

Automatic Landmarking of Magnetic Resonance brain Images

Camille Izard^{*a,b}, Bruno M. Jedynak^{a,b} and Craig E.L. Stark^c

^aDépartement de Mathématiques,
Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France

^bCenter for Imaging Science, Johns Hopkins University, Baltimore, MD

^cDepartment of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD

ABSTRACT

Landmarking MR images is crucial in registering brain structures from different images. It consists in locating the voxel in the image that corresponds to a well-defined point in the anatomy, called the landmark. Example of landmarks are the apex of the head (HoH) of the Hippocampus, the tail and the tip of the splenium of the corpus colosum (SCC). Hand landmarking is tedious and time-consuming. It requires an adequate training. Experimental studies show that the results are dependent on the landmarker and drifting with time.

We propose a generic algorithm performing semi-automated detection of landmarks. The first part consists in learning from a training set of landmarked images the parameters of a probabilistic model, using the EM algorithm. The second part inputs the estimated parameters and a new image, and outputs a voxel as a predicted location for the landmark. The algorithm is demonstrated on the HoH and the SCC. In contrast with competing approaches, the algorithm is generic: it can be used to detect any landmark as soon as a collection of hand-landmarked images is provided for training.

1. INTRODUCTION

An anatomical landmark is a point of the anatomy as well as a location (pixel or voxel) in an image. For one thing, landmark data is useful on its own. It can be processed and statistically analyzed to perform morphometrics.¹ On the other hand, landmarks provide a unique tool to obtain a coarse registration between anatomical shapes, eventually preceding a finer registration technique.²

A motivating application is functional neuroimaging where the central difficulty is that the location, size, and shape of brain structures varies considerably from person to person. Techniques for whole-brain alignment exist, but sacrifice accuracy of regional alignment for global accuracy. Additionally, typical whole brain alignment techniques (e.g., the Talairach transformation or the related MNI transformation) are not designed to respect neuroanatomical boundaries or landmarks and are therefore often quite poor in their alignment (here, we note variability of approximately 10mm following Talairach alignment). The spatial blur introduced by poor alignment of these boundaries or landmarks will result in reduced precision in localization of functional results and reduced statistical power in cross-participant analyzes of functional data.³

Landmarking brain images consists in selecting voxels whose location is characterized by the surrounding anatomy and which are distinguishable in all participant. Examples of such landmarks are the anterior commissure (AC) and the posterior commissure (PC), identified by Tournoux and Talairach⁴ to define a common coordinate system, the head of the hippocampus, the tail of the hippocampus and other remarkable points on structures.

Although it is tough, time-consuming and error-prone, neuroscientists have been landmarking manually brain images. Studies show that the landmarking precision drifts with time and that the accuracy depends on the landmarker. Using automatic methods would improve the current situation and allow large scale quantitative studies.

Point landmarks can be defined more generally, as prominent points in the anatomy, that are distinguishable and sparsely distributed in the image. 3D operators are used to detect that kind of landmarks automatically.^{5,6} The detected points may correspond but not necessarily to manually defined landmarks, used by neuroscientists. Some efforts have been done on specific landmarks, like AC or PC, using scene analysis⁷ or template matching⁸

Further correspondence to Camille Izard, camille@cis.jhu.edu, Bruno Jedynak, bruno.jedynak@jhu.edu, Center for Imaging Science, Clark 301, Johns Hopkins University, 3400N Charles Street, Baltimore, MD 21218-2686, USA.

to automate their detection. More recently, the efforts focused more specifically on tip-like, saddle-like and sphere-like structures landmarks, using deformable template⁹ or 3D parametric intensity model.¹⁰ All these approaches require prior knowledge about the landmark type.

The proposed method is generic, so that it could be applied to any point landmark as soon as a landmarked training set of images can be provided. It will be test on a tip-like landmark, the splenium of the corpus colosum (SCC) and the apex of the head of the hippocampus (HoH). It is anatomically speaking the tip of the hippocampus but the low intensity variations along the structure boundary (Figure 1) make its detection a challenging task, even for trained experts.

At first, the algorithm is learning the geometry of the images using the training set that has been manually landmarked by an expert. We define a probabilistic model of the image intensity, conditionally to the landmark location. The parameters are estimated by the EM-algorithm. The first procedure outputs are a probability map of the brain, conditionally to the sought landmark location and a set of informative locations. Then providing a new image and using the learned parameters, the algorithm predicts the location of the landmark.

Details of the proposed method are described in the section 2, and the method is demonstrated on HoH and SCC in section 3.

2. METHOD

2.1. Description of the dataset

We use 23 MR brain images acquired on a Philips-Intera 3-Tesla scanner, with resolution 1mm^3 , encoded in gray-level intensity from 0 to 1076. Brains were first transformed into standardized Talairach space using Analysis of Functional Neuroimages (AFNI) to provide a canonical orientation (anterior and posterior comisures (AC and PC) made co-linear) and approximate alignment. All the images have the same size after the transformation: $161 \times 191 \times 151$ voxels or mm. Brains were viewed in continuously synchronized sagittal, axial, and coronal planes about a common point in space. One of us, referred in the rest of the paper as "the expert" landmarked the image set. For both the head of the hippocampus and the splenium of the corpus colosum, visualization

