Available online at www.sciencedirect.com
ScienceDirect
Neurocomputing 71 (2008) 2046–2054
NEUROCOMPUTING
www.elsevier.com/locate/neucom

A new sub-pixel mapping algorithm based on a BP neural network with an observation model

Liangpei Zhang*, Ke Wu, Yanfei Zhong, Pingxiang Li

State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Available online 4 March 2008

Abstract

The mixed pixel is a common problem in remote sensing classification. Even though the composition of these pixels for different classes can be estimated with a pixel un-mixing model, the output provides no indication of how such classes are distributed spatially within these pixels. Sub-pixel mapping is a technique designed to use the output information with the assumption of spatial dependence to obtain a sharpened image. Pixels are divided into sub-pixels, representing the land cover class fractions. This paper proposes a new algorithm based on a back-propagation (BP) network combined with an observation model. This method provides an effective method of obtaining the sub-pixel mapping result and can provide an approximation of the reference classification image. With the upscale factor, the model was tested on both a simple artificial image and a remote sensing image, and the results confirm that the proposed mapping algorithm has better performance than the original BPNN model.

Keywords: Neural networks; Classification; BPNN model; Multi-resolution; Observation model

1. Introduction

Remotely sensed images usually contain both pure and mixed pixels. Hard classification techniques assign every pixel (often incorrectly) to a single class; a loss of information is inevitable during this process, because mixed pixels can be composed of different classes in varying combinations. Soft classification techniques (or sub-pixel classifiers) were introduced to compensate for this loss. They assign pixel fractions to the land cover classes corresponding to the areas represented inside a pixel. A soft classification yields a number of fraction images equal to the number of land cover classes. However, the assignment to these classes renders no information about the location of these fractions inside the pixel. Atkinson [3] stated that it is possible to assign the fractions spatially to so-called ''sub-pixels'', where sub-pixels are a finer resolution representation of a parent pixel. Sub-pixels can be either pure or mixed. In this work, sub-pixels will be assumed pure, belonging to one class only. Sub-pixel mapping can be seen as a technique that uses the information present in a soft classification to attain a higher resolution representation. The aim is to accord fractions from a sub-pixel classification spatially to the sub-pixels that are parts of a pixel.

Many different techniques [1,3,6,10,11,13,18,20,21] have been proposed to tackle the sub-pixel mapping issue. These techniques are based on the concept of spatial dependence, which refers to the tendency of proximate sub-pixels to be more alike than those located far apart. In all these approaches, the detail and accuracy of the super-resolution map were greater than the corresponding hard-classified images. Although these super-resolution mapping techniques were successful to varying degrees, there is a limit to the detail and accuracy of the resulting mapping image, since these techniques are based only on the soft-classified proportion data at the pixel level and the spatial dependence assumption. Therefore, it is suggested that additional noise produced during the process of fractioning an image could not be reduced. This noise could be detrimental to the location of sub-pixels inside a pixel because it leads to inaccurate traces and blurs the final sub-pixel mapping image. Therefore, it is still necessary to develop new super-resolution mapping

*Corresponding author. Tel.: +86 27 68778452; fax: +86 27 68778229.
E-mail address: zlp62@public.wh.hb.cn (L. Zhang).

0925-2312/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2007.08.033

methods to eliminate the additional noise and enhance the precision of the sub-pixel mapping.

Neural networks have proven their modeling capabilities in many instances and can learn relations from examples, without making assumptions about data distribution or the nature of the relation between inputs and outputs [4,16]. The building blocks of artificial neural networks (ANNs) are artificial neurons. These neurons, like biological neurons, are connected to many other neurons. To compute the output of a neuron, each input is multiplied by a weight factor. The sum of these weighted inputs is called the neuron activation. The neuron output is calculated as a function of the activation. It is well known that neural networks [2,7,14] can be applied to sub-pixel (soft) classification. However, sub-pixel mapping is clearly distinct and it can use the sub-pixel classification as an input. The networks used in this paper are restricted to feed-forward networks, and feed-back connections are not possible. Using sub-pixel mapping examples, the network is trained to learn the most appropriate locations. A common procedure to find the appropriate network weights is the standard back-propagation (BP) algorithm, it has a slow rate of convergence and the convergence is confronted with locally optimal phenomenon. So the algorithm with momentum coefficient solution is presented here. At the same time there is also a drawback that neural networks are often considered to be black box models. Even though the sub-pixel mapping result is not precise, it is difficult to improve the result from the structure directly because of the limitation of the realization of the complex model. In order to overcome this drawback, a new observation model is proposed in this paper that is applied after the BP neural network is processed. The observation model describes the relationship between the fraction image and the spatial distribution of sub-pixels. In fact, the fractions yielded by the soft classification technology can be regarded as the low-resolution images, while the sub-pixels representing the land cover class fractions can be regarded as the high-resolution images. An iterative algorithm based on a constrained least-squares (CLS) solution is presented in the model to lessen the noise produced by the sub-pixel mapping with the neural network. Experiments and comparisons show that the improved BP model presented in this paper is an efficient approach in sub-pixel mapping.

The rest of this paper is organized as follows. Section 2 provides a detailed description of the proposed BPNN sub-pixel mapping model. Section 3 gives the experimental results of the improved algorithm, which are compared with hard classification and the original BPNN algorithm. Finally, a conclusion is drawn in Section 4.

