Ewing's Sarcoma Tumours Segmentation in MR Images

Joanna Czajkowska, *Silesian University of Technology*

Abstract

This paper presents the first stage of a semi-automatic method for the segmentation of nodules of the skeleto-muscular system from magnetic resonance (MR) imaging. The method suggested is not sensitive to the location of tumour in human body. It is based on adaptive filtering, a robust kernelised version of Fuzzy C-Means clustering (FCM)-KFCM and Gaussian Mixture Models (GMM). The adaptive filtering algorithm applied improved the results of both pixels classification methods, KFCM as well as GMM [5]. For each MR study, two pixel classification methods mentioned are applied and the segmentation result consists of an area with the highest probability values calculated on the basis of these methods. The suggested algorithm was evaluated on three T2-weighted series of MR images of different parts of the body, where Ewing's sarcomas were indicated by a radiologist.

1. Ewing's Sarcoma

Ewing's sarcoma is a type of cancer usually found in children and young adults between the age of 10 and 20 [1][2]. It occurs in about one teenager in 50000 and it is the most lethal bone tumour. It can arise anywhere in the body, in the bone or close to the bone, it also may involve muscles and soft tissues surrounding the tumour. The most common sites are pelvis, thigh, lower leg, upper arm and rib. The diagnostics of this tumour is still a challenge for medicine, but medical imaging like computed tomography (CT) and magnetic resonance (MR) brought a considerable improvement. MR imaging is the modality of choice for diagnosing bone, muscular system or soft tissue lesions [3]. The diagnostics of sarcomas is based on several types of MR images, especially fast-recovery fast spin-echo T2-weighted or fast spin-echo T1-weighted images with contrast.

The most important task in radiological diagnosis, which is dealt with in this study, is to determine tumour changes between the follow-up studies. Comparing tumour volumes in different patient studies, the radiologists can decide if the therapy applied is effective. Therefore 3-D segmentation of the tumourously changed tissue is required.

2. Methods

2.1. Adaptive filtering

KFCM_S as well as GMM methods are sensitive to image artifacts and noise. Pre-processing filtering is then applied in order to avoid this problem.

Most of the sarcomas tissue is characterized by the pixels density higher than the density of healthy tissue. However, the edges of tumour tissue are often not well defined. Commonly used low pass filters, like median or mean filters can often additionally blur the edges. One of the solutions to this problem is to apply adaptive filtering. In the suggested approach anisotropic diffusion is used as adaptive filtering. It is described in detail in [6].

2.2. KFCM clustering

FCM is a clustering method based on the minimisation of an objective function. It divides a finite collection of elements into k clusters with respect to some given criterion. The number of clusters k is usually predefined [7].

The standard FCM objective function for partitioning a dataset $\{x_i\}_{i=1}^{N}$ into k clusters can be written as

$$J_\beta = \sum_{k=1}^{c}\sum_{i=1}^{N} u_{ki}^{\beta} \left\| x_i - v_k \right\|^2, \qquad (1)$$

β being a weighted exponent on each fuzzy membership.

In each iteration of the FCM algorithm, for each cluster k, fuzzy cluster centre [7]

$$\{v_k^t\} = \frac{\sum_{i=1}^{N} (u_{ki})^{\beta} x_i}{\sum_{i=1}^{N} (u_{ki})^{\beta}} \qquad (2)$$

is calculated using a membership function $U^{(t)}$. In the first iteration, $U^{(t)}$ is set to random values. In the next step

$$U_{ki}^{(t+1)} = \frac{1}{\sum_{j=1}^{c} (d_{ki}/d_{ji})^{\frac{2}{(\beta-1)}}}. \quad (3)$$

If $\left\|U^{(t+1)} - U^{t}\right\| \le \varepsilon$, the algorithm is finished and the analysed data x_i belongs to the class k with maximum u_{ki} value.

