Spline-Based Probabilistic Model for Anatomical Landmark Detection

Camille Izard[1,2]⋆, Bruno Jedynak[1,2], and Craig E.L. Stark[3]

[1] Laboratoire Paul Painlevé, Université des Sciences et Technologies de Lille, France
[2] Center for Imaging Science, Johns Hopkins University, Baltimore, MD
[3] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD
{camille.izard,bruno.jedynak,ctark}@jhu.edu

Abstract. In medical imaging, finding landmarks that provide biologically meaningful correspondences is often a challenging and time-consuming manual task. In this paper we propose a generic and simple algorithm for landmarking non-cortical brain structures automatically. We use a probabilistic model of the image intensities based on the deformation of a tissue probability map, learned from a training set of hand-landmarked images. In this setting, estimating the location of the landmarks in a new image is equivalent to finding, by likelihood maximization, the "best" deformation from the tissue probability map to the image. The resulting algorithm is able to handle arbitrary types and numbers of landmarks. We demonstrate our algorithm on the detection of 3 landmarks of the hippocampus in brain MR images.

1 Introduction

Anatomical landmarks are well-defined points in the anatomy that experts use to establish biologically meaningful correspondences between structures [1]. Such correspondences are commonly used by registration algorithms, as initialization and/or as constraints [2,3,4]. Landmarks also provide a local shape description useful for anatomical shape comparison [5].

However, locating landmarks on biological structures is a challenging and time-consuming task, even for experts. This has motivated the development of several methods for automatic landmarking. Previously proposed methods use either 3D filters to detect, for example, high curvature points and corners in the image [6], or a geometric model of the image intensities to detect, for example, the tip of a structure [7,8]. Both techniques rely on local intensity variations, but some landmarks are not detectable using only intensity information. This is the case of the head of the hippocampus whose intensity is similar to the surrounding amygdala. Notice also that these techniques are designed for the detection of independent landmarks. They cannot integrate information from the previous detections, although it should be helpful to know, for example, the location of the head of the hippocampus to locate its tail.

In this paper, we propose a probabilistic approach[4] to landmark brain structures such as the hippocampus, on which one needs to locate the head, the tail and the extremity of the hippocampal uncus. We identify the location of the landmarks in an image with a unique deformation of the underlying 3D space. The set of deformations considered in this paper is a set of Gaussian interpolating splines where the landmarks act as control points. We then build a probabilistic model for an image, given the landmark locations. The estimation of the model parameters consists of learning a local tissue probability map, using a training set of hand-landmarked images. In new images, landmarks are identified using a gradient ascent algorithm on the likelihood function.

The resulting algorithm is generic and specializes automatically to the structure or region of interest during the learning of the tissue probability map. Therefore the algorithm is able to specialize to arbitrary types and numbers of landmarks.

In section 2, we describe the generative model of the image intensities. In section 3, we show how to estimate the parameters of this model, including the photometric parameters of the tissue types and the local tissue probability map. In section 4, we show how the likelihood function can be maximized to estimate the position of the landmarks in a new image. Finally, in section 5, we test the method on the simultaneous detection of 3 landmarks of the hippocampus.

2 Generative model of the image intensities

In this section, we describe the generative model of the intensities of image i, $\boldsymbol{X}^{(i)}$. Let $\boldsymbol{L}^* = (L_1^*, \cdots, L_K^*)$ and $\boldsymbol{L} = (L_1, \cdots, L_K)$, be two sets of K landmarks in $\mathbb{R}^3$ and $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ a small deformation such that $\phi(\boldsymbol{L}^*) = \boldsymbol{L}$. Since many deformations verify this condition, we restrict the set of deformations to a Gaussian spline interpolation of the landmarks displacements, see equation (3). We fix a standard configuration $\boldsymbol{L}^*$, the center of mass of landmark locations in the training set. Hence finding the position of the landmarks $\boldsymbol{L}$ is equivalent to estimating the "best" deformation from $\boldsymbol{L}^*$ to $\boldsymbol{L}$. By the Bayes' formula the joint distribution becomes $P(\boldsymbol{X}^{(i)}, \phi) = P(\boldsymbol{X}^{(i)}|\phi)P(\phi)$. We choose a uniform prior so that maximizing the joint distribution is equivalent to maximizing the conditional distribution $P(\boldsymbol{X}^{(i)}|\phi)$.