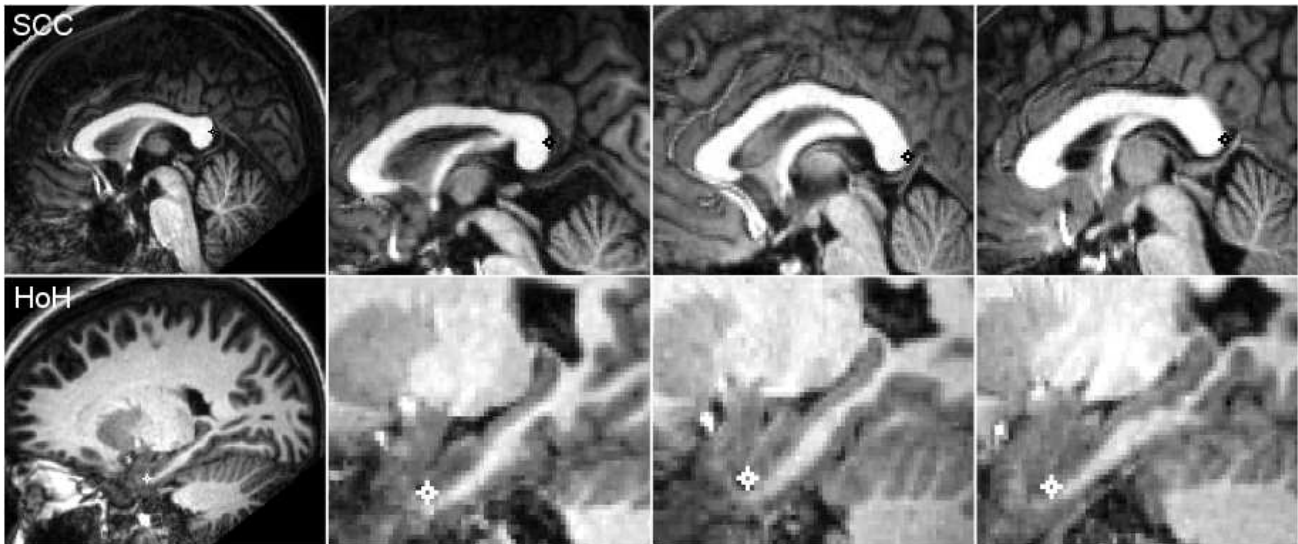


Figure 1. Location of landmarks in different brain images. The central voxel of the cross represents the exact location of the landmark, as defined by the expert. The leftmost images represent a general view of the brain. The other zoom on the landmark for different patients. On the first line, SCC: the posterior extremity of the splenium of the corpus colosum. The shape of the structure varies a lot and even the geometry of the splenium itself varies. On the second line, HoH: the apex of the head of the hippocampus. The angle of the hippocampus, the geometry and the size of the head are patient-dependent.

was best in the sagittal and coronal planes. For the splenium of the corpus colosum (SCC), the mid-sagittal

plane was first identified. The splenium was then marked as the most posterior extent of the corpus collosum. In several instances, the corpus collosum’s most posterior extent did not lie along the mid-sagittal slice, but 1mm to either the left or the right of the midline. In these cases, the splenium was identified as the most posterior extent in this slice, as represented in Figure 1. For the head of the hippocampus (HoH), an imaginary line following the principal axis of the hippocampus was drawn diagonally in the sagittal plane. The head was identified as the furthest extent of this line in three dimensional space (in the anterior and inferior directions). Some examples are represented in Figure 1

Once landmarked the dataset is split in a training set composed of 14 images and a 9-image testing set. The prior location of the landmarks in the Talairach space is evaluated thanks to the instances of the training set.

2.2. Model description

We denote S the set of voxels of an image, x_s is the gray-level at the voxel s , y the location of the sought landmark. We adopt a probabilistic modeling approach. Y , a random variable, is the unknown location of the landmark in the image, $(X_s)_{s \in S}$ is a collection of random variables modeling the voxel values. $A \subset S$ is a subset of selected voxels, we write $x_A = \{x_s, s \in A\}$ the collection of voxel values in the set A . The selection method will be detailed in section 2.6. What should be the cost of landmarking y when the true landmark is at y' ? The answer might depend on the application. In what follows we use the squared Euclidian distance. The bayesian prediction is then the expectation of the landmark location, given the observations at the selected voxels, x_A . The formula is given below and the rest of this section will be devoted to the exploration of the different components.

$$E[Y|X_A = x_A] = \sum_{y \in \mathcal{Y}} y P(Y = y | X_A = x_A) \quad (1)$$

$$= \sum_{y \in \mathcal{Y}} y \frac{P(X_A = x_A | Y = y) P(y)}{\sum_{y' \in \mathcal{Y}} P(X_A = x_A | Y = y') P(y')} \quad (2)$$

(2) is obtained using Bayes’ formula. $P(y)$ is the prior distribution of the landmark. We choose it to be the Uniform distribution over a box $\mathcal{Y} \subset S$, that is guarantied to contain the landmark. We choose a $6 \times 9 \times 10$ -voxel box for the SCC and a $12 \times 12 \times 12$ -voxel box for the HoH. Next, we model the probability to observe x_A given the location of Y , that is $P(X_A = x_A | Y = y)$.

First, we assume that the intensities are independent given the landmark location y . This is a standard simplifying assumption. A more precise model would involve modeling Markov dependencies. The conditional probability $P(X_A = x_A | Y = y)$ is then written as a product:

$$P(X_A = x_A | Y = y) = \prod_{s \in A} P(X_s = x_s | Y = y)$$

Hence, we need to define $P(X_s = x_s | Y = y)$.

