

**Military Technical College
Kobry El-Kobbah,
Cairo, Egypt**



**6th International Conference
on Electrical Engineering
ICEENG 2008**

Accurate affine image registration using radial basis neural networks

By

Haldun Sarnel and Yavuz Senol*

Abstract:

Neural network-based image registration using global image features is relatively a new research subject and the schemes devised so far use a feedforward neural network to find the geometrical transformation parameters. In this work, we propose to use a radial basis function neural network instead of feedforward neural network to overcome lengthy pre-registration training stage. This modification has been tested on a typical neural network-based registration method using discrete cosine transformation features in the presence of noise. The proposed scheme does not only speed up the training stage enormously, but also increases the accuracy and robustness against additive white noise owing to the better generalization ability of the radial basis function neural networks.

Keywords:

Image registration, affine transformation, radial basis function neural network and discrete cosine transform features

* Electrical and Electronics Engineering, Dokuz Eylül University, Izmir, Turkey

1. Introduction:

Image registration is a procedure to determine the spatial best fit between two images that overlap the same scene, and a fundamental stage in many image processing applications such as medical image analysis, remote sensing, image matching-based vehicle guidance [1], [2] and super-resolution [3]. To register two images, a transformation must be found so that each point in one image can be mapped to a point in the second. Registration is often a complicated task and includes a wide range of problems to deal with such as image distortions, scene dependency and determining a suitable geometric transformation model.

The most commonly used registration transformation is the affine transformation. In this work, global affine transformation that is composed of the Cartesian operations of a scaling, a translation, and a rotation is assumed. Correlation-based methods or frequency domain methods are disadvantageous for this type of transformation when computational complexity is concerned. On the other hand, local feature-based or control points methods cannot be relied on when image is noisy or feature extraction is problematic due to scene content.

Image registration based on neural networks is a relatively new approach and requires further consideration and research. At the beginning, this approach was used in aligning a set of landmark (control) points extracted from the images to be registered, such as [4]. Later on, the approach was also employed in the schemes [5], [6] that extract global image features and feed them into a feedforward neural network (FNN) to find the affine transformation parameters. In [5], a FNN is used to estimate the affine transformation parameters of a test image with respect to a reference image. Discrete cosine transform (DCT) features are extracted as inputs to the network and estimated parameters are obtained at the output. In a pre-registration phase, extracted DCT features form a set of translated, rotated and scaled images of the same scene are employed to train a FNN. Although their method gives fairly accurate results for noisy images, its main drawback is the long period of iterative learning process of FNN. To overcome such a lengthy pre-registration training stage, replacing the FNN with a radial basis function neural network (RBF NN) is the fundamental object of this work. The proposed scheme here does not only speed up the training stage, but also increases the accuracy and robustness against noise owing to the better generalization ability of RBF NN.

2. Radial Basis Function Neural Networks:

RBF neural networks with their structural simplicity and training efficiency are good candidate to perform a nonlinear mapping between the input and output vector spaces. RBF NN is a fully connected feedforward structure and consist of three layers namely, an input layer, a single layer of nonlinear processing units, and an output layer. The network structure is shown in Figure (1). Input layer is composed of input nodes that are equal to the dimension of the input vector x . The output of the j th hidden neuron with Gaussian transfer function can be calculated as

$$h_j = e^{-\|x-c_j\|^2/\sigma^2} \quad (1)$$

where h_j is the output of the j th neuron, $x \in \mathcal{R}^{n \times 1}$ is an input vector, $c_j \in \mathcal{R}^{n \times 1}$ is the j th RBF center, σ is the center spread parameter which controls the width of the RBF, and $\|\cdot\|^2$ represents the Euclidean norm. The output of any neuron at the output layer of RBF network is calculated as

$$y_i = \sum_{j=1}^k w_{ij} h_j \quad (2)$$

where w_{ij} is the weight connecting hidden neuron j to output neuron i and k is the number of hidden layer neurons.

The mapping properties of the RBF NN can be modified through the weights in the output layer, the centers of the RBFs, and spread parameter of the Gaussian function. The simplest form of RBF network training can be obtained with fixed number of centers. If the number of centers is made equal to the number of input vectors, namely exact RBF, then the error between the desired and actual network outputs for the training data set will be equal to zero. In this work, exact RBF NN was used. The number of RBF centers was made equal to the number of input vectors (e.g. 256, see section 3).