We make the simplifying assumption that the voxel intensities are independent given the transformation ϕ. That is, $P(\boldsymbol{X}^{(i)}|\phi) = \prod_{v \in V} P(X_v^{(i)} = x|\phi)$, with $X_v^{(i)}$ the intensity at voxel v in image i and V the set of voxels in the image. The image intensities are modeled with a mixture of 6 gaussian distributions, corresponding to the following tissues: CSF, CSF+GM, GM, GM+WM, WM and skull+blood vessels. We denote by Z_v the discrete random variable representing the tissue type at voxel v. We assume that the intensity $X_v^{(i)}$, given the tissue type Z_v is independent of the deformation and write the conditional probability of the ith image given the deformation as

$$P(\boldsymbol{X}^{(i)}|\phi) = \prod_{v \in V} \sum_{j=1}^{6} P(X_v^{(i)} = x|Z_v = j)P(Z_v = j|\phi). \tag{1}$$

---

[4] In [9,10], probabilistic models of the image intensities have also been proposed to segment and register brain MRI.
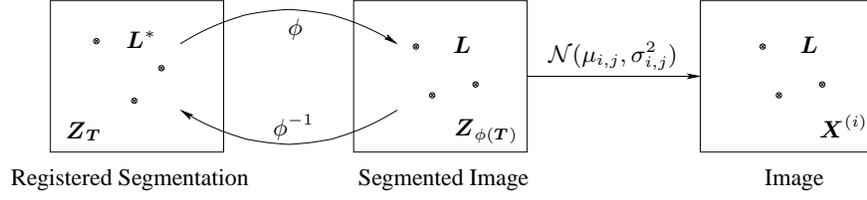
Fig. 1. To generate a new image $\boldsymbol{X}^{(i)}$ (rightmost): draw a random segmentation (leftmost) based on the distribution of the tissue probability map; apply a random deformation ϕ to find the new image segmentation $\boldsymbol{Z}_{\phi(\boldsymbol{T})}$ (middle) and assign an intensity chosen with the corresponding Gaussian distribution.

The first term characterizes the photometry of each tissue in the image, while the second term encodes the geometry. It gives the probability of observing each tissue type at each location in the brain. Figure 1 illustrate the generative model.

Let β be the parameters of the transformation ϕ, the estimation of the landmark position in a new image is $\hat{\boldsymbol{L}} = \phi_{\hat{\beta}}(\boldsymbol{L}^*)$ with $\hat{\beta} = \arg\max_{\beta} \ln P(\boldsymbol{X}|\phi_{\beta})$. We denote by $\boldsymbol{T} = \{P(Z_t = j), 1 \leq j \leq 6\}_{\forall t \in \boldsymbol{T}}$ the tissue probability map, i.e. the probability of observing tissue j at each voxel. Algorithm 1 summarizes the algorithm for automatic landmarking which will be described in detail in the following sections.

Algorithm 1 : Automatic Landmarking

Learning step

Given N training images with K landmarks $\boldsymbol{L}_i = (L_{i,1}, \cdots, L_{i,K}), 1 \leq i \leq N$:

1. *Identify the standard configuration:* $\boldsymbol{L}^* = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{L}_i$,
2. *Learn the photometry parameters:* $\mu_{i,j}$ and $\sigma_{i,j}^2$ for each tissue type j and image i,
3. *Register the images of the training set:*
    (a) for each image i, find the unique ϕ_{β_i}, such that $\phi_{\beta_i}(\boldsymbol{L}^*) = \boldsymbol{L}_i$,
    (b) for each t in the probability map, extract the intensities $x_{\phi_{\beta_i}(t)}^{(i)}$ for all i, $1 \leq i \leq N$,
4. *Estimate the tissue probability map:*
    for each t in the probability map, estimate the proportions of each tissue in $(x_{\phi_{\beta_1}(t)}^{(1)}, \cdots, x_{\phi_{\beta_1}(t)}^{(N)})$, using the EM algorithm.