2. Methods

2.1. Sub-pixel mapping

The key problem in sub-pixel mapping is determining the most likely locations of the fractions of each land cover class within the pixel. This can be accomplished by assuming spatial dependence, i.e., the tendency for spatially proximate observations of a given property to be more alike than more distant observations. Land cover is spatially dependent both within and between pixels on the condition that the intrinsic scale of variation is not smaller than the sampling scale imposed by the image pixels [3].

The large pixels are divided into smaller ones and the land cover is allocated to the latter, in such a way that spatial dependence is maximized. The main advantage of applying this technique is that it will avoid losing important information. There are S^2 sub-pixels per pixel, with S representing the scale factor in the row and column directions. From the coarse resolution soft classification, the values of the sub-pixels are computed. The neighboring pixel values take an important role and express spatial dependence. Since each sub-pixel is surrounded by neighboring ones, they are affected by each other, possibly belonging to the same land cover class. A simple representation of the problem and two possible scenes are given in Fig. 1. There is a 3×3 coarse spatial resolution set of pixels, with associated proportions of one land cover class. A single coarse resolution pixel is to be divided into 16 sub-pixels, each corresponding to 1/16 of the coverage of the coarse resolution pixel. The value for a sub-pixel (x_{ij}) was calculated as:

$$x_{ij} = \begin{cases} 1 & \text{if sub-pixel } j \text{ is assigned to land cover class } i \\ 0 & \text{otherwise} \end{cases}$$

Summing x_{ij} for all S^2 sub-pixels returns the spatial dependence for that specific configuration of the sub-pixels inside the pixel (Eq. (1)):

$$\eta_i = \sum_{j=1}^{S^2} x_{ij} \tag{1}$$

In order to reach a certain fraction value η_i, there are two different arrangements in Fig. 1: the number of sub-pixels assigned to the land cover class can both correspond to the indicated proportion. But in fact, when enumerating all possible configurations with C possible land cover classes, there are $C^{(S^2)}$ possible configurations for every pixel through permutation 2003. To find the optimal configuration of the sub-pixels, the image was initialized with a random configuration for every pixel, yet respecting the fraction values.

2.2. Neural network model

2.2.1. Traditional BP network

ANNs are powerful tools for the prediction of non-linearities. These mathematical models comprise individual processing units called neurons that resemble neural activity. Each processing unit sums weighted inputs, and then applies a linear or nonlinear function to the resulting sum to determine the outputs. The neurons are arranged in

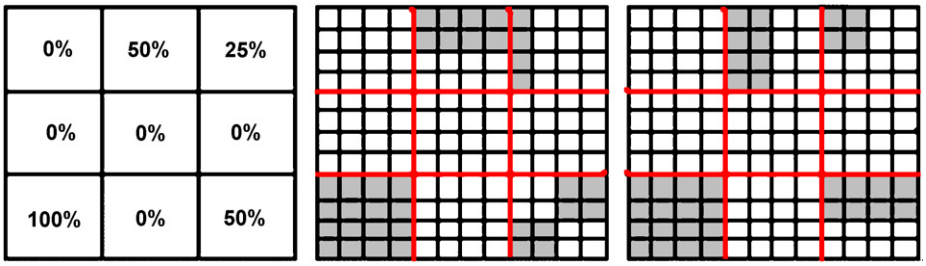

Fig. 1. A raster grid of 3×3 coarse pixels, each divided into 4×4 sub-pixels.

layers and are combined through excessive connectivity. The BP networks are extensively employed in back-analysis because of their simplicity and power to extract useful information from patterns (i.e., samples). It allows specification of multiple input criteria and the generation of multiple output recommendations, without pre-assumptions regarding the form of functions related to input and output variables. The BP model eliminates the limitations of the traditional regression methods, and accurately establishes the mapping between the input and output variables. It can approximate an arbitrary nonlinear function with better precision. It is a monitored learning method, and the training course is stopped when the error function reduces to below a given tolerance. Then, the fixed structure of a BP model is obtained. The sigmoid function is widely employed as the activation function:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

2.2.2. Improved BP network algorithm

BP network has a good diagnostic capability, but they too have left some problems unsolved, including those of local and slow convergence in training. There are many learning algorithms or modifications of the BP algorithm in the literature [5,15,17,19], momentum coefficients aimed to decrease the BP network's sensitivity to small details in the error surface, accelerate the convergence, normally with momentum coefficient a network can slide through some shallow local minima and the convergence speed can be improved. The adjustments of the learning rate and momentum coefficient are given as follows.

Let x_i ($i = 1, 2, \ldots, m$) be the inputs to the network, y_j ($j = 1, 2, \ldots, p$) the outputs from the hidden layer, o_k ($k = 1, 2, \ldots, q$) the output layer and w_{jk} the connection weight from the jth hidden node to the kth output node. By adding an additional term proportional to previous weights into the current weights $w_{fk}(n)$ based on the BP method, new weights $w_{fk}(n+1)$ can be produced as follows:

$$w_{jk}(n+1) = (1-\alpha)\eta\delta_k y_j + \alpha w_{jk}(n) \tag{3}$$

$$\delta_k = (d_k - o_k)o_k(1 - o_k) \tag{4}$$

where n is the training epoch, η denotes the learning rate, α is the momentum coefficient, d_k is the target output value, and δ_k is the error derivative of BP. When the iteration converges to an acceptable error, learning ends and the weights are determined. In the training process, improved BPNN is applied to construct a local sub-pixel mapping model that describes the relationship between fractions in the local window and the spatial distribution of sub-pixels assigned to the target in the central coarse pixel [22]. For convenience, it is explained as a simple expression (5). Let the zoom factor as 2, under this condition, every neuron represented coarse pixel can be affected by the eight neighbors and the application of the influence is different. The nine neurons in a fractional image are taken as input of the network, according to centered neuron, there are four sub-pixels in the model. The relationship expressed can be described as:

$$\begin{bmatrix} x_{i-1,j-1} & x_{i-1,j} & x_{i-1,j+1} \\ x_{i,j-1} & x_{i,j} & x_{i,j+1} \\ x_{i+1,j-1} & x_{i+1,j} & x_{i+1,j+1} \end{bmatrix} \rightarrow \begin{bmatrix} y_{ij}^1 & y_{ij}^2 \\ y_{ij}^3 & y_{ij}^4 \end{bmatrix} \tag{5}$$

where

$$y_{ij}^k = \begin{cases} 1 & \text{if it is corresponding to target class} \\ 0 & \text{otherwise} \end{cases}$$
$$k = 1, 2, 3, 4 \tag{6}$$

After the training process is finished, the two groups of the weight values can be acquired. In the predicting process, the input pattern is normalized to (0–1) to ensure that similar training samples can be obtained for as many different images as possible. Obviously the number of the output nodes in the network is determined by the factor size, but the number of the input nodes is changeless. Each sub-pixel in the sub-pixel mapping result is an output pattern, exported in the form of $Y = y_{ij}^k$ ($k = 1, 2, 3, 4$), in this way the outputs Y are not integral, but the probability of sub-pixel in the high-resolution image cannot be dissatisfied. The land covers of this sub-pixel are needed to label.

Two ways of deriving the sub-pixel composition from the network output are considered. One method simply assigns each sub-pixel to the class with the highest output value. This method, however, can only be used for small images with few land cover classes and a small up-scaling factor S. Another method maintains the fractions of the different classes. Using the second method, the output fraction was

then weighted in respect to the sum of the outputs in the set for that sub-pixel. Differences with the first method were found when the highest value belonged to the class whose total number of sub-pixels to be assigned (defined by the fraction values) had already been exhausted [12].

2.2.3. Architecture of the network

A three-layer network was used. The input layer consisted of nine neurons, corresponding to a central pixel and its eight surrounding neighbors. This implied that sub-pixels could not be calculated for border pixels, as part of the information about their surroundings was missing.

The input values were the membership values of these pixels for a certain class. The (unknown) spatial configuration was assumed equal for all classes. Consequently, the network was trained using membership values regardless of their originating class. In this way, only one network for all classes had to be trained, instead of having to train a separate network for each land cover class. The number of neurons in the hidden layer was set at 15. In this training algorithm, an input pattern is first propagated through the network in the so-called feed-forward phase. Afterwards, the difference between the calculated and the desired output is back-propagated from the output neurons to the first layer of the network, thereby adjusting the network weights in the opposite direction of the derivative of the network error with respect to each individual network weight. The learning rate was set at 0.1, and a momentum coefficient of 0.2 was applied. It is believed that at that point network generalization is maximal. By repeating this multiple times for each pattern in a training set, the network can be taught to map the inputs on the correct outputs. The network architecture, together with the input and output representation, is illustrated in Fig. 2.

2.3. Sub-pixel mapping algorithm based on an observation model

BP is a popular algorithm employed for training multilayer connectionist learning systems with nonlinear activation functions (sigmoid). However, its drawback has been partly attributed to the following: (1) there are some flat-spot problem regions, where the derivative of the sigmoid activation function approaches zero, and the weight changes become negligible; (2) because the gradient descent algorithm is used here, the constructed high-resolution image is not very good, as indented shapes appear in the boundary, which makes the algorithm inefficient. To resolve this problem, an observation model is introduced in every fraction image after the BP.

Fryer and McIntosh, proposed a super-resolution image reconstruction method (FM) [8]. In their method, the observing model must be established in which the relationship between the low-resolution images and the high-resolution images is strictly geometrical counterpoint, that is, each pixel in the low-resolution image must be "mapped" onto the high-resolution pixel coordinate system, thus determining which high-resolution pixels are affected by each individual low-resolution one [9]. The observation model was incorporated after the BP worked in order to overcome the problems mentioned before. This method is defined as BPFM in this paper. For example, in

