Optimization Neural Networks for the Segmentation of Magnetic Resonance Images

S. C. Amartur, D. Piraino, and Y. Takefuji

Abstract— **Segmentation of the images obtained from magnetic resonance imaging (MRI) is an important step in the visualization of soft tissues in the human body. The multispectral nature of the MRI has been exploited in the past to obtain better performance in the segmentation process. The new emerging field of artificial neural networks promises to provide unique solutions for the pattern classification of medical images. In this preliminary study, we report the application of Hopfield neural network for the multispectral unsupervised classification of MR images. We have used winner-take-all neurons to obtain a crisp classification map using proton density-weighted and T_2-weighted images in the head. The preliminary studies indicate that the number of iterations to reach "good" solutions was nearly constant with the number of clusters chosen for the problem.**

I. Introduction

MAGNETIC resonance imaging (MRI) has an unique advantage over other modalities in that it can provide images of tissues with a variety of contrast based on a simple adjustment of parameters which define the experiment. In a sense the images obtained from MRI resemble the multiband or multispectral images of the earth obtained from the remote sensing satellites. The unique advantages of multicontrast images obtained from MRI for the automated segmentation of classification of tissues is well known and has been successfully employed in the past (also known in the literature as sensor fusion) [2]–[8]. The analysis of such multicontrast images can be achieved by using parametric or nonparametric supervised methods or with unsupervised methods [1]. The supervised classification methods employ a training set consisting of features with known *a priori* probability distributions to train the algorithm (Bayesian approach). The unsupervised methods classify samples without the aid of a training set. The algorithm finds natural groupings or homogeneous clusters in a pattern set based on a criteria defining some property such as intensity or texture. The disadvantage of the supervised methods is the need for the generation of training sets and the associated human interaction and that of the unsupervised methods is its ignorance of the relationships between different objects and a tendency to generate too many clusters or groupings which can be misleading. A hybrid scheme which uses an unsupervised algorithm to form clusters in the first stage and use the statistical properties within the clusters provided in the first stage to train a supervised algorithm in the second stage has been reported recently [8].

Artificial neural networks (ANN) are relatively new computing systems whose architecture is made of massive number of densely interconnected simple analog processing elements. The processing is done in parallel either in a synchronous or asynchronous mode. The architecture of ANN is modeled after the human nervous system with some unique processing capabilities which are not found in the conventional, sequential computing systems. One such processing task in which the ANN excels is in the area of pattern recognition. An excellent review of this new technology can be found in [9]. A few papers on the application of ANN for medical image processing have appeared in the recent past [10]–[12]. All the reported papers have used a specific ANN architecture known in the literature as generalized perceptrons (GP) or back propagation networks (BPN). These networks use a training set very similar to the conventional supervised methods mentioned earlier with the exception that, no *a priori* probabilistic knowledge is required (similar to the conventional nonparametric methods). Another neural architecture which has attracted considerable attention in the recent past was proposed by Hopfield [13]–[16]. This type of network has been proposed for the unsupervised classification of patterns [17]–[18]. In [18], the images are modeled as a Markov random field and assigns region labels or classes to pixels by minimizing the expected percentage of misclassified pixels on a Hopfield network.

In this paper we will outline a scheme for the crisp unsupervised classification of MRI images based on energy minimization using the Hopfield network. An energy function will be formulated which allows for the hyperellipsoidal cluster distribution in the pattern space [1]. We present results from subjects with a normal and an abnormal physiological condition. The paper concludes with some discussion and comments and a brief summary of future work.

II. Hopfield Neural Network

Hopfield network for the optimization application consists of many interconnected neuron elements. The network minimizes an energy function of the form [13]–[16]

$$E = \sum_{k=1}^{N}\sum_{l=1}^{N} T_{kl} V_k V_l - \sum_{k=1}^{N} I_k V_k \qquad (1)$$

where N is the number of neurons, V_k is the output of kth neuron, I_k is the bias term, and T_{kl} is the interconnection

Manuscript received June 8, 1991; revised November 22, 1991.

