A Neural Network Approach to Medical Image Segmentation and Three-Dimensional Reconstruction

Vitoantonio Bevilacqua, Giuseppe Mastronardi, and Mario Marinelli

Dipartimento di Elettrotecnica ed Elettronica, Polytechnic of Bari,
Via E. Orabona, 4, 70125, Bari, Italy
{bevilacqua, mastrona}@poliba.it

Abstract. Medical Image Analysis represents a very important step in clinical diagnosis. It provides image segmentation of the *Region of Interest* (ROI) and the generation of a three-dimensional model, representing the selected object. In this work, was proposed a neural network segmentation based on Self-Organizing Maps (SOM) and a three-dimensional SOM architecture to create a 3D model, starting from 2D data of extracted contours. The utilized dataset consists of a set of CT images of patients presenting a prosthesis' implant, in DICOM format. An application was developed in Visual C++, which provides an user interface to visualize DICOM images and relative segmentation. Moreover it generates a three-dimensional model of the segmented region using Direct3D.

1 Introduction

Image Processing, aimed to interpret and classify image content, has attracted research attention since the birth and spread of computers. Technology improvements have brought more applications to image classification such as *Image* and *Scene Analysis*, *Image Understanding*, *Object Recognition* and *Computer Vision*. Recognition, description and automatic classification of structures inside images represent an important issue in a huge set of scientific and engineering subjects.

The medical research has been quite receptive of image processing like X-ray CT and Magnetic Resonance. The output of these techniques, an image of the patient's body (*slice*), allows physicians examining and diagnosing without the need of surgery. In order to perform these complex medical processes, some operations over the images have to be performed.

A standard was introduced by the American College of Radiology (*ACR*) and the National Elecrical Manufactoring Association (*NEMA*) called DICOM (1993) [3] which is now considered the world standard for transferring and managing medical images.

Image Analysis represents an image processing application directed to extract significant informations for describing internal structures. In particular, image segmentation represents a fundamental step in image analysis since it allows extracting the region of interest to recognize and classify. So this is a critical step: precision and quality of the result can heavily affect next processes. Segmentation aim [4] is to decompose an image in distinct parts or ROIs: each ROI is homogeneous to a particular characteristic and it is separate from adjacent regions. The segmentation of an

D.-S. Huang, K. Li, and G.W. Irwin (Eds.): ICIC 2006, LNCS 4113, pp. 22 – 31, 2006.

image can be carried out by different techniques that are based on the discontinuity and similarity of the grey levels of an image. Some techniques proposed in literature [5, 6] are based on artificial neural networks algorithms. Artificial Neural Networks [7] try to simulate a structure similar to the one that is believed the human brain has got. The area of the brain is organized into several sensory modalities such as speech or hearing. The engineering approach of neural networks develops hardware or software inspired by the brain's structure.

The segmented regions can be processed to generate a three-dimensional model of the selected object. Three-dimensional visualization, in fact, is now a precious tool in biomedical imaging. The medical applicability of 3D image processing is very important, the field of applications are being very large, from the medical statistics to computer assisted tomography [8].

2 Self-Organizing Maps (SOM)

Kohonen's Self-Organizing Maps (SOM) [1, 2] are simple analogies of the human brain's way of organizing information in a logical manner. Kohonen's method emulates the unsupervised learning in an elegant, extremely easy manner. It consists of one layer of neurons and uses the method of competitive learning with "*winner takes all*" logic.

Kohonen's Self-Organizing Maps consist of one layer of neurons organized in one, two and multi-dimensional arrays (fig.1).

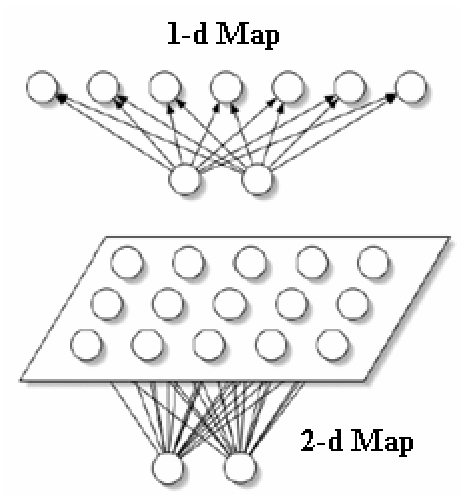

Fig. 1. One-dimensional and two-dimensional Kohonen's maps

Each neuron has as many input connections as there are number of attributes to be used in the classification. All the neurons in the map are interconnected and each neuron influences others which are into its neighbourhood. The lateral coupling of the neurons is thought as a function of the distance in two ways: excitatory and inhibitory.

