Im. Proc.*, vol. 60, no. 4, pp. 233–260, 1998.

[8] R. Phellan, A. X. Falcão, and J. K. Udupa, "Medical image segmentation via atlases and fuzzy object models: Improving efficacy through optimum object search and fewer models," *Med. Phys.*, vol. 43, no. 1, pp. 401–410, 2016.

[9] M. R. Sabuncu, S. K. Balci, M. E. Shenton, and P. Golland, "Image-driven population analysis through mixture modeling," *IEEE Trans. Med. Imag.*, vol. 28, no. 9, pp. 1473–1487, 2009.

[10] A. Guimond, J. Meunier, and J.-P. Thirion, "Average brain models: A convergence study," *Comput. Vis. Image Underst.*, vol. 77, no. 2, pp. 192–210, 2000.

[11] J. C. Mazziotta, A. W. Toga, A. Evans, P. Fox, and J. Lancaster, "A probabilistic atlas of the human brain: theory and rationale for its development the international consortium for brain mapping (icbm)," *NeuroImage*, vol. 2, no. 2PA, pp. 89–101, 1995.

[12] J. K. Udupa, D. Odhner, L. Zhao, Y. Tong, M. M. Matsumoto, K. C. Ciesielski, A. X. Falcão, P. Vaideeswaran, V. Ciesielski, B. Saboury, S. Mohammadianrasanani, S. Sin, R. Arens, and D. A. Torigian, "Body-wide hierarchical fuzzy modeling, recognition, and delineation of anatomy in medical images," *Med. Image Anal.*, vol. 18, no. 5, pp. 752–771, 2014.

[13] P. A. V. Miranda, A. X. Falcão, and J. K. Udupa, "Clouds: A model for synergistic image segmentation," in *ISBI*, Paris, France, 2008, pp. 209–212.

[14] J. Liu and J. K. Udupa, "Oriented active shape models," *IEEE Trans. Med. Imaging*, vol. 28, no. 4, pp. 571–584, 2009.

[15] R. Gauriau, R. Ardori, D. Lesage, and I. Bloch, "Multiple template deformation application to abdominal organ segmentation," in *ISBI*, Prague, Czech Republic, April 2015, pp. 359–362.

[16] P. A. V. Miranda, A. X. Falcão, G. C. S. Ruppert, and F. A. M. Cappabianco, "How to fix any 3d segmentation interactively via image foresting transform and its use in mri brain segmentation," in *ISBI*, Chicago, USA, March 2011, pp. 2031–2035.

[17] P. A. V. Miranda, A. X. Falcão, and G. C. S. Ruppert, "How to complete any segmentation process interactively via image foresting transform," in *SIBGRAPI*, Gramado, Brazil, 2010, pp. 309–316.

[18] T. V. Spina and A. X. Falcão, "Robot users for the evaluation of boundary-tracking approaches in interactive image segmentation," in *ICIP*, Paris, France, Oct 2014.

[19] P. E. Rauber, T. V. Spina, P. Rezende, and A. X. Falcão, "Interactive segmentation by image foresting transform on superpixel graphs," in *SIBGRAPI*, Arequipa, Peru, 2013.

[20] K. McGuinness and N. E. O'Connor, "Toward automated evaluation of interactive segmentation," *Comput. Vis. Image Underst.*, vol. 115, no. 6, pp. 868–884, 2011.

[21] P. Kohli, H. Nickisch, C. Rother, and C. Rhemann, "User-centric learning and evaluation of interactive segmentation systems," *Int. J. Comput. Vis.*, vol. 100, no. 3, pp. 261–274, 2012.

[22] R. Gauriau, D. Lesage, M. Chiaradia, B. Morel, and I. Bloch, "Interactive multi-organ segmentation based on multiple template deformation," in *MICCAI*, München, Germany, 2015, pp. 55–62.

[23] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. P. W. Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, 2010.

[24] P. A. V. Miranda and A. X. Falcão, "Links between image segmentation based on optimum-path forest and minimum cut in graph," *J. Math. Imaging Vis.*, vol. 35, no. 2, pp. 128–142, Oct 2009.

