Automatic Landmarking of Magnetic 
Resonance brain Images

Bruno M. Jedynak,* Camille Izard*and
Craig Stark †

August 10, 2004

Presentation Preference: Oral presentation

Bruno Jedynak  received his doctorate in Statistics and Stochastic Models
from the Université Paris Sud, in France, in February 1996. His disserta-
tion was performed at INRIA under the supervision of Professor Donald
Geman. He then spent a year as post-doc in the Department of Statistics
at the University of Chicago. Dr. Jedynak was appointed Professor at the
Université Scientifique et Technique de Lille in France in September 1997.
He is currently a Visiting Associate Professor in the Department of Applied
Mathematics and Statistics and is working at the Center for Imaging Science
within the Johns Hopkins University. His research interests are focused in
statistical modeling at large. He has specialized in computer vision. Appli-
cations include road, face, and human skin detection. He is recently involved
in langage modeling and biomedical imaging.

*Département de Mathématiques, Université des Sciences et Technologies de Lille,
France, and Center for Imaging Science, The Johns Hopkins University. Mailing Adress :
Clark 301, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD, 21218-
2686. Tel : (410) 516-7734 Fax : (410) 516-4594 Emails: bruno.jedynak@jhu.edu and
camille@cis.jhu.edu web page : http://www.cis.jhu.edu/~bruno/
†Department of Psychological and Brain Sciences, The Johns Hopkins University. Mail-
ing Adress : Ames 135, The Johns Hopkins University, 3400 N. Charles Street, Baltimore,
MD, 21218-2686. Tel : (410) 516-7813 Fax : (410) 516-4478 Email :cstark@jhu.edu
Abstract:

An anatomical landmark in the brain is a well defined point of the anatomy of the brain. Locating a landmark in a Magnetic Resonance brain Image, or landmarking, consists in selecting a particular voxel in the image, corresponding to the anatomical landmark in the imaged brain. This voxel, like an anchor, is a precious piece of information for performing measurements and registration of brain structures. The most well known landmarks in the brain are the anterior commissure (AC) and the posterior commissure (PC) used to define the Talairach proportional coordinate system. Others, based on brain structures, are the “apex of the Head of the hippocampus” (HoH) and the “Tail of the hippocampus”. We will be more specifically interested in these. Landmarking can be a tedious manual procedure, expensive and time consuming. It might be erroneous, difficult to assess, dependent on the scanner and on the landmarker. We present in this paper a generic algorithm that permits to partially automate the landmarking process. The algorithm has two components. One is an off-line procedure, the other is on-line. The former is a system that estimates the parameters of a probabilistic model from a training set of landmarked images using the Estimation Maximization (EM) algorithm. The later inputs an image as well as the parameters previously estimated and outputs a tentative location for the landmark as well as a covariance metric that assesses the remaining uncertainty. The selected location can then be validated or corrected manually. The system will be demonstrated using the HoH. The images are acquired on a Philips Intera 3-Tesla scanner with resolution 1mm³. The training set is made of 14 images. The mean Euclidian distance between the HoH and the prediction is less than 3.5mm as measured on a disjoint set of images. This value is shown to be comparable to the Euclidian distance between two hand-landmarking of the same images.

Description of purpose:

A central difficulty in functional neuroimaging studies 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 in their alignment (here,
we note variability of approximately 10 mm 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 analyses of functional data [5].

Identification of landmarks within the medial temporal lobe (e.g., the apex of the head of the hippocampus) will facilitate the process of aligning multiple participant’s brains while respecting neuroanatomical structure as these landmarks can be used to guide initial alignment.

Methods:
Let’s notate $S$ the set of voxels of an MRI, $x_s$ the gray-level at voxel $s$ and $z_s$ the type of matter at voxel $s$, which can be Cerebrospinal fluid (CSF), grey matter (GM) or white matter (WM). Finally, let’s notate $y$ the voxel in $S$ corresponding to the landmark. The probabilistic model describing the landmark together with the entire set of gray-levels in the image is

$$p(y, x) = p(y) \prod_{s \in S} \sum_{j=1}^{3} \pi_{s-y}(j) g_{\theta_j}(x_s)$$

$p(y)$ is a prior distribution on the landmark location. In the case of the right HoH, we have chosen a Uniform distribution over a cube of size $11 \times 12 \times 12 \text{mm}^3$. The center of the cube is estimated using training data. More generally, the prior could be provided by a probabilistic atlas of the brain. The three densities $g_{\theta_j}$, for $1 \leq j \leq 3$ are Gaussian. They model the gray-levels observed for each matter type $j$. Their means and variances are image dependent. Finally, the family of discrete probabilities $\pi_{s-y}$ describe the probability that one will find a particular matter (CSF or GM or WM) at voxel $s$ given that the landmark is at location $y$. It is a crucial simplifying assumption to consider that this probability depends on $y$ only through the displacement $s - y$.