S. C. Amartur is with the Department of Radiology, University Hospitals of Cleveland, Case Western Reserve University, Cleveland, OH 44106.

D. Piraino is with the Department of Radiology, Cleveland Clinic Foundation, Cleveland, OH 44106.

Y. Takefuji is with the Department of Electrical Engineering and Applied Physics, Case Western Reserve University, Cleveland, OH 44106.

IEEE Log Number 9106786.

weight between kth and lth neuron. The minimization is achieved by solving a set of motion equations satisfying

$$\partial U_i/\partial t = -\partial E/\partial V_i \tag{2}$$

where U_i is the input of ith neuron. If the input–output function of the neuron is monotonically increasing and the system satisfies (2), the energy continuously reduces as a function of time and the system converges to a local minimum (in a stochastic sense), also known as an attractor basin or a fixed point. The energy landscape in general has more than one local minimum due to the nonconvex nature of the energy surface. (The Hessian matrix H of E: $H_{ij} = \partial^2 E/\partial V_i\,\partial V_j$, in indefinite.) The energy function defined in (1) is slightly different from the one formulated by Hopfield in that, the decay term has been removed for convergence reasons, see the Appendix for a proof using purely deterministic arguments. If a problem can be cast in the form of or mapped to a minimization of Hopfield energy function, a neural network can be realized to obtain a reasonably "good" solution (deep in the energy landscape). To ensure this, one may have to find solutions from different initial states and pick the best among them. Only under some special situations the global minimum can be reached from arbitrary initial states [19].

III. ANN for Classification

A. Mapping

In our study, we follow the ideas proposed in reference [17]. Our objective here is to assign N pixels of P features among M classes ($M^N/M!$ possible ways) such that the assignment of the pixels minimizes a criterion function. The criterion function can be based on a metric measure between a pixel to a class or a nonmetric similarity measure indicating how close a pixel is with other members within a class. The conventional K-means algorithm has a time complexity of $O(NMP)$. A multicomputer version of the algorithm has been proposed with a time complexity of $O(P + \log NM)$ and a space complexity of $O(NM)$ [20]. The generalized distance measure between the kth pixel and the centroid of class l is given by

$$R_{kl} = \left\|\boldsymbol{X}_k - \overline{\boldsymbol{X}}_l\right\|_{\boldsymbol{A}_l^{-1}} \tag{3}$$

where $\boldsymbol{X}_k$ is the P-dimensional feature vector of kth pixel, $\boldsymbol{A}_l$ is a $P \times P$ positive definite weighting matrix, $\overline{\boldsymbol{X}}_l$ is the P-dimensional centroid of class l, and $\|\boldsymbol{X}\|_{\boldsymbol{A}} = \boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}$. If $\boldsymbol{A}_l$ is an identity matrix, the distance measure is Euclidian and the resulting clusters will be restricted to hyperspherical distributions [1]. If $\boldsymbol{A}_l = \Sigma_l$, the covariance matrix of class l, the distance measure is Mahalanobis, which allows for hyperellipsoidal cluster distributions. The weighted inner-product norm also provides invariance to linear transformation (scaling) of features in the images [1]. The overall energy or cost function E to be minimized is given by

$$E = 1/2 \sum_{k=1}^{N}\sum_{l=1}^{N} R_{kl}V_{lk}^2. \tag{4}$$

Note that the distance measure R_{kl} is an implicit function of V_{kl} and is changing at each iteration with the new estimation of the weighting matrix $\boldsymbol{A}_l$. By applying the relation in (2), we get a set of equations for the neuronal dynamics given by

$$dU_{kl}/dt = -R_{kl}V_{kl} \tag{5}$$

where U_{kl} and V_{kl} are the input and output of klth neuron, respectively. The input–output function for the kth row (to assign a label n to the kth pixel) is given by

