CONTINUOUS LABEL BAYESIAN SEGMENTATION, APPLICATIONS TO MEDICAL BRAIN IMAGES

*Lars Aurdal†, Isabelle Bloch†, Henri Maître†, Christine Graffigne‡ and Catherine Adamsbaum**

†Département IMA ENST, CNRS URA 820, 46, rue Barrault, 75634 Paris Cedex 13. France.
‡Université René Descartes, 45, rue des Saints Pères, Paris, France.
*Hôpital St. Vincent de Paul, 87, avenue Denfert Rochereau, 75014 Paris, France.

ABSTRACT

Continuous label segmentation approaches have recently attracted much interest as they provide a formalism for handling image artifacts due to the partial volume effect which is common in for instance medical images. In this article we propose a new approach to this type of segmentation. Our work represents an extension of the now classic Markovian Bayesian discrete label segmentation approaches and provides good results on synthetic images simulating the presence of partial volumes as well as on real patient MR images.

1. INTRODUCTION

A problem with virtually all medical imaging devices is the partial volume effect: if the support of a pixel covers the boundary between two or more tissues, then the measured intensity value for this pixel will stem from a mixture of partial contributions from all the involved tissue types.

One way of segmenting these images while taking the partial volume effect into account is to label each pixel with a continuous label vector $\vec{\xi} = (\xi_1, \xi_2, \ldots, \xi_{N_c})$ where N_c is the number of classes and under the conditions that $\sum_{n=1}^{N_c} \xi_n = 1$ and that each element $\xi_n \in [0, 1]$. This can be interpreted as assigning a vector to each pixel in the segmented image describing the percent content of each tissue in this pixel.

Most of the work on segmentation of medical images has been done using algorithms that make 'hard decisions' concerning tissue type, i.e. that label a pixel as being of the most predominant tissue type in the pixel. This can reduce precision of morphometric measurements especially in thick-slice acquisitions where the partial volume effect can be considerable. Numerous methods aiming at solving the partial volume problem have therefore been proposed.

Clark et al. [3] use a fuzzy segmentation approach based on the fuzzy c-means algorithm. Choi et al. [2] model the partial volume effect by the *mixel model* and propose a MAP estimator for segmentation (see also [9]). Johnston et al. [6] have extended an approach similar to that of Choi et al. to 3D. Santago et al. [8] use a statistical model for the partial volumes together with a finite mixture density description of the tissues (see also [7]). Finally, Vincken et al. [10] propose a multi-scale approach in which partial volume voxels are analyzed at a sub-voxel resolution.

We start this paper by repeating the fundamentals of Bayesian discrete label segmentation methods, we then show how a slight change of these methods allows for continuous label segmentation. Finally, we present results on synthetic and real images.

2. BACKGROUND

Let $I = i$ be a particular realization of a process I, the image to be segmented, and $L = l$ a particular realization of a process L, the segmented image. Let s be any particular site in an image. Given $I = i$, we seek the segmented image, $L = l$, which maximizes the posterior conditional probability $P(L = l|I = i)$. Using Bayes' rule, this maximization can be performed by maximizing:

$$P(L = l|I = i) \propto P(I = i|L = l)P(L = l) \quad (1)$$

with respect to l where the first term $P(I = i|L = l)$ is the probability of a particular image $I = i$ given its segmentation $L = l$ and the second term $P(L = l)$ is the a priori probability of the segmented image $L = l$.

Under certain assumptions (see [5] for details) such as the conditional independence between sites, the presence of a spatially non-correlated, white, zero-mean Gaussian noise and that pixels in each region in the image to be segmented have Gaussian distributed gray-levels with means μ_{l_s} and variances σ_{l_s}, it can be shown that the first term in

Partially funded by the Norwegian Research Council.
E-mail: aurdal@ima.enst.fr
Web: http://www-ima.enst.fr/~aurdal/

0-8186-8183-7/97 $10.00

eq. 1 is equivalent to:

$$P(I=i|L=l)=\prod_s \frac{1}{\sqrt{2\pi}\sigma_{l_s}} \exp - \left[\frac{(i_s-\mu_{l_s})^2}{2\sigma_{l_s}^2}\right] \quad (2)$$

Assuming furthermore that the Markovian hypothesis holds for $P(L=l)$, we may use the Hammersley-Clifford theorem (see [5]) to write the a priori probability for L as a Gibbsian distribution, thus:

$$P(L=l)=\frac{1}{Z}\exp\left[-U(l)\right], \quad U(l)=\sum_{c\in\mathcal{C}} V_c \quad (3)$$

where $\mathcal{C}$ is the set of cliques of order higher than one defined on L and V_c a clique potential function. From eqs. 1, 2 and 3 we obtain the following expression that should be minimized with respect to l_s:

$$\sum_s \frac{(i_s-\mu_{l_s})^2}{2\sigma_{l_s}^2}+\sum_{c\in\mathcal{C}} V_c \quad (4)$$

The two terms of this equation can be interpreted as describing the clique potential terms defined on singletons and higher order cliques respectively. We shall refer to the first term as the data attachment term E as it is this term that links gray-levels i_s observed in $I=i$ to the means μ_{l_s} of the segmented regions. We will refer to the second term as the clique attachment term W as this term is the sum of all the clique potentials excluding the singletons. The sum of the terms E and W is often referred to as the energy corresponding to a particular segmentation.