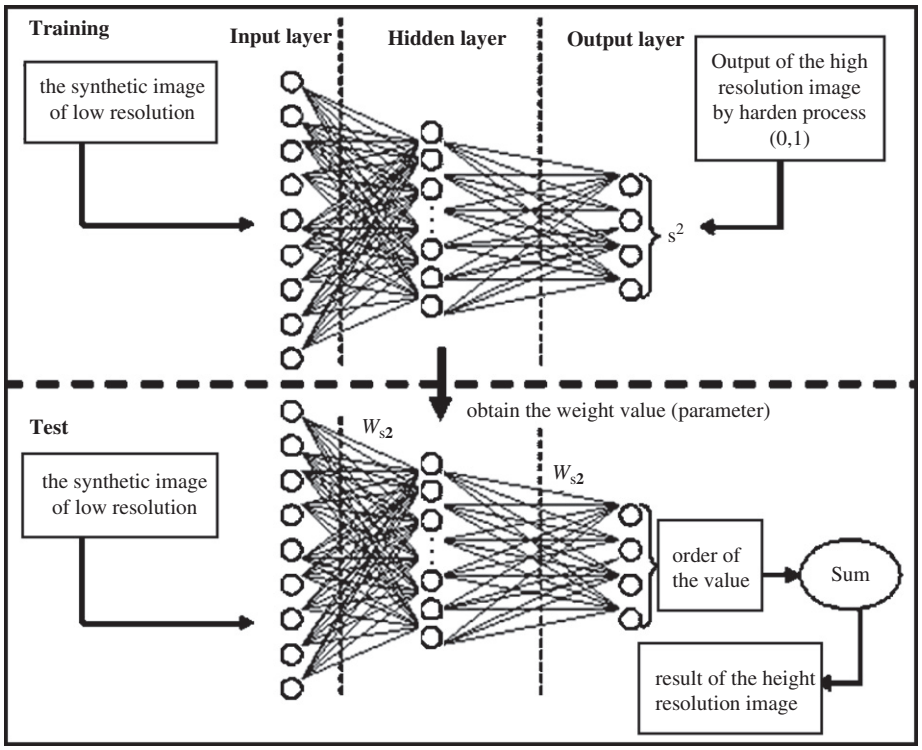

Fig. 2. Neural network architecture.

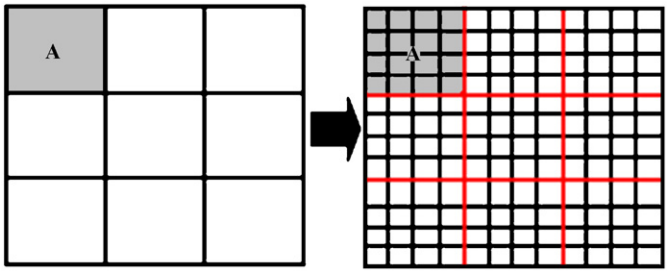

Fig. 3. Observation model: low-resolution pixels mapped on the high-resolution grids.

Fig. 3, for pixel A, we can obtain the following function:

$$y = \frac{[x_1 + x_2 + \cdots + x_{16}]}{\rho^2} \tag{7}$$

where y is the gray value of the low-resolution image, x_i represents the gray values of the high-resolution image, and ρ is the enhancement ratio, which in this case is 4.

Therefore, the observation model can be expressed in matrix notation as:

$$y = Ax + n \tag{8}$$

where y is a vector composed of the gray values of the low-resolution image pixels, x contains the required high-resolution image pixels, A is the coefficients matrix, and n is the noise matrix that is produced in the process of BP sub-pixel mapping. From Section 2, the sub-pixel mapping result $\mathbf{x}_{BP}$ based on the BP algorithm has been obtained when the value of y is unchangeable; the smaller $||x - x_{BP}||$ is, the better the result will be. Assuming a constrained least-squares (CLS) solution, only the noise matrix n is considered minimal, and we have the solution:

$$\mathbf{x} = \text{argmin}||\mathbf{n}||^2 = \text{argmin}||\mathbf{y} - \mathbf{A}\mathbf{x}||^2 \tag{9}$$

The above is considered as a constrained optimization, just corresponding to:

$$\boldsymbol{x} = \text{argmin}\{||\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}||^2 + \alpha||\boldsymbol{C}\boldsymbol{x}||^2\} \tag{10}$$

where α is the selective parameter controlling the terms, C is a Laplacian operator, which can smooth the noise:

$$\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y) \tag{11}$$

We produce the function:

$$\boldsymbol{x} = [\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} + \alpha\boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}]^{-1} \times \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \tag{12}$$

An iterative reconstruction algorithm is proposed in this process; the sequence of iterations is generated by:

$$\mathbf{x}_k = \mathbf{x}_{BP} + [\mathbf{A}^{\mathrm{T}}\mathbf{Y} - (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \alpha_k\mathbf{Q}^{\mathrm{T}}\mathbf{Q})\mathbf{x}_{BP}] \tag{13}$$

The criterion used to terminate the iterations is defined as:

$$\frac{||\mathbf{x}_k - \mathbf{x}_{BP}||^2}{||\mathbf{x}_{BP}||^2} \leqslant \mathrm{d} \tag{14}$$

where d is the iteration termination coefficient.

3. Experiments

3.1. Synthetic imagery

The sub-pixel mapping algorithm should be applied on fraction images of a spatial resolution. The result can then be compared to a hard classification of a finer resolution image. Following this approach, errors due to co-registration and poor soft classification are introduced. In these experiments, the algorithm was developed and tested only on synthetic imagery. In that way it was possible to concentrate solely on errors introduced by the sub-pixel mapping process. Synthetic imagery can be created by degrading hard classifications to coarser spatial resolutions by applying an averaging filter. The source material is either real or artificial. The resulting images are interpreted as fraction images.