$$\begin{aligned} V_{kn}(t+1) &= 1 \quad \text{if} \quad U_{kn} = \text{MAX}[U_{kl}(t);\ \forall\, l] \\ V_{kl}(t+1) &= 0 \quad \text{otherwise} \end{aligned} \tag{6}$$

where t is the iteration step. The use of maximum neuron provides a crisp classification as opposed to some alternative methods which provide a fuzzy classification of the data. The monotonicity property of the maximum neuron follows from the fact that it is equivalent to a MaCulloch–Pitts neuron with a dynamic threshold equal to U_{kn}. The rapid convergence property of this type of neuron has been reported earlier [16]. Another advantage of using maximum neuron is the elimination of gain parameters in the motion equation and hence the need for *ad hoc* selection procedures (the energy function has only one component). The neural architecture consists of a grid of $N \times M$ neurons with each column representing a class and each row representing a pixel thus, having a space complexity of $O(MN)$. The system inputs are initialized to random values (mean zero and a standard deviation of 0.1) and the dynamical equations (5) are allowed to evolve as a function of time. The network is said to have reached convergence if all the neurons have negligible input-output activity ($dU/dt \approx 0$). A factor of four saving in space can be attained by working with an image half the resolution as the original.

The algorithm was tested on a sequential digital computer (microVAX, Digital Equipment Corporation, Inc.) and an array processor (Sky computers, Inc., Lowell, MA) as follows.

1) Initialize the input of neurons to random values (which amounts to a random assignment of N pixels to M classes).
2) Apply the input-output relation given in (6) to obtain the new output values for each neuron, establishing the assignment of pixels to classes.
3) Compute the centroid (or class means) $\overline{X}_l$ and the covariance matrix Σ_l for class l as follows:

$$\overline{\boldsymbol{X}}_l = \left[\sum_{k=1}^{N} \boldsymbol{X}_k V_{kl}\right] \Big/ n_l \tag{7}$$

$$\Sigma_l = \left[\sum_{k=1}^{N} V_{kl}\left(\boldsymbol{X}_k - \overline{\boldsymbol{X}}_l\right)\left(\boldsymbol{X}_k - \overline{\boldsymbol{X}}_l\right)^T\right] \Big/ (n_l - 1) \tag{8}$$

where n_l is the number of pixels in class l. The covariance matrix is normalized by dividing each of its element with $(\det[\Sigma_l])^{1/P}$.

4) Solve the set of differential equations in (5) using Euler's approximation to update the inputs to each neuron