2.1. Data attachment terms, discrete case

As we see from eq. 4, the potential function for the singletons is Gaussian and if all the assumptions allowing the deduction of eq. 4 were to hold, this would effectively be the optimal shape for these functions. However, these assumptions are rarely realized in real world images and many authors have experimented with other forms for these functions. We have previously reported using piecewise linear potential functions for simultaneous fusion and discrete label segmentation of multi-echo MR images in [1] where the data attachment term E in site s is expressed by:

$$E_{\{s\}}(l_s)=\sum_{m=1}^{N_i} \lambda_{m,l_s}\Phi_{(m,l_s)}(g_m(s)) \quad (5)$$

where N_i is the number of echoes to be fusioned, the terms λ_{m,l_s} are confidence factors indicating the confidence we have in echo m representing class l_s, the functions Φ are the piecewise linear potential functions (see fig. 1) and $g_m(s)$ is the gray-level in site s in echo m.

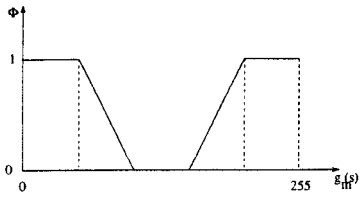

Figure 1: Potential function $\Phi_{(m,l_s)}(g_m(s))$. The lower the value of this function for a certain gray-level $g_m(s)$, the better the correspondence to the corresponding class l_s.

2.2. Clique attachment terms, discrete case

The calculation of the potential functions for the higher order cliques is typically restricted to the second order 8-connected cliques to which site s belongs, and the clique attachment term in site s can, in the discrete case, be calculated by:

$$W_{\{s\}}=\sum_{c|s\in\mathcal{C}_{2,8}} d(l_s,l_r) \quad (6)$$

where $\mathcal{C}_{2,8}$ is the set of all second order, 8-connected cliques on the set S of sites, l_s and l_r are the discrete labels in sites s and r respectively and d is some distance measure (in [1] we use $d(l_s,l_r)=\delta(l_s\neq l_r)$).

3. PROPOSED ALGORITHM

The proposed algorithm is an extension of the continuous label Bayesian segmentation approach described above. We now discuss the necessary modifications of the data and clique attachment terms in order to perform continuous label segmentation.

3.1. Data attachment terms, continuous case

In the continuous case, the data attachment term E in site s is expressed by:

$$E_{\{s\}}=\sum_{m=1}^{N_i}\sum_{n=1}^{N_c} \lambda_{m,n}\left|1-\Phi_{(m,n)}(g_m(s))-\xi_n\right| \quad (7)$$

where N_c is the number of classes, ξ_n is the nth label assigned to site s and the other parameters as defined before.

Calculating $E_{\{s\}}$ in this way therefore amounts to comparing the value of each element ξ_n of the label vector $\vec{\xi}$ with the value of the corresponding potential function $1-\Phi_{(m,n)}$. If the vectors differ from what is expected based on the $\Phi_{(m,n)}$ functions, this increases the value of the data attachment term and thus represents an unfavorable constellation.

Notice that if $\xi_n=1$ for a particular n, that is, only one single tissue is present in the corresponding site, and if this corresponds to a consistent segmentation, then eq. 7 becomes eq. 5 which clearly

shows that eq. 7 is an extension to the continuous case of eq. 5.

3.2. Clique attachment terms, continuous case

In the continuous case, the clique attachment term W in site s is expressed by:

$$W_{\{s\}} = f\left[\sum_{n=1}^{N_c} \sum_{c|s\in\mathcal{C}_{2,s}} d(\xi_{n_s}, \xi_{n_r})\right] \qquad (8)$$

where ξ_{n_s} and ξ_{r_s} are the n-th components of the label vector $\vec{\xi}$ in sites s and r respectively and d a distance measure given by $d(\xi_{n_s}, \xi_{r_s}) = |\xi_{n_s} - \xi_{r_s}|$ (other distance measures could obviously be used, this one seems to provide good results in practice). The argument of the function f is thus simply the sum of measures of the distances between the label vectors in each clique to which s belongs.