The excitatory is in a short range up to a certain radius, and the inhibitory surrounds the excitatory area up to a bigger radius. A cluster or bubble around one particular node of the network is formed due to the lateral coupling around a given cell. The primary input determines a "winner" node, which will have a certain cluster, and then, following the input, the winner node with its surrounding cluster or neighbourhood will adapt to the input.

The training procedure consists of finding the neuron with weights closest to the input data vector and declaring that neuron as the winning neuron:

$$\left|w_i - T\right| \leq \left|w_k - T\right| \qquad \forall k = 1 \ldots N$$

where:

w_i is winner neuron
w_k is any neuron
T is the input vector to classify
N is the number of neurons

After finding the winner, the next step is to update the weights' vector of the neurons, as:

$$w_k = w_k(t-1) + \mu \cdot G(i,k)\left(T - w_k(t-1)\right) \quad \forall k = 1 \ldots N$$

where:

$\mu \in [0, 1]$ is the learning rate

$$\begin{cases} t & \text{actual state} \\ t-1 & \text{previous state} \end{cases}$$

The factor $G(i,k)$ is the Gauss function related to the dimension of neuron's neighbourhood. It considers the influence between the winner neuron i and a neuron k in its vicinity. It is defined as:

$$G(i,k) = e^{-\frac{\left|w_k - w_i\right|^2}{2\sigma^2}}$$

where σ is the bubble's radius. At the beginning this parameter is set to a certain value which is decreased at each iteration of the training procedure.

So the training algorithm is iterative and this procedure is repeated a certain number of times. The stop criterion of the algorithm is determined by one of the following:

- σ or μ go down under an arranged value;
- the maximum number of epochs is reached;
- weights' values are stable.

At the end of the training algorithm, the network is seen as a "black box" which associates to each sample of the input pattern the index or weights of a neuron in the map. In this way, the net classifies and labels each element in the input pattern.

3 Medical Image Segmentation

Image segmentation techniques are based on image properties such as intensity, colour or geometry to identify homogeneous ROIs and their contours. There are two main approaches: the first is based on regions and the second on contours. These are based on two properties: similarity and discontinuity.

We have used two main techniques based on the first approach: *Thresholding* and a *Neural Network segmentation* based on Kohonen's Self-Organizing Maps.

Thresholding is the most popular segmentation technique, due to its conceptual and computational simplicity. It uses the property of local similarity of image's pixels to define a certain grey level T, called *binarization threshold*, which permits to subdivide an image in significant regions.

Given a couple of grey values (g_0, g_1), once the threshold value T was defined, the segmentation result of an image $I(x,y)$ is a binary one $I_b(x,y)$ which satisfies the following relation:

$$I_b(x,y) = \begin{cases} g_0 & \text{if } I(x,y) < T \\ g_1 & \text{if } I(x,y) \geq T \end{cases}$$

In the case of medical images in DICOM format, we can use a CT Number, expressed in Hounsfield units, as threshold value instead of its associated grey level. This creates a more efficient segmentation, because the Hounsfield scale is larger than the grey one. The limited range of the grey scale, equal to 255, compresses the Hounsfield scale, up to 4000, with consequent loss of voxels' informations. This makes more difficult to distinguish regions among which there is a small difference in Hounsfield values that becomes almost null in the grey scale.

This segmentation process can be made automatic by a neural network approach based on **Kohonen's self-organizing maps**. Their ability of self-organization permits to classify image's pixels basing on some features of the regions to find out. The objective is to label each pixel grouping them in a certain number of clusters.
In particular, the input pattern is represented by the set of slices of the acquired volume, while the output layer represents the regions to be found.

The SOM network is organized as follows (fig. 2):

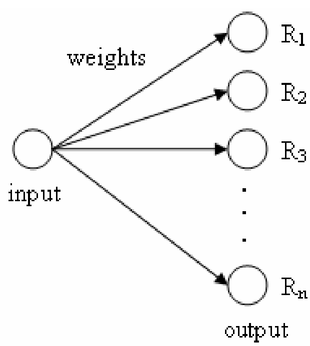

Fig. 2. Kohonen's Self-Organizing Map

- one input neuron represented by the voxel's CT Number;
- n output neurons representing the number of ROIs to extract from the original image.

The map is one-dimensional and the neurons are set on a virtual line, so the distance between two consecutive neurons is unitary.

