Dipole Source Localization of MEG by BP Neural Networks

Yohsuke Kinouchi^, Gen Ohara^, Hirofumi Nagashino^, Tetsuro Soga*, Fumio Shichijo*, and Keizo Matsumoto*

Summary: The purpose of this study was to examine the usefulness of BP neural networks for source localization of MEG. Since the performance of this method does not depend on the complexity of brain parameters and source models, a homogeneous brain model and a single current dipole source are assumed for convenience. Localization accuracy was examined in relation to the configuration and scale of the network. As a result, average error for position and moment estimations was within 2%, while the maximum error was about 5%. It was therefore concluded that the neural network method was useful for MEG source localization, though some improvements are still necessary.

Key words: MEG; Source Localization; Localization accuracy; Neural network; Error back propagation algorithm.

Introduction

Source localization of MEG is very useful and important for studying brain function and for diagnostic purposes (Ueno and Iramina 1990). Localizing sources from both MEG and EEG data can be thought as an "inverse problem". As the clinical applications of MEG increase, it will be necessary to solve the inverse problems in real time, in spite of the complexity of brain and source models, in real time.

A BP neural network (using an error Back Propagation algorithm) is used extensively for solving inverse problems. If an inverse problem has a unique solution, the network can solve the problem by only one forward calculation, without iterative convergence. Since the source localization problems for EEG and MEG are considered to have unique solutions, provided there are a finite number of sources, the neural network approach can be applied to the problem. The neural network method also results in a quick solution, regardless of the complexity of the model, so long as the training patterns are adequate. Therefore, the neural network may offer a promising and practical method of source localization. We have already reported on the usefulness of the neural network for EEG source localization in the case of single and two dipole sources (Abeyratne et al. 1991a,b).

The purpose of this study was to examine the usefulness of a BP neural network for source localization of MEG. In order to install the desired inverse function on the neural network, it needs to be trained. Training patterns may be obtained theoretically or experimentally, according to the complexity of the brain model and the field measurement system. However, once the patterns have been obtained, the training procedure and the resulting performance of the network do not depend on the complexity. A simple model is therefore used here for the purpose of training, assuming a single dipole source. Localization accuracy is mostly examined in relation to the structure and scale of the network because this is of primary importance to the actual use of the network.

Neural Network Localization

The geometry used for the head model was a hemisphere, as shown in figure 1. It is essentially the same as the three-concentric-shell model of Rush and Driscoll (1968) which has also been used for source localization of EEG by neural networks (Abeyratne et al. 1991). 39 sensors (SQUID) were placed uniformly on the surface of the scalp to measure the radial component B_r of the magnetic flux density (assuming usual measurement procedures for a SQUID). The number and location of the sensors are the same as those for neural network localization of EEG (Abeyratne et al. 1991), since we will be comparing both

^Department of Electrical and Electronic Engineering, The University of Tokushima, Tokushima, Japan.

*Department of Neurological Surgery, The University of Tokushima, Tokushima, Japan.

Accepted for publication: November 6, 1995.

Correspondence and reprint requests should be addressed to Yohsuke Kinouchi, Department of Electrical and Electronic Engineering, The University of Tokushima, Minami-Josajima, Tokushima, 770 Japan.

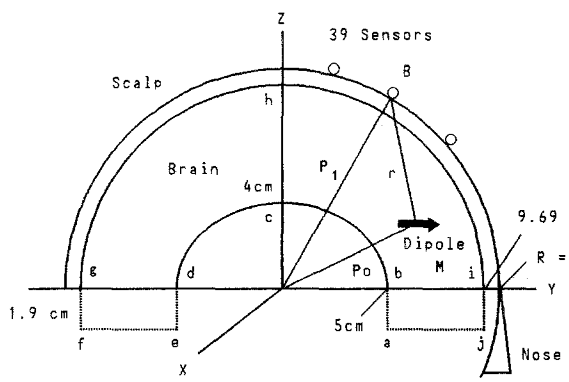

Figure 1. Hemisphere model of the head. 39 magnetic sensors are placed uniformly on the surface of the scalp. Position $P_0(x,y,z)$ and moment $M(Mx,My,Mz)$ of a single current dipole in the testing region b-c-d-g-h-I-b are estimated.

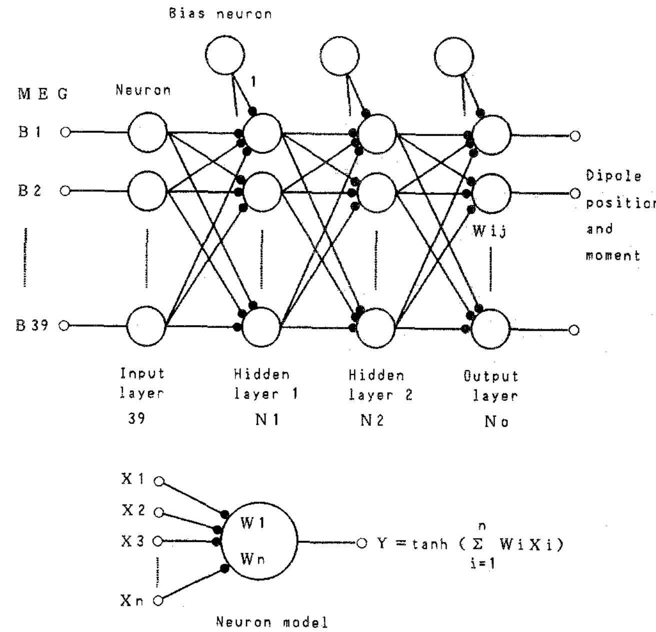

Figure 2. Configuration of the neural network consisting of input, output and two hidden layers. The input to the network is the radial field strength, B_1 to B_{39} and the output is dipole position P(x,y,z,) and moment M(Mx,My,Mz).

EEG and MEG. We used a current dipole source in a homogeneous substance. A current dipole with moment M yields B given by (Hosaka 1976):

$$B=\frac{\mu_0}{4\pi R r^3}\left\{P_1\cdot(M\times P_0)\right\} \qquad (1)$$

where P_0 = Position vector of a current dipole
P_1 = Position vector of a sensor
$R = |P_1|$ = Radius of the head (11.14 cm)
$r = |P_1 - P_0|$
Radius of the brain = 9.69 cm

Configuration of the neural network used is shown in figure 2. If an MEG pattern is fed into the network, the output will be an estimation of the position of the dipole, and the dipole moment. The network consists of four layers (input layer, 2 hidden layers, and output layer) with feed-forward connections. The input layer has 39 neurons receiving flux densities from 39 sensors, i.e., B_1 - B_{39}. Let us denote N_1, N_2 and N_0 the numbers of neurons in the first and second hidden layers and the output layer respectively. The network has bias neurons with a constant output value of 1, to adjust the threshold of each neuron. Two types of network are used here.

Network A: $N_1 = 60$, $N_2 = 40$ and $N_0 = 6$
Network B: Six networks of $N_1 = 30$, $N_2 = 15$ and $N_0 = 1$

In network B, the output layer consists of only one neuron representing one of the components of dipole position P_0, i.e., (x, y, z) or moment M, i.e., (Mx, My, Mz). Hence, the six networks are used in parallel for position and moment estimations. That is, network B makes use of the independence of the six source parameters. Neurons in the input layer have a linear input-output function, while other neurons have a sigmoid function (tanh). Synaptic connection to neurons are modifiable except for those to neurons in the input layer. The scale of a network is measured by the number of modifiable synapses which is 5086 or 10086 for networks A and B respectively.

To acquire the desired input-output transformation, the network should be trained and tested. Training and testing MEG patterns are generated by using expression (1), from dipoles put randomly in the regions a-b-c-d-e-f-g-h-i-j-a and b-c-d-g-h-i-b of figure 1 respectively. These regions are called the training region and the testing region (or localization region), respectively. The central region of the brain is excluded from the localization region because the MEG signal generated by a dipole in the central region is much smaller than that generated in the surface region, as shown by expression (1). If it is necessary, there are other neural networks that can be used for localization in the central region. The training region is chosen to be somewhat larger than the testing region in order to achieve high localization accuracy for the testing patterns. The number of training and testing patterns is 10,000 for each. Each MEG pattern is scaled so that the effective value (rms. value) of the set B_1 to B_{39} is 0.6, so as to obtain a good localization. At the same time, the corresponding dipole moment is scaled by the same factor. Then, the six source parameters of all patterns are scaled again by a common factor so as to be within 0.8 (below saturation of a neuron output), e.g.,

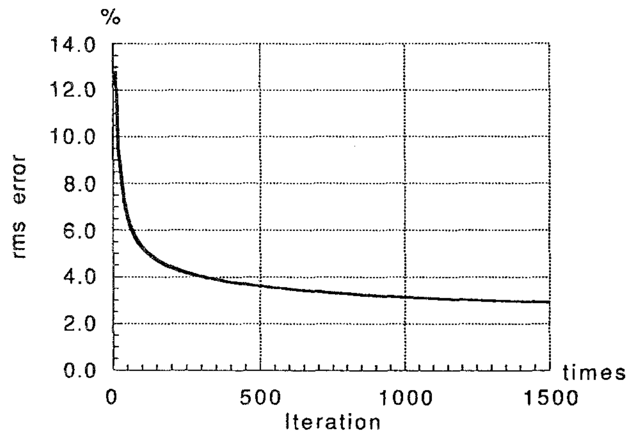

Figure 3. Convergence process during training of neural network A. Before training, synaptic connections Wij have random values, hence the network yields a large error. After training with 1500 iterations, the average rms. error ε_{rms} becomes as small as 2.89%.

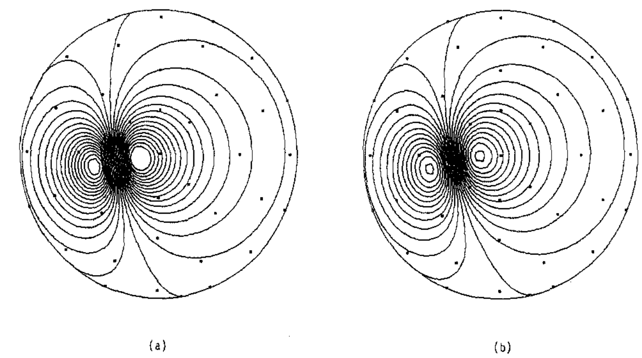

Figure 4. Comparison of an original MEG pattern(i.e. measured MEG)(a) and another (b) estimated by neural network B in the case of ε_{rmsp} = 1.89% (average). Both patterns give a dipolarity of 99.1%, which may be high enough to allow the practical use of a BP neural network.

position P_0 = (x, y, z) [cm] is scaled to $(0.8/9.69)P_0$ where 9.69 [cm] is the radius of the brain model.

To train the network, the typical error back propagation algorithm is used (Rumelhart et al. 1986). The error energy function is defined as follows:

$$E_p = \frac{1}{2}\sum_{i=1}^{6}\left(O_{pi} - T_{pi}\right)^2$$

where O_{pi} is a network output and T_{pi} is a target value (true value) at the *i*th output neuron for an MEG pattern (number p). The relative rms error for the p-th MEG pattern is:

$$\varepsilon_{rmsp} = \frac{1}{0.8}\sqrt{\frac{1}{6}\sum_{i=1}^{6}\left(O_{pi} - T_{pi}\right)^2}$$

Average rms error for all MEG patterns:

$$\varepsilon_{rms} = \frac{1}{N}\sum_{p=1}^{N}\varepsilon_{rmsp}$$

where N is the number of patterns (10,000). In the training phase, the change of synaptic connections δWji (i-th to j-th neurons) for a MEG pattern is given by

$$\delta W_{ij} = -\eta\frac{\partial E_p}{\partial W_{ij}}$$

where :

Learning rate $\eta = \eta_0\left(1 + \frac{1}{0.05}\varepsilon_{rmsp}\right)$; $\eta_0 = 0.01$

That is, the larger the error ε_{rmsp} is, the bigger the learning rate η becomes. This is useful for making error distribution narrow after training.

Results and Discussion

Figure 3 shows the convergence process of network A during 1500 training iterations (one iteration means a training cycle using all of the 10000 training patterns). After convergence, ε_{rms} = 2.89 % and the maximum of ε_{rmsp} = 8.87 % for training patterns, while ε_{rms} = 3.03 % and max ε_{rmsp} = 11.16 % for testing patterns. This figure therefore indicates that the network has acquired the desired inverse transformation with an accuracy of about 97 %, on average. The average errors are smaller than those for the neural network localization of EEG (about 4%) (Abeyratne et al. 1991). However, the range of error ε_{rmsp} is still quite large.

After the same training sequence, network B resulted in more accurate localization, with ε_{rms} = 1.89 % and max ε_{rmsp} = 9.26 % for the training patterns. This improvement may have been due to the dividing of the output layer and an increase in the network scale. Figure 4 shows the comparison of two MEG patterns in the case of ε_{rmsp} = 1.89 % (average error). One was derived by the forward calculation of expression (1) for a dipole, and the other was derived from the dipole estimated by feeding the former pattern into the neural network. In this simulation, the former pattern is assumed to be a measured MEG and the latter is assumed to be the estimated MEG. Both patterns are very similar and hence result in a very high dipolarity of 99.1%. Here, dipolarity D is defined by:

$$D = \sqrt{1 - RV}$$

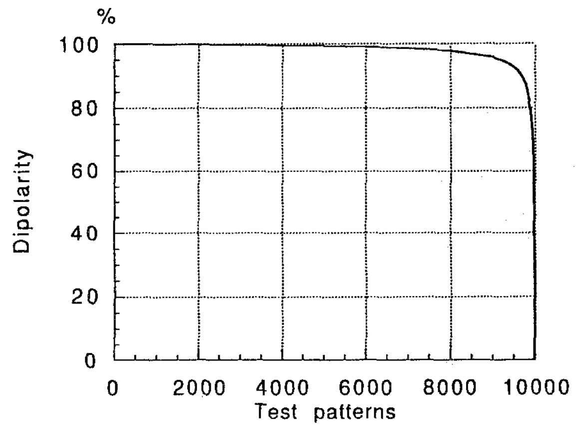

Figure 5. Distribution of dipolarity for 10,000 testing patterns. The testing patterns are arranged in order of dipolarity. Dipolarities were more than 90% for 97% of the testing patterns.

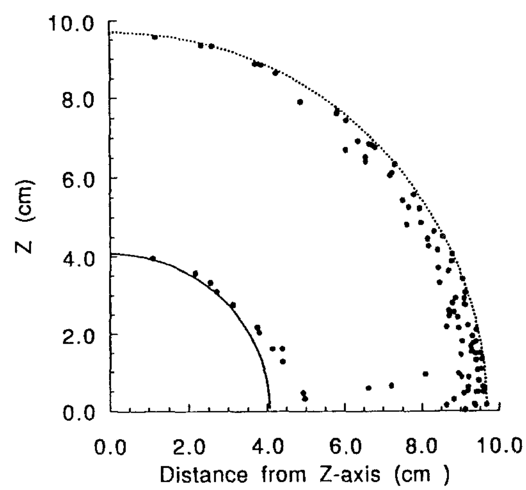

Figure 6. Spatial distribution of 100 dipoles yielding large errors. They are located near the boundary of the testing region. Border dipoles are difficult to localize because of the small number of training patterns.

$$RV = \frac{\sum_{i=1}^{39}\left(B_{m,i} - B_{e,i}\right)^2}{\sum_{i=1}^{39}\left(B_{m,i}\right)^2}$$

where $B_{m,i}$ and $B_{e,i}$ are the flux densities of the i-th sensor for measured and estimated MEGs respectively. Figure 5 shows the distribution of dipolarity D for 10,000 testing patterns, and it is found that D is larger than 90% for 97% of the testing patterns.

We investigated the patterns with large ε_{rmsp} (near max ε_{rmsp}) in more detail in order to determine the cause of the error. Figure 6 shows the spatial distribution throughout the brain of 100 dipoles with large error, that were chosen out of the 10,000 dipoles used for generating the test patterns. It was found that the large error resulted from dipoles situated near the boundary. Therefore, we restricted the locations of testing dipoles into a region: 4.93 cm $\leq$ | PO | $\leq$ 9.00 cm and z $\geq$ 1 cm (inside the testing region of figure 1). Figure 7 shows the distribution of ε_{rmsp} in this case, which gives ε_{rms} = 1.50% (average error) and max ε_{rmsp} = 4.89%. Thus, localization accuracy became considerably high except in the region near the boundary.

In order to achieve high accuracy for all of the testing region, the training region has been expanded to include a region: 4.5 cm $\leq$ | PO | $\leq$ 10.2 cm and z $\geq$ -1 cm. At the same time, the sensor locations have been expanded onto the surface of a sphere with a radius of 12 cm to avoid the influence of strong fields emanating from the boundary dipoles. Only the x-position variable is obtained for dipole localization in this area, by using a neural network with N1 = 60 and N2 = 40 (i.e., 39-60-40-1 neurons in each layer), which is about three times larger in scale than network B is for the calculation of x-position. Figure 8

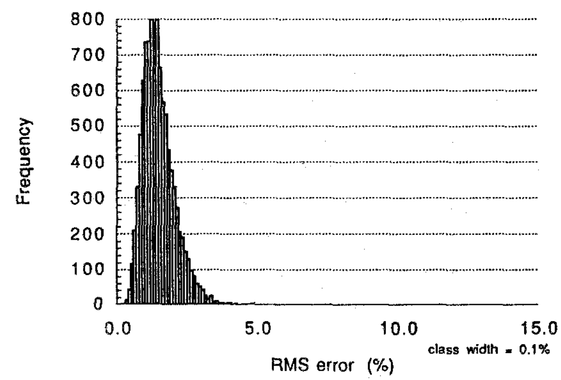

Figure 7. Distribution of relative rms. error ε_{rmsp} for a restricted testing region. The average rms. error ε_{rms} is considerably small as (as low as 1.50%).

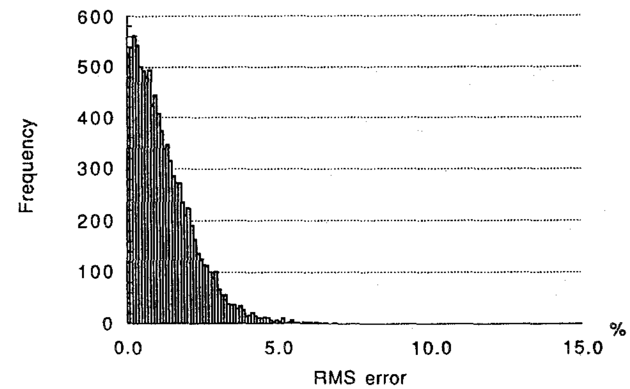

Figure 8. Distribution of relative rms. error ε_{rmsp} when only the x-position was obtained from the new network, which has three times as many neurons as network B. The average position error is as small as 1.23% owing to a scaling up of the network.

shows the distribution of ε_{rmsp} for test patterns which have an $\varepsilon_{rms} = 1.23\%$ (average error) and max $\varepsilon_{rmsp} = 6.9\%$. Therefore, on average, the localization accuracy becomes higher with increasing network scale.

In practical applications of this method, it may be necessary to decrease the ε_{rmsp} (in other words, the variance of the error distribution), which can be achieved by controlling the learning rate. There is also another approach whereby the training and testing regions are restricted to certain regions of interest in the brain. Such a zooming function is easily incorporated into the network by preparing a special training data set. This method is also quite robust when working with noisy MEG data.

Conclusion

We conclude that the neural network method for source localization of MEG is useful. Several different configurations of the neural network were investigated so as to improve localization accuracy. The average error was within 2%, while the maximum error was about 5% for single dipole localization. Reducing the maximum error is important for increasing the performance of the network. Further, the network should be evaluated for the presence of noise in the MEG patterns, as well as for multiple dipole localization.

References

Abeyratne, U.R., Kinouchi, Y., Oki, H., Okada, J., Shichijo, F. and Matsumoto, K. Artificial Neural Networks for Source Localization in the Human Brain. Brain Topography, 1991, 4: 3-21.

Abeyratne, U.R., Kinouchi, Y., Soga, T., Okada, J., Shichijo, F. and Matsumoto, K. Artificial Neural Networks for Multiple Dipole Localization: A Feasibility Study. In: Abstracts of the Second International Congress on Brain Electromagnetic Topography, Toronto, Canada, 1991: 33.

Hosaka, H. On the Basic Study of Magnetic Detection of Electrical Activity in the Heart. Japanese Journal of Medical and Biological Engineering, 1976, 14: 34-40.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. Learning Internal Representation by Error Propagation. In: D.E. Rumelhart, J.L. McClelland, and the PDP Research Group (Eds.), Parallel Distributed Processing. MIT Press, Cambridge, 1986: 318-362.

Ueno, S. and Iramina, K. Modeling and Source Localization of MEG Activities. Brain Topography, 1990, 3: 151-165.