$$U_{kl}(t+1) = U_{kl}(t) + dU_{kl}/dt. \tag{9}$$

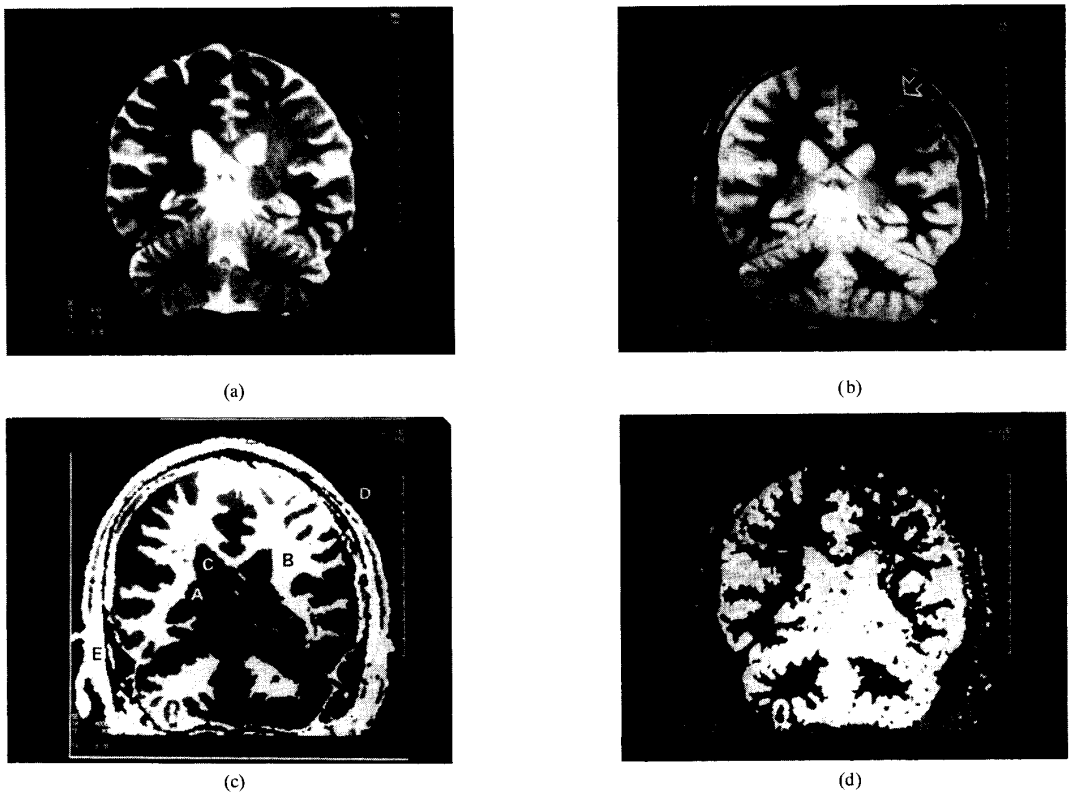


Fig. 1. MR head images of a normal subject (a) T_2-weighted at level 1; (b) proton density-weighted at level 1; (c) classification with 5 clusters; (d) classification with 7 clusters; (e) T_2-weighted image at level 2; (f) proton density-weighted image at level 2; (g) classification with 5 clusters. An arrow in (b) shows a region with low signal due to inhomogeneity of the radiofrequency coil. The tissues identified in (c) and (g) are (A) gray matter; (B) white matter; (C) cerebral spinal fluid; (D) background region; and (E) subcutaneous fat.

5) If there is sufficient activity (significant change in the inputs as determined by a preset threshold) in the input of each neuron then, repeat from step 2) else, terminate.

If the number of clusters chosen is too large, the algorithm may tend to overclassify the image data into many disjoint regions. This may require a merging of subregions or form a super-clusters based on a similarity or a dissimilarity criterion. One such dissimilarity criterion which has been suggested is defined as [21]

$$\text{Disimilarity} = \left[n_i \left\| \overline{\boldsymbol{X}}_i - \overline{\boldsymbol{X}}_{ij} \right\| + n_j \left\| \overline{\boldsymbol{X}}_j - \overline{\boldsymbol{X}}_{ij} \right\| \right] / (N-1) \tag{10}$$

where N is the total number of pixels, n_i and n_j are the number of pixels in classes i and j, and $\overline{\boldsymbol{X}}_{ij}$ is the combined mean for the two classes given by

$$\overline{\boldsymbol{X}}_{ij} = \left(n_i \overline{\boldsymbol{X}}_i + n_j \overline{\boldsymbol{X}}_j\right) / (n_i + n_j). \tag{11}$$

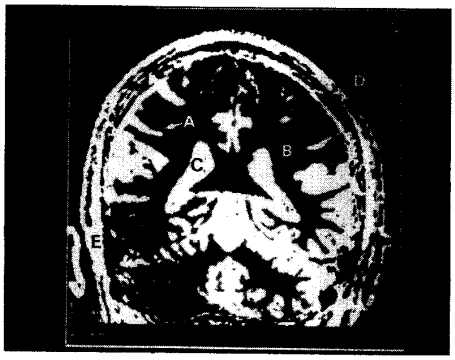
To ensure that a pixel is assigned to utmost one label, the energy function (4) is usually augmented with a constraint relation [13]. This function is automatically handled by the structure of the maximum neuron used in our study which does not violate this constraint. The other conflicting situation of a class being completely vacant or no pixels assigned to a class, never occurred in our study, supporting the observation made in [17] thus, the solution at each iteration is a valid solution. Notice that except for the input–output equation (6), all other equations are amenable for implementation on a fast pipe-line array processor. The parallel hardware realization of winner-take-all neuron is an important issue in neural networks. Even though the algorithm was simulated on a sequential machine, the neural network implementation of the algorithm is fully parallel and can be realized with analog processing elements hence, is expected to provide very rapid solutions to the problem.

