Neural Network Technique for Merging Radar and Visible Band Images

By George Lemeshewsky

Open-File Report 96-301

U.S. Department of the Interior
U.S. Geological Survey
National Mapping Division

Neural Network Technique For Merging Radar And Visible Band Images

George Lemeshewsky
U.S. Geological Survey
521 National Center
Reston, VA 22092
Open-File Report 96-301

ABSTRACT

Merging images from multiple sensor systems produces a composite image that may provide more information than either image alone. Spatial features of SPOT panchromatic and interferometric synthetic aperture radar (IFSAR) images were merged to produce an improved IFSAR composite. Existing multiresolution analysis techniques were used to extract spatial edge features from the image pair. A neural network, trained to implement a logical OR-like function, was used to combine these features. The composite, or merged, image was reconstructed from the neural network combined edge features. These results were compared with those from an existing multiresolution merging technique that combines edge information by using a maximum amplitude criterion. Simple visual comparisons of the reconstructed, merged IFSAR images from these two techniques showed many similarities; however, the neural network merging produced improved, noise-free edge boundaries. Also, there were promising results for IFSAR speckle-noise reduction, based on an existing soft-threshold technique.

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

I. INTRODUCTION

Integrating radar and visible band images can produce a composite image with information from both images and this composite may be more easily interpreted than either separate image. This study used a known multiresolution merging process to first decompose two images into multiresolution edge feature images and then merge the edge samples by using a maximum amplitude criterion. The composite image was obtained by reconstructing the image from its merged edge information. For comparison, a second merging method was developed. For this new method, a neural network (NN) was trained to approximate a logical OR-like function when given a small neighborhood of edge samples from each image. That is, a multilayer feedforward NN was trained as a function approximator by using artificial examples of the desired input-to-output data mapping.

All merging examples were for a SPOT panchromatic (pan) band, 10-m sample distance image and a C-band interferometric synthetic aperture radar image (IFSAR). Reconstructing the IFSAR image from merged edge images gives the composite.

Following the premise that excessive speckle noise in the IFSAR image is objectionable for some interpretation tasks, we tested a preliminary method for speckle reduction; this method was based on a variation of a reported soft-threshold modification of multiresolution derived detail images. Speckle is due to constructive and destructive interference of radar signals. This causes radar image pixels in homogeneous regions of the scene to have large differences in brightness (Richards, 1993). Speckle noise-reduced IFSAR images were used in the merging experiments.

The following sections briefly describe these initial tests in merging and noise reduction. They include discussions of the existing and new NN merging methods, the NN training process, and the noise reduction method. Image examples illustrate noise reduction and merging. The last section gives conclusions and recommendations.

II. MULTIRESOLUTION MERGING

Ogden and others (1985) described multiresolution image pyramids and methods for merging two images. There are three steps: (1) each of the images is decomposed into Laplacian image pyramids, (2) Laplacian image samples of the image to be enhanced are replaced on the basis of a maximum amplitude selection rule, and (3) the merged, composite image is reconstructed from its modified Laplacian image pyramid.

Briefly, image pyramid techniques transform an image into a completely reversible (without loss) multiresolution representation. For complete information, see Burt and Adelson (1983). Image decomposition and reconstruction by pyramid techniques can be easily described (Ogden and others, 1985) by defining the processes known as reduce and expand. Reduce (RE) is the combination of low-pass spatial filtering followed by image down sampling (or subsampling) by a factor of 2 in each dimension. This report used the Gaussian-like filter kernel (1 4 6 4 1) (Burt, 1985). Expand (EX) involves the two processes: upsampling by a factor of 2 in each dimension,

followed by interpolation by low-pass spatial filtering.