In the initialization step neurons' weights are set to zero, for assuring the same association order to all slices. In the training procedure the weights will adapt to the CT Numbers of the presented slice. This procedure is repeated for each slice of the acquired volume, while the initialization is performed only for the first slice. In this way the weights will adapt to all CT Numbers of the entire volume. At the end of the training procedure, the net is able to assign an index to each voxel representing its own region.

After segmenting all images and extracting the region of interest, the next step is the Edge Detection. This process creates binary images in which activated pixels identify contours of the extracted regions using a particular operator. In the considered case the most significant is the *gradient* operator. It determines the boundary pixels between two different regions. This process is a simple task if we consider the segmented regions for their definition.

Edge detection is the final step before the next three-dimensional reconstruction starting from 2D contours.

4 Three-Dimensional Reconstruction

The aim of three-dimensional reconstruction is to generate a 3D model of the object extracted by the segmentation techniques which were described previously. The main problem is to generate a polygonal mesh starting from two-dimensional contours placed on different planes. This consists of establishing an order among contours' points in order to generate a well connected mesh which will represent the external surface of the object. Some techniques were proposed in literature [9] which were based on a geometrical and mathematical approach. We have developed an alternative approach based on artificial neural networks.

This approach consists of a particular SOM architecture that is based on a three-dimensional topology. In this network, neurons represent vertices of the triangular mesh to generate, so weights represent three-dimensional coordinates of the mesh's vertices.

The input pattern contains three-dimensional coordinates of the contours extracted from two contiguous slices, so the SOM input layer is composed of three input neurons representing these three coordinates.

The neurons' weights in the three-dimensional self-organizing map are initialized to create a cylindrical mesh which is the first approximation of the external surface (fig. 3). This is an important step because, in this way, the training procedure and so the algorithm convergence is faster.

The activation bubble, in this case, is seen as a sphere which radius' length is defined by the neighbourhood dimension of a neuron in the map. This sphere must contain only vertices placed on a single slice, because the influence on neurons, placed on a different plane, may cause a wrong mesh's deformation along z-axis. This influence can be reduced by increasing the distance between two contiguous slices.

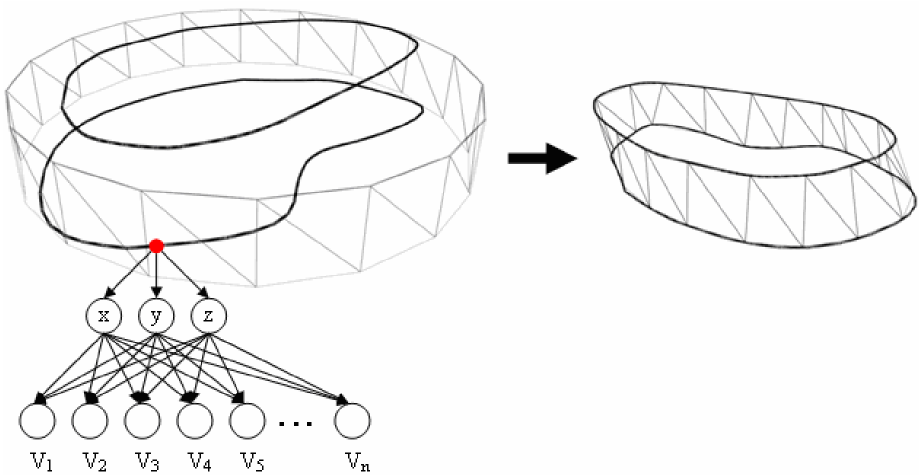

Fig. 3. Weights' initialization to create a first cylindrical mesh and resulting mesh at the end of the training procedure

In the training procedure, each contours' point is randomly presented to the network as input which activates the nearest neuron.

So winning neuron's weight vector is updated to bring the corresponding mesh's vertices closer to the point's input data. In order to maintain similarity of topologically closed neurons, weights of all neurons within a selected radius are adjusted.
In the case where one neuron is continually the winning neuron, its computed distance is also modified by some amount to allow other neurons winning. At each iteration, the radius of correction is also gradually reduced.

As the training procedure progresses, mesh's vertices will adapt to the contours' points generating an ***adaptive triangular mesh***, approximating the external surface**.**

The training procedure ends when one of the following stop criteria is satisfied:

- the maximum number of epochs is reached;
- the approximation coefficient is reached.

