Segmentation of Echocardiographic Data. Multiresolution 2D and 3D Algorithm Based on Grey Level Statistics

Djamal Boukerroui¹, Olivier Basset¹, Atilla Baskurt², and Alison Noble³

¹ CREATIS, CNRS Research Unit (UMR 5515) and affiliated to INSERM
INSA 502, 69621 Villeurbanne cedex, France

{djamal.boukerroui,olivier.basset}@creatis.insa-lyon.fr

² LIGIM (EA 1899), Claude Bernard University, Lyon, France
abaskurt@ligim.univ-lyon1.fr

³ Departement of Engineering Science
University of Oxford, Oxford OX1 3PJ, UK
noble@robots.ox.ac.uk

Abstract. We propose in this paper a robust adaptive region segmentation algorithm of dirty images in a Bayesian framework. A multiresolution implementation of the algorithm is performed using a wavelets basis. The algorithm can process both 2D and 3D data. In this work we focus on the adaptive character of the algorithm and we discuss how global and local statistics can be taken into account in the segmentation process. Results of segmentation performed on echocardiographic sequences (2D+T) and an evaluation of the performance of the proposed algorithm are presented.

1 Introduction

The first step toward automatic analysis or evaluation of a given image is generally considered to be successful segmentation. In recent years, many authors have applied Bayesian estimation techniques for image segmentation. These statistical approaches have improved the segmentation results of different image modalities (natural scenes [2], texture images [4], ultrasonic images [1,10], ...). The segmentation results depend largely on the estimation of the model region parameters that can be based on local/global statistics and can be adaptive. In previous work [1], we have presented an adaptive segmentation algorithm based on texture characterisation. This paper proposes an improvement on the adaptivity by introducing an enhancement to control the adaptive properties of the segmentation process.

The paper is organised as follows. Section 2 describes, briefly, the mathematical framework of our approach and the proposed improvement to take into account both global and local statistics of the data in the optimisation process. A brief description of the algorithm and its implementation is given in Sect. 3. Section 4 presents the segmentation results on echocardiographic sequence images and a comparison of the computer-generated boundaries to the hand-outlined by a medical expert. The paper's conclusions are summarised in Sect. 5.

2 Modelling

We assume that the observed data Y is a random field defined on a 2D (3D) rectangular grid S . Y_s denotes the value of Y at the site $s \in S$. A segmentation of the image (volume) into regions will be noted by X , where $X_s = i$ means that the pixel (voxel) at s belongs to region i . The probability $P(X = x)$ is written as $P(x)$. The number of different regions in X is k . The conditional density function of Y given X is assumed to exist and to be strictly positive and is denoted by $P(y|x)$. The image may be segmented by estimating the pixel classification X given the observed image Y using the maximum *a posteriori* (MAP) estimation of X . Hence, once the distributions of $P(y|x)$ and $P(x)$ are defined, the problem of segmenting an image is reduced to that of minimising an energy function.

We use a Markov Random Field to model the region process X , due to its restriction to local interaction. With this assumption, according to Hammersley-Clifford theorem [9], and for a given neighbourhood system, the prior density $P(x)$ can be written as a Gibbs density where:

$$\ln P(x) \cong \sum_{c \in C} V_c(x) , \quad (1)$$

$$\text{and } V_c(x) = V_{\langle s, t \rangle} = \beta_{\langle s, t \rangle} (1 - \delta(x_s, x_t)) . \quad (2)$$

Here, $V_c(x)$ are the clique potentials and δ is the Kronecker delta function.

We use a 8-connected (6-connected) spatial neighbourhood for 2D (3D) lattice. $\beta_{\langle s, t \rangle} = \beta$ if the clique $\langle s, t \rangle$ is horizontal or vertical and $\beta_{\langle s, t \rangle} = \beta/\sqrt{2}$ if the clique $\langle s, t \rangle$ is right or left diagonal. The Gibbsian parameter β is positive, so that two neighbouring pixels are more likely to belong to the same class than to different classes. Increasing β value increases region size and leads to excessive smoothing of boundaries.