[25] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, 2006.
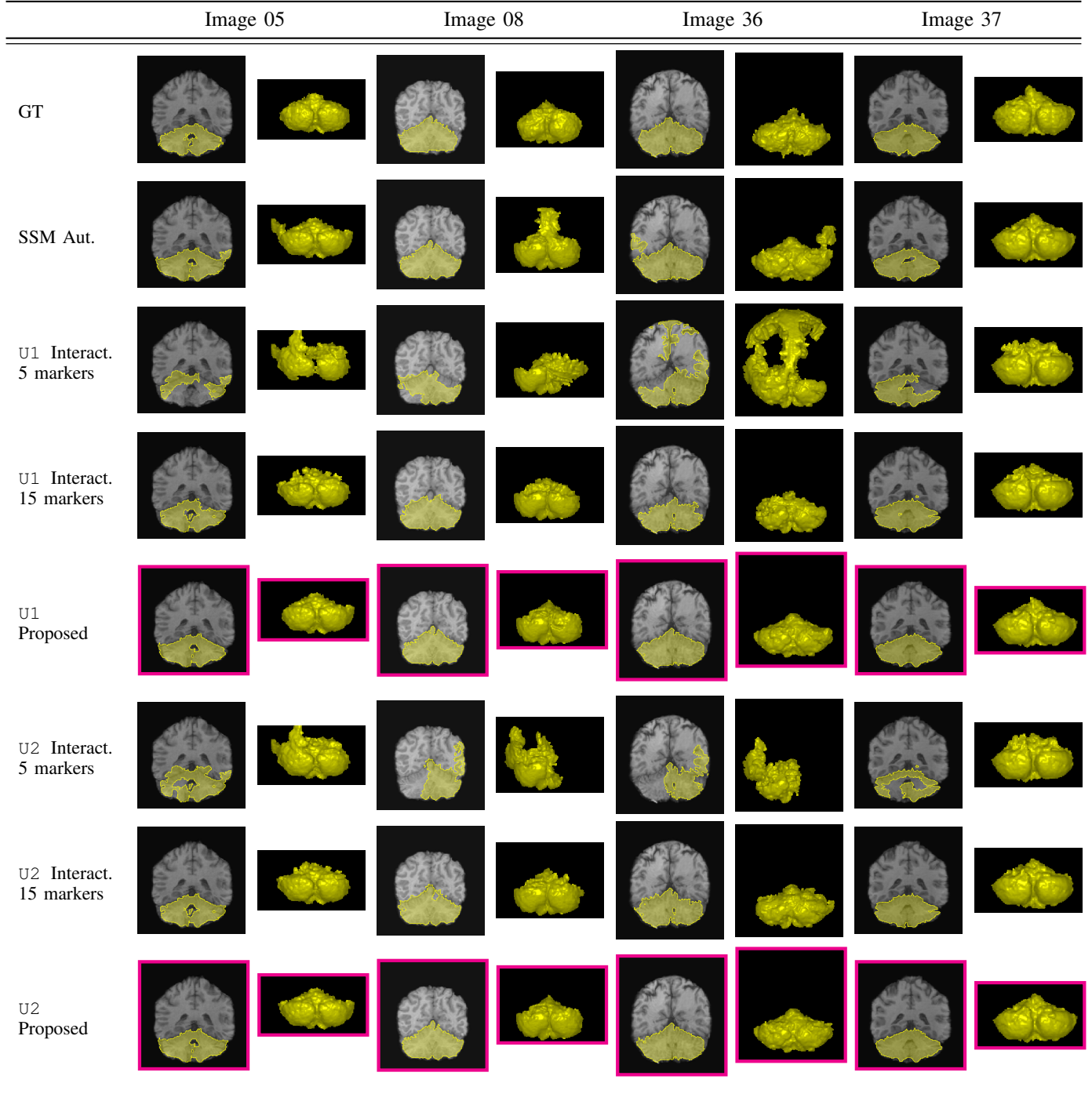
Fig. 5. Segmentation results of the cerebellum for four images from Table I. We depict for each experiment and for both users U1 and U2 an anterior-posterior coronal slice of the original 3D image with the segmentation result overlayed and the corresponding 3D rendition of the result, rotated for better visualization of errors (i.e., depicting the posterior-anterior view). The figures refer to the ground truth images (first row), the automatic result of SSM with errors (second row), the interactive result obtained with 5 and 15 markers (rows 3/4 and 6/7), and the proposed approach with 5 markers for correction (rows 5 and 8 with images surrounded by magenta frames).