In the following, subscripts denote the image pyramid resolution level of Laplacian (L) or Gaussian (G) images. Image G_0 at level 0 is the original image at full resolution. Image L_0 is computed as follows:

$$G_1 = \mathrm{RE}(G_0) \quad (1)$$

$$G_2 = \mathrm{RE}(G_1) \quad (2)$$

where G_0, G_1, G_2 comprise a three-resolution level Gaussian image pyramid. Then

$$L_0 = G_0 - \mathrm{EX}(G_1). \quad (3)$$

Similarly, image L_1 is

$$L_1 = G_1 - \mathrm{EX}(G_2). \quad (4)$$

Rearranging equation 3 and combining with equation 4 shows, for example, that the original full-resolution image, G_0, can be reconstructed exactly from its three-level (0, 1, 2) Laplacian (L_0, L_1, G_2) pyramid as follows:

$$\begin{aligned} G_0 &= L_0 + \mathrm{EX}(G_1) \\ &= L_0 + \mathrm{EX}(L_1 + \mathrm{EX}(G_2). \end{aligned} \quad (5)$$

This study used the multiresolution image decomposition and reconstruction indicated by equations 1, 2, and 3; that is, two resolution levels of L images. The L images are comparable to the result of band pass-like filtering the same level G image (Lim, 1990).

The method for combining the edge pattern images ($L_{k,a}$ and $L_{k,b}$) derived from two images (a, b) at resolution level k (Ogden and others, 1985) is the following maximum amplitude selection rule:

if $|L_{k,b}(x, y)| > |L_{k,a}(x, y)|$ then

$$L_{k,a}(x, y) = L_{k,b}(x, y)$$

else

$$L_{k,a}(x, y) = L_{k,a}(x, y)$$

end if. (6)

The merged image is obtained by reconstructing (per equation 5) modified, merged image(s) $L_{k,a}$.

In contrast to merging samples by equation 6, the following section describes a new NN-based technique that was developed to combine samples from images L_a, L_b.

III. NOISE REDUCTION

A technique for noise reduction of SAR images was reported by Lang and others, (1995). The technique started with a shift variant discrete wavelet transform and then computed shift invariant discrete wavelet transform coefficients. These coefficient values were then modified by the soft-threshold denoising method reported by Donoho. Following this concept, we tested a method for IFSAR image noise reduction using multiresolution image pyramids. Samples of the L image pyramid were merely modified, using the soft-threshold method, but without consideration for shift variant effects caused by the subsampling steps of pyramid generation. Basically, shifted versions of the original image will have different L image decompositions, and thus different

results to soft threshold noise reduction. This is the subject of further study.

The reported soft threshold function, $T_s(L(x, y), t)$, of Donoho (given by Lang and others, 1995), as applied here to L image samples $L(x, y)$ and threshold t is

$$T_s(L(x, y), t) = \begin{cases} sign\,(L(x, y)(|L(x, y)| - t), & |L(x, y)| \geq t \\ 0, & |L(x, y)| < t\,. \end{cases} \quad (7)$$

IFSAR denoising examples are given in a later section.

IV. NEURAL NETWORK PATTERN MERGING

An NN method was developed for merging L image samples. Figure 1 shows the two merging techniques: (1) maximum rule (equation 6) and (2) new NN method. Implementation of the NN merging required two steps: (1) training the NN's to produce a mapping of input to desired output and then (2) using the trained NN's to merge the L images.The following is a very brief description of a multilayer, feedforward NN used for merging edge pattern samples. For further information on NN see Rumelhart and others (1986), Pao (1989), and Wasserman (1993). An artificial NN, or NN, is any computing architecture that consists of massive parallel interconnections of simple "neural" processors (Lau and Widrow, 1990). Often (as in this study) multilayer NN's are computer simulations of interconnected "neural" processors or artificial neurons modeled as the net sum of the input signals to the neuron times the respective interconnection weight, followed by a nonlinear mapping of net sum to produce the neuron output. Net sum also includes a bias value: bias weight times +1.0 input signal. The nonlinear transfer function in this application was the widely used sigmoid: $f(net\ sum) = 1 / (1 + e^{-(\text{net sum})})$.