The model is based on the gray-level values of the image. Since the gray-level intensity histogram is different from one image to another, the same component in the brain appears at different intensities in two different images. To deal with this difficulty, we introduce for each voxel s , the hidden discrete variable Z_s , corresponding to the matter of the brain at the voxel s . The brain is composed of 3 matters: the CerebroSpinal Fluid (CSF), the Gray Matter (GM) and the White Matter (WM), characterized by distinct intensities, respectively the lowest, the intermediate and the highest ones. Some mixed voxels have intensities that do not correspond to one matter but to a mixture of two matters. Because of the brain configuration only 2 mixtures are observable: CSF-GM and GM-WM. It is important to model them too, since they are present at the boundaries of structures, locations that will turn out to be informative, see section 2.6. Hence, the gray-level intensities are modeled by a 5-component mixture. Using Bayes’ formula,

$$P(X_s = x_s | Y = y) = \sum_{j=1}^5 P(Z_s = j | Y = y) P(X_s = x_s | Z_s = j, Y = y)$$

Second, we assume that the gray-level intensity at a given voxel s , given the matter j , is independent of the landmark location y . This is a mild simplification. Consequently, the second term of the product becomes $P(X_s = x_s | Z_s = j)$ and contains only the information about the gray-level intensities whereas the first term $P(Z_s = j | Y = y)$ characterizes the geometry. Modeling the quantity $P(X_s = x_s | Z_s = j)$ is a fairly standardized problem in photometric modeling of MR images. We use a 5-component mixture of gaussian densities.¹¹ Finally, we model the geometry, using a translation model: the probability to observe the matter j at the voxel s depends only on the translation between s and the landmark y . Hence,

$$P(Z_s = j | Y = y) = \pi_{s-y}(j)$$

This is clearly an over-simplification for the specification of the anatomy of the brain. However it leads to a tractable model which is worth investigating. The conditional probability is then:

$$P(X_A = x_A | Y = y) = \prod_{s \in A} \sum_{j=1}^5 \pi_{s-y}(j) P(X_s = x_s | Z_s = j) \quad (3)$$

We need to learn offline the parameters of this model to apply it online to new images.

2.3. Algorithm description

We describe now the general procedure of the algorithm.

The learning procedure, also called the off-line algorithm, consists in 3 successive tasks:

1. For each image of the training set, estimate the 5×2 parameters of the gray-level gaussian densities
2. Compute the probability map $\pi_{s-y}(j)$: the proportions of the matter j at the voxel s , given that the landmark is at y , which makes a total of 4 parameters per voxel
3. Select the informative voxels in the probability map

The on-line procedure predicts the location of the landmark in a new image. It is made of 2 parts:

1. Estimate the 5×2 parameters of the gray-level gaussian densities in the new image
2. Compute the landmark location, using eq.(2) and eq.(3)

2.4. Learning the gray-level distribution

Modeling the inter-scan intensity inhomogeneities with a mixture of Gaussian densities is a common procedure in MRI segmentation and is usually combined with the EM algorithm to estimate the density parameters.^{12, 13} The 2-step EM-algorithm¹⁴ iteratively alternates between estimation of intensity model parameters and voxel classification. We choose to model the intensities as a 5-component mixture of Gaussian densities, each of them corresponding to one matter (CSF, CSF-GM, GM, GM-WM, WM). We made the assumption that the intensity parameters are constant on the images, i.e. we neglect the bias field produced by the scanner and the patient interactions. Therefore, it is possible to learn the parameters of the mixture in a cube $S_\Omega \subset S$. The cube is chosen centered on AC, since this point has already been detected for the Talairach transformation, independently of the landmark. The size of S_Ω is $41 \times 41 \times 41$ voxels.

The EM algorithm consists in maximizing the likelihood over the parameters of the model, considering the matter z_s as missing data. The E-step computes the posterior probability, given the estimated parameters of the previous step $(\alpha_j^{(n)}, \mu_j^{(n)}, \sigma_j^{2(n)})$:

$$P^{(n+1)}(Z_s = j | X_s = x_s) = \frac{\alpha_j^{(n)} N_j^{(n)}(x_s)}{\sum_{j'=1}^5 \alpha_{j'}^{(n)} N_{j'}^{(n)}(x_s)} \text{ with } N_j^{(n)}(x_s) = \frac{1}{\sqrt{2\pi\sigma_j^{2(n)}}} \exp\left(-\frac{|x_s - \mu_j^{(n)}|^2}{2\sigma_j^{2(n)}}\right)$$

The M-step maximizes over $(\alpha_j^{(n+1)}, \mu_j^{(n+1)}, \sigma_j^{2(n+1)})$, $j \in \{1, \dots, 5\}$ the log-likelihood:

$$\sum_{s \in S_\Omega} \log \sum_{j=1}^5 \alpha_j N_j(x_s) \quad (4)$$

leading to:

$$\begin{aligned} \alpha_j^{(n+1)} &= \frac{\sum_{s \in S_\Omega} P^{(n)}(Z_s = j | X_s = x_s)}{|S_\Omega|} \\ \mu_j^{(n+1)} &= \frac{\sum_{s \in S_\Omega} P^{(n)}(Z_s = j | X_s = x_s) x_s}{\sum_{s' \in S_\Omega} P^{(n)}(Z_{s'} = j | X_{s'} = x_{s'})} \\ \sigma_j^{2(n+1)} &= \frac{\sum_{s \in S_\Omega} |x_s - \mu_j^{(n)}|^2 P^{(n)}(Z_s = j | X_s = x_s)}{\sum_{s' \in S_\Omega} P^{(n)}(Z_{s'} = j | X_{s'} = x_{s'})} \end{aligned}$$

Since the EM algorithm is sensitive to the initialization, the estimation is performed 3 times with random initialization. We keep the parameters which give the best approximation, measuring the χ^2 -distance between the estimated histogram and the observations. The user may have to validate visually the fitted Gaussian

