Study of Medical Image Segmentation using a Statistical Framework

Anca Loredana (Ion) Udristoiu*, Stefan Udristoiu*

**University of Craiova, Faculty of Automation, Computers and Electronics, Romania*
(phone: 40-251-438-198; fax: 40-251-438-198; e-mail: aion@software.ucv.ro).

Abstract: This paper is a part of a complex study for automatic detection of medical image diagnostic. Thus, a method based on Gaussian statistical model is developed for medical image segmentation. Due to its computation complexity, we used also a method for the Gaussian model simplification using the hierarchical clustering. At the end of this process, an image is represented as a mixture of Gaussian components. The experiments were realized on a medical database, which contains images obtained through different medical proceedings: endoscopy, radiology, magnetic resonance imaging, etc.

Keywords: Image segmentation, image color, image texture, Gaussian mixture model.

1. INTRODUCTION

The success of methods for medical image analysis depends on the quality of segmentation process. The image segmentation represents the partitioning process of the image space in homogenous non-overlapping regions. So, the objective of the current work is to improve the performance of existing approaches in the diagnostication of medical images. In this paper, the first step of this complex study, which means the representation the medical images as a collection of Gaussian mixture components, is realized. We use the Gaussian mixture model, because is widely used in the unsupervised pixels classification of an image, due to its capability of resolving the uncertainty of the mixed pixels.

Thus, a lot of researches were developed to investigate automated techniques for extracting the low-level features that could generate semantic descriptions of the medical image content. Among these techniques are the methods based on machine learning that manually annotate the test image datasets. Algorithms that recognize specific organs with different structures of the medical images are studied in (Hong et al., 2006). FIRE (Deselaers et al., 2004) application and IRMA (Lehmann et al., 2004) use with good results the sub-symbolic processing of images. Though, the actual methodologies of medical image analysis are not generically sufficient for interpreting different diseases.

The medical applications with automatic diagnosis capacity imply unique challenges, but at the same time new opportunities. In some way we understand an image from nature and in another way a medical image, if we are not physicians. On the other hand, there are a lot of formal representations of the medical knowledge that could be exploited to realize the automation of the medical diagnosis in any medical domain.

2. IMAGE REPRESENTATION

2.1 Color feature

Color is a very important feature in many image domains and is the most used feature in the content- based image retrieval systems, because the color characteristic is easy to be detected from images and objects. More, the color is invariant to orientation and scaling and the analysis by color is intuitive. The performance and efficiency of the color feature for characterizing the perceptual similitude of the color is strongly influenced by the selection of the color space and its quantization.

The HSV color space offers an intuitive representation of color and approximates the way in which humans visualize and manipulate color. The transformation from *RGB* to *HSV* color space is nonlinear, but irreversible (Smith et al., 1996) and is realized by (1), (2), (3).

Let be $v_c = (r,g,b)$ a pixel in RGB colour space and $w_c = (h,s,v)$ a pixel in HSV color space. Then:

$$v = \max(r,g,b) \tag{1}$$

$$s = \frac{v - \min(r,g,b)}{v}. \tag{2}$$

Let be,

$$r' = \frac{v - r}{v - \min(r,g,b)},\ g' = \frac{v - g}{v - \min(r,g,b)},$$

$$b' = \frac{v - b}{v - \min(r,g,b)}$$

Then the hue is:

$$\begin{aligned} &h=5+b' \text{ then } r = \max(r,g,b) \text{ and } g = \min(r,b,g)\\ &h=1-g' \text{ then } r = \max(r,g,b) \text{ and } g \neq \min(r,b,g)\\ &h=1+r' \text{ then } g = \max(r,g,b) \text{ and } b = \min(r,b,g)\\ &h=3-b' \text{ then } g = \max(r,g,b) \text{ and } b \neq \min(r,b,g)\\ &h=3+g' \text{ then } b = \max(r,g,b) \text{ and } r = \min(r,b,g) \end{aligned} \tag{3}$$

h=5-r' otherwise.