The off-line algorithm consists in three steps. During the first step, each training image is manually registered in Talairach proportional coordinate system. A box of dimension $40^3 \text{mm}^3$, centered at the anterior commissure is extracted and the parameters of the three Gaussian densities are estimated using the standard EM algorithm, see figure 1. This might require a human validation.

In the second step, the training images are aligned by translation such that the landmarks of the different images sit at the same voxel, say $y_0$. 

3
Figure 1: The EM algorithm is used to fit a mixture of three Gaussian densities to the gray-level histogram of an MRI image. The x-axis is the gray-level. The y-axis is the number of voxels with this gray-level.

Then, an EM algorithm is performed at each voxel to estimate by maximum likelihood the values $\pi_{s-y_0}(j)$, for $1 \leq j \leq 3$.

![Brain images](image)

Figure 2: **Left:** A sagittal slice of the brain. The Head of the Hippocampus (HoH) is shown in red. **Center:** Probability for matter type in false color. Red channel : CSF, Green channel : GM, Blue channel : WM. **Right:** in blue : most informative voxels. in red : least informative voxels. The scale corresponds to the value (2) in mm.

Finally, in a third step, a subset $A \subset S$ of informative voxels is selected. This is done by ranking the voxels according to the trace of the expected conditional variance of the location of the landmark given the matter type at $s$

$$E \| Y - E(Y|Z_s) \|^2$$

and then selecting sequentially the locations $s$ for which this quantity is the smallest. See figure 2.
The on-line algorithm consists also in three steps. Given a new image, the first step is the same as for the off-line algorithm. Then, in the second step, using the stored values $\pi_{s \rightarrow t_0}$ as well as the set $A$, the algorithm computes, according to (1), the conditional or posterior mean

$$E(Y|x_A) \text{ where } x_A = \{x_s; s \in A\}$$

and $x_s$ are the observed gray-levels of the image. The value (3) is then proposed as a tentative location for the landmark. The third step consists in computing the remaining uncertainty as the $3 \times 3$ covariance metric of $Y$ given $x_A$ according (1) as before.

**Results:**

Human performances in HoH landmarking are difficult to measure accurately. One of us, refereed as the “expert” in what follows has landmarked the HoH in 23 images of 23 different patients acquired with the same scanner. The same procedure was repeated several weeks after for a subset of 5 images. The Euclidian distance between the measurements made the first time and the second has mean 0.93mm and standard deviation (std) 0.96mm. This is, we believe, a gold standard for these scans. In a second step, we have taught a graduate student to landmark the HoH. As for the expert, the same images have been landmarked twice, leading to a mean distance of 3.92mm ($std = 1.98mm$). Finally, we have measured the distance between the landmarks of the students and those of the expert. The mean distance is 3.26mm ($std = 0.98mm$). Even if we believe that with further training the student would have, eventually, reached the performances of the expert, this experiment shows how the HoH landmarking task is very challenging.

We describe now the experiments performed with the automatic system. The training, consisting of running the off-line algorithm, was made with 14 out of the 23 landmarked images. Then, testing, consisting in running the on-line algorithm, was performed for each image. The results, measured with Euclidian distance between the expert landmark and the prediction output by the algorithm is presented in Figure 3.

The mean error decreases when the number of voxels used to do the prediction increases. According to the training set results, the error is already stabilized
Figure 3: Euclidian distance in $mm$ between expert landmark and automatic prediction for each image. A bar corresponds to a single image. First group : 186 voxels are selected during training (off-line algorithm). Second group : 2230. Third group : 6000. Left : results for the images used during training. Right : Results for new images.

with 2230 voxels. The best accuracy achieved on the new images is $3.49 mm$ (std=1.93). This error is comparable with the error made by the student in our experimental study. We believe that increasing the size of the training set will permit to increase the performances on never-seen images by allowing the system to capture more of the geometric variability of the anatomical structures nearby the landmark.

**New or breakthrough work to be presented:**

In the computer vision literature, landmarks are usually defined as a noticeable point in an image [3],[2]. We proceed differently and define a landmark as “a mark in the land” that is a well defined point of the anatomy as in [4]. In [6] and [1], methods are presented to automatically locate AC and PC. Our work differ from these in several aspects. First, we are primarily concerned with internal landmarks chosen to identify fine structures of the brain. These landmarks are much more difficult to locate than AC and PC. Secondly, we present a generic system that is not dedicated to a specific landmark. Our system specializes to a particular landmark using a training set of landmarked images. We were unable to find in the relevant literature a comparable design.
Conclusions:
We have presented a method for automating the landmarking of the apex of the head of the Hippocampus in Magnetic Resonance brain Images. First results, achieved with a small training set, are comparable with “student performances” and hence encouraging. The algorithms developed are generic and can be used for other landmarks when provided with an appropriate training set of landmarked images.

Acknowledgments:
The authors thank Laurent Younes for his helpful suggestions. 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.

This work is not and has not been submitted for publication elsewhere.

References