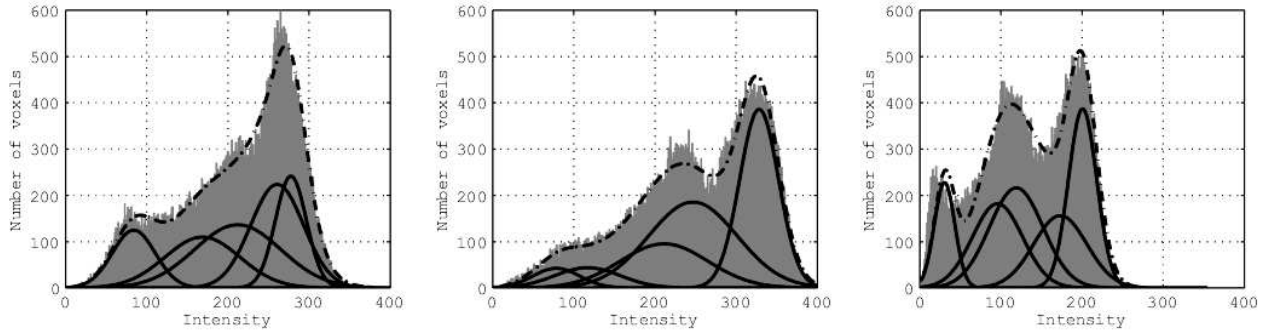


Figure 2. Examples of Gaussian mixture parameters estimation on the 41-by-41-by-41 voxel box around AC. The plain lines are the 5 estimated Gaussian densities and the dot line represents the estimated histogram.

densities. Figure 2 gives some examples of intensity histograms and fitted Gaussian densities. Some images are more difficult to model with a Gaussian mixture, because of the intensity peaks in the histogram and because there are no negative intensities. However, since we just need a probabilistic classification of the voxels, we are satisfied even if the model is imperfect.

2.5. Learning the geometry

Learning the geometry consists in estimating the probability $\pi_{s-y}(j)$ to observe at the location $s-y$ the matter j . We firstly superimpose the images i , applying a translation such that the landmark sits at the same voxel for each image. Then we use the EM-algorithm transversally to estimate the proportions of each matter j in the mixture observed at a given voxel s . The resulting image is called the probability map.

Here the location of s is fixed and the gray-levels are observed in the different images, using the estimated parameters of the gaussian densities from the section 2.4. The E-step computes the posterior probability, given the estimated parameters $\pi_{(s-y)}^{(n)}(j)$ estimated at the previous step.

$$P_i^{(n+1)}(Z_s = j | X_s = x_s^{(i)}) = \frac{\pi_{(s-y)}^{(n)}(j) N_{ij}(x_s^{(i)})}{\sum_{j'=1}^5 \pi_{(s-y)}^{(n)}(j') N_{ij'}(x_s^{(i)})} \text{ with } N_{ij}(x_s^{(i)}) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left(-\frac{|x_s^{(i)} - \mu_{ij}|^2}{2\sigma_{ij}^2}\right)$$

The M-step maximizes over $\pi_{(s-y)}(j), j \in \{1, \dots, 5\}$ the proportions of the mixture of matters at the fixed voxel s :

$$\sum_i \log \sum_{j=1}^5 \frac{\pi_{s-y}(j)}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left(-\frac{|x_s^{(i)} - \mu_{ij}|^2}{2\sigma_{ij}^2}\right) \quad (5)$$

leading to:

$$\pi_{s-y}(j)^{(n+1)} = \frac{\sum_{i=1}^{14} P_i^{(n+1)}(Z_s = j | X_s = x_s^{(i)})}{14}$$

The whole probability map is obtained, by running one EM-algorithm per voxel.

Figure 3 shows slices of the probability maps obtained when the training images are superimposed on the SCC