B. Postclassification Filtering

Due to the spectral variability in the MR data, the classified image may exhibit a salt-and-pepper appearance. A majority filter was used to reduce this artifact [22]. The filter consists of a moving window or mask which is passed over the classified image. The new assignment of label for the center pixel is the label associated with the majority class within the window. Multiple iterations can be employed to control the amount of smoothness at the expense of losing some small structures. As an alternative to the postprocessing, the smoothing operation described above can be incorporated within the classification process with suitable augmentation of the energy function.

(e)

(f)



(g)

Fig. 1. *Continued*

IV. Results and Discussion

We present results obtained with images acquired on a patient with normal physiology and one example of a patient diagnosed with a known pathology. Fig. 1(a) and (b) are the T_2-weighted (TR = 4000 ms, TE = 90 ms) and a proton density-weighted (TR = 4000 ms, TE = 15 ms) image in the brain of a normal volunteer. The cerebral spinal fluid (CSF) appear bright in the T_2-weighted image and dark in the proton-weighted image. The proton density-weighted image on the other hand shows good differentiation between the gray and white matter with the gray matter slightly brighter when compared with the white matter. The number of clusters or the cluster size in a given image set is usually not known *a priori* in practice (cluster validity). A measure for the "goodness" of classification is the cluster size which provides the minimum E when the network reaches equilibrium (or flattens out). Fig. 1(c)–(d) show the classification results with cluster sizes equal to 5 and 7, respectively. The classified images were filtered using a majority filter with one iteration. The different class labels have been assigned separate gray levels in the classification map. Table I gives the value of E at equilibrium for different cluster sizes.

From Table I, a cluster size of 7 gives the least energy at equilibrium and hence should be optimum. The corresponding classification map [Fig. 1(d)] provides an accurate grouping of different homogeneous regions. However, there is a need for a merger of group members belonging to the same tissue. From a human observer point of view, a choice of 5 clusters [Fig. 1(c)] seem to provide a more meaningful classification without the need for a merging step. The anatomical identification of regions [Fig. 1(c)] corresponding to the assigned labels are the background, CSF, gray matter, white matter, and subcutaneous fat. Nearly symmetric nature of the gray-white matter tissues in the classified image is indicative of the normal physiological condition of the subject. The algorithm has misclassified tissues in some regions of the image due to the nonstationary response of the image acquisition (signal gain variation). Fig. 1(e) and 1(f) are the T_2-weighted and proton density-weighted images of the same patient at a different level in the brain. The corresponding classification map with 5 clusters is shown in Fig. 1(g).

Fig. 2(a) and (b) are the T_2-weighted (TR = 2500 ms, TE = 90 ms), and proton density-weighted (TR = 2500 ms,