2.2 Texture feature

The texture is another important characteristic taken into consideration for classifying and recognizing the sick regions of medical images. The frequency elements are very important characteristics for texture analysis. The local frequency of the image regions can be analysed by Fourier transform that offers a frequency/space representation of the image (Zhang et al., 2000).

So, the bi-dimensional image *p(x,y)* is multiplied by the window function *q(x,y)*, followed by the Fourier transform as in (4):

$$W(x0, y0, u,v) = \\ = \int_{-\infty}^{\infty} p(x,y)q(x-x0,y-y0)e^{-2\pi i(ux+vy)}dxdy = \\ = e^{-2\pi i(ux0+vy0)}[p(x0,y0)*m_{u,v}(x0,y0)]. \quad (4)$$

where u, v are the horizontal and vertical frequencies, and (x_0, y_0) is the image location where the frequency is computed.

The Gabor filter (Zhang et al., 2000) is of form as in (5), (6):

$$m_{f,\varphi}(x,y) = \frac{1}{2\pi\tau^2\lambda} e^{\frac{-\frac{1}{2}x'^2/\lambda^2 + y'^2}{\tau^2}} e^{2\pi i f x'}. \quad (5)$$

$$M_{f,\varphi}(u,v) = e^{-2\pi\tau^2[(u'-f)^2\lambda^2 + v'^2]}. \quad (6)$$

with the center frequency $f = \sqrt{u_0^2 + v_0^2}$ and with the rotation coordinates:

$$(x', y') = (x\cos\varphi + y\sin\varphi, -x\sin\varphi + y\cos\varphi)$$

After the multiplication of the Fourier transform of the image P(u, v) and of the Gabor filter $M_{f,\varphi}(u, v)$ (Zhang et al., 2000), the inverse Fast Fourier Transform is applied as in (7).

$$g_{f,\varphi}(x,y) = FFT^{-1}\{P(u,v)\cdot M_{f,\varphi}(u,v)\}. \quad (7)$$

The texture characteristics are computed by adding the quadratic filter outputs over all the spatial coordinates x and y, as in (8):

$$G_{f\varphi} = \sum_{x,y} g_{f,\varphi}^2(x,y). \quad (8)$$

We interpret the hue and saturation channels like polar coordinates to allow the direct use of the HSV color space for Fourier transform (Palm et al., 2000). This technique is used for the extraction of the Gabor characteristics for color texture.

The color space HSV is a non-linear transformation of the RGB color space. The H, S, V components closely correspond to the human color perceptions. Starting from the representation of the HSV color space, we may represent the color in complex.

The affix of any point from the cone base can be computed as: $z_M = S\,(\cos H + i \sin H)$.

So, the saturation is interpreted as the magnitude and the hue as the phase of the complex value b; the value channel is not included. The advantages of this representation of complex color are the simplicity due to the fact that the color is now a scalar and not a vector, and the combination between channels is done before filtering.

So, the color can be represented in complex (Palm et al., 2000):

$$b(x,y) = S(x,y)\cdot e^{iH(x,y)}. \quad (9)$$

The computation of Gabor characteristics for the image represented in the HS-complex space is similar with the one for the unichrome Gabor characteristics, because the combination of color channels is done before filtering:

$$C_{f,\varphi} = (\sum_{x,y} (FFT^{-1}\{P(u,v)\cdot M_{f,\varphi}(u,v)\}))^2. \quad (10)$$

The Gabor characteristics vector is created using the value $C_{f,\varphi}$ computed for 3 scales and 4 orientations.

3. GAUSSIAN MIXTURE MODEL FOR IMAGE SEGMENTATION

3.1 Image segmentation

To resolve the problem of image segmentation, each region is represented as a parameterized distribution, like Gauss (continuous) or Poisson (discrete) distribution.

A Gaussian mixture model is a powerful framework to estimate the probability density function of a variable. It was widely used in statistic, image and signal processing, physic, biology, finance, information extraction (Udristoiu et al., 2008), etc.

The image is modelled as a mixture of Gaussian distribution, where an individual distribution is used to specify the region of pixels. Thus, the image is modelled as "random field" (Greenspan, 2002), being composed from two collections of two random variables Y and X. The values of the first variable correspond to the classes/regions, while the values of the second variables correspond to "measurements" or "observations" of the pixels. The problem of segmentation consists in determining Y, knowing X. A general method for statistical segmentation is to represent the probability of the density function as a mixture because the data is a combination of individual density of the components which correspond to regions (Bishop, 2006).