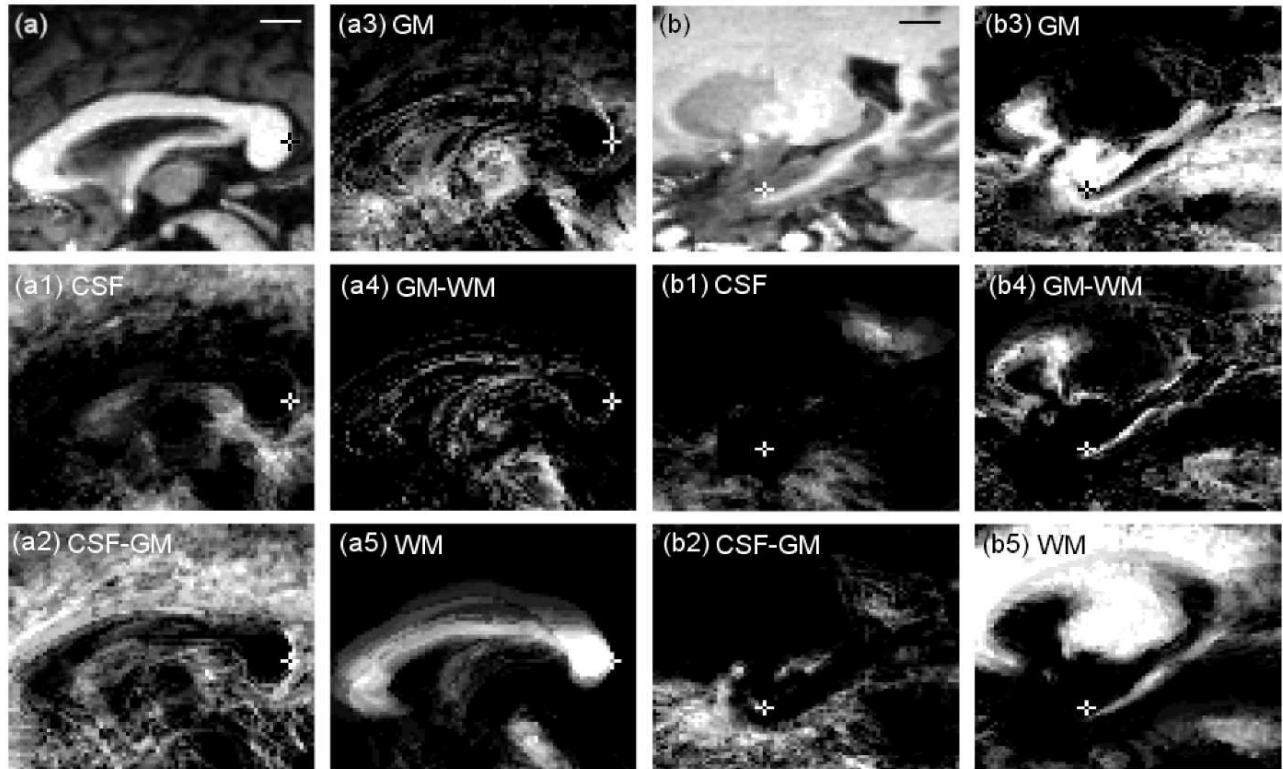


Figure 3. Slices of the probability maps obtained when the images are superimposed on the landmark SCC **(a)** or on the landmark HoH **(b)**. The cross represents the location of the landmark. The images (a) and (b) are examples of the brain structures corresponding to the part of the probability map shown in the images a1-a5 and b1-b5. The images from 1 to 5 represent both for (a) and (b) the probability $\pi_{s-y}(j)$ given the location of the landmark y of respectively the CSF, the CSF-GM, the GM, the GM-WM, the WM. The white and black lines are 10mm long.

(a) or on the HoH (b). The images 1 to 5 represent the different matters: respectively CSF, CSF-GM, GM, GM-WM, WM.

For example, in (a5), one visualizes the probability that a voxel taken from a random image, registered at the SCC, would correspond to white matter (WM). Notice that some structures appear with sharp boundaries, often close to the landmark. In the case of the SCC probability map, the corpus collosum is identifiable. Along the shape the contour evaluates to be sharper near the landmark. One expects it, since the images are registered at this point.

We refer the reader to the Figure 4 to understand the details of the anatomy around the HoH, which are observable in the probability map Figure 3. As for the HoH probability map, the hippocampus appears in the gray matter with a high probability, surrounded by an elongated structure of gray matter below it. This is the

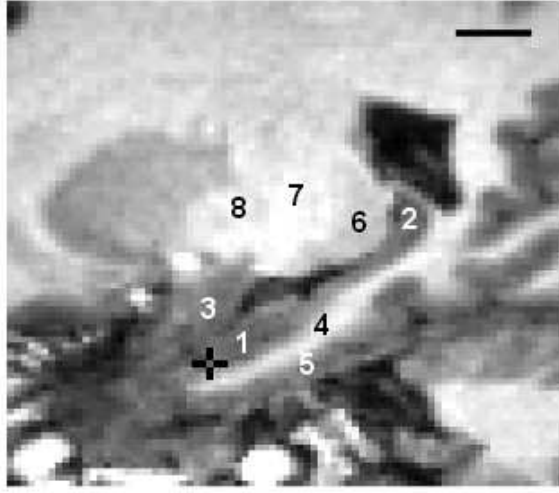


Figure 4. Details of the anatomy around the head of the hippocampus. The black bar gives the scale: 10mm. The cross represents the HoH. 1, head of the hippocampus; 2, tail of the hippocampus; 3, amygdala; 4, white matter of the parahippocampal gyrus; 5, gray matter of the parahippocampal gyrus; 6, pulvinar; 7, posterior limb; 8, globus pallidus.

gray matter of the parahippocampal gyrus. The white matter of this structure appear clearly in the image b5 of the Figure 3. In addition on this same image, one notices the large white structure composed of the posterior limb, the pulvinar and partially the globus pallidus.¹⁵

2.6. Selecting the informative voxels

The probability map obtained as demonstrated in section 2.4, is almost as large as the initial images: $161 \times 191 \times 151$ voxels. Computing the conditional expectation with all the points of the probability map is untractable and inefficient. Indeed most of the voxels contain no information about the landmark location. For example, the points included in large structures with constant color are useless to locate the landmark. In contrast, contours of brain structures should intuitively be informative.

We choose to measure the amount of information contained in a voxel by the expected variance reduction of the landmark location, Y , when the gray-level is observed at the voxel s :

$$E_{Y, X_s} \|Y - E_Y(Y|X_s = x_s)\|_2^2 \quad (6)$$

This quantity ranks the voxels and allows us to select only the most informative ones. However, because of the gray-level variations from one image to another one cannot easily compute (6). Consequently we choose to consider the expected reduction of variance when the matter at the voxel s is given, that is:

$$E_{Y, Z_s} \|Y - E_Y(Y|Z_s = j)\|_2^2 \quad (7)$$

One can prove that (6) and (7) would be equivalent, leading to the selection of the same voxels, if the 5 densities $P(Z_s = j|X_s = x_s)$ had disjoint supports.

The information at a voxel s , defined by (7) depends only on the prior distribution and the probability map. Using the definition of the variance and the model of the matters in the image, the information at the voxel s can be expanded as follows:

$$\begin{aligned} E_{Y, Z_s} \|Y - E_Y(Y|Z_s = j)\|_2^2 &= \sum_{k=1}^3 \sum_{j=1}^5 [E_{Y_k}(Y_k^2|Z_s = j) - E_{Y_k}(Y_k|Z_s = j)^2] \\ &= \sum_{k=1}^3 \sum_{j=1}^5 \left[\sum_{y_k \in \mathcal{Y}_k} y_k^2 \frac{\pi_{s-y_k}(j)}{\sum_{y_j'' \in \mathcal{Y}_k} \pi_{s-y_j''}(j)} - \left(\sum_{y_k \in \mathcal{Y}_k} y_k \frac{\pi_{s-y_k}(j)}{\sum_{y_j'' \in \mathcal{Y}_k} \pi_{s-y_j''}(j)} \right)^2 \right] \end{aligned}$$

where $Y = (Y_1 \ Y_2 \ Y_3)^T$ and $y = (y_1 \ y_2 \ y_3)^T$, and $\mathcal{Y} = \mathcal{Y}_1 \otimes \mathcal{Y}_2 \otimes \mathcal{Y}_3$. The lower the expected variance is, the more informative are the corresponding voxels. We called the resulting image the information map.