Starting with the inputs, a multilayer, fully connected NN consists of consecutive layers of neurons, and at each layer each neuron output is connected to all inputs of the next layer.
NN's were trained as function approximators by using artificial combinations of pan image data that represented a logical OR-like function mapping of input to output edge patterns. The pan, rather than the IFSAR image, was selected simply because it had less noise and thus might speed up training. Training with combinations of pan and IFSAR images has not been tested. The OR-like mapping was achieved by training, or adjusting, the NN weights: that is, presenting the NN with numerous and repeated examples of input to desired output and then adjusting the network weights, as a function on the error between desired and actual output such that the NN output approximates (least squares sense) the desired output. Since hidden layer(s) have no direct connection to the output, the output error is propagated back to the hidden layer. Backpropagation training by the generalized rule with momentum (Rumelhart and others, 1986) was used.

NN inputs associated with one output sample were 5 by 5 neighbor samples from the two L images, respectively; thus 50 inputs. During training, therefore, the output is assumed to be derived from a small local area of edge samples instead of a single location.

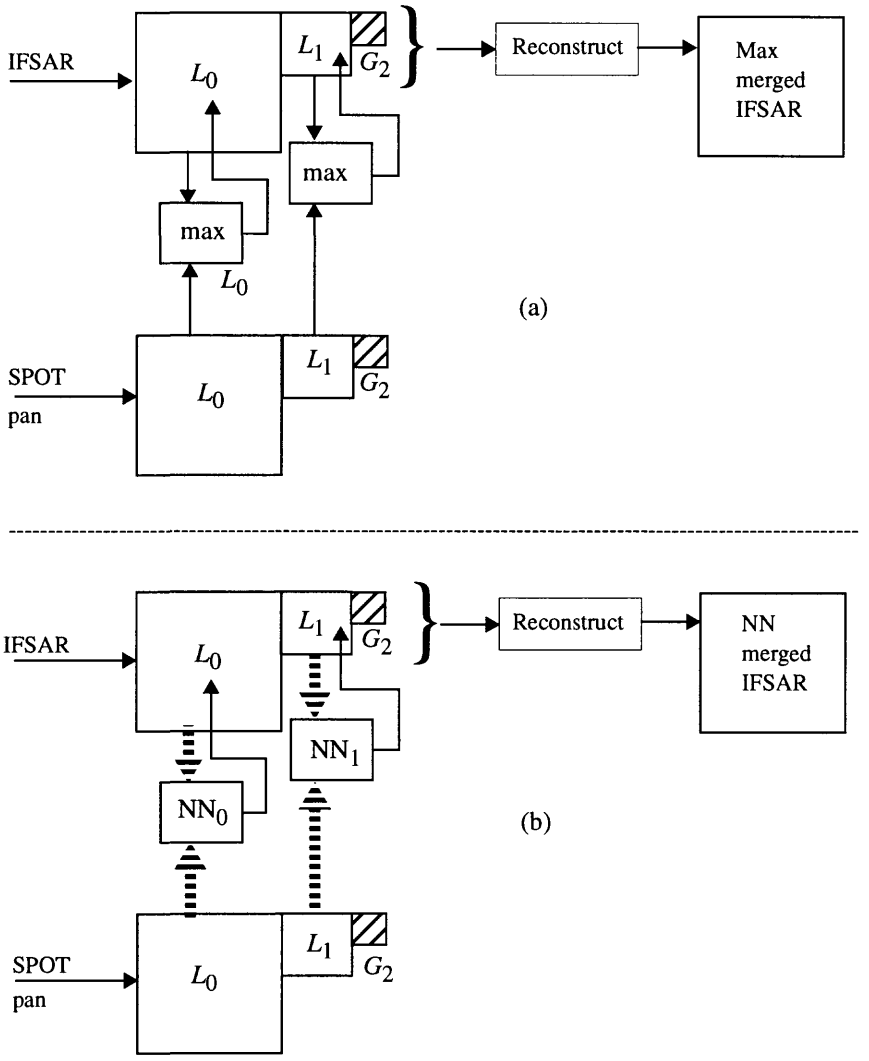

Figure 1. Two methods for multiresolution merging.

