Coupling Statistical Segmentation and PCA Shape Modeling

Kilian M. Pohl[1], Simon K. Warfield[2], Ron Kikinis[2], W. Eric L. Grimson[1], and William M. Wells[2]

[1] Computer Science and Artificial Intelligence Lab, http://www.csail.mit.edu, Massachusetts Institute of Technology, Cambridge MA, USA., {kpohl,welg}@csail.mit.edu

[2] Surgical Planning Laboratory, http://www.spl.harvard.edu, Harvard Medical School and Brigham and Women's Hospital, 75 Francis St., Boston, MA 02115 USA, {warfield,kikinis,sw}@bwh.harvard.edu

Abstract. This paper presents a novel segmentation approach featuring shape constraints of multiple structures. A framework is developed combining statistical shape modeling with a maximum a posteriori segmentation problem. The shape is characterized by signed distance maps and its modes of variations are generated through principle component analysis. To solve the maximum a posteriori segmentation problem a robust Expectation Maximization implementation is used. The Expectation Maximization segmenter generates a label map, calculates image intensity inhomogeneities, and considers shape constraints for each structure of interest. Our approach enables high quality segmentations of structures with weak image boundaries which is demonstrated by automatically segmenting 32 brain MRIs into right and left thalami.

1 Introduction

For many age or disease related brain studies large quantities of Magnetic Reasoning Images (MRI) have to be accurately segmented into anatomical regions. Achieving high quality brain MRI segmentation is quite challenging for automatic methods so researchers often have to rely on labor intensive, manual delineation. The task is challenging because some structures have very similar intensity characteristics, such as substructures in the cortical gray matter, while others have only weakly visible boundaries (e.g. thalamus). Recent methods using enhanced anatomical knowledge have greatly improved the quality of automatically generated results. We briefly summarize methods that incorporate shape constraints into the segmentation process. A promising approach [1,2,3] is based on level set functions. It characterizes shape based signed distance maps in combination with the Principle Component Analysis (PCA) [4] . Generally, PCA finds the largest modes of variation among the signed distance maps. Besides level sets, deformable model methods have used many different shape representations, such as spherical harmonics [5], point based models [4], skeleton or medial representations [6], and finite element models [7].

The novel approach presented in this paper is most closely related to work by Tsai and Leventon [1,2]. While PCA based segmentation methods are very robust they are also
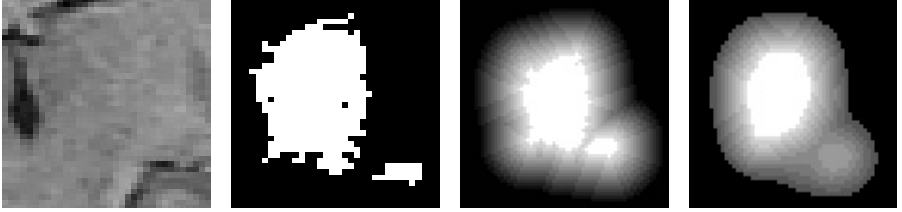
Fig. 1. Example of a left thalamus and corresponding segmentation, related signed distance map, and structure's mean where the voxel's brightness corresponds to the value in the distance map.

constraint in the degrees of freedom of the shape variations allowed. We therefore couple the PCA based shape modeling with a maximum a posteriori estimation problem which will be solved through an Expectation Maximization (EM) implementation developed by Pohl et al. [8]. This allows the system to accommodate shapes that differ some what from those modeled by the PCA. Additionally, the method can segment multiple objects and estimate intensity inhomogeneities in the image.

2   Method

This section discusses the integration of shape constraints into an EM segmentation algorithm. First, the shape variations across subjects are captured through PCA [9]. Afterwards, the shape constraints are added to the parameter space of an EM-based segmentation algorithm [8].