Given the data, the task of image segmentation is to identify a set of pixels in it and provide a model for each of the pixels. This is realized using the Expectation Maximization (EM) algorithm, which is an effective and popular technique for estimating the mixture model parameters. It iteratively refines an initial cluster model to better fit the data and terminates at a solution that is locally optimal for the underlying clustering criterion (Dempster et al., 1977).

In our case, the image is represented in the color-texture space meaning 3 components for color and 12 components for texture. For example, we consider an image pixel represented as a vector of dimension *n*, $X=(x_1....x_n)$, $X \in R^n$. If we want to attach a pixel *x* to one of the clusters/components $z_1....z_k$, we have to determine the conditional probability $p(z|x)$.

In conformity with Bayes theorem (Bishop, 2006), the conditional probability is as in (11):

$$p(z|x) = \frac{p(x|z)p(z)}{p(x)}. \quad (11)$$

The pixels could be in one of the clusters with the initial probabilities: $w_1=p(z_1)$, $w_2=p(z_2)....w_k=p(z_k)$.

The conditional probability of *x*, for a given z_k is modelled by a Gaussian distribution parameterized by two parameters μ_k and V_k:

$$p(x \mid z_k) = N(x \mid \mu_k, V_k). \quad (12)$$

In conformity with the product rule (Bishop, 2006), the joint probability is as in (13):

$$p(x,z) = p(z|x)p(x). \quad (13)$$

In conformity with the sum rule (Bishop, 2006), the marginal probability is as in (14):

$$p(x) = \sum_z p(x,z) = \sum_z p(z \mid x)\, p(x) = \sum_z p(z)\, p(x \mid z). \quad (14)$$

Thus, using the sum rule, the mixture density function is:

$$p(x) = \sum_z p(z)p(x|z) = \sum_{k=1}^{K} w_k N(x|\mu_k, V_k). \quad (15)$$

where, μ_k is the mean of the Gaussian mixture and V_k is the covariance matrix of the Gaussian mixture.

The likelihood function is:

$$p(X \mid w, \mu, V) = \prod_{n=1}^{N} \sum_{k=1}^{K} w_k N(x_n|\mu_k, V_k). \quad (16)$$

Where,

$$N(x \mid \mu, V) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|V|^{1/2}} \cdot \exp\{-\frac{1}{2}(x-\mu)^T V^{-1}(x-\mu)\}. \quad (17)$$

The function $p(X \mid w, \mu, V)$ has to be maximized using the expectation-maximization (EM) algorithm. An advantage of expectation-maximization method is that it is capable for handling uncertainties due to mixed pixels and helps in designing multivalued recognition systems (Greenspan et al., 2008; Dempster et al., 1977; Bishop, 2006).

In the next steps, the expectation algorithm for the Gaussian mixture model is described:

- the Gaussian mixture model is given
- the likelihood function is maximized

Step 1: Initialize the means μ_k, co-variances V_k and evaluate initial value of likelihood function.

Step 2: E-step - Evaluate the conditional probability of z_k given x_n, $p(z_k \mid x_n)$ as in (18):

$$\gamma(z_{nk}) = p(z_{nk} \mid x_n) = \frac{w_k N(x_n \mid \mu_k, V_k)}{\sum_{k=1}^{K} w_k N(x_n \mid \mu_k, V_k)}. \quad (18)$$

Step 3: M-step-Re-estimate parameters:

$$\mu_k^{new} = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{nk}) x_n$$

$$V_k^{new} = \frac{1}{N_k} \sum_{n-1}^{N} \gamma(z_{nk})(x_n - \mu_k^{new})(x_n - \mu_k^{new})^T \quad (19)$$

$$w_k^{new} = \frac{N_k}{N};\ N_k = \sum_{n=1}^{N} \gamma(z_{nk}).$$

Step 4: Evaluate the log-likelihood from (20) and check for convergence of either the parameters or the log likelihood. If the convergence criterion is not satisfied, return to step 2:

$$\ln p(X \mid w, \mu, V) = \sum_{n=1}^{N} \ln\{\sum_{k=1}^{K} w_k N(x_n \mid \mu_k, V_k)\}. \quad (20)$$