This shows the two methods for merging pan and IFSAR images after multiresolution decomposition to Laplacian image pyramids, L_0, L_1, G_2. (a) is the reported sample replacement merging using maximum amplitude criterion (equation 6). (b) is the new neural network method. Inputs to neural networks NN_0, NN_1 for one output sample 1 value are 5 by 5 sample values (larger arrows) from L_{pan} and L_{IFSAR} images, respectively. The merged image is obtained by reconstructing (equation 5) the IFSAR L image pyramid after sample replacement.

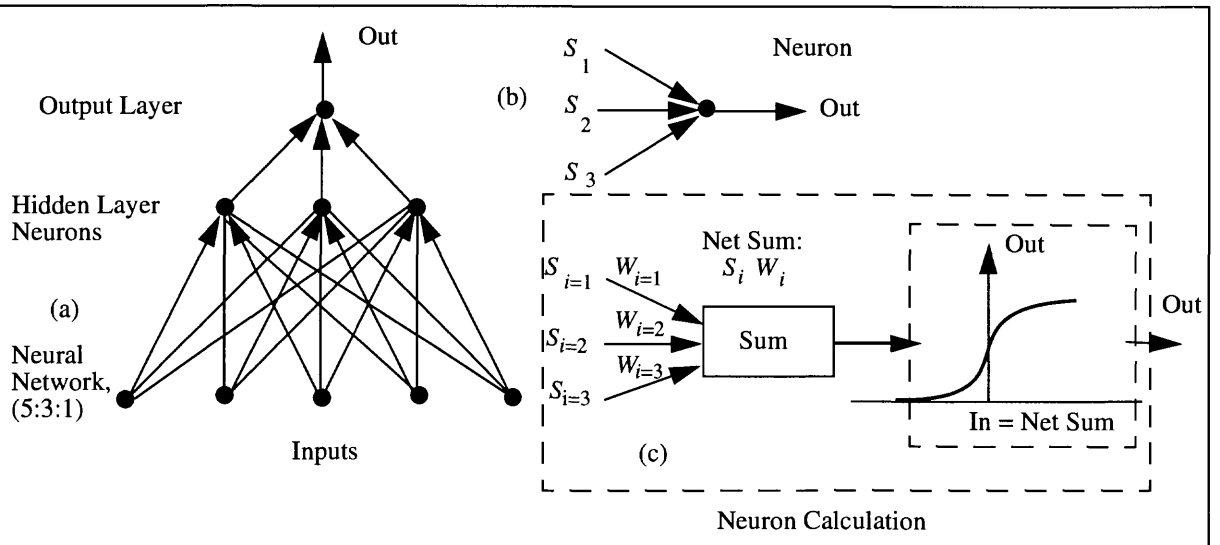

Figure 2. Simple one hidden layer, feedforward neural network, (NN).

(a) shows a multilayer, feedforward, fully connected NN with three 'neurons' or processing elements, (b) in the hidden layer. (c) shows the calculation performed by each processing element (b): the net sum of input, S_i times connection weight, w_i, followed by a nonlinear mapping of net sum to neuron output. (5:3:1) denotes the number of input, hidden layer, and output neurons.

Training data, which establish the mapping of input edge patterns to output edge samples, were created from several artificial OR-like combinations of pan L image data as shown in table 1. In table 1, Input_1 and Input_2 refer to the 5 by 5 neighborhood of samples (including the center) associated with a given sample location of the L_{PAN} image. The single L_{PAN} sample at this location is the desired NN output, OUT of table 1. To better simulate actual data, inputs denoted as '0' were small random values, with absolute value less than 0.50 gray level out of a nominal range of -128 to 128. Note that L images have an average value of zero. The training (and test) data were from a small area of the L_{PAN} image that included edge boundaries and regions of constant contrast.

Table 1: Neural network training data

INPUT_1	INPUT_2	OUT
L_{pan}	'0'	L_{pan}
'0'	L_{pan}	L_{pan}
L_{pan}	L_{pan}	L_{pan}

Because of their generalization property (similar inputs produce similar outputs), trained NN's function as adaptive estimators for the remaining image data