2.1   Shape Representation

Various shape representations have been explored in medical imaging. For our work, we chose signed distance maps due to their robustness. The structure's shape variations are captured by PCA. To apply PCA to the training data we first align all training sets using the affine registration method developed by Warfield [10]. Then, each data set i is transferred into structure specific signed distance maps $\mathcal{D}_a^{(i)}$, where a represents the structure of interest (see also Figure 1). In these distance maps positive values are assigned to voxels within the boundary of the object, while negative values indicate voxels outside the object. By taking the average over all these distance maps $\mathcal{D}_a^{(i)}$ we define the mean distance map $\overline{\mathcal{D}}_a := \frac{1}{n} \sum_i \mathcal{D}_a^{(i)}$ and the mean corrected signed distance maps $\tilde{\mathcal{D}}_a^{(i)} := \mathcal{D}_a^{(i)} - \overline{\mathcal{D}}_a$. The input for PCA is the vector $\tilde{\mathcal{D}}^{(i)} := (\tilde{\mathcal{D}}_1^{(i)^T}, \cdots, \tilde{\mathcal{D}}_N^{(i)^T})^T$ defined by the mean corrected signed distance maps of the N structures of interests. Therefore, PCA is applied to all structures at once. This analysis defines the shape constraints of the entire image which is represented by the eigenvector or modes of variation matrix U, eigenvalue matrix Λ, and $\overline{\mathcal{D}} := (\overline{\mathcal{D}}_1^T, \cdots, \overline{\mathcal{D}}_N^T)^T$ (see also Figure 2). To reduce the computational complexity for the EM implementation, U and Λ will only be defined by the first K eigenvectors and eigenvalues, where K represents 99 % of the eigenvalues' energy.

The shapes in a specific brain image will be captured by the expansion coefficients of the eigenvector representation which we call shape parameters $\mathcal{S} = (\mathcal{S}_1, \cdots, \mathcal{S}_K)$.
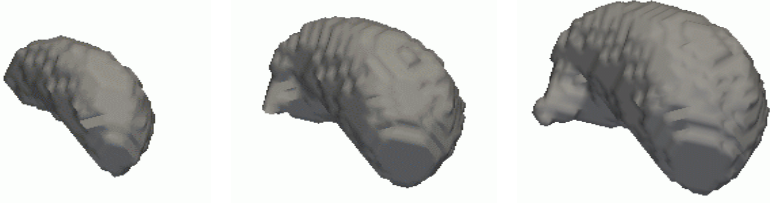
Fig. 2. These are the results of PCA applied to a training set of manually segmented thalami. As clearly visible from the images the first mode of variation, i.e. the deformation along the eigenvector with the largest eigenvalue, defines the size of the structure.

$\mathcal{S}$ relates to the distance maps by $\mathcal{D}_\mathcal{S} = \overline{\mathcal{D}} + U \cdot \mathcal{S}$. We will refer to the shape parameter generated distance map of a specific structure a as $\mathcal{D}_{\mathcal{S},a} = \overline{\mathcal{D}}_a + U_a \cdot \mathcal{S}$, where U_a are just the entries in U that refer to structure a.

The probability distribution over the shape parameters $p(\mathcal{S})$ is now defined by the Gaussian distribution

$$p(\mathcal{S}) = \frac{1}{\sqrt{(2\pi)^K |\Lambda|}} \exp\left(-\frac{1}{2}\mathcal{S}^T \Lambda^{-1} \mathcal{S}\right)$$

where K is the dimension of eigenvalue matrix Λ.

2.2 Estimating Intensity Inhomogeneities and Shape

The algorithm proposed in this chapter is based on an EM-based segmentation algorithm by Pohl et al. [8] which uses probability atlases to define the spatial distribution of structures . Expanding this approach, we will not only approximate the maximum a posteriori estimate (MAP) of the image intensity inhomogeneities $\mathcal{B}$ but also the MAP estimate of the shape parameters $\mathcal{S}$. In this framework the MAP estimates of the parameter space, i.e. $\mathcal{B}$ and $\mathcal{S}$, depend on the partition of the image in anatomical regions $\mathcal{T}$ (the hidden data), the log intensities of the input image $\mathcal{I}$ (the observed data), and previous estimations of the inhomogeneities $\mathcal{B}'$ as well as the shape parameter $\mathcal{S}'$. Therefore, our approach tries to solve the following problem:

$$(\mathcal{B}'', \mathcal{S}'') = \arg\max_{\mathcal{B},\mathcal{S}} Q(\mathcal{B}, \mathcal{S}|\mathcal{B}', \mathcal{S}') = \arg\max_{\mathcal{B},\mathcal{S}} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{B}, \mathcal{S}|\mathcal{T}, \mathcal{I}))$$
$$= \arg\max_{\mathcal{B},\mathcal{S}} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{I}|\mathcal{T}, \mathcal{S}, \mathcal{B}) + \log p(\mathcal{S}|\mathcal{T}, \mathcal{B}) + \log p(\mathcal{B}|\mathcal{T}))$$
$$= \arg\max_{\mathcal{B},\mathcal{S}} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{I}|\mathcal{T}, \mathcal{B}) + \log p(\mathcal{S}|\mathcal{T}, \mathcal{B}) + \log p(\mathcal{B}|\mathcal{T}))$$
$$\tag{1}$$

where $E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{B}, \mathcal{S}|\mathcal{T}, \mathcal{I})) := \sum_\mathcal{T} p(\mathcal{T}|\mathcal{I}, \mathcal{B}', \mathcal{S}') \cdot \log p(\mathcal{B}, \mathcal{S}|\mathcal{T}, \mathcal{I})$ and we assume independence of $\mathcal{S}$ in $p(\mathcal{I}|\mathcal{T}, \mathcal{S}, \mathcal{B})$. If we further assume independence between $\mathcal{B}$ and $\mathcal{S}$, and $\mathcal{B}$ and $\mathcal{T}$ than the maximization problem can be simplified to:[1]

$$\mathcal{B}'' = \arg\max_{\mathcal{B}} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{I}|\mathcal{T}, \mathcal{B})) + \log p(\mathcal{B}) \tag{2}$$
$$\mathcal{S}'' = \arg\max_{\mathcal{S}} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{S}|\mathcal{T})) \tag{3}$$

---

[1] $p(\mathcal{S}|\mathcal{T}, \mathcal{B}) = \frac{p(\mathcal{S},\mathcal{T},\mathcal{B})}{p(\mathcal{T},\mathcal{B})} = \frac{p(\mathcal{S},\mathcal{T})p(\mathcal{B})}{p(\mathcal{T})p(\mathcal{B})} = p(\mathcal{S}|\mathcal{T})$ and $p(\mathcal{B}|\mathcal{T}) = p(\mathcal{B})$.

To solve these two equations the EM algorithm iterates between the Expectation Step (E-Step) and the Maximization Step (M-Step). The E-Step first updates $\mathcal{B}'$ and $\mathcal{S}'$ with $\mathcal{B}''$ and $\mathcal{S}''$. Then it calculates the expected value of the two functions based on $\mathcal{B}'$ and $\mathcal{S}'$. The M-Step approximates separately the MAP estimates $\mathcal{B}''$ and $\mathcal{S}''$ based on the results of the E-Step. For a general overview of EM we refer the reader to [11].

In the remainder of this section we will first discuss the two MAP estimation problems separately and then integrate these two MAP estimation problems into the EM framework.

Estimating the Intensity Inhomogeneities

To find the MAP estimate of $\mathcal{B}$ we assume statistical independence of the voxel location x for $\mathcal{B}$ and $\mathcal{I}$. Therefore, Equation (2) simplifies to:

$$0 = \frac{\partial}{\partial \mathcal{B}_x} \left(E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'} \left(\log p(\mathcal{I}_x | \mathcal{B}_x, \mathcal{T}_x) \right) \right) + \frac{\frac{\partial}{\partial \mathcal{B}_x} p(\mathcal{B})}{p(\mathcal{B})} \qquad (4)$$

The conditional intensity distribution is modeled by a Gaussian distribution:

$$p(\mathcal{I}_x | \mathcal{T}_x = e_a, \mathcal{B}_x) \overset{2}{:=} \frac{1}{\sqrt{(2 \cdot \pi)^n |\sigma_a|}} e^{-\frac{1}{2}(\mathcal{I}_x - \mathcal{B}_x - \mu_a)^T \cdot \sigma_a^{-1} \cdot (\mathcal{I}_x - \mathcal{B}_x - \mu_a)}$$

where n is the number of input channels, and (μ_a, σ_a) define the intensity distribution of structure a. '$\overset{x}{=}$' refers to footnote x for further explanation. Let's define

$$A_x(a) := \frac{\partial}{\partial \mathcal{B}_x} p(\mathcal{I}_x | \mathcal{T}_x = e_a, \mathcal{B}_x) = \sigma_a^{-1} \cdot (\mathcal{I}_x - \mathcal{B}_x - \mu_a)$$