For analysis of convenience, the artificial images are imitated from reference [13] to aid in the design and development of the algorithm. Their usefulness is determined by their size, resulting in a considerable reduction of computation time. Visual checking of the algorithm's performance is easy on these images as they represent simple geometric figures. The real imagery used in this work is from a Landsat Thematic Mapper (TM) image of the northern part of Wuhan in Hubei Province (October 11, 1998; 30 m resolution). The primary objective of the survey was to discriminate various objects. Hence, four major land cover classes can be distinguished: river, urban city, vegetation, and lake. These images are used as input for a sub-pixel mapping. The aforementioned algorithm was tested on two different images. Experiments were conducted to test its performance. Consistent comparisons among BPFM, traditional MLC (Maximum Likelihood Classification) algorithms and BP neural network are completed. The estimation of classification accuracy for the different methods is provided.

3.2. Experiment 1: artificial imagery

The accuracy is calculated in terms of percent correctly classified (PCC) and the Kappa coefficient (Cohen's Kappa) of agreement. In addition, there is a new assessable method called PCC′, which is different from PCC. In order to take out the effect of the accuracy assessment brought by the existing un-mixed pixels, it is only concerned with the right proportion of mixed pixels in the imagery, as opposed to all undistinguished pixels. Visual assessment of the results, however, remains very important.

The tricuspoid image was used as the training set, while the test set contained the concentric circles membership values. The trained network was then used to predict the spatial pattern for the concentric circles. The original concentric circles image, illustrated in Fig. 4(a), has two classes. The BP network was trained on a tricuspoid image, which together with the derived fraction image is shown in Fig. 4(b), and (c). The hard classification of the degraded artificial image and the sub-pixel mappings using the BP and BPFM algorithm are shown in Fig. 4(d)–(f).

Visual assessment leads to the conclusion that sub-pixel mapping makes sense, and the proposed method is efficient. The hardened soft classification result leads to blurry concentric circles with scale 4. The two sub-pixel mapping methods compensate for this. In some regions, there are many mixed pixels. The result of the BP algorithm is not accurate, as indented holes are introduced into the characters, which is not a desirable feature, always giving a blurry scene. The proposed method suffers less from this side-effect and makes the border of the classification image

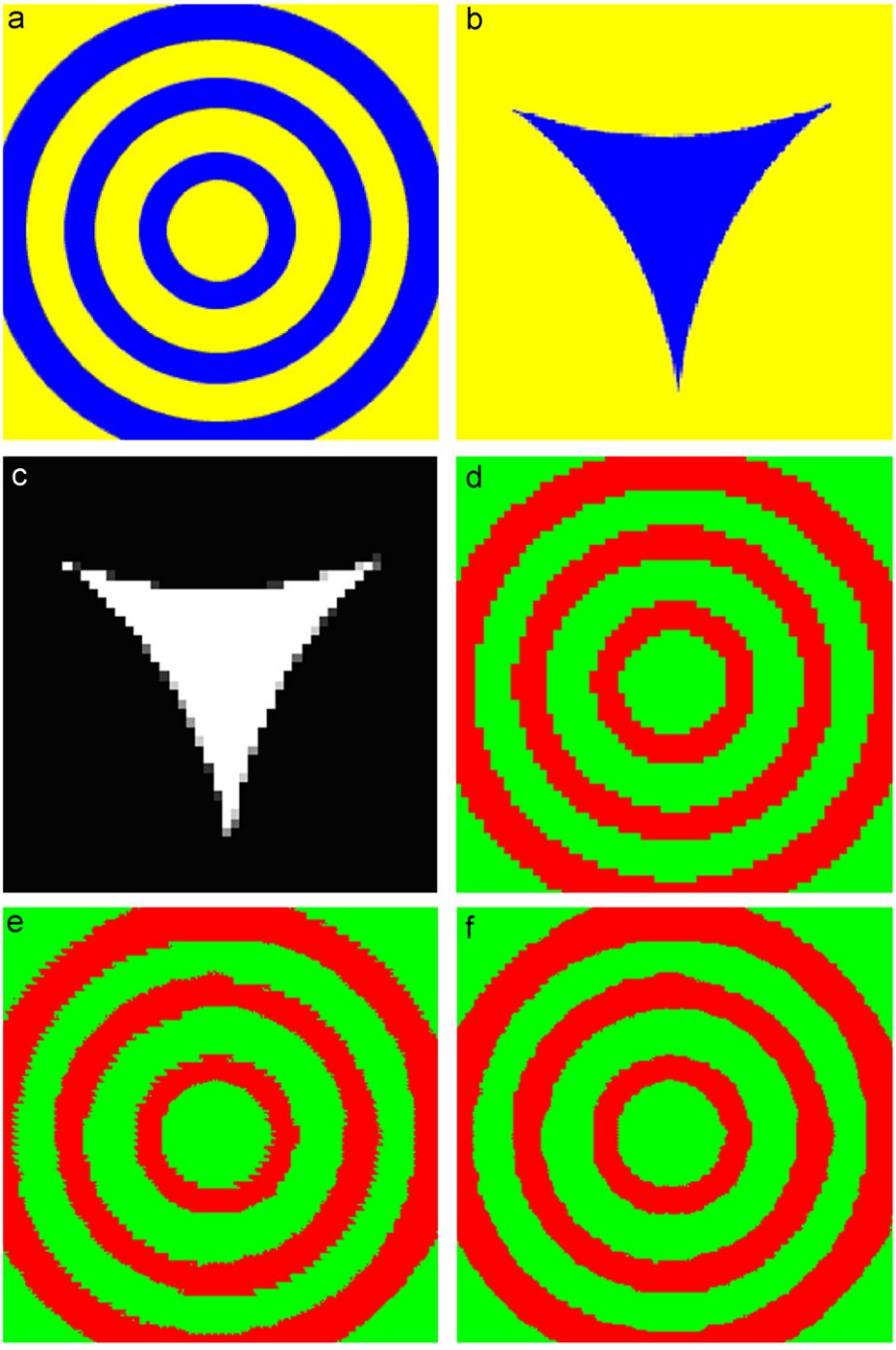

Fig. 4. Concentric circles image scale 4. (a) Original artificially image, (b) tricuspoid image for training, (c) fraction image class 1, (d) hard classification, (e) sub-pixel mapping for BP, and (f) sub-pixel mapping for BPFM.