TABLE I

Cluster size	E at equilibrium
3	1869
4	1678
5	1662
6	1454
7	1304
8	1479

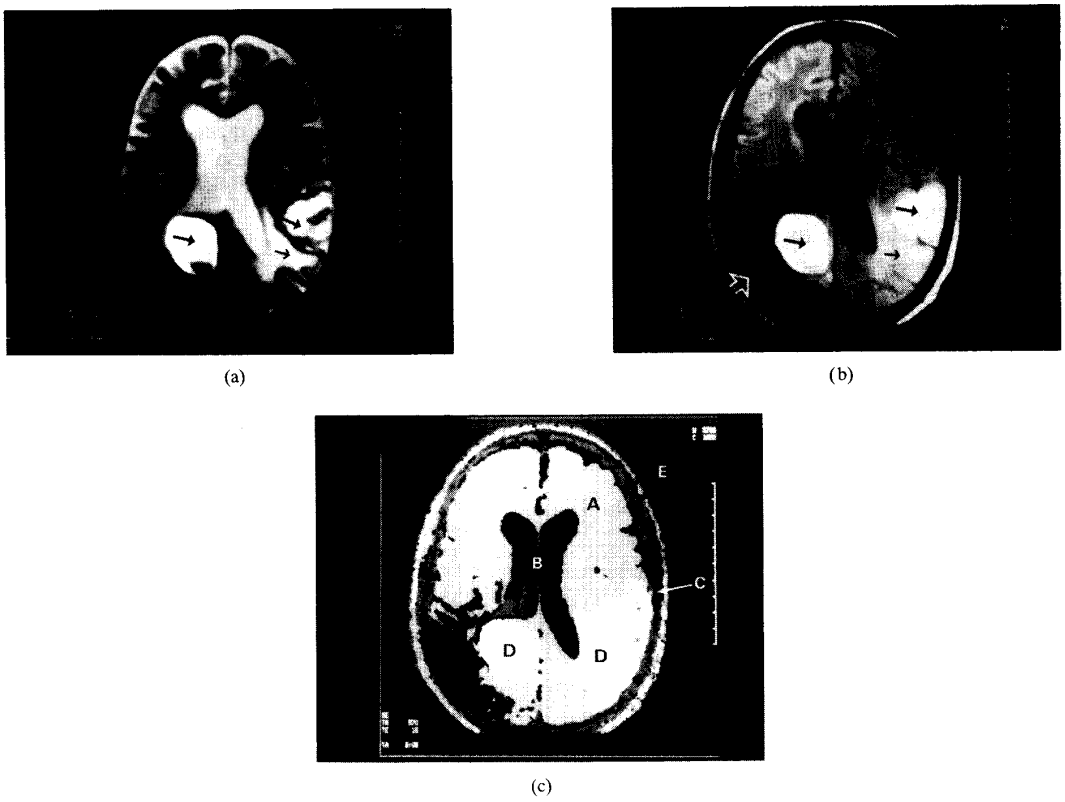


Fig. 2. MR head images of a subject with a known pathology (a) T_2-weighted image; (b) proton density-weighted image; (c) classification with 5 clusters. Metastases are shown with long arrows and the edema is shown with small arrows in (a) and (b). An open arrow in (b) shows a region with low signal due to inhomogeneity of the radiofrequency coil. The tissues identified in (c) are (A) white matter; (B) cerbral spinal fluid; (C) subcutaneous fat; (D) metastases and edema; (E) background region.

TE = 15 ms) images, respectively, from a patient diagnosed with a melanoma metastatic to the brain. The metastases are seen as two hyperintense masses in both images. The proton density weighted and T_2-weighted images also show an abnormally bright region in the white matter between the two metastatic lesions, thought to be due to edema caused by the mass effect. Note also the signal variation within the metastases in proton density-weighted image. Fig. 2(c) is the classification map with 5 clusters. From the map we can see that the metastatic regions have been well delineated along with the white matter and CSF in the lateral ventricles. The algorithm has failed to pick the gray matter seen clearly in the proton density-weighted image as a separate cluster. This may be attributed to the relatively small number of pixels associated with the gray matter and the possible nearness of its cluster in the feature space with another cluster belonging to a different tissue with a large numerical superiority [1]. Increasing the cluster size followed by a merging operation may alleviate this problem. Notice also the large space classified as CSF [Fig. 2(c)], not evident in the feature images. This may be attributed to the intensity attenuation in this region (note that attenuation of intensity from the top right corner towards the bottom left corner in the proton density-weighted image) due to inhomogeneity or change in sensitivity of the radiofrequency coil. This imaging artifact may have to be corrected before the application of neural net classifier.