and the weights $\mathcal{W}_x(a) := E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\mathcal{T}_x(a))$ so that Equation (4) turns into

$$0 = \frac{\partial}{\partial \mathcal{B}_x} \left(E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'} \left(\log p(\mathcal{I}_x | \mathcal{T}_x, \mathcal{B}_x) \right) \right) + \frac{\frac{\partial}{\partial \mathcal{B}_x} p(\mathcal{B})}{p(\mathcal{B})}$$

$$= E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\mathcal{T}_x) \cdot \frac{\partial}{\partial \mathcal{B}_x} \log p(\mathcal{I}_x | \mathcal{T}_x, \mathcal{B}_x) + \frac{\frac{\partial}{\partial \mathcal{B}_x} p(\mathcal{B})}{p(\mathcal{B})} = \mathcal{W}_x^T \cdot A_x + \frac{\frac{\partial}{\partial \mathcal{B}_x} p(\mathcal{B})}{p(\mathcal{B})}$$

As Wells shows [12] the above problem can be approximated by a low pass filter H applied to the weighted residual $\bar{R}$: $\mathcal{B} \approx H\bar{R}$. Now, we will explicitly define the weights $\mathcal{W}_x(a) := E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\mathcal{T}_x(a))$:

$$\mathcal{W}_x(a) := E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\mathcal{T}_x(a)) \overset{3}{=} E_{\mathcal{T}_x|\mathcal{I}_x,\mathcal{B}'_x,\mathcal{S}'}(\mathcal{T}_x(a))$$

$$= 0 \cdot p(\mathcal{T}_x(a) = 0 | \mathcal{I}_x, \mathcal{B}'_x, \mathcal{S}') + 1 \cdot p(\mathcal{T}_x(a) = 1 | \mathcal{I}_x, \mathcal{B}'_x, \mathcal{S}') \qquad (5)$$

$$= p(\mathcal{T}_x(a) = 1 | \mathcal{I}_x, \mathcal{B}'_x, \mathcal{S}') \overset{4}{=} \frac{p(\mathcal{I}_x | \mathcal{T}_x(a) = 1, \mathcal{B}'_x) \cdot p(\mathcal{T}_x(a) = 1 | \mathcal{S}')}{p(\mathcal{I}_x | \mathcal{B}'_x, \mathcal{S}')}$$

We will model $p(\mathcal{T}_x(a) = 1 | \mathcal{S})$ as a measure of agreement among the shape $\mathcal{S}$ an the label map $\mathcal{T}$. This is achieved by transforming the distance maps $\mathcal{D}_\mathcal{S}$ produced by $\mathcal{S}$ into binary maps through $\mathcal{H}$:

---

[2] e_a has a 1 at position a and 0 otherwise

[3] Bayes' rule: $\sum_{\mathcal{T}(i,a_j)} p(\mathcal{T}(1, a_1), \cdots, \mathcal{T}(n, a_m) | \mathcal{I}, \mathcal{B}') \cdot \mathcal{T}_x(a) = p(\mathcal{T}_x(a) = 1 | \mathcal{I}, \mathcal{B}')$

[4] Based on previous independence assumption

$$\mathcal{H}_{\mathcal{S}}(x,a) := \begin{cases} 1 & \text{, if } \mathcal{D}_{\mathcal{S},a}(x) \geq 0 \\ 0 & \text{, if } \mathcal{D}_{\mathcal{S},a}(x) < 0 \end{cases}$$

where $\mathcal{H}_{\mathcal{S}}(x,a)$ is the Heaviside function for structure a. $p(\mathcal{T}|\mathcal{S})$ penalizes any disagreement between $\mathcal{T}_x$ and $\mathcal{H}_{\mathcal{S}}(x) = (\mathcal{H}_{\mathcal{S}}(x,1), \cdots, \mathcal{H}_{\mathcal{S}}(x,N))^T$:

$$p(\mathcal{T}|\mathcal{S}) := \frac{1}{Z} e^{-\frac{1}{2}\sum_x d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x)) + \log(f(\mathcal{T}_x))}$$