For a new image

1. *Learn the photometry parameters:* μ_j and σ_j^2 for each tissue type j,
2. *Estimate the transformation and predict the landmark location:*
    (a) Maximize the likelihood with respect to β using a gradient method,
    (b) Compute the transformation $\phi_{\hat{\beta}}$ and the predicted landmark location $\hat{\boldsymbol{L}} = \phi_{\hat{\beta}}(\boldsymbol{L}^*)$.

---

3 Estimation of the model parameters

In this section, we show how to estimate the parameters of the probabilistic model. Since the photometry and the geometry are assumed to be independent in our model, we can estimate the photometric parameters and the tissue probability map separately. Let us start with the photometry.

3.1 Learning the photometric parameters

The intensity of an image $\boldsymbol{X}^{(i)}$ is modeled as a mixture of Gaussian distributions, assuming independence of the voxels, as it is commonly done in literature [11]. Thus,

$$P(\boldsymbol{X}^{(i)}) = \prod_{v \in V} \sum_{j=1}^{6} g_{i,j}(x_v^{(i)})\alpha_{i,j}, \text{ with } g_{i,j} \sim \mathcal{N}(\mu_{i,j}, \sigma_{i,j}) \text{ and } \sum_{j} \alpha_{i,j} = 1. \quad (2)$$

Given the segmentation, it would be straightforward to estimate the parameters of the model. However, here the segmentation is unknown, also we use the Expectation-Maximization (EM) algorithm [12] to maximize $\ln P(\boldsymbol{X}^{(i)})$ with respect to $\mu_{i,j}$, $\sigma_{i,j}^2$, and $\alpha_{i,j}$, $1 \leq j \leq 6$. In the case of a Gaussian mixture, both the E-step and the M-step can be written in closed form and convergence to a local maximum of the likelihood function is guaranteed.

3.2 Learning the tissue probability map

The standard atlas for registration in neuroimaging is that of Talairach and Tournoux, [2]. Tissue probability maps have been estimated based on images aligned by linear transformation to that atlas [13]. In the problem we consider the images are already manually transformed into Talairach space but we try to reduce the residual variability (we note variability of 10mm for the head of the hippocampus [14]). Also we use the correspondences of the landmarks of the training set to improve the quality of the alignment of the structure of interest and consequently learn a locally more precise probability map. The estimation of the tissue probability map is obtained as follows: first, register the training images so that the landmarks lie at the same location in all the images and then estimate the proportions of each tissue type at each location.

Registration of the training set: Since the images have been previously aligned by a linear transformation, we deal only with small deformations. More specifically we use a spline interpolation ϕ_β based on the landmark matching constraint.

$$\forall t \in \mathbb{R}^3, \phi_\beta(t) = t + \sum_{k=1}^{K} \frac{\beta_k}{\sqrt{2\pi}^3 \sigma_k^3} \exp\left(-\frac{\|t - L_k^*\|^2}{2\sigma_k^2}\right), \quad (3)$$

where $\beta_k \in \mathbb{R}^3$ is the displacement vector at the landmark L_k^* and $\sigma_k \in \mathbb{R}$ is the smoothing parameter. This type of deformation has a simple analytical form, easy to

generalize to more landmarks. It also has a simple interpretation and an almost local support so that computation will be easier. Registering the training images consists of finding the unique transformation ϕ_β that transforms the landmarks of the tissue probability map to those of the image, i.e. $\phi_\beta(\boldsymbol{L}^*) = \boldsymbol{L}$. The choice of the smoothing parameters σ_k is done manually, so that we ensure the invertibility of the deformation. These parameters could also be learned during the training phase. Once σ_k are fixed, registering one training image is equivalent to solving for β the linear system given by the landmark matching constraints. Other deformations could be used to register the images which in addition to satisfying the landmark matching constraints, either minimize the bending energy (Thin-Plate Splines) or ensure invertibility (Geodesic Interpolating Splines) [15], at some computational cost.