The *approximation coefficient* is defined as the mean value of distances between each net input and its corresponding neuron (mesh's vertex). A smaller value of this coefficient will bring the adaptive mesh to a better approximation of the external surface.
This SOM algorithm is repeated for each set of corresponding contours and for each couple of slices.

5 Application and Results

The idea of developing an own application arose from analysis of commercial softwares such as MATLAB and Etdips 2.0.

The application was developed in Visual C++ using Windows APIs to create a graphical and interactive user interface. It is based on a MDI (*Multiple-Document*

Interface) application in order to allow the user working with multiple documents within a single application [10].

The application provides functions for reading medical images in DICOM format, making the image thresholding and SOM segmentation, extracting regions' contours and finally generating and visualizing the 3D model of the extracted objects.

The considered dataset contains CT images of a patient presenting a prosthesis' implant. In particular, it consists of a set of 264 DICOM images (512x512 pixels) representing slices which are 0.8mm thick.

Thresholding technique is defined by two values: a minimum and a maximum threshold. In our dataset, the thresholding values for prosthesis identification are set as follows: $T_{min} = 2000$, $T_{max} = 3000$, according to the Hounsfield scale.

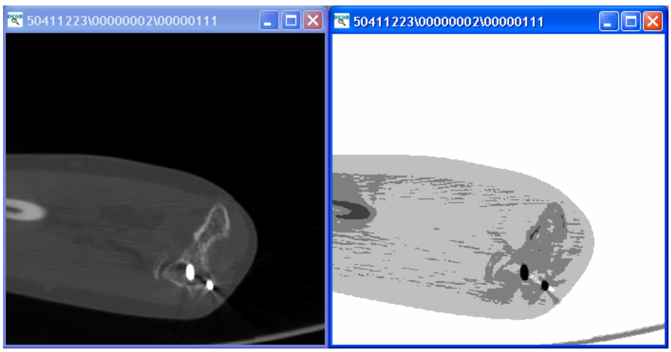

Fig. 4. SOM segmentation result after training procedure

The SOM segmentation technique automates this process through the definition of network parameters: number of neurons (regions to identify) and number of epochs (training iterations). After training and labeling each pixel in the dataset, the result is shown in figure 4: the prosthesis is identified with black colour. The streaks in the segmented images are due to the presence of metal streak artifacts in the original images.

After selecting the desired region (prosthesis), edge detection extracts the region's boundary pixels generating the binary image shown in figure 5.