where d is a correlation metric between $\mathcal{T}_x$ and $\mathcal{H}_{\mathcal{S}}(x)$. Here $d(v_1, v_2) := (v_1 - v_2)^T(v_1 - v_2)$, which means d is zero when v_1 and v_2 agree, and 1 or greater when they disagree. $f(\mathcal{T}_x)$ represents a prior probability on $\mathcal{T}_x$ defined by a probability atlas [8]. We therefore can ignore f in the normalizing function Z with m being the number of voxels in the image

$$Z := \sum_{\mathcal{T}'} e^{\sum_x -\frac{1}{2}d(\mathcal{T}'_x, \mathcal{H}_{\mathcal{S}}(x))} = \prod_x \sum_{\mathcal{T}_x} e^{-\frac{1}{2}d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x))}$$

$$\overset{5}{=} \prod_x (1 + (N-1) \cdot e^{-1}) = (1 + (N-1) \cdot e^{-1})^m$$

If $p(\mathcal{T}_x|\mathcal{S}) := (1 + (N-1) \cdot e^{-1})^{-1} \cdot e^{-\frac{1}{2}d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x))} \cdot f(\mathcal{T}_x)$ defines the local conditional probability than

$$p(\mathcal{T}|\mathcal{S}) := (1 + (N-1) \cdot e^{-1})^{-m} \cdot e^{\sum_x -\frac{1}{2}d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x)) + \log(f(\mathcal{T}_x))} = \prod_x p(\mathcal{T}_x|\mathcal{S})$$

Estimating the Shape Parameters $\mathcal{S}$

As mentioned in Section 2.1 statistical independence among the coefficients of $\mathcal{S} = (\mathcal{S}_1, \cdots, \mathcal{S}_N)^T$ is assumed. Therefore, Equation (3) is solved for each component of $\mathcal{S}$:

$$0 = \frac{\partial}{\partial \mathcal{S}_i} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{S}|\mathcal{T})) = \frac{\partial}{\partial \mathcal{S}_i} E_{\mathcal{T}|\mathcal{I},\mathcal{B}',\mathcal{S}'}(\log p(\mathcal{T}|\mathcal{S}) + \log p(\mathcal{S}))$$

$$= \left(\sum_x \frac{\partial}{\partial \mathcal{S}_i} (E_{\mathcal{T}_x|\mathcal{I}_x,\mathcal{B}'_x,\mathcal{S}'}(\log p(\mathcal{T}_x|\mathcal{S}))) \right) + \log \frac{\partial}{\partial \mathcal{S}_i} p(\mathcal{S}) \qquad (6)$$

$$= \sum_x E_{\mathcal{T}_x|\mathcal{I}_x,\mathcal{B}'_x,\mathcal{S}'}(\mathcal{T}_x)^T \frac{\partial}{\partial \mathcal{S}_i} \log p(\mathcal{T}_x|\mathcal{S}) - \Lambda_i^{-1} \mathcal{S}_i$$

where $\frac{\partial}{\partial \mathcal{S}_i} \log p(\mathcal{T}_x|\mathcal{S})$

$$= \frac{\partial}{\partial \mathcal{S}_i} -\frac{1}{2}(\mathcal{T}_x - \mathcal{H}_{\mathcal{S}}(x))^2 = \frac{\partial \mathcal{H}_{\mathcal{S}}(x)}{\partial \mathcal{S}_i}$$

$$\cdot (\mathcal{T}_x - \mathcal{H}_{\mathcal{S}}(x)) \overset{6}{=} (\delta(\mathcal{D}_{\mathcal{S}}(x))^T U_i(x)) \cdot (\mathcal{T}_x - \mathcal{H}_{\mathcal{S}}(x))$$

---

[5] $\sum_{\mathcal{T}_x} e^{-\frac{1}{2}d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x))} = |\mathcal{H}_{\mathcal{S}}(x)|^2 e^{-\frac{1}{2}(|\mathcal{H}_{\mathcal{S}}(x)|^2 - 1)} + (N - |\mathcal{H}_{\mathcal{S}}(x)|^2)e^{-\frac{1}{2}(|\mathcal{H}_{\mathcal{S}}(x)|^2 + 1)}$. If we assume each voxel is part of only one shape then $|\mathcal{H}_{\mathcal{S}}(x)| = 1$ and $\sum_{\mathcal{T}_x} e^{-\frac{1}{2}d(\mathcal{T}_x, \mathcal{H}_{\mathcal{S}}(x))} = 1 + (N-1)e^{-1}$