However, the above presented algorithm is sensitive to artifacts and noise. The authors of [4] and [5] suggested some modifications, which deal with this problem. One of them-kernel based fuzzy c-means KFCM is used in this work. The algorithm proposed in [5] uses kernel-induced distance metric to replace the squared-norm one applied in FCM. In this modification, the objective function is given as [5]

$$J_{\beta}^{\phi} = \sum_{k=1}^{c}\sum_{i=1}^{N} u_{ki}^{\beta} \left\|\phi(x_i) - \phi(v_k)\right\|^2, \quad (4)$$

where ϕ is a nonlinear transformation into a higher dimensional feature space F and

$$\left\|\phi(x_i) - \phi(v_k)\right\| = K(x_i, x_i)K(v_k, v_k) - 2K(x_i, v_k). \quad (5)$$

Using Gaussian RBF kernel $K(x,y) = \exp(-\left\|x-y\right\|^2/\sigma^2)$, $K(x,x) = 1$, equation (4) can be simplified to

$$J_{\beta}^{\phi} = 2\sum_{k=1}^{c}\sum_{i=1}^{N} u_{ki}^{\beta}(1 - K(x_i, v_k)) \quad (6)$$

and equations (2) and (3) can be written as

$$\{v_k^t\} = \frac{\sum_{i=1}^{N} (u_{ki})^{\beta} K(x_i, v_k) x_i}{\sum_{i=1}^{N} (u_{ki})^{\beta} K(x_i, v_k)} \quad (7)$$

$$U_{ki}^{(t+1)} = \frac{(1 - K(x_i, v_k))^{\frac{-1}{\beta-1}}}{\sum_{j=1}^{c} (1 - K(x_j, v_k))^{\frac{-1}{\beta-1}}} \quad (8)$$

2.3. Gaussian Mixture Model

Mixture model assumes that each group (component) of the data is generated by a probability distribution [7].

For the Gaussian mixtures, each component density has a normal probability distribution

$$p(x/\Theta_m) = \frac{\exp\{-\frac{1}{2}(x-\mu_m)^T C_m^{-1}(x-\mu_m)\}}{\sqrt{2\pi \det(C_m)}} \quad (9)$$

For the mixture model with M (M>1) components

$$p(x) = \sum_{m=1}^{M} \alpha_m p(x/m), \quad (10)$$

where $\alpha_m \in [0,1]$ $(\forall m = 1,...,M)$ and

$$\sum_{m=1}^{M} \alpha_m = 1, \quad (11)$$

α_m denotes the mixing proportion coefficient.

Then, to get $\Theta = (\alpha_1,...,\alpha_M, \Theta_1,...,\Theta_M)$, where $\Theta_m = (\mu_m, C_m)$ equation (10) can be written as follows

$$p(x/\Theta) = \sum_{m=1}^{M} \alpha_m p(x/\Theta_m). \quad (12)$$

The model parameters are estimated by the expectation maximisation (EM) algorithm. It consists of two alternating steps, the Expectation step (E) and the Maximisation step (M).

At each iteration $(t+1)$, in the E step complete data log-likelihood function is computed as follows [7]

$$Q(\Theta/\Theta^t) = \sum_{k=1}^{N}\sum_{m=1}^{M} \{\log \alpha_m \, p(x_k/\Theta_m)\} P(m/x_k; \Theta^{(t)}) \quad (13)$$

where

$$P(m/x_k;\Theta^{(t)}) = \frac{\alpha_m^{(t)} p(x_k/\Theta_m^{(t)})}{\sum_{l=1}^{M} \alpha_l^{(t)} p(x_k/\Theta_l^{(t)})}. \quad (14)$$

In the M step the $(t+1)$ th estimation $\Theta^{(t+1)}$ of Θ is found:

$$\alpha_m^{(t+1)} = \frac{1}{N}\sum_{k=1}^{N} P(m/x_k;\Theta^{(t)}) \quad (15)$$

$$\mu_m^{(t+1)} = \frac{\sum_{k=1}^{N} x_k P(m/x_k;\Theta^{(t)})}{\sum_{k=1}^{N} P(m/x_k;\Theta^{(t)})} \quad (16)$$

$$C_m^{(t+1)} = \frac{\sum_{k=1}^{N} P(m/x_k;\Theta^{(t)})\{(x_k - \mu_m^{(t+1)})(x_k - \mu_m^{(t+1)})^T\}}{\sum_{k=1}^{N} P(m/x_k;\Theta^{(t)})} \quad (17)$$

The estimated model is then used for data clustering. The data are assigned to the group with the maximum conditional probability.

3. Experimental Results

The segmentation algorithm proposed was evaluated on 3 series of fast-recovery fast spin-echo T2-weighted images. The results for one of them are described in detail below. Fig.1 (left) shows, one of the analysed images from MR examination with Ewing's sarcoma visible on the right leg. The tumour is the high-density area around and on the shinbone

or on the kneecap. The image presented was previously subjected to an adaptive filtering procedure using anisotropic diffusion [6]. The result is shown in Fig.2. For each MR image series, pixel intensity is a distinguishing feature that can be used to segment (cluster) the images.