The conditional density distribution $P_s(y_s | x_s = i)$ of the observed grey level intensity at a site s is assumed to be Gaussian, with mean μ_s^i and variance $(\sigma_s^i)^2$. The local class mean μ_s^i is a slowly varying function of s . These assumptions lead to the following energy function:

$$U(x|y) = \sum_s \left(\ln(\sigma_s^{x_s}) + \frac{1}{2(\sigma_s^{x_s})^2} (y_s - \mu_s^{x_s})^2 \right) + \sum_{\langle s, t \rangle} V_{\langle s, t \rangle}(x) . \quad (3)$$

This function has two components. The first term constrains the region intensity to be close to the data and the second is a regularisation term, which imposes a smoothness constraint. Computation of the exact minimum of the energy function is time consuming. As an alternative to Simulated Annealing [11], the *Iterated Conditional Mode* algorithm [8] has been used, which does not guarantee a global minimum of the energy function but is a fast deterministic algorithm. Starting from an initial segmentation x_0 the algorithm updates the label of sites in x in order to maximise the conditional density function at each site knowing the labels values at its neighbourhood and the observation y .

2.1 Parameter Estimation

A rectangular window is used for the estimation of the local class means and variances for all region types i and at each site s . Given the region labels x , μ_s^i and σ_s^i are the average and the standard deviation of pixels (voxels) of the region i inside a window of width W centred at s . This is achieved using the Maximum Likelihood Estimation. Initially, for robust estimation, the window size W is equal to the whole image (volume) and then, as the algorithm progresses, the segmentation becomes better and smaller windows give more reliable and accurate estimations. Thus, the algorithm fits progressively to local characteristics of each region. In our implementation, as in [2], the window size is reduced by a factor of two until a final value $W_{\min}$. This gives the adaptive character to the algorithm. A major difficulty of this algorithm, as it stands, is that it takes into account the global statistics of the regions only in the first few iterations. Indeed, when the algorithm starts from a bad initial solution or when the statistics of image regions are not Gaussian (due to the presence of multiplicative noise, gradient of intensity, outliers, ...) the adaptive properties decrease the robustness of the algorithm and detain it to converge to the optimal solution.

2.2 The Proposed Energy Function

We propose, in this paper, to weight the energy function to take into account the global statistics of the image. Then the energy function becomes:

$$U(x|y) = \sum_s \frac{\varphi(W_s^{x_s})}{N_s^\varphi} \left(\ln(\sigma_s^{x_s}) + \frac{1}{2(\sigma_s^{x_s})^2} (y_s - \mu_s^{x_s})^2 \right) + \sum_{\langle s,t \rangle} V_{\langle s,t \rangle}(x), \quad (4)$$

where $W_s^{x_s}$ is a weighting coefficient which measures the similarity between the local statistics of the region x_s and the global ones. φ is a transformation function and N_s^φ is a normalised constant of the function φ at the site s defined by the following condition:

$$\sum_{x_s} \varphi(W_s^{x_s}) = k. \quad (5)$$

Here, as already noted, k is the number of regions.

In the case of a segmentation problem, the similarity measure that is largely used to compare two distributions is the Kolmogorov-Smirnov distance [3,4,5]. So we take $W_s^{x_s} = D(y(x_s), y(s, x_s))$; where D is the Kolmogorov-Smirnov distance between the cumulative distribution functions of the whole sites of the class x_s and the sites of the same class in the analysed window W centred at s . The introduction of the transformation function φ aims to control the adaptive term and the weighted one. We define φ as:

$$\varphi(W_s^{x_s}) = A_s + (W_s^{x_s})^b \quad (6)$$

$$\text{and } \frac{A_s}{\max_{\text{all class}} [(W_s^{x_s})^b] - \min_{\text{all class}} [(W_s^{x_s})^b]} = a \quad \forall s, \quad (7)$$

where a is a positive constant.

Thus, at each site s , the constant a controls the contribution of the adaptive term A_s (local statistics) and the weight $(W_s^{x_s})^b$ which, takes into account the global statistics. The influence of the control parameters a and b on the φ function is illustrated Fig. 1 (For example for $a \gg 1$, (4) is equivalent to (3)). We use $b = 1$ in the following.

Fig. 1. The influence of the control parameters a and b .