To understand the link between the informative points and the landmark we use synthetic images on which we compute the information map. In order to simplify we consider only three matters corresponding to the three major matters of the brain (CSF, GM, WM). A set of synthetic data is created using the gray-level parameters estimated on the images of the training set. The synthetic images are composed of a sphere A rigidly related to the landmark location, and of a sphere B whose center location is fixed, independently of the landmark location. The radii of both the small sphere and the large sphere are constant. We compute the probability map of the superimposed images and the information map. Figure 5 shows the result of the experiment. The probability

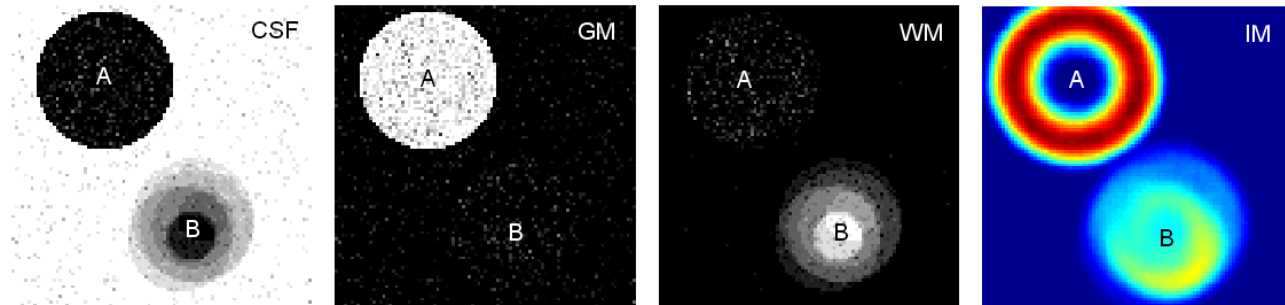


Figure 5. Probability map and information map obtained on synthetic images. The images are composed of three matters, the sphere A in GM, the sphere B in WM and the background in CSF. The sphere A is rigidly related to the sought landmark. The center of B is independent of the landmark. CSF, GM, WM are the three components of the probability map which are visualized in the three leftmost images. The red voxels in the information map (IM), in the rightmost image, are the most informative to locate the landmark, the blue are the least informative.

map, Figure 5 (CSF, GM, WM), presents a sharp boundary for the sphere A whose center is fixed with respect to the landmark location, whereas the sphere B has a blurred boundary. The information map, Figure 5 rightmost image, shows that the boundary of the sphere A is the location of the informative voxels, in red, as intuition suggests.

We now compute the information map for the HoH and SCC. Figure 6 presents one slice of the result. The algo-

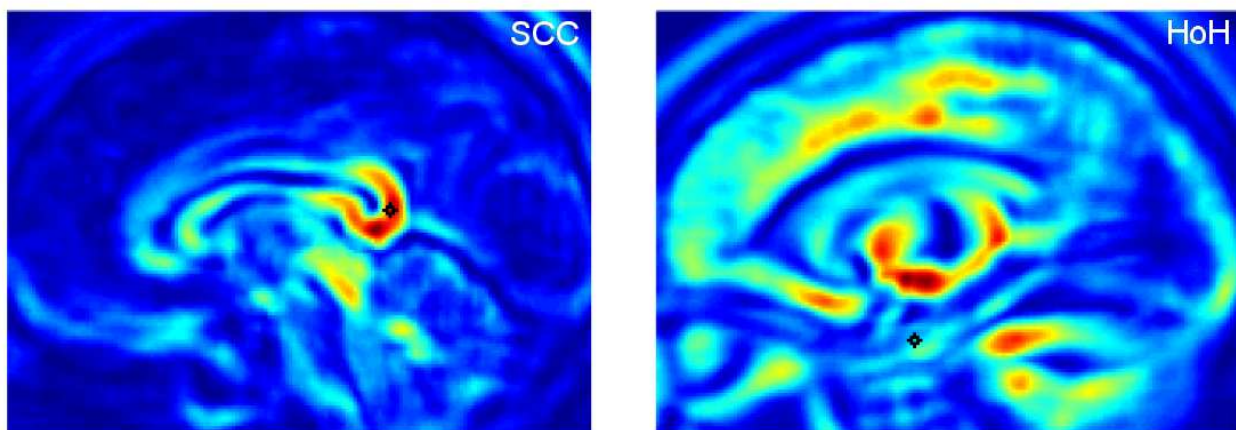


Figure 6. Information map obtained from the probability map of SCC on the left and HoH on the right. The landmark is represented by the cross. The most informative voxels are in red.

gorithm selects automatically the informative locations in the probability map. For the SCC, the most informative

voxels are located on the boundary of the splenium itself. In this case, the algorithm selects automatically the voxel located on the "corner". The current best method¹⁰ to detect tip-like landmark uses an intensity model for corner, fitted on the shape. The selection creates here automatically a tip-like probabilistic template.

The results are more difficult to interpret for the HoH. The informative voxels are located on a boundary between gray matter and white matter. This shape appeared already clearly in the probability map, Figure 3. On the other hand, one notices that the voxels corresponding to the parahippocampal gyrus are not selected even if sharp white and gray structures appeared in the probability map, Figure 3. We believe that this is due to the thinness of the structure. Indeed the computation of the information is a spatial average on a box as large as the prior box \mathcal{Y} . Because of the surrounding constant matter, the expected information of the voxels on the hippocampal gyrus is erased. The following experiment confirms this intuition: the information map is computed with decreasing prior box, so that the spatial average is more localized. While the box size decreases, the information of the voxels of the parahippocampal gyrus increases.