is zero unless $\mathcal{T}_x(a) \neq \mathcal{H}_\mathcal{S}(a)$ for a structure a and voxel x is located at the border of the shape of a. Thus, if Ω is the set of voxels at the boundaries of $\mathcal{H}_\mathcal{S}$ Equation (6) simplifies to :

$$
\begin{aligned}
0 &\overset{7}{=} \left(\textstyle\sum_{x \in \Omega} E_{\mathcal{T}_x | \mathcal{I}_x, \mathcal{B}'_x, \mathcal{S}'}(\mathcal{T}_x)^T \cdot \left(\delta^0(\mathcal{D}_\mathcal{S}(x))^T U_i(x) \right) \cdot (\mathcal{T}_x - \mathcal{H}_\mathcal{S}(x)) \right) - \Lambda_i^{-1} \mathcal{S}_i \\
&= \left(\textstyle\sum_{x \in \Omega} \mathcal{W}_x^T \cdot \left(\delta^0(\mathcal{D}_\mathcal{S}(x))^T U_i(x) \right) \cdot (\mathcal{T}_x - \mathcal{H}_\mathcal{S}(x)) \right) - \Lambda_i^{-1} \mathcal{S}_i \\
\Rightarrow \mathcal{S}_i &= \Lambda_i \cdot \textstyle\sum_{x \in \Omega} \left(\delta^0(\mathcal{D}_\mathcal{S}(x))^T U_i(x) \right) \cdot \mathcal{W}_x^T (\mathcal{T}_x - \mathcal{H}_\mathcal{S}(x))
\end{aligned}
$$

From the above equation the updated shape parameter $\mathcal{S}_i$ is defined by the weighted sum of its eigenvector values located at borders and scaled by the ith eigenvalue. In other words, the eigenvector values $U_i(x)$ defines the 'direction of change' for parameter $\mathcal{S}_i$ and the $\mathcal{W}_x^T(\mathcal{T}_x - \mathcal{H}_\mathcal{S}(x))$ control the 'speed of change'.

2.3   The Shape Constraint EM Algorithm

The EM Algorithm is now defined by the E-Step who generates the structure posterior probabilities $\mathcal{W}$, called *weights*, based on the constraints imposed by shape, intensity, image inhomogeneities, and location (see Equation (5))

$$
\mathcal{W}_x(a) = \frac{p(\mathcal{I}_x | \mathcal{T}_x(a){=}1, \mathcal{B}'_x) p(\mathcal{T}_x(a){=}1 | \mathcal{S}')}{p(\mathcal{I}_x | \mathcal{B}'_x, \mathcal{S}')} = \frac{p(\mathcal{I}_x | \mathcal{T}_x(a) = 1, \mathcal{B}'_x) p(\mathcal{T}_x(a) = 1 | \mathcal{S}')}{\sum_{a'} p(\mathcal{I}_x | \mathcal{T}_x(a){=}1, \mathcal{B}'_x) p(\mathcal{T}_x(a) = 1 | \mathcal{S}')}
$$

The M-Step calculates the image inhomogeneities $\mathcal{B}$ and shape parameters $\mathcal{S}$ based on the newly updated weights $\mathcal{W}$. $\mathcal{B} = H \cdot \bar{R}$ is approximated by a simple low pass filter H and the weighted residuum $\bar{R}_x = \sum_a \mathcal{W}_x(a) \sigma_a^{-1}(\mathcal{I}_x - \mu_a)$ (see also [12]).

Table 1. Summary over 32 cases of the Dice comparison between the results of EM implementations and expert segmentations. The minimum and maximum list the worst and best Dice measure over all cases. As clearly by the numbers the new approach of this paper, EM-Shape, outperformed the other two methods

DICE Measure over 32 cases

Method	Mean	Variance	Minimum	Maximum
EM-Rigid	0.755	0.0221	0.449	0.883
EM-NonRigid	0.715	0.0149	0.34	0.883
EM-Shape	0.82	0.0117	0.625	0.909

---

[6] $\frac{\partial \mathcal{H}_\mathcal{S}(x,a)}{\partial \mathcal{S}_i} = \delta(\mathcal{D}_{\mathcal{S},a}(x)) \cdot \frac{\partial \mathcal{D}_{\mathcal{S},a}(x)}{\partial \mathcal{S}_i} = \delta(\mathcal{D}_{\mathcal{S},a}(x)) \cdot U_{a,i}(x)$ where δ is the Dirac's delta function and the Eigenvector matrix U_a was defined in Section 2.1