Estimation of the probability tissue map: The registration of the training set provides us a vector of intensities at each location $t \in \boldsymbol{T}$, $\left(x^{(1)}_{\phi_{\beta_1}(t)}, \ldots, x^{(N)}_{\phi_{\beta_N}(t)} \right)$, where $x^{(i)}_{\phi_{\beta_i}(t)}$ is the gray level of the image i, $1 \leq i \leq N$, at the location $\phi_{\beta_i}(t)$. Learning the geometry consists of estimating the proportion of each tissue type j at each location t of the probability map, based on these observations. Since both the tissue types (we observe only the gray level) and their distribution are unknown, we need to use once more the EM-algorithm to estimate the proportions of the mixture. The EM algorithm maximizes the following quantity with respect to $P(Z_t = j)$:

$$\sum_{i=1}^{N} \ln \sum_{j=1}^{6} \frac{P(Z_t = j)}{\sqrt{2\pi\sigma_{i,j}^2}} \exp\left(-\frac{(x^{(i)}_{\phi_{\beta_i}(t)} - \mu_{i,j})^2}{2\sigma_{i,j}^2} \right). \tag{4}$$

We use the photometric parameters $(\mu_{i,j}, \sigma_{i,j}^2)$ estimated previously on each one of the images of the training set. Figure 2 shows one sagittal slice of the tissue probability map. Notice that the anatomy is sharply described by the tissue probability map in the vicinity of the landmarks while it is more diffuse at longer distances.

4   Landmark detection on a new image

Consider a new image $(x_v)_{v \in V}$. Since the likelihood of the intensities is a function of β, the deformation parameters, we look for the deformation ϕ_β that maximizes the conditional likelihood $l(\beta)$. That is, $\hat{\beta} = \arg\max_\beta l(\beta)$ with

$$l(\beta) = \sum_{v \in V} \ln \sum_{j=1}^{6} g_j(x_v) P(Z_{\phi^{-1}(v)} = j). \tag{5}$$

The estimated landmark location is simply given by $\hat{\boldsymbol{L}} = \phi_{\hat{\beta}}(\boldsymbol{L}^*)$.

However, ϕ_β^{-1} cannot be expressed in a simple analytic form, so we make the change of variables $v = \phi_\beta(t)$. The likelihood becomes,

$$l(\beta) = \sum_{t \in \phi_\beta^{-1}(V)} \ln \sum_{j=1}^{6} g_j(x_{\phi_\beta(t)}) P(Z_t = j) |J_{\phi_\beta}(t)|, \tag{6}$$
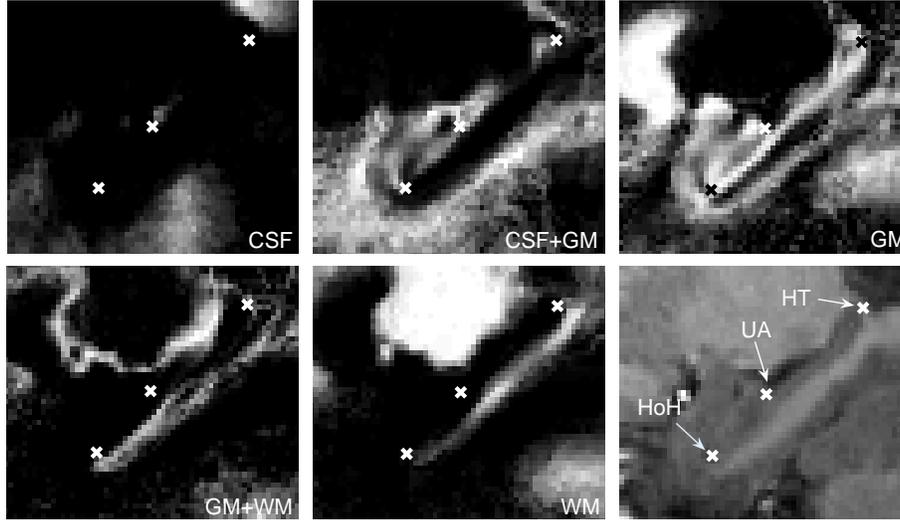
Fig. 2. Tissue probability map obtained when the deformation is estimated on three landmarks: HoH the head of the Hippocampus, HT the hippocampus tail and UA, the posterior apex of the hippocampal uncus. The slice corresponds to the sagittal section containing HoH, HT. UA is represented on the same slice although it lies in the next sagittal slice. We represent the 5 main tissues of the brain (the sixth one is not present in this region of the image). White regions have a high probability to belong to the corresponding tissue. The rightmost bottom image is an example of a registered image of the training set.