Table 1
The accuracy statistics of the classification results with MLC, BP, and BPFM

	Concentric circle			Real imagery		
	MLC	BP	BPFM	MLC	BP	BPFM
Kappa coefficient	0.813	0.832	0.891	0.718	0.806	0.859
PCC (%)	90.7	96.1	98.6	75.0	88.2	91.1
PCC′ (%)	70.2	97.2	99.0	75.4	85.7	88.6

smooth. It is considered better than the BP network techniques.

This is also confirmed by the accuracy measures. Table 1 shows the Kappa coefficient, PCC and PCC′ values for sub-pixel mapping of the synthetic shapes. BPFM improves the Kappa coefficient from 0.813 to 0.891, an improvement of 0.078, and BPFM exhibits the best overall classification accuracy and the best percentage of correctly classified pixels among all the test pixels considered, with a gain of 7.9% and 2.5% over the MLC and BP algorithms,

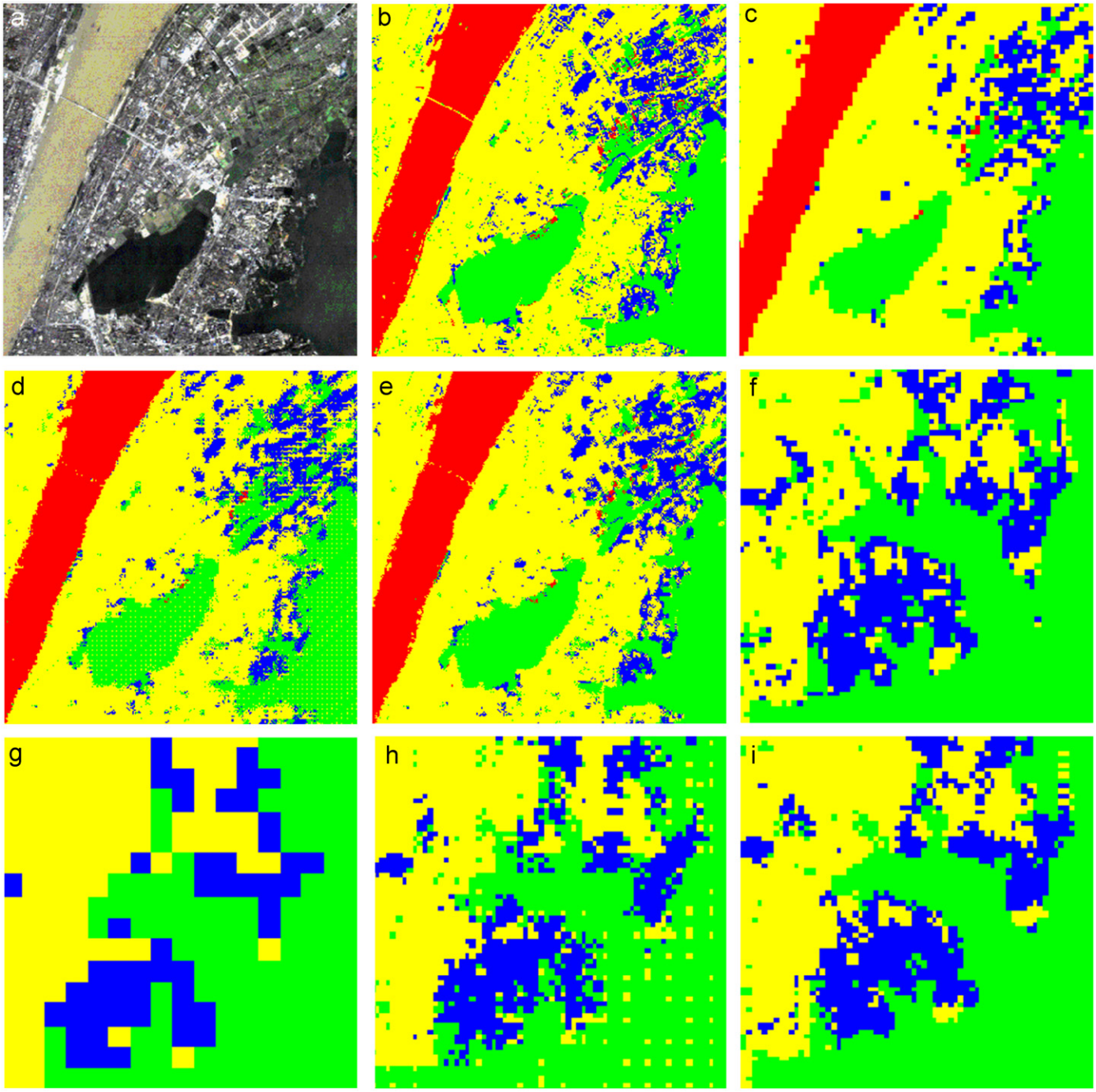

Fig. 5. Real image for Wuhan TM scale 4. (a) Original real image for TM, (b) reference image, (c) hard classification, (d) sub-pixel mapping for BP, (e) sub-pixel mapping for BPFM, (f) zoom 6 for reference image, (g) zoom 6 for hard classification, (h) zoom 6 for BP, and (i) zoom 6 for BPFM) (, River; , Urban; , Vegetation; , Lake).