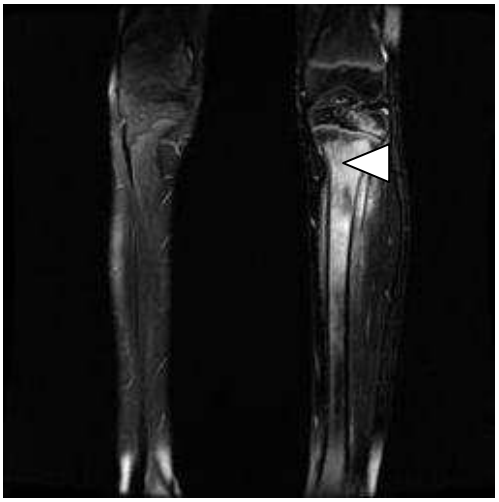

Fig.1. T2-wieghted MR image of Ewing's sarcoma on the right leg.

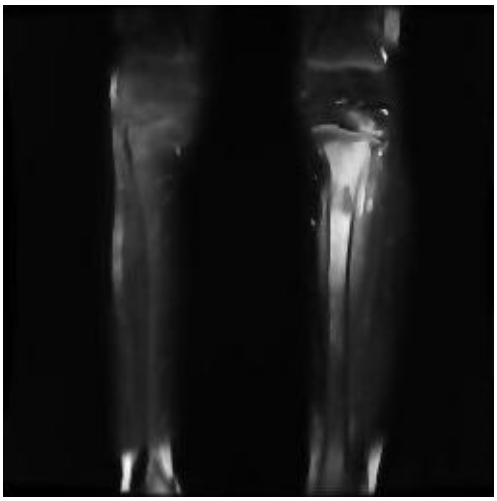

Fig.2. T2-weighted MR image of Ewing's sarcoma after anisotropic diffusion filtering.

As described below, KFCM as well as GMM methods require a priori knowledge about the number of clusters. The tests carried out indicate, that depending on the examination, the methods suggested in different configurations alternately give visually better results. The statistical results presented in Fig.3 show that some of the Gaussian models are similar as far as component parameters and even the number of components are concerned, and the number of components higher than four does not contribute any useful information. It only results in dividing the group of the highest density (with the tumour area) into smaller parts. Therefore, only the results with three and four clusters are taken into account in further analysis. Then, the results of GMM analysis for three and four mixture components are calculated and on this basis the σ value for KFCM algorithm is calculated as a standard derivative of pixel intensity of the segmented areas. The KFCM algorithm with k=3 and k=4 clusters is applied. Probability maps (Fig.4) are created relying on the results obtained. Each pixel has now a value (0.25, 0.5, 0.75, 1) depending on the number of cases, in which it belonged to the group with the highest intensity. Pixels with value >=0.5 create the segmented area. The area of interest is interactively selected on the one (arbitrarily selected) slice, by way of indicating one pixel in it. The results are shown in Fig.5. The selected region is transferred to the 3-D space, as areas adherent to regions indicated on or transferred to the adjoining slices.

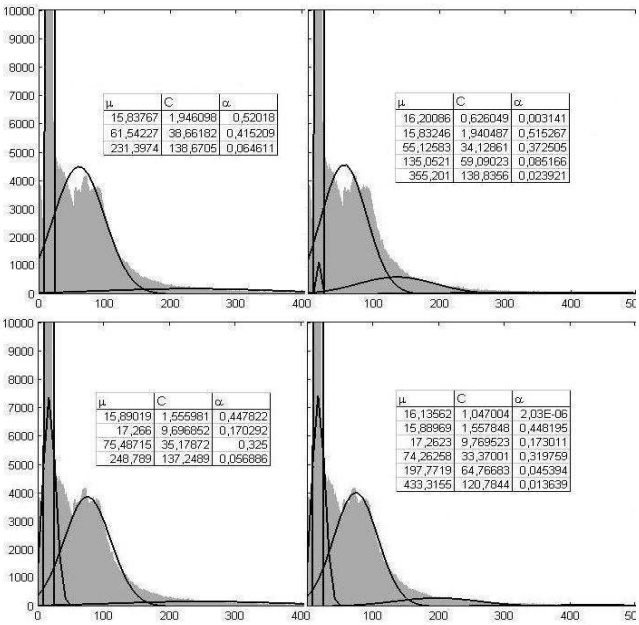

Fig.3. Histograms and Gaussian Mixture Model parameters for different number of mixture components (3,5,4,6).