3.2 Kullback-Leibler divergence

In the GMM-KL framework, the distance measure between two images is defined as a distance measure between the two Gaussian mixture distributions obtained from the images. The matching between images is treated as a distribution matching task, using the information-theoretic motivated Kullback-Leibler (KL) distance (Dempster et al., 1977; Greenspan, 2002). Denote the

Gaussian mixture models computed from the two images by GMM_i and GMM_j:

$$D(GMM_i, GMM_j) = \int GMM_i(x) \log \frac{GMM_i(x)}{GMM_j(x)} dx. \quad (21)$$

If, for instance, GMM_i and GMM_j are two multivariate Gaussian distributions parameterized by their means, μ_i and μ_j and by their covariance matrices, V_i and V_j, the equation leads to a closed form expression of the KL distance (Greenspan et al., 2008):

$$D(GMM_i, GMM_j) = \frac{1}{2} \log \frac{|V_j|}{|V_i|} + \frac{1}{2} tr(V_j^{-1} V_i) +$$

$$+ \frac{1}{2} (\mu_j - \mu_i)^T V_j^{-1} (\mu_j - \mu_i) - \frac{n}{2}. \quad (22)$$

3.3 Hierarchical Clustering

Although Gaussian mixture models are used in many research domains from image processing to machine learning, this statistical mixture modelling is usually complex and need to be simplified (Goldberger et al., 2008; Garcia, 2010). In this paper, we present a GMM simplification method based on a hierarchical clustering algorithm. This algorithm provides a hierarchical representation of the initial Gaussian mixture model and experiments on medical image processing are reported.

Given a set of k Gaussian distribution $GMM_1 \ldots GMM_k$ to be reduced and a k*k similarity matrix, the basic process of hierarchical cluster (Johnson, 1967) is this:

1. Start by assigning each Gaussian distribution to a cluster, so we now have N clusters, each containing just one item. The distances (similarities) between the clusters are the same as the distances (similarities) between the items they contain.
2. Find the closest (most similar) pair of clusters and merge them into a single cluster, so that now you have one cluster less.
3. Compute distances (similarities) between the new cluster and each of the old clusters $D(GMM_i, GMM_j)$.
4. Repeat steps 2 and 3 until all items are clustered into k cluster.

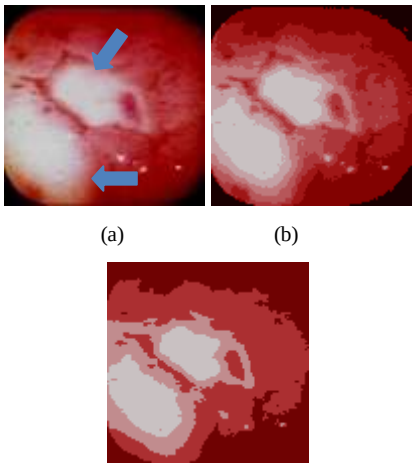
Step 3 can be done in different ways, using *single-linkage* or *complete-linkage* or *average-linkage* clustering. In *single-linkage* clustering, we consider the distance between one cluster and another cluster to be equal to the shortest distance from any member of one cluster to any member of the other cluster. In *complete-linkage* clustering (also called the *diameter* or *maximum* method), we consider the distance between one cluster and another cluster to be equal to the greatest distance from any member of one cluster to any member of the other cluster. In *average-linkage* clustering, we consider the distance between one cluster and another cluster to be equal to the average distance from any member of one cluster to any member of the other cluster. In this paper we use the single-linkage clustering of Gaussian distributions.

3. EXPERIMENTS AND CONCLUSION

The image collections used in our experiments were taken from free repositories on the Internet (Gatrolab, 2010; Gastroenterology, 2010; Radiological images, 2010). The experiments were carried out on images diagnosed with: duodenal ulcer, gastric ulcer, gastric cancer, esophagitis, breast cancer, brain tumours, etc. Through our experiments, we considered 8 and 4 components for the mixture model to observe their results.