respectively. The PCC′ gives an improvement over the MLC and BP algorithms of 19.8% and 1.9%. This is due to the existence of the mixed pixels; the hard classification method cannot correctly classify these pixels. On the other hand, BPFM can overcome this problem.

3.3. Experiment 2: real imagery

Sub-pixel mapping of the real world case is considered a more complex problem. With more land cover types, the sub-pixel mapping algorithm has to solve a more complex problem, as the classes of sub-pixels per pixel increase.

Here we tested the proposed algorithm using the 30-m resolution multi-spectral Landsat TM image shown in Fig. 5(a). There is an original reference image, which is classified by the MLC method, and shown in Fig. 5(b); it is regarded as the truth data. Fig. 5(c) illustrates the hard classification of the degraded reference image. The neural networks for prediction of the real imagery were trained on an analogical classified TM image, its primary makeup are still river, urban city, vegetation, and lake, which is in the upper east part of Wuhan and holds the same assumption of spatial dependence. These two different classifications are chosen to demonstrate the same spatial dependence present among features in both images. It is believed that the more similar the images are, the higher the accuracy will be [13]. Using this trained image, the sub-pixel mapping images of the degraded real imagery using the BP model and the BPFM methods are shown in Fig. 5(d) and (e).

The visual comparisons of the three classifications in Fig. 5(b), (d), and (e) suggest varying degrees of accuracy of pixel assignment. It can be seen that the BP model and BPFM method have similar sub-pixel mapping results and they can provide much better location of the sub-pixels inside a pixel than the direct hard classification. Comparison with the original classified image makes it hard to differentiate between these two methods. In Fig. 5(f)–(i), an area in the southeast of the image has been chosen for display with a zoom factor of 6. This area clearly illustrates the differences between the separate images. Fig. 5(f) and (g) show the real reference image and its hardened degraded version. It can be observed that a considerable amount of detail is lost in the degradation process. Comparison of the images (h) and (i) leads to the conclusion that more isolated pixels, which can be taken as noise, are preserved using the BP model than when using the BPFM method. Taking a closer look at the lake in the east of the BP network result shows that a set of scattered pixels distributed in the lake are incorrectly classified. The BPFM sub-pixel mapping results both show a smoothing result between most, but not all, of the pixels, illustrating the utility of a sub-pixel mapping.

The accuracy measures for the real imagery are still acceptable, but worse compared to the synthetic imagery results. Table 1 shows the Kappa coefficient and PCC values for sub-pixel mapping of the real imagery compared to the synthetic shapes. This could be expected, due to the higher complexity of the real imagery. The confusion matrix (Table 2), still the core of today's classification accuracy assessment, shows that the BPFM model produces better classification results than other methods. In the more detailed verification of the results, the BPFM model exhibits the best overall classification accuracy, i.e., the best percentage of correctly classified among all the testing pixels considered, with a gain of 16.1% and 2.9% over the MLC and BP algorithms, respectively. PPC' improves with a gain of 13.2% and 0.9%. The Kappa coefficient was improved from 0.718 to 0.859, an improvement by 0.141. In Table 2, vegetation was mainly confused with urban and lake in quite a number of cases. The hard classification, on the other hand, contained both quantity and location errors.

Table 2
Comparison of three methods of classification by confusion matrix

Methods		River	Lake	Vegetation	Urban
MLC	River	89	9	7	2
	Lake	8	86	8	6
	Vegetation	3	4	66	21
	Urban	0	1	19	71
BP	River	95	3	0	2
	Lake	2	90	6	8
	Vegetation	2	6	82	10
	Urban	1	1	12	79
BPFM	River	98	3	0	4
	Lake	2	92	4	8
	Vegetation	0	4	89	5
	Urban	0	1	7	83

4. Conclusions

Sub-pixel mapping aims at increasing the use, and preserving the enhanced information content, of a soft classification mapping. This paper investigated the BP neural network, and proposed a new approach based on an observation model to improve the accuracy of sub-pixel mapping technology. Through dealing with the two synthetic images, the superiority of the BPFM over the original BPNN technique is proved with the assumption of the additional factor. The accuracy and the visual effect of the resulting image are both improved. Tests show that BPFM is an efficient approach in remotely sensed sub-pixel mapping. Future research will focus on further improvement of the proposed techniques and sub-pixel mapping accuracy assessment.

Acknowledgements

This work was supported by the 863 High Technology Program of the People's Republic of China under Grant No. 2007AA12Z148 and the National Natural Science Foundation of China under Grant No. 40771139.

References

[1] P. Aplin, P.M. Atkinson, Sub-pixel land cover mapping for per-field classification, Int. J. Remote Sens. 22 (2001) 2853–2858.