Fig.4. Probability maps.

At this stage of the presented study incorrectly segmented areas, not containing the tissue of interest, have to be interactively excluded, also by way of indicating one pixel in them.

To sum it up, the method described was tested on 46 MR images. The tumorous tissue is visible on 35 of them. On each of these 35 images the area segmented contains the whole tumorous tissue.

Nevertheless, there appeared also pixels with the same intensity values as the tumour and belonging to the healthy tissue. They should be excluded from the segmented regions during the next steps of processing, so that these results may be used for calculating the tumour volume.

4. Conclusions

In this paper a preliminary step of Ewing's sarcoma tissue segmentation is presented. The segmented regions contain all the tumourous areas and results of combinations of KFCM and GMM algorithms are better than those obtained using only GMM, KFCM or FCM clustering. Nevertheless, the suggested algorithm requires some improvements. The areas, like vessels, with the same pixel intensity values as the tumourous tissue should be removed. Transformations into the 3-D space should also be fully automatic. According to radiologists' opinion, complete diagnostics would also require a juxtaposition of the results obtained from T1-weighted and T2-weighted imaging series and a segmentation of the tumour area on the basis of images acquired in other planes. However, the achieved results are promising and encourage to further develop this method.

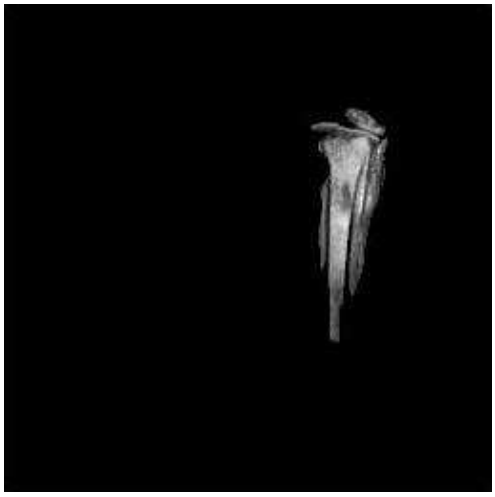

Fig.5. Segmented area.

Acknowledgements

The author would like to thank the medical staff of the Helimed Diagnostic Imaging Centre, Katowice for providing the images.

Bibliography

[1] University of California, San Francisco Children's Hospital, Medical Service, *Sarcomas,* http://www.ucsfhealth.org/childrens/medical_services/cancer/sarcomas/conditions/ewings/signs.html

[2] http://www.cancerindex.org/ccw/guide2e.htm#Menu

[3] Mark J. Kransdorf, Mark D. Murphey: *Imaging of Soft Tissue Tumours*, Lippincott Williams & Wilkins.

[4] Songcan Chen, Daoqiang Zhang: *Robust Image Segmentation Using FCM With Spatial Constraints Based on New Kernel-Induced Distance Measure*, IEEE Transactions on Systems, Man and Cybernetics-part B, Cybernetics, vol. 34, No. 4, August 2004.

[5] Dao-g Zhang, Sgcan Chen, Wi-Song Pan, Ke-Ren Tan: *Kernel-Based Fuzzy Clustering Incorporating Spatial Constrains for Image Segmentation,* Proceedings of the Second International Conference on Machine Learning and Cybernetics, Wan, 2-5 November 2003.

[6] Michael J. Black, Guillermo Sapiro, David H. Marimont: *Robust Anisotropic Diffusion,* IEEE Transactions on Image Processing, vol. 7, No. 3, March 1998

[7] Faiza Mekhalfa, Aïcha Baya Goumeïdane: *Unsupervised Algorithm for Radiographic Image Segmentation Based on the Gaussian Mixture Model*, EUROCON 2007 The International Conference on "Computer as a Tool" Warsaw, September 9-12.

Authors:

Joanna Czajkowska, MSc
Politechnika Śląska
ul. Akademicka 16
44-100 Gliwice
tel. (032) 237 19 88
fax (032) 237 22 25

email: joanna.czajkowska@polsl.pl