In this section, we present experiments realized on images diagnosed with duodenal ulcer, colon cancer, and breast cancer.

The duodenal ulcers can come in different shapes, sizes, and textures (Gastroenterology, 2010), increasing the complexity to diagnose them. For example, the image from Figure 1 shows a single, white-based ulcer. The segmentation results on this image can be observed in Figure 1.



(a) (b)

(c)

Fig.1. Results of segmentation on an image diagnosed with duodenal ulcer: (a) Original image; (b) Image segmented with 8 regions; (c) Image segmented with 4 regions.

In the case of images segmented into 8 components, the interested regions are observed in Figure 2. There are 4 extracted regions of interest in different hues.

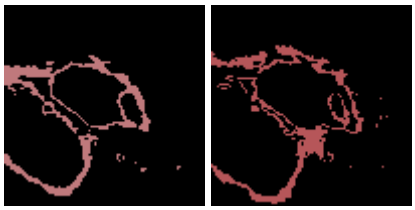


Fig.2. Regions of interest extracted from the image segmented with 8 components.

In the case of images segmented into 4 components, the interested regions are observed in Figure 3. There are only 2 extracted regions of interest in different hues, because the mixture model was reduced to 4 components.

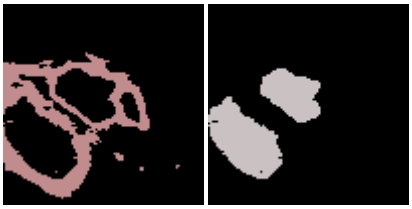
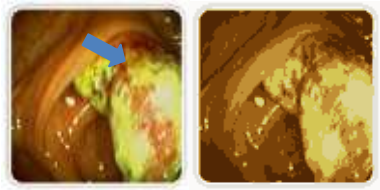


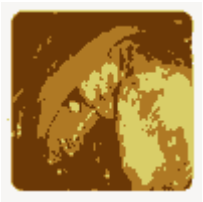
Fig.3. Regions of interest extracted from the image segmented with 4 components.

The image from Figure 4 shows an advanced cancer in the right colon (Gastroenterology, 2010) and the sick region come in different yellow hues.

The segmentation results on this image can be observed in Figure 4.



(a) (b)



(c)

Fig.4. Results of segmentation on an image diagnosed with colon cancer: (a) Original image; (b) Image segmented with 8 regions; (c) Image segmented with 4 regions.

In the case of images segmented into 8 components, the interested regions are observed in Figure 5:

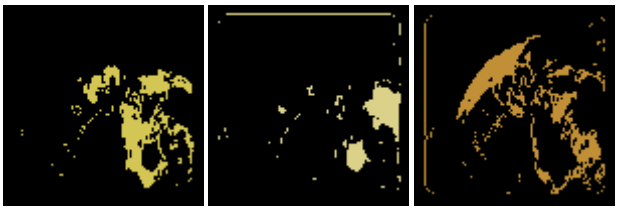


Fig.5. Regions of interest extracted from the image segmented with 8 components.

In the case of images segmented into 4 components, the interested regions are observed in Figure 6:

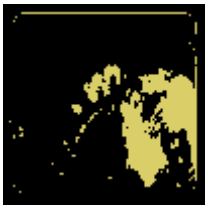
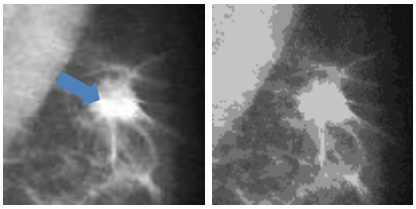
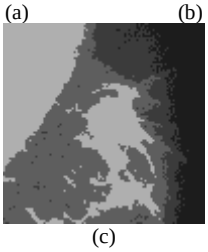


Fig.6. Regions of interest extracted from the image segmented with 4 components.

The image from Figure 7 shows a breast cancer (Gastroenterology, 2010).