[7] where δ^0 is the null function with $\delta^0(0) = 1$, $\delta^0(x) = 0$ for $x \neq 0$, and $\delta^0(X) := (\delta^0(X(1)), \cdots, \delta^0(X(n)))^T$ for a vector X
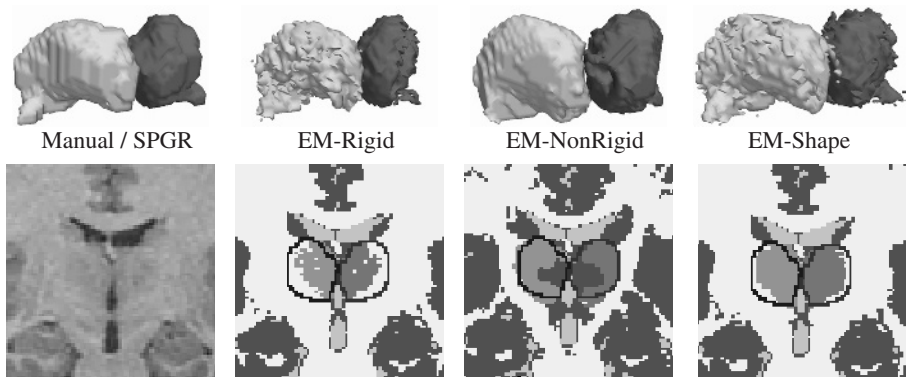
Manual / SPGR   EM-Rigid   EM-NonRigid   EM-Shape

Fig. 3. Segmentation results from different EM implementation. As clearly visible in the 2D images the shape constraint approach (EM-Shape) is closest to the expert's segmentation indicated by the black lines. EM-Shape was also the only method who properly captured the hypothalamus (see 3D models), while EM-NonRigid is too smooth and EM-Rigid underestimated the structure.

The shape parameters $\mathcal{S} = (\mathcal{S}_1, \cdots, \mathcal{S}_N)^T$ are updated in the M-Step by:

$$\mathcal{S}_i = \Lambda_i \cdot \sum_{x \in \Omega} \left(\delta^0 (\mathcal{D}_\mathcal{S}(x))^T U_i(x) \right) \cdot \mathcal{W}_x^T (\mathcal{T}_x - \mathcal{H}_\mathcal{S}(x))$$

The EM algorithm iterates between E- and M-Step until the cost function $Q((\mathcal{B},\mathcal{S}),(\mathcal{B}',\mathcal{S}'))$ of Equation (1) converges to a local maximum, which is guaranteed by the EM framework if the iteration sequence has an upper bound [11].

3   Validation

We validate our approach by segmenting 32 test cases into white matter, grey matter, cortical spinal fluid, and the left and right thalamus. The study uses segmentations from one expert which are restricted to the right and left thalamus, which this study regards as gold standard. To introduce no bias into the segmentation approach we only generated shape atlases for those two structures (see also Section 2.1). The shape atlases are produced for each test case by applying PCA to the remaining 31 cases. From the analysis we use the first five modes of variations, which corresponds to 99% of the eigenvalues' energy. Furthermore, we manually calibrate the EM segmentations by comparing one automatic segmentation result to an expert's segmentation. Especially for structures like the thalamus, where borders are not clearly visible, large variations of the experts' opinion about structure's boundary exist. Therefore, this manual calibration is essential so that automatically generated results meet the experts' expectations. To measure the robustness of the method (EM-Shape) we compare the automatic with the expert segmentations using the volume overlap measure Dice [13]. We then compare the experts segmentations to the results of two different EM implementations. The first algorithm (EM-Rigid) uses rigid alignment of atlas information and no shape constraints. The second implementation (EM-NonRigid) also does not incorporate shape constraints but uses non-rigid registration for the initial alignment and models neighborhood relationships through Markov Random Field approximation [8].

Generally, EM-Shape outperformed the other two method (see also Table 1). It had the highest mean average value of agreement, the lowest variance, the highest minimum Dice measure over all cases, and the highest maximum dice measure. Of the three methods, EM-Shape relies the least on the initial registration of the atlas to the patient. The new shape constraints allow a better adjustment of the EM parameters to the specific brain images during the segmentation process. It can capture subtle difference in the shape as the hypothalamus which is underrepresented in both EM-Rigid and EM-NonRigid (see 3D images in Figure 3).