with $|J_{\phi_\beta}(t)|$ the absolute value of the determinant of the Jacobian of the change of variable. We approximate it by computing the quantity on the tissue probability map support T instead of $\phi_\beta^{-1}(V)$. In addition, since the support of the deformation ϕ_β is almost local, the computation can be reduced to the points around the landmarks, saving memory usage and time.

The derivative of the likelihood with respect to β can be computed exactly. The likelihood is maximized by simple gradient ascent in $\mathbb{R}^{3K}$. Conjugate gradient method did not improve the experimental results.

5   Experiments: landmarking the hippocampus in brain MRI

The training set is composed of 38 T1-weighted MR brain images acquired on a Philips-Intera 3-Tesla scanner, with resolution $1mm^3$. Brains were first manually transformed into standardized Talairach space using Analysis of Functional Neuroimages (AFNI) to provide a canonical orientation (anterior and posterior commissures (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. An expert located the apex of the Head of the Hippocampus (HoH) and define on the same sagittal slice the Tail of the Hippocampus (HT) and he located the posterior apex of the hippocampal uncus (UA). (see Figure 2 bottom right)

We apply Algorithm 1 to predict simultaneously the location of HoH, HT and UA in the training set and in the testing set, composed of 9 images acquired in the same experimental setting as the training set and transformed into Talairach space. We experimented with different values of σ but the results were comparable. We present in more details the case $\sigma_{HoH} = \sigma_{HT} = \sigma_{UA} = 5$.

After learning the model as described in section 2, we estimate the transformation parameters $\beta \in \mathbb{R}^9$ for each image independently. We assess the quality of the prediction by computing the Euclidean distance between the true landmarks L and the estimated position $\hat{L}$. The hand-landmarking procedure defines HoH and HT so that they lie on the same sagittal slice. This is enforced in the optimization algorithm by expressing the constraints on β. For computational efficiency, we restrict the domain of computation to the set of voxels in the probability map within 3σ of each landmarks, since the gradient will be null at further distance of the landmarks. Table 1 presents the mean error over the images of the training and the testing set. For comparison,

Table 1. Mean prediction error in mm obtained by automatic landmarking for HoH, HT and UA on the training set (38 images) and on the testing set (9 images).

	mean error (mm) on the training set	mean error (mm) on the testing set
HoH	2.27 (1.58)	2.96 (1.17)
HT	2.49 (1.25)	2.57 (1.20)
UA	2.20 (1.30)	2.78 (1.59)

the specialist's intra-variability for HoH is 1.22mm ($\sigma = 0.92$), while the non-expert intra-variability on the same images for the same landmark is 3.58mm ($\sigma = 0.98$). The resulting inter-observer variability for HoH is 3.26mm ($\sigma = 0.98$). The algorithm reaches performance comparable to the inter-observers variability and offers for most of the images a reliable prediction of the landmark position.