A number of choices exist for the functional form for the function f of eq. 8. We have chosen to use the form $f(x) = 1 - \frac{1}{1+|x|}$ suggested in [4]. This type of function has the advantage of preserving discontinuities due to its asymptotic behavior as the argument becomes large.

3.3. Generation of random vectors

Performing the energy minimization by simulated annealing makes it necessary to draw new vectors $\vec{\xi}$ at random to be used in the minimization process.[1]

This poses a problem as these vectors must satisfy the conditions $\sum_{n=1}^{N_c} \xi_n = 1$ and $\xi_n \in [0,1]$. The problem therefore reduces to generating vectors $\vec{\xi}$ that are randomly distributed on a part of the hyperplane given by the equation $\sum_{n=1}^{N_c} \xi_n = 1$. In practice, we seek vectors $\vec{\xi}$ that are uniformly distributed on this part of the hyperplane.

Several methods exist for generating such vectors, the one we have implemented makes use of the conditional probabilities of each element ξ_n knowing the values of the elements ξ_m, $m = 1, ..., n-1$. It is possible to show that these conditional distribution functions are given by the expression:

$$p_{\Xi_n|\Xi_1,\ldots,\Xi_{n-1}}(\xi_n|\xi_1,\ldots,\xi_{n-1}) = \begin{cases} \frac{(N_c-n)}{(1-\sum_{j<n}\xi_j)^{(N_c-n)}}(1-\sum_{j<n}\xi_j-\xi_n)^{N_c-1-n}, \\ \text{if } \xi_n \in [0, 1-\sum_{j<n}\xi_j] \\ 0, \text{ otherwise} \end{cases}$$

In practice we generate the first $N_c - 1$ elements of $\vec{\xi}$ according to their conditional distributions given above. Element ξ_{N_c} is given by $\xi_{N_c} = 1 - \sum_{n=1}^{N_c-1} \xi_n$, the validity of this last operation is easily proved.

[1]The authors would like to thank Olivier Catoni of the Ecole Normale Supérieure, Paris, France for his invaluable help on this problem.

4. RESULTS AND DISCUSSION

In this section we discuss the results we obtain using the proposed method. Fig. 2 shows a synthetically generated projection of a half-sphere that we use for this study as well as its mesh representation. Fig. 3 shows the results we obtain segmenting this

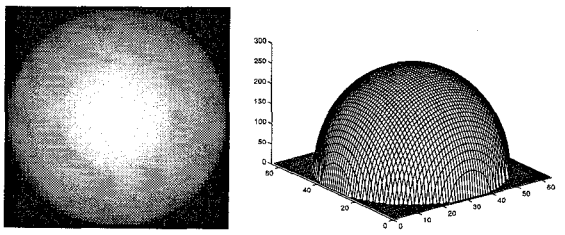

Figure 2: Synthetic image data and its mesh representation.

image into two classes, *background* and *half-sphere*.

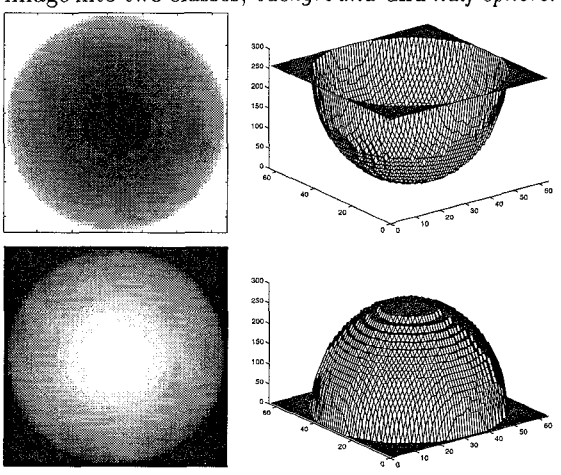

Figure 3: Segmentation result on the synthetic image data. The upper row of images in this figure shows the segmentation result for the *background* class and its mesh representation. The bottom row shows the same results for the class *half-sphere*. Labels range from 0 to 255 indicating the percent content of a particular tissue in a particular site s, high label values indicate high tissue content. Notice that for the class *background* this content is high on the image borders and low as we approach the central parts of the image. The converse is true in the result for the class *half-sphere*.

Fig. 4 shows a subregion of a multi-echo MR acquisition of the brain of a patient suffering from (2) adrenoleukodystrophy, the pathological region appears bright in the rightmost echo. Segmenting these images into three classes, *brain*, *cerebrospinal fluid* and *pathology* by fusing the information present in each of them, we obtain the results shown in fig. 5 where we show only the results for the classes *cerebrospinal fluid* and *pathology*. Notice that whereas central parts of the tissues all are segmented as belonging completely to their respective classes, the border zones have labels indicating that these re-

gions contain numerous partial volumes. Notice also that the transitions in these border zones are smooth indicating gradual changes between classes. The use of continuous labels in this case allows for a more precise estimation of true tissue volumes than could be obtained using a discrete label segmentation.