(a) (b)



(c)

Fig.7. Results of segmentation on an image diagnosed with breast cancer: (a) Original image; (b) Image segmented with 8 regions; (c) Image segmented with 4 regions.

In the case of images segmented into 8 components, the interested regions are observed in Figure 8.

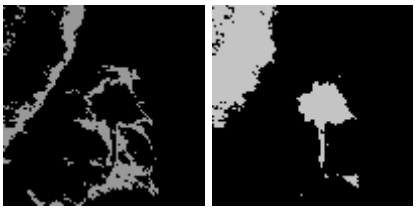


Fig.8. Regions of interest extracted from the image segmented with 8 components.

In the case of images segmented into 4 components, the interested regions are observed in Figure 9.

Fig.9. Regions of interest extracted from the image segmented with 4 components.

By analysing the results of medical images segmentation using the Gaussian mixture model, good results can be observed. If the sick region has different hues then the model with 8 components is indicated, otherwise the one with 4 components can be used.
In the future work, we intend to develop a framework to compare the results of segmentation manually and automatically done.

ACKNOWLEDGMENT

This work was supported by the strategic grant POSDRU/89/1.5/S/61968, Project ID61968 (2009), co-financed by the European Social Fund within the Sectorial Operational Program Human Resources Development 2007-2013.

REFERENCES

Deselaers, T., Keysers, D. and Ney, H. (2004). FIRE - flexible image retrieval engine: ImageCLEF 2004 evaluation. *Multilingual Information Access for Text, Speech and Images*, vol. 3491, pp. 688-698.

Hong, W., Georgescu, B., Zhou, X.S., Krishnan, S., Ma, Y., and Comaniciu, D. (2006). Database-guided simultaneous multi-slice 3D segmentation for volumetric data. *Proceedings of 9th European Conference on Computer Vision*, Graz, Austria, pp. 397-409.

Lehmann, T., Güld, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohne, M., Schubert, H. and Wein, B. (2004). Content-based image retrieval in medical applications. *Methods Inf Med*, vol. 43, no.4, pp. 354-361.

Smith, J.R., Chang, S-F. (1996). VisualSEEk: a fully automated content-based image query system. *Proceedings of the Fourth ACM International Multimedia Conference and Exhibition*. Boston, MA, USA, 1996. pp. 87-98.

Zhang, D., Wong A., Indrawan, M., Lu, G. (2000). Content-based Image Retrieval Using Gabor Texture Features. *The First IEEE Pacific-Rim Conference on Multimedia*, Sydney, 13-15 December 2000, pp 392-395.

Palm, C., Keysers, D., Lehmann, T., Spitzer, K. (2000). Gabor Filtering of Complex Hue/Saturation Images for Color Texture Classification. *Proc. 5 th Joint Conference on Information Science (JCIS2000) 2*, Atlantic City, USA, pp. 45-49.

Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society B*, vol. 39, pp.1–38.

Udristoiu, S. and Ion, A. L. (2010). Image Annotation by Learning Rules from Regions Patterns. *Proceedings of 4th International Conference on Complex, Intelligent and Software Intensive Systems CISIS*, Krakow, Poland, pp.124-131 .

Greenspan, H. (2002). Probabilistic models for generating, modelling and matching image categories. *Proceedings of the International Conference on Pattern Recognition*.

Goldberger, J., Greenspan, H., Dreyfuss, J. (2008). Simplifying Mixture Models Using the Unscented Transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence archive*, vol. 30(8).

Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*, New York: Springer-Verlag.

Garcia, V., Nielsen, F. (2010). Simplification and hierarchical representations of mixtures of exponential families. *Signal Processing*, vol. 90(12), pp. 3197-3212

Johnson, S. C. (1967). Hierarchical Clustering Schemes. *Psychometrika*, vol. 2, 1967, pp. 241-254

Gatrolab (2010). *http://www.gastrolab.net/*.

Gastroenterology (2010). *http://gicare.com/Endoscopy-Center/ Endoscopy-images.aspx*.

Radiological images (2010). *http://www.meddean.luc.edu/lumen/MedEd/Radio/curriculum/Harrisons/Harrisons_f.htm*