The selection step reduces drastically the computation, but there is a risk to lose some informative voxels too. Once ranked, we picked only the 100, 500, 2000 and 4500 first points, corresponding to less than one per thousand voxels. In Figure 3 only the red voxels are selected.

2.7. Predicting the location of the landmark in a new image

The first part of the online algorithm is identical to the gray-level estimation detailed in section 2.4. Then it consists in computing the conditional expectation of Y , as detailed in the equations (2) and (3).

3. EXPERIMENTAL RESULTS

In this section we firstly evaluate the human performance in landmarking brain images. Then we detailed the experimental results obtained on the two landmarks SCC and HoH.

3.1. Manual performance in detecting landmarks

Landmarking manually MR brain images is a tough task. Results vary depending on the landmarker and the experiment. To evaluate his consistency, the expert landmarked a first time SCC and HoH on all the images. Then several weeks later, he landmarked a second time in similar conditions 15 among a total of 23 images. For SCC, the mean Euclidian distance between the landmarks is 0.71mm with standard deviation of 0.61mm. For HoH, the mean distance is 1.22mm and the standard deviation 0.92mm. We consider these results as gold standard for the considered scans.

A second experiment has been performed on the HoH. The expert taught a graduate student to landmark this point. As the expert did, the student landmarked a first time the images, and after several weeks relandmarked them. The mean Euclidian distance between his two landmarking is 3.58mm with standard variation of 0.98mm. In addition we measure the Euclidian distance between student's and expert's landmarks. The mean distance is 3.26mm with standard deviation of 0.98mm. Even if we think that the student would eventually achieve the expert performance with additional training, this experiment shows how difficult it is to landmark consistently the images and to teach another person to do the same.

3.2. Algorithm results

We demonstrate the algorithm for the detection of SCC and of HoH, using 14 images in the training set, and 9 images for the test.

The estimation of the gray-level densities is common to the two experiments. We performed it only once and used the same estimated parameters. The algorithm is implemented in Matlab and runned with a Xeon 2.2GHz processor. The first EM-algorithm takes 10 seconds per images, when the box size is $41 \times 41 \times 41$. However, because of its sensitivity to the initialization, the EM-algorithm has to be runned several times on each image, in order to define the best parameters.

Learning the geometry, consists in learning $161 \times 191 \times 151 \times 4$ parameters, using 14 values at each location of the probability map. It takes around 30 minutes to compute it completely.

Finally the last step of the off-line algorithm selects the informative voxels. The computation takes around 20 minutes.

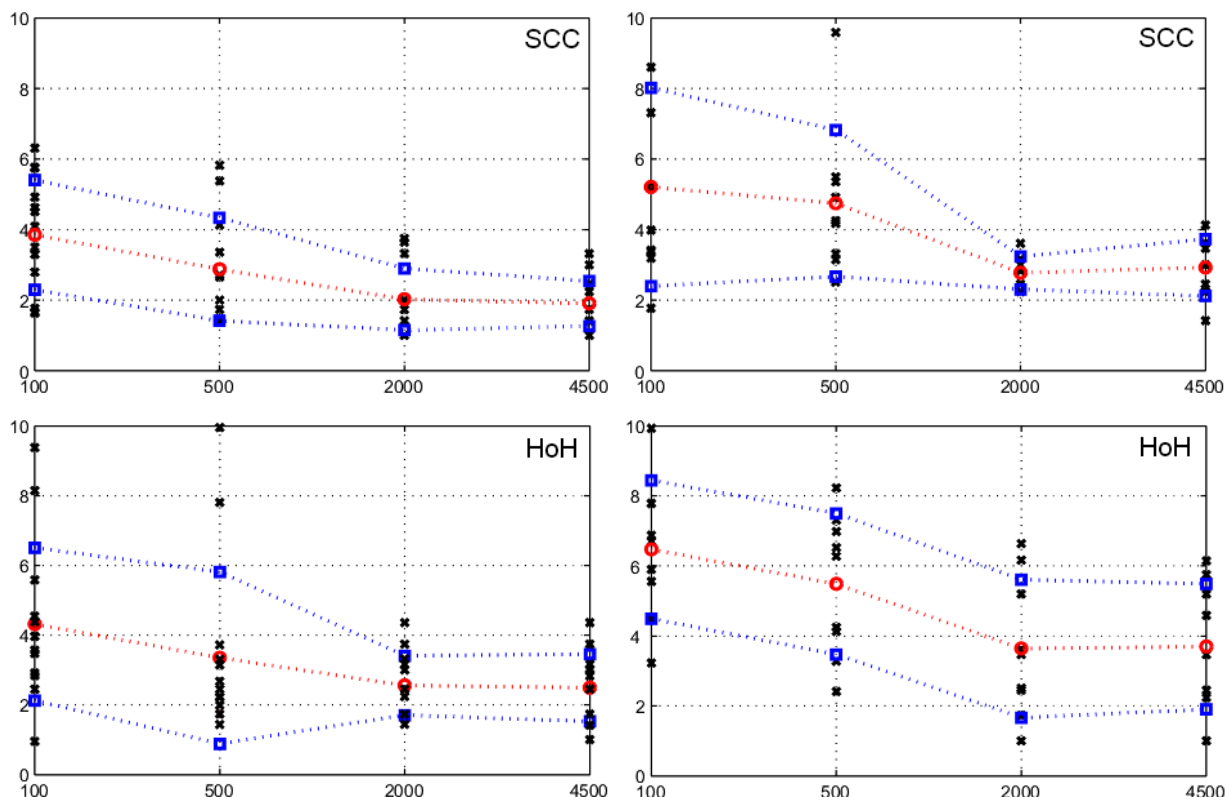


Figure 7. Prediction error when the number of selected informative voxels increases. X-axis: number of points used, Y-axis: Euclidian distance between the prediction and the ground-truth in mm. Each cross corresponds to one image, the dot line with circle marker represents the mean error and the dot lines with square markers represent the mean \pm one standard deviation. On the left: results on the training set, On the right: results on the testing set. First line: SCC, Second line: HoH.