3 Algorithm

In Sect. 2, the conditional density distribution of the observed grey intensity at a site s is assumed to be Gaussian. Clearly this assumption is not true in many cases, especially for displayed ultrasound data. We make use of the Central Limit Theorem as in [10], which states that the distribution of the mean of a large number of independent random observations tends toward a Gaussian distribution centred to their collective mean. This assumption is reasonably acceptable for low-pass filtered and decimated images that are originally governed by non-Gaussian statistics and leads us to a multiresolution implementation of our algorithm.

Hence, a hierarchical structure has been used to implement our algorithm [1,6]. Starting from the highest resolution image, a multiresolution pyramid is built using the Discrete Wavelet Transform (DWT) approach. The coarsest resolution is initially segmented and the result of the segmentation is passed on to the next finer level of resolution and so on, until the finest resolution image is segmented. So a lattice point at one resolution will correspond to four (eight) points at the next finer resolution in the case of 2D (3D) data. An initial solution of the minimisation problem, at the coarsest resolution, is obtained with the K-means clustering algorithm. Starting from this segmentation, the

algorithm alternates between the estimation of region labels and model parameters and is stopped when no further changes in the labels occur. The Gibbsian parameter at each scale is derived from the full resolution one according to the multiscale approach proposed in [7]. In this work, the number of regions k and the Gibbsian parameter β at the full resolution are not estimated. For more detailed description of the hierarchical implementation see [1,7].

4 Results

In this section we present the results of applying our new algorithm on a long-axis echocardiographic sequence of 106 images (2D+T). In the segmentation process the sequence is considered as a 3D volume to enable temporal continuity to be use in the segmentation process. All the results presented in the following are obtained for 3 classes, $\beta = 1$ at the full resolution and for 3 levels of decomposition of the Discrete Wavelet Transform.

Figure 2 shows the segmentation results (superimposed on the original data) for 4 frames (1–4) and for 3 different values of the control parameter a (A, B, C). Images A were obtained for $a = +\infty$ i.e using the energy function in (3). B and C are obtained with $a = 10$ and $a = 1$ respectively. A deviation with respect to the actual boundaries is observed in the left part of the images (right ventricle) when the adaptive character is emphasized (Fig. 2 A), whereas a good detection was obtained when global statistics are included in the segmentation process (Fig. 2 C). As expected, the segmentation is better when both global and local statistics are included in the segmentation process. Our formulation of the problem allows us to control the effective contribution of each statistics by varying the parameter a . There is little difference between results A and B, probably because in experiment B the global statistic is taken with only 10 percent in comparison with the local one. This allows us to have grown some insight into the importance of the parameter a .

An evaluation of the computer detection of the boundary of the left ventricle has been made. The validation of a segmentation algorithm is an important and

Table 1. Perimeter and area error measures of the computer-generated segmentation of the left ventricle relative to measurement made by hand by a medical expert. The minus sign indicates that our estimation is larger than the medical expert's.

	$a = +\infty$	$a = 10$	$a = 1$
Perimeter (pixel size)	$\mu = -28.83 \%$ $\sigma = 8.22 \%$	$\mu = -25.38 \%$ $\sigma = 7.17 \%$	$\mu = -19.52 \%$ $\sigma = 4.99 \%$
Area (pixel size) ²	$\mu = -6.29 \%$ $\sigma = 6.91 \%$	$\mu = -6.03 \%$ $\sigma = 6.92 \%$	$\mu = -7.59 \%$ $\sigma = 6.71 \%$

difficult task. First we computed 2 classical measures (Perimeter and Area) on the 106 frames of the computer-generated and the hand-outlined boundaries. The results are summarised in Table 1. We can see that the perimeter shows us that a little improvement is made when global statistics are involved. However it is quite difficult to make a conclusion using only these two measures.

Recently, a more applicable methodology for evaluation of boundary detection algorithm has been proposed [12]. This methodology uses essentially two distance measures: The Hausdorff Distance (HD) and the Average Distance (AD). The HD between two curves is defined as the maximum of the Distance to the Closest Point's (DCP) between the two curves. The DCP associates each point on both curves to a point on the other curve, and the HD finds the largest distance between the associated points. The AD is the average of all distances between the correspondence points between the curves. This allow us to analyse the regional differencies between the curves.