The EM-NonRigid heavily relies on the initial non-rigid registration. Even though it produced excellent results for the superior temporal gyrus [14], it performed worse on the thalamus, because the initial alignment process cannot detect the thalamus' weakly visible boundaries. It produces very smooth segmentations due to the Mean Field approximation which models neighborhood dependencies within an image. On the downside, it also smoothed over subtle differences within small gyri and the thalamus, which are better captured by EM-Shape and EM-Rigid.

4   Discussion

A novel shape constraint segmentation approach has been presented. Embedded in an EM segmentation framework, the algorithm deals with multiple brain structures as well as estimates the intensity inhomogeneities. It generates high quality segmentations of structures with weakly visible boundaries. The approach is not restricted to the modes of variations presented in the shape model but models patient specific abnormalities. Furthermore, we have documented its robustness by segmenting 30 different cases and comparing them to other EM-like methods as well as manual segmentations. In the future we would like to include more complex conditional probabilities that better model the dependencies between label maps and the shape of the object. We also would like to couple pose and labeling of the objects because their solution depend on each other.

References

1. M.E. Leventon, W.E.L. Grimson, O.D. Faugeras, "Statistical shape influence in geodesic active contours," in *CVPR*, pp. 1316 – 1323, 2000.
2. A. Tsai, A. Yezzi, W. Wells III, C. Tempany, D. Tucker, A. Fan, W. Grimson, A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transaction on Medical Imaging*, vol. 22, no. 2, pp. 137 – 154, 2003.
3. R. D. M. Rousson, N. Paragios, "Active shape models from a level set perspective," Tech. Rep. 4984, Institut National de Recherche en Informatique et en Automatique, Sophia-Antipolis, 2003. ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-4984.pdf.
4. T.F. Cootes, A. Hill, C.J. Taylor, J. Haslam, "The use of active shape models for locating structures in medical imaging," *Imaging and Vision Computing*, vol. 12, no. 6, pp. 335–366, 1994.

5. A. Kelemen, G. Szekely, G. Gerig, "Elastic model-based segmentation of 3-d neuroradiological data sets medical imaging," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 828 – 839, 1999.

6. Stephen M. Pizer and Guido Gerig and Sarang Joshi and Stephen R. Aylward, "Multiscale medial shape-based analysis of image objects," in *Proceedings of the IEEE, Special Issue on: Emerging Medical Imaging Technology*, vol. 91, pp. 670 – 679, 2003.

7. Xenophon Papdemetris and Albert J. Sinusas and Donald P. Dione and R. Todd Constable and James S. Duncan, "Estimation of 3-d left ventricular deformation form medical images using biomechanical models," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 786 – 800, 2002.

8. Kilian M. Pohl, Sylvain Bouix, Ron Kikinis, W. Eric L. Grimson, "Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework," in *ISBI*, pp. 81–84, 2004.

9. T.F. Cootes, G.J. Edwards, C.J. Taylor, "Active appearance model," in *ECCV*, vol. 2, pp. 484–498, 1998.

10. S. K. Warfield, J. Rexilius, P. S. Huppi, T. E. Inder, E. G. Miller, W. M. Wells, G. P. Zientara, F. A. Jolesz, R. Kikinis, "A binary entropy measure to assess nonrigid registration algorithm," in *MICCAI*, pp. 266–274, Oct. 2001.

11. Geoffrey J. McLachlan, Thriyambakam Krishnan, *The EM Algorithm and Extensions*. John Wiley and Sons, Inc., 1997.

12. W.M. Wells III, W.E.L Grimson, R. Kikinis, F.A Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 429–442, 1996.

13. L.R.Dice, "Measure of the amount of ecological association between species," *Ecology*, vol. 26, pp. 297–302, 1945.

14. K. M. Pohl and W. M. Wells and A. Guimond and K. Kasai and M. E. Shenton and R. Kikinis and W. E. L. Grimson and S. K. Warfield, "Incorporating non-rigid registration into expectation maximization algorithm to segment MR images," in *MICCAI*, pp. 564–572, 2002.