The online algorithm starts with the estimation of the gray-level densities and computes the conditional expectation $E[Y|X_A = x_A]$. Experimental results are obtained for increasing number of voxels in the set A : 100, 500, 2000 and 4500. The Euclidian distance between the prediction and the expert landmark measures the prediction error. Figure 7 presents the evolution of the mean error when the number of selected informative voxels increases.

Table 1. Mean prediction error obtained with our algorithm for SCC and HoH using the training set (14 images) and the testion set (9 images).

Landmark	prior box (in voxels)	initial mean error	mean error (training set)	mean error (testing set)
SCC	$6 \times 9 \times 10$	6.01mm	1.90mm ($\sigma = 0.64$ mm)	2.92mm ($\sigma = 0.81$ mm)
HoH	$12 \times 12 \times 12$	8.48mm	2.49mm ($\sigma = 0.96$ mm)	3.70mm ($\sigma = 1.79$ mm)

As shown in Figure 7, the mean error of the prediction decreases when the number of selected voxels increases. It is observable for both the training set and the testing set, for the SCC prediction as well as for the HoH prediction. However the improvement is stabilized and it seems useless to add more points. The performances of the automatic landmarking algorithm on the test are very close to the ones of the trained student, but still not comparable with the expert's ones. The computing time is about 1 minute per image.

4. CONCLUSION

The proposed method is a generic algorithm to detect a landmark in MR images. It is demonstrated on the HoH, the head of the hippocampus, and the SCC, the splenium of the corpus colosum. We were unable to find similar generic approach, since the competing approaches use geometrical properties of the landmark and hence need to be tailored for each type of landmark. Our algorithm is able to locate the informative points in the image and design automatically a probabilistic template used on new images. It gives a tentative location, comparable to the work of a non-expert landmarker. The expert could use it as an automatic way to provide an initial guess.

ACKNOWLEDGEMENTS

The authors thank Laurent Younes for his helpfull sugestions. This research was partially supported by ARO DAAD19/-02-1-0337 and general funds from the Center for Imaging Science at The Johns Hopkins University.

REFERENCES

1. F. Bookstein, *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, February 1992.
2. U. Grenander and M. Miller, "Computational anatomy: An emerging discipline," in *Quarterly of Applied Mathematics*, pp. 617–694, LVI, December 1998.
3. C. Stark and Y. Okado, "Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition," *J. of Neurosci.* **23**, pp. 6748–6753, 2003.
4. J. Talairach and P. Tournoux, *Co-planar stereotaxic Atlas of the Human Brain*, Thieme Medical Publishers, 1988.
5. J.-P. Thirion, "New feature points based on geometric invariants for 3d image registration," *Int. J. of Computer Vision* **18:2**, pp. 121–137, 1996.
6. K. Rohr, "On 3d differential operators for detecting point landmarks," *Image and Vision Computing* **15:3**, pp. 219–233, 1997.
7. L. Vérard, P. Allain, J.-C. B. Jean-Marcel Travère, and D. Bloyet, "Fully automatic identification of ac and pc landmarks on brain using scene analysis," *IEEE Trans. Med. Imag.* **16**, October 1997.
8. Y. Han and H. W. Park, "Automatic brain mr image registration based on talairach reference system," in *Proc. ICIP'2003 International Conference on Image Processing, I*, pp. 1097–1100, 2003.
9. S. Frantz, K. Rohr, and H. Stiehl, "Localization of 3d anatomical point landmarks in 3d tomographic images using deformable models," in *Proc. MICCAI'2000, Pittsburg, Pennsylvania/USA, Oct. 11-14, 2000*, S. Delp, A. DiGioia, and B. Jaramaz, eds., *Lecture Notes in Computer Science* **1935**, pp. 492–501, Springer-Verlag, (Berlin Heidelberg), 2000.
10. S. Wörz and K. Rohr, "3d parametric intensity models for the localization of different types of 3d anatomical point landmarks in tomographic images," in *Proc. 25rd Symposium Pattern Recognition (DAGM'2003), Sept. 10.-12., 2003, Magdeburg/Germany*, B. Michaelis and G. Krell, eds., *Lecture Notes in Computer Science* **2781**, pp. 220–227, (Springer Berlin Heidelberg), 2003.
11. J. Ratnanather, K. Botteron, T. Nishino, A. Massie, R. Lal, S. Patel, S. Peddi, R. Todd, and M. Miller, "Validating cortical surface analysis of medial prefrontal cortex," *NeuroImage* **14**, pp. 1058–1069, 2001.
12. K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based bias field correction of mr images of the brain," in *IEEE Trans. Med. Imag.*, **18**, pp. 885–896, Oct. 1999.
13. W. Wells, R. Kikinis, W. Grimson, and F. Jolesz, "Adaptive segmentation of mri data," in *IEEE Trans. Med. Imag.*, **15**, pp. 429–442, 1996.
14. A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *J. R. Stat. Soc.* **39**, pp. 1–38, 1977.
15. H. M. Duvernoy, *The Human Hippocampus: Functional Anatomy, Vascularization and Serial Sections with MRI*, Springer, 2nd ed., 1998.