It was experimentally found that the network reaches equilibrium or convergence for the two cases presented here within 100 iterations for cluster sizes up to 15. Hence, the time for convergence is the order of 100 τ where τ is the time constant of each analog processing element (microseconds). The near constancy of convergence time with the problem size is an important property which indicates that the effective time complexity is close to $O(1)$. The performance of the network is mainly dictated by the choice of the energy function or cost criterion. The criterion employed in this study provides "good" classification for images with features which are well separated and exhibit nonoverlapping clusters. The feature elements in the acquired data set may in general be highly correlated. Techniques are available to de-correlate the feature elements for better classification results [1].

V. Conclusion

This preliminary study has demonstrated the applicability of Hopfield net for the tissue classification in MRI. The quantitative performance study was made difficult by the lack of a gold standard to compare with nor the availability of ground truth for verification of the results. We have relied upon the human experts for the performance evaluation. The

technique is at present not robust to handle images corrupted by non-stationary sensitivity of the image acquisition and partial volume effects. The availability of dedicated neural net hardware in the future is expected to provide the luxury of running the network with different initial states to obtain the "best" classification. The search for global minimum using the Hopfield neural networks is still an open problem.

The future work will focus in the following directions: a) account for the non-stationary response of the image acquisition; b) model tissues as mixture of many classes with the use of fuzzy segmentation; c) use of regions from the classified map obtained from the technique presented in this paper as a training pattern set for a second stage consisting of a back propagation network; d) incorporate additional constraints to the energy function.

APPENDIX

Hopfield formulated the Liapunov energy function [13]–[16]:

$$E = 1/2\sum_{k=1}^{N}\sum_{l=1}^{N} T_{kl}V_kV_l - \sum_{k=1}^{N} I_kV_k + 1/\lambda\sum_{i=1}^{N}(1/R_i)\int_0^{V_i(t)} g^{-1}(V(t)\,dV(t)) \quad (12)$$

$$= E_D + 1/\lambda\sum_{i=1}^{N}(1/R_i)\int_0^{V_i(t)} g^{-1}(V(t)\,dV(t)) \quad (13)$$

where $g(\cdot)$ is the input-output function for the neuron. The motion equation for the ith neuron is given by

$$dU_i/dt = \partial E_D/\partial V_i - U_i/\tau \quad (14)$$

where τ is a time constant. Using the chain rule for differentiation

$$dE/dt = \sum_{i=1}^{N}(dU_i/dt)(dV_i/dU_i)(\partial E/\partial V_i). \quad (15)$$

For very large λ (an assumption made to simplify the mapping), $\partial E/\partial V_i = \partial E_D/\partial V_i$. Hence,

$$dE/dt = -\sum_{i=1}^{N}(dU_i/dt)^2 dV_i/dU_i - \sum_{i=1}^{N}(U_i/\tau)\cdot(dV_i/dU_i)(dU_i/dt) \quad \text{using (14)}. \quad (16)$$

$$= -\sum_{i=1}^{N}(dU_i/dt)^2(dV_i/dU_i) - \sum_{i=1}^{N}(1/2\tau)(dV_i/dU_i)\big(dU_i^2/dt\big). \quad (17)$$

The first summation term is always positive due to the monotonic increasing property of the input–output function. The second summation term in (17) which arises due to the decay term or the last term in (13), can be either positive or negative. Hence, the overall change in E can be either positive or negative. However, if the decay term is removed from the energy function, the energy change is strictly negative, proving the convergence with the use of modified Hopfield energy function.

REFERENCES

[1] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis.* New York: Wiley, 1973.

[2] M. W. Vannier, D. L. Rickman, D. M. Jordon, W. A. Murphy, and P. R. Biondetti, "Multispectral magnetic resonance image analysis," *CRC Critical Rev. Biomed. Eng.*, vol. 12, pp. 117–144, 1987.

[3] D. A. Ortendahal, N. M. Hylton, L. Kaufman, and L. E. Crooks, "Automated tissue characterization with NMR imaging," in *Proc. 8th Int. Conf. Imaging,* F. Deconinick, Ed., The Netherlands: Nijhoff, The Hague, 1984.