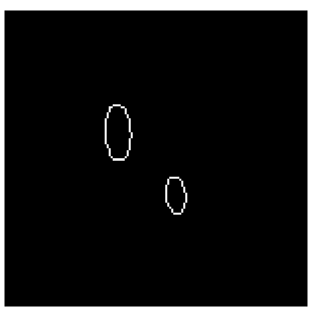

Fig. 5. Particular of the binary image after edge detection

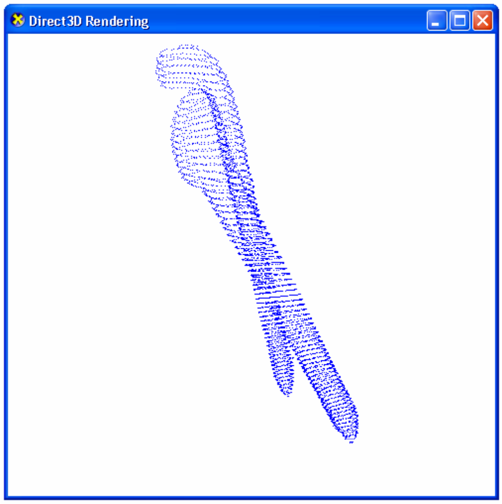

Fig. 6. 3D visualization of the extracted contours

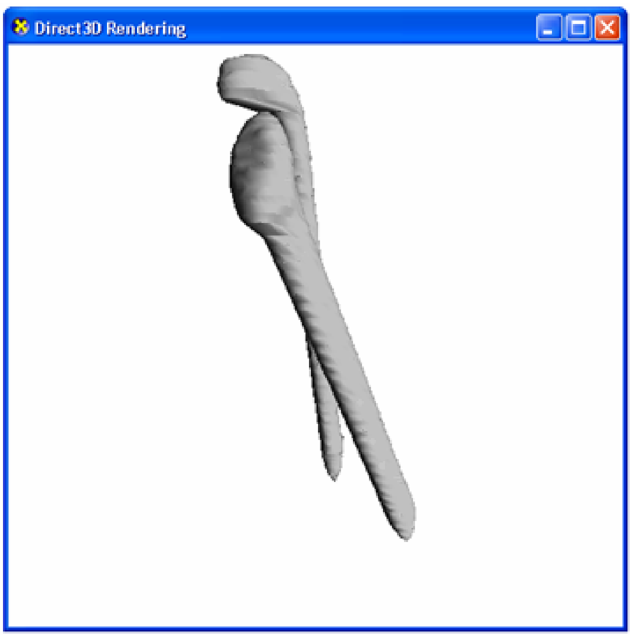

Fig. 7. 3D visualization of the resulting triangular mesh

The final step is to generate a three-dimensional model of the prosthesis starting from the set of the extracted contours. We have used Direct3D APIs to render the 3D model, instead of the traditional OpenGL used by most of the commercial clinical applications. A first 3D model consists of rendering the entire set of contours which gives a first idea about the object to be represented (fig. 6).

The second 3D model consists of rendering the adaptive triangular mesh obtained with the proposed neural algorithm. The resulting mesh is shown in figure 7. The mesh consists of 6096 vertices equal to 1/2 of contours' pixels. The application of lights and texture makes the scene more realistic. The visualization is also interactive, so the user can rotate the object in all the directions using mouse pointer.

To evaluate our application's elaboration timings we must consider the dataset's size, the size of the object to be rendered and the resolution of the 3D model. Running our application on a Pentium III (667 MHz) with 256 megabytes of main memory and considering our dataset, data reading and edge detection timings are negligible, while thresholding technique takes few seconds. SOM thresholding method timings depend on epoch's number: it takes about 15 minutes with 200000 epochs for each slice. 3D reconstruction timings mainly depend on the number of neurons (mesh vertices): it takes from few seconds for low mesh resolution (1/5 of contours' pixels) to about 20 minutes for high mesh resolution (1/2 of contours' pixels).

6 Conclusions and Future Work

Artificial Neural Networks represents an optimal alternative to some complex algorithms for data analysis and particularly for three-dimensional reconstruction. SOM neural networks, in particular, represents an important tool which can be applied to a lot of different problems thanks to their self-organizing capability. In this work we have proposed two different SOM applications: medical image segmentation and three-dimensional reconstruction. The application of this algorithm to 3D model reconstruction is resulted simpler and more intuitive with respect to some other algorithms which are based on mathematical and geometrical approaches.

With the considered dataset, we have noticed that the presence of metal streak artifacts into original images doesn't permit to make a correct segmentation of all the regions in the acquired volume. So only the prosthesis was correctly identified.

Moreover, Direct3D has represented an optimal alternative to OpenGL for making a fast rendering which has permitted to create an interactive 3D visualization.

Further research will be directed to make the SOM segmentation fully automatic using ART (Adaptive Resonance Theory) networks. It will also provide the implementation of three-dimensional segmentation techniques to identify correctly all the regions inside an acquired volume, characterized by images presenting a great presence of streak artifacts.

References

1. Kohonen, T.: Self-Organization and Associative Memory, Springer-Verlag, Heidelberg, (1988)
2. Kohonen, T.: Self-Organizing Maps: Optimization and Approaches, ICANN, Espoo, Finland, (1991)
3. Digital Imaging and Communications in Medicine (DICOM), National Electrical Manufacturers Association, (2004)

4. Fu, K.S., Mui, J.K.: A Survey of Image Segmentation. Pattern Recognition, vol. 13, n.1,(1981) 3-16
5. Gelenbe, E., Feng, Y., Krishnan, K. R. R.: Neural Network Methods for Volumetric Magnetic Resonance Imaging of the Human Brain. IEEE, (1996) 1488–1496
6. Reddick, W. E., Glass, J. O., Cook, E. N., Elkin, T. D., Deaton, R. J.: Automated Segmentation and Classification of Multispectral Magnetic Resonance Images of Brain Using Artificial Neural Networks. IEEE Trans. Medical Imaging, 16(1997)911–918
7. L. Fausett.: Fundamentals of Neural Networks – Architectures, Algorithms and Applications
8. Stănciulescu, V.: 3D Tomography: Computerized Modeling of 2D scanned images. Results. Advantages. General Inspectorate for Communications and Information Technology Timişoara, Romania
9. Galin, E., Akkouche, S.: Fast Surface Reconstruction from Contours using Implicit Surfaces, Laboratoire d'Informatique Graphique Image et Modèlisation
10. Petzold, C.: Programming Windows, Microsoft Programming Series, (1998)