Since locating landmarks in the image is aimed at reducing the local variability of the alignment, we compute the average gray-level across the images of the testing set and compare the initial alignment to the alignment after registration, using the automatic landmarks or the manual landmarks. For this experiment, we use the same registration technique as in section 3.2. Figure 3 presents the results. The improvement between (a) and (b) are concentrated around the landmarks. Around HT (top-right cross) the average location of the tail is more consistent and accurate than in (a). The head of the hippocampus (bottom-left cross) is visible in (b) while it is not in (a). The overall alignment of the hippocampus is better in (b) than in (a).

6   Conclusion

We have proposed a simple and generic algorithm for automating the detection of landmarks on anatomical structures of the brain. The algorithm belongs to the class of generative models using a training set of manually located landmarks to specialize to the structure or region of interest. It adapts easily to various numbers and types of anatomical landmarks. Experimental results on brain MRI for locating three landmarks in the hippocampus are promising with an error comparable to the variability between landmarkers.
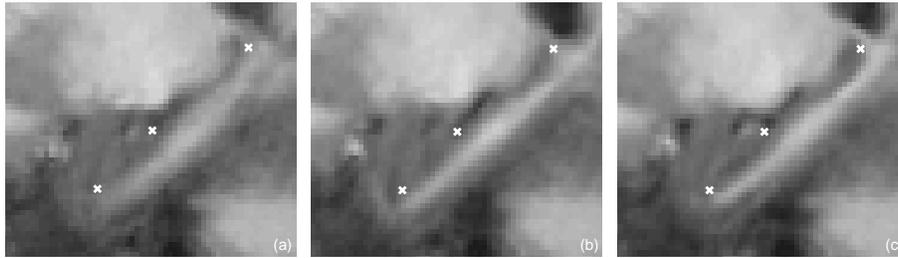
Fig. 3. Alignment experiment results on the testing set composed of 9 images initially aligned in the Talairach space. (a) represents the initial alignment without using landmarks, (b) represents the alignment obtained when the registration is made using the landmarks detected automatically, (c) alignment based on the expert's landmarks. The crosses represent HoH, HT and the projection of UA as it lies in the next sagittal slice.

References

1. Dryden, I.L., Mardia, K.V.: Statistical shape analysis. Wiley (1998)
2. J.Talairach, Tournoux, P.: Co-planar stereotaxic Atlas of the Human Brain. Thieme Medical Publishers (1988)
3. Joshi, S., Miller, M.: Landmark matching via large deformation diffeomorphisms. In: IEEE Trans. in Image Processing. Volume 9. (2000) 1357–1370
4. Rohr, K., Stiehl, H., Sprengel, R., Buzug, T., Weese, J., Kuhn, M.: Landmark-based elastic registration using approximating thin-plate splines. IEEE Trans. Med. Img. **20**(6) (2001) 526–534
5. Bookstein, F.: Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press (1992)
6. Thirion, J.P.: New feature points based on geometric invariants for 3d image registration. Int. J. of Computer Vision **18:2** (1996) 121–137
7. Wörz, S., Rohr, K.: Localization of anatomical point landmarks in 3d medical images by fitting 3d parametric intensity models. Medical Image Analysis **10**(1) (2006) 41–58
8. Rohr, K.: Landmark-based Image Analysis using Geometric and Intensity Models. Kluwer Academic, Dordrecht (2001)
9. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage (26) (2005) 839–851
10. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33** (2002) 341–355
11. Wells, W., Kikinis, R., Grimson, W., Jolesz, F.: Adaptive segmentation of mri data. In: IEEE Trans. Med. Imag. Volume 15. (1996) 429–442
12. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the em algorithm. J. R. Stat. Soc. **39** (1977) 1–38
13. Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C.: Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. Journal of Computer Assisted Tomography **18** (1994) 192–205
14. Stark, C., Okado, Y.: Making memories without trying: Medial temporal lobe activity associated with incidental memory formation during recognition. J. of Neurosci. **23** (2003) 6748–6753
15. Camion, V., Younes, L.: Geodesic interpolating splines. In: Proceedings of Energy Minimisation in Computer Vision and Pattern Recognition (EMMCVPR). Volume 2134., Springer-Verlag (2001) 513–527