[4] M. O'Donnel, J. C. Gore, and W. J. Adams, "Toward an automated analysis system for nuclear magnetic resonance imaging—II—initial segmentation algorithm," *Med. Phys.*, vol. 13, no. 3, pp. 293–297, 1986.

[5] M. I. Kohn, N. K. Tanna, G. T. Herman, S. M. Resnick, P. D. Mozley, R. E. Gur, A. Alavi, R. A. Zimmerman, and R. C. Gur, "Analysis of brain and cerebrospinal fluid volumes with MR imaging Part I. Methods, reliability, and validation," *Radiology,* vol. 178, pp. 115–122, 1991.

[6] H. E. Cline, W. E. Lorensen, R. Kikinis, F. Jolesz, "Three-dimensional segmentation of MR images of the head using probability and connectivity," *J. Comp. Assist. Tomogr.*, vol. 14, pp. 1037–1045, 1990.

[7] R. H. Lee, "Multi-sensor image segmentation algorithms," in *Proc. SPIE Sensor Fusion,* vol. 1306, 1990, pp. 11–17.

[8] R. L. DeLaPaz and R. Bernstein, "Computerized analysis and information extraction of magnetic resonance images," in *Proc. SPIE,* vol. 902, 1988, pp. 151–154.

[9] R. P. Lippman, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, pp. 4–22, Apr. 1987.

[10] A. Gronquist and R. Lenz, "Detection of blood vessels in 3-D MR-images," *IEEE Int. Conf. Neural Networks,* vol. 1, 1988, pp. 145–149.

[11] K. Oshio and M. Singh, "Automatic segmentation of MR head images using neural nets," in *8th Annu. Meet. Soc. Magnet. Resonance Med.*, Amsterdam, Netherlands, 1989.

[12] R. P. Velthuizen, L. P. Clarke, J. W. Flieder, and D. T. Glennon, "MR 3D image segmentation with artificial neural nets for tissue classification," *J. Magnet. Resonance Imaging,* vol. 1, p. 200, 1991.

[13] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in *Proc. Nat. Acad. Sci.*, vol. 79, 1982, pp. 2554–2558.

[14] Y. Takefuji and K. C. Lee, "A near-optimum parallel planarization algorithm," *Science,* vol. 245, pp. 1221–1223, 1989.

[15] H. H. Szu, "Reconfigurable neural nets by energy convergence learning principle based on extended MaCulloch–Pitts neurons and synapses," in *IEEE INNS Int. Joint Conf. Neural Nets.* vol. 1, 1988, pp. 485–495.

[16] K. Lee, Y. Takefuji, and N. Funabiki, "A parallel improvement algorithm for the bipartite subgraph problem," Case Western Reserve Univ., CAISR Tech. Rep. TR91-105, 1991.

[17] B. Kamagar-Parsi, J. A. Gualtieri, J. E. Devaney, and B. Kamagar-Parsi, "Clustering with neural networks," *Biological Cybernet.*, vol. 61, pp. 201–208, 1990.

[18] E. Rignot and R. Chellapa, "Segmentation of synthetic aperture radar complex data," *J. Opt. Soc. Amer. A,* vol. 8, no. 9, pp. 1499–1509, 1991.

[19] S. Park, "Signal space interpretations of Hopfield neural network for optimization," in *IEEE ISCAS Int. Symp. Circuits Syst.*, 1989, pp. 2181–2184.

[20] S. Ranka and S. Sahni, "Clustering on a hypercube multicomputer," in *10th IEEE IPAR Int. Conf. Pattern Recogn.*, vol. 1, pp. 532–536, 1990.

[21] J. P. Strong, "Computations on the massively parallel processor at the Goddard space flight center," *Proc. IEEE,* vol. 79, pp. 548–558, 1991.

[22] T. M. Lillesand and R. W. Kiefer, *Remote Sensing and Image Interpretation.* New York: Wiley, 1987.