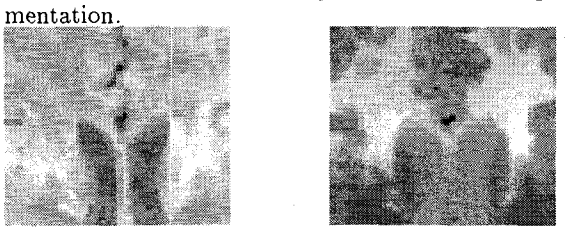

Figure 4: Real patient MR image data

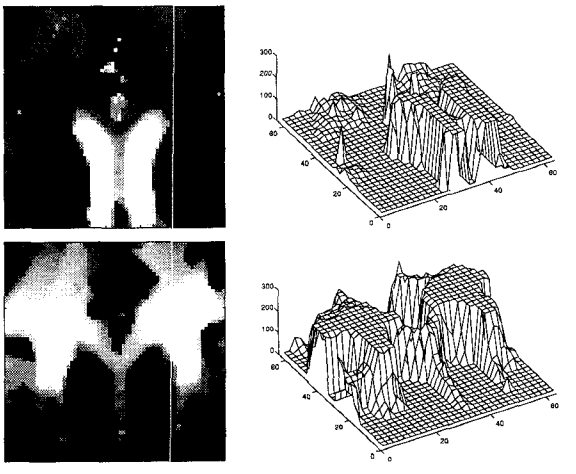

Figure 5: Segmentation result on real patient image data.

5. CONCLUSIONS

In this article we have presented a novel approach to continuous label segmentation. Compared with fuzzy segmentation approaches such as the fuzzy c-means algorithm, it has the great advantage of being more robust to noise due to the regularization inherent to this algorithm. Comparing our algorithm with that of Choi et al., with which it has much in common, our method is seen to have several advantages. Whereas Choi used Gaussian potential functions, we use piecewise linear potential functions making it easier to adapt our algorithm to real world problems. Choi used the ICM algorithm for minimization whereas we use simulated annealing, this minimization method has the advantage of guaranteeing theoretical convergence to a *global* minimum of the energy. We are able to use the simulated annealing approach due to the scheme we propose for rapidly generating random vectors $\vec{\xi}$. Finally, we introduce the function f in the clique attachment terms which allows for limiting the penalty incurred to abrupt gray-level changes, this has the effect of preserving edges in the images. The proposed method provides good results on synthetic and real patient MR image data.

6. REFERENCES

[1] L. Aurdal, X. Descombes, H. Maître, I. Bloch, C. Adamsbaum, and G. Kalifa. Fully Automated Analysis of Adrenoleukodystrophy from Dual-Echo MR Images. In *Proceedings CAR-95, Berlin*, pages 35–40, June 1995.

[2] H. Choi, D. Haynor, and Y. Kim. Partial Volume Tissue Classification of Multichannel Magnetic Resonance Images- A Mixel Model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, September 1991.

[3] M. Clark, L. Hall, D. Goldgoff, L. Clarke, R. Velthuizen, and M. Silbiger. MRI Segmentation Using Fuzzy Clustering Techniques. *IEEE Engineering in Medicine and Biology*, pages 730–742, November 1994.

[4] D. Geman and G. Reynolds. Constrained Restoration and the Recovery of Discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3):367–383, March 1992.

[5] S. Geman and D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, November 1984.

[6] B. Johnston, M. Atkins, B. Mackiewich, and M. Anderson. Segmentation of Multiple Sclerosis Lesions in Intensity Corrected Multispectral MRI. *IEEE Transactions on Medical Imaging*, 15(2):154–169, April 1996.

[7] T. Langenberger, R. Meuli, P. Schroeter, P. Maeder, J. Vesin, and P. Schnyder. Automatic Segmentation and Volume Measurements of Brain with MRI. In *Proceedings CAR-96, Paris*, pages 219–224, June 1996.

[8] P. Santago and H. Gage. Quantification of MR Brain Images by Mixture Density and Partial Volume Modeling. *IEEE Transactions on Medical Imaging*, 12(3):566–574, September 1993.

[9] A. Solberg, V. Bosnes, and G. Storvik. Tissue Classification in MR Images based on a Mixed Pixel Model. In *Procedings of the 9th Scandinavian Conference on Image Analysis, Uppsala, Sweden*, pages 113–119, June 1995.

[10] K. Vincken, A. Koster, and M. Viergever. Probabilistic Segmentation of Partial Volume Voxels. *Pattern Recognition Letters*, pages 477–484, May 1994.