Table 2 shows the minimum, the maximum, the mean and the standard deviation of the two distances over the whole range of the data. Here also we can see that the HD does not show any significant difference between the performance of the three algorithms. However the AD shows clearly that the third algorithm is better than the two others ($p < 0.0001$ for the Friedman's rank test [12]).

5 Conclusion

In this paper, we have proposed an adaptive region segmentation algorithm based on global and local statistics in a Bayesian framework. The enhancement we developed can be regarded as a generalisation of previous work [1,10]. The formulation of the segmentation problem allows us to control the effective contribution of each statistic. Segmentation results performed on a long-axis echocardiographic sequence (2D+T) has been presented. A comparison of the computer-generated boundaries of the left ventricle to the hand-outlined by a medical expert has been made. Our experiments indicate that taking into account the global statistic, in an adaptive algorithm, improves the segmentation results.

Table 2. Direct comparison of the computer-generated boundaries of the left ventricle to the hand-outlined by a medical expert.

		$a = +\infty$	$a = 10$	$a = 1$
Hausdorff Distance	(min , max)	(8.06 , 23.53)	(8.06 , 23.53)	(8.48 , 22.56)
	(μ , σ)	(14.21 , 3.91)	(13.63 , 3.74)	(13.64 , 3.48)
Average Distance	(min , max)	(2.49 , 17.13)	(2.45 , 15.80)	(1.82 , 11.75)
	(μ , σ)	(8.78 , 3.09)	(8.19 , 2.94)	(5.96 , 2.20)

The first results obtained with our algorithm on echocardiographic data gives satisfying segmentation. However, a comparison of the segmentation results with boundaries outlined by multiple observers will be an important future step to confirm this conclusion. It would be also interesting to compare or to associate our approach with high-level model-guided segmentation, using "deformable-templates" for instance, which includes *a priori* knowledge about the object structure is in the segmentation process.

The proposed method is general and can be applied to other adaptive image segmentation problems that can be posed in a Bayesian framework. Other similarity measures and φ functions could be used.

References

1. D. Boukerroui, O. Basset, N Guérin, and A. Baskurt. Multiresolution texture based adaptive clustering algorithm for breast lesion segmentation. *European Journal of Ultrasound*, 8(2):135–144, 1998. 516, 519, 520, 521
2. T. N. Pappas. An adaptive clustering algorithm for image segmentation. *IEEE Transactions on Signal Processing*, SP-40(4):901–914, 1992. 516, 518
3. R. Muzzolini, Y.-H. Yang, and R. Pierson. Multiresolution texture segmentation with application to diagnostic ultrasound images. *IEEE Transactions on Medical Imaging*, 12(1):108–123, 1993. 518
4. C. Kervrann and F. Heitz. Segmentation non supervisée des images naturelles texturées: approche statistique. *Traitement du signal*, 11(1):31–41, 1994. 516, 518
5. D. Geman. Random fields and inverse problems in imaging. *Ecole d'été de probabilités de Saint-Flour XVIII, Hennequin, P.L., (eds.) (1988), Vol. 1427. Springer-Verlag, Lecture Notes in Mathematics, Berlin*, 161–172, 1990. 518
6. B. C. Vemuri, S. Rahman, and J. Li. Multiresolution adaptive K-means algorithm for segmentation of brain MRI. *Intl. Compu. Sci. Conf. On Image Analysis and Comp. Graphics, Hong Kong*, 347–354, 1995. 519
7. F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59(1):125–134, 1994. 520
8. B. J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society B*, 48(3):259–302, 1986. 517
9. B. J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, 26(2):192–236, 1974. 517
10. E. A. Ashton and K. J. Parker. Multiple resolution Bayesian segmentation of ultrasound images. *Ultrasonic Imaging*, 17(2):291–304, 1995. 516, 519, 521
11. S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983. 517
12. V. Chalana and Y. Kim. A methodology for evaluation of boundary detection algorithm on medical images. *IEEE Transactions on Medical Imaging*, 16(1):642–652, 1997. 521

Fig. 2. Segmentation results for 3 different values of the control parameter a : from left to right $a = +\infty$, $a = 10$ and $a = 1$; frames 1 and 2 : systole, frames 3 and 4 : diastole.