[2] P. Atkinson, A. Tatnall, Neural networks in remote sensing, Int. J. Remote Sens. 18 (4) (1997) 699–709.

[3] P.M. Atkinson, M. Cutler, H.G. Lewis, Mapping sub-pixel proportional land cover with AVHRR imagery [J], Int. J. Remote Sens. 18 (4) (1997) 917–935.

[4] J. Benediktsson, P. Swain, O. Ersoy, Neural network approaches versus statistical methods in the classification of multisource remote sensing data, IEEE Trans. Geosci. Remote Sens. 28 (1990) 540–552.

[5] I. Christian, H.U. Michael, Improving the rprop learning algorithm, in: Proceedings of the Second International Symposium on Neural Computation, ICSC Academic Press, US, 2000, pp. 115–121.

[6] G.M. Foody, Sharpening fuzzy classification output to refine the representation of sub-pixel land cover distribution, Int. J. Remote Sens. 19 (1998) 2593–2599.

[7] G.M. Foody, Advances in Remote Sensing and GIS Analysis, Wiley, Chichester, 1999 (Chapter 3).

[8] J. Fryer, K. McIntosh, Enhancement of image resolution in digital photogrammetry [J], Photogramm. Eng. Remote Sens. 67 (6) (2001) 741–749.

[9] J. Fryer, G. Scarmana, Algorithm development for the enhancement of photogrammetric digital images to improve DEM generation, 〈http://www.isprs〉.

[10] H.N. Gross, J.R. Schott, Application of spectral mixture analysis and image fusion techniques for image sharpening, Remote Sens. Environ. 63 (1998) 85–94.

[11] K.C. Mertens, L.P.C. Verbeke, E.I. Ducheyne, R.R. De Wulf, Using genetic algorithms in sub-pixel mapping, Int. J. Remote Sens. 24 (2003) 4241–4247.

[12] K.C. Mertens, L.P.C. Verbeke, R.R.D. Wulf, Sub-pixel mapping and sub-pixel sharpening using neural network predicted wavelet coefficients[J], Remote Sens. Environ. 91 (2004) 225–236.

[13] K.C. Mertens, L.P.C. Verbeke, R.R.D. Wulf, Sub-pixel mapping with neural networks: real-world spatial configurations learned from artificial shapes, in: Proceedings of 4th International Symposium on Remote Sensing of Urban Areas, Regensburg, Germany.

[14] A. Moody, S. Gopal, A. Strahler, Artificial neural network response to mixed pixels in coarse-resolution satellite data, Remote Sens. Environ. 58 (1996) 329–343.

[15] S.C. Ng, C.C. Cheung, S.H. Leung, Magnified gradient function with deterministic weight modification in adaptive learning, IEEE Trans. Neural Netw. 15 (2) (2004) 1411–1423.

[16] J. Paola, R. Showengerdt, A detailed comparison of backpropagation neural network and maximum likelihood classifiers for urban land use classification, IEEE Trans. Geosci. Remote Sens. 33 (1995) 981–996.

[17] S. Perantonis, D. Karras, An efficient constrained learning algorithm with momentum acceleration, Neural Networks 8 (2) (1995) 237–249.

[18] C.R. Pinilla, F.J.L. Ariza, Restoring SPOT images using PSF derived deconvolution filters, Int. J. Remote Sens. 23 (2002) 2379–2391.

[19] K.C. Tan, H.J. Tang, New dynamical optimal learning for linear multilayer FNN, IEEE Trans. Neural Netw. 15 (1) (2004) 1562–1570.

[20] A.J. Tatem, H.G. Lewis, P.M. Atkinson, M.S. Nixon, Super-resolution land cover pattern prediction using a Hopfield neural network, Remote Sens. Environ. 79 (2002) 1–14.

[21] J. Verhoeye, R. De Wulf, Land cover mapping at sub-pixel scales using linear optimization techniques, Remote Sens. Environ. 79 (2002) 96–104.

[22] L. Wang, Y. Zhang, J. Li. BP neural network based sub-pixel mapping method. International Conference on Intelligent Computing, 2006.

Liangpei Zhang received the B.S. degree in physics from Hunan Normal University, Chang-Sha, China, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1982, 1988, and 1998, respectively. From 1997 to 2000, he was a Professor in the School of the Land Sciences, Wuhan University. In August 2000, he joined the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, as a Professor and head of the Remote Sensing Section. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence. He is Associate Editor of the *Geo-spatial Information Science Journal*. Dr. Zhang has served as Co-Chair of the SPIE Series Conferences on Multispectral Image Processing and Pattern Recognition (MIPPR), the Conference on Asia Remote Sensing in 1999, Editor of the MIPPR01, MIPPR05 Symposia, and the Chinese National Committee for the International Geosphere-Biosphere Programme.

Ke Wu received the B.S. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002. He is currently pursuing the Ph.D. degree in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan. His current research interests focus on multi- and hyperspectral image processing, artificial neural network, and pattern recognition.

Yanfei Zhong received the Ph.D. degree in Photogrammetry and Remote Sensing from Wuhan University, Wuhan, China, in 2007. He is with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan. His current research interests focus on multi- and hyperspectral image processing, artificial intelligence, and pattern recognition.

Pingxiang Li received the B.S., M.S., and Ph.D. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1986, 1994, and 2003, respectively. Since 2002, he has been a Professor in the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. His research interests include photogrammetry and SAR image processing.