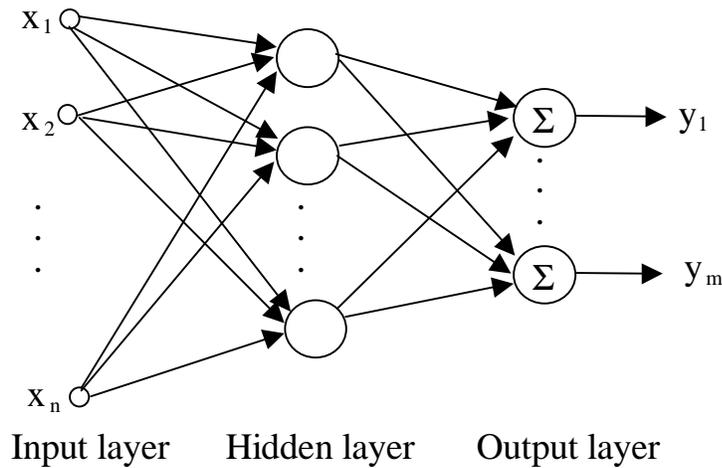


Figure (1): Structure of RBF NN

3. Experimental Work and Results:

An aerial image of an urban area is used to produce a set of translated, rotated and scaled images. Table (1) shows the affine transformation parameters values used for this purpose. A total of 256 images, each of which is 128 by 128 pixels size, were first added with white Gaussian noise. The noisy images then underwent to DCT to obtain frequency domain coefficients. A region of 6 by 6 coefficients in the lowest frequency band in the DCT plane was cut out and used as a feature vector for each affine transformed image. A matrix of 35 by 256 (zero frequency coefficient is discarded) coefficients was obtained and used to train an exact RBF NN. An optimal spread value of the transfer function for the neurons was found empirically. To evaluate the performance of the RBF NN-based registration method, 81 test images with the transformation parameter values given in Table (1) was also created and added with noise of the same strength as that in the training set. Features from the test images were obtained exactly in the same manner as explained for training data. Mean of the absolute value of registration errors resulted by the network was computed for each affine parameter, and given in Table (2). Experiments were carried out for noise-free and noisy images at two different signal to noise ratios (SNR), namely, 20 dB and 5 dB.

For a comparison, mean errors from a FNN with a 20 neurons in one hidden layer by using the same training and test data were also computed and given in the Table (3). The FNN had a tangent-sigmoid transfer function for the hidden layer neurons and a linear function for the output layer neurons, and was trained using the Levenberg-Marquardt method. In order the FNN to learn the training data reasonably well, we had to repeat

the training phase many times until the network gave both a small output training error and a small registration error for the test data. During the time consuming experimental work with the FNN several training attempts failed due to the nonconverged or wrongly converged networks. The registration error results given in Table (3) were obtained with the best networks found after a reasonably persistent search effort.

Table (1): Affine transformation parameter values used in experiments

| transform parameter | values used for training set | values used for test set |
|----------------------------|-------------------------------------|---------------------------------|
| scale | 0.9, 0.965, 1.035, 1.1 | 0.93, 1, 1.07 |
| rotation | -5, -2, 2, 5 (degrees) | -3, 1, 4 (degrees) |
| vertical translation | -5, -2, 2, 5 (pixels) | -4, 0, 3 (pixels) |
| horizontal translation | -5, -2, 2, 5 (pixels) | -3, 1, 4 (pixels) |

Table (2): Mean absolute registration errors of RBF NN-based scheme

| transform parameter | Noise-free | 20 dB SNR | 5 dB SNR |
|----------------------------|-------------------|------------------|-----------------|
| scale | 0.0002 | 0.0006 | 0.002 |
| rotation | 0.014 | 0.06 | 0.18 |
| vertical translation | 0.007 | 0.03 | 0.10 |
| horizontal translation | 0.021 | 0.05 | 0.15 |

Table (3): Mean absolute registration errors of FNN-based scheme

| transform parameter | Noise-free | 20 dB SNR | 5 dB SNR |
|----------------------------|-------------------|------------------|-----------------|
| scale | 0.0015 | 0.0023 | 0.005 |
| rotation | 0.046 | 0.09 | 0.3 |
| vertical translation | 0.03 | 0.06 | 0.15 |
| horizontal translation | 0.05 | 0.09 | 0.3 |

The experimental results clearly show that the RBF NN-based registration scheme is accurate and more robust to noise, especially in estimating the scale and rotation parameters of the affine transformation. For images with relatively high SNR, this performance difference becomes more apparent in favour of the RBF NN-based registration scheme. With the data size chosen in the experiments, training a RBF NN takes much less than one second on an average personal computer. On the other hand, training a FNN with the same data takes from many seconds to several minutes.

4. Conclusions:

This paper proposes to use a RBF NN in NN-based image registration approach fed by global image features such as DCT coefficients. That schemes, only a few known so far, all use a FNN to estimate the registration parameters. On the other hand, It is shown here, experimentally, that employing a RBF NN instead of FNN to estimate affine registration parameters gives more accurate and robust results in the presence of noise. This performance superiority of the RBF NN-based scheme can be accounted for its better affine transformation generalizations over the parameter space. More importantly, the proposed scheme is fast and easy to implement as a result of avoiding the disadvantages of FNN-based scheme, such as lengthy iterations and nonconvergence problem encountered during the network training stage. Only parameter that has to be determined to well-train an exact RBF NN is the spread parameter of the Gaussian function. Although there is an optimal spread value depending on the training data for a network in the training stage, the experiments also show that any suboptimal spread value can be easily estimated and used without decreasing the performance drastically. The proposed scheme can also be applied to any other image features than the DCT coefficients. As a future work, developing schemes to use RBF NN to estimate registration parameters from a much wider range will be investigated and their performance tests will be carried on.

References:

- [1] L.G. Brown, *A Survey of Image Registration Techniques*, ACM Computing Surveys, Vol. 24, No. 4, P. 325-376, 1992.
- [2] B. Zitová and J. Flusser, *Image registration methods: a survey*, Image and Vision Computing, Vol. 21, No. 11, P. 977-1000, 2003.
- [3] D. Capel and A. Zisserman, *Computer vision applied to super-resolution*, IEEE Signal Processing Magazine, Vol. 20, No. 3, P. 75-86, 2003.
- [4] Z. Qian, and J. Li, *Use of Hopfield Neural Network for Complex Image Registration*, Proceedings of the 9th International Conference on Tools with Artificial Intelligence, P. 204-207, 1997
- [5] I. Elhanany, M. Sheinfeld, A. Beck, *et al.*, *Robust Image Registration Based on Feedforward Neural Networks*, IEEE International Conference on Systems, Man and Cybernetics, No. 12, P. 1507-1511, 2000.
- [6] J. Wu and J. Xie, *Zernike Moment-based Image Registration Scheme Utilizing Feedforward Neural Networks*, Fifth World Congress on Intelligent Control and Automation (WCICA 2004), Vol. 5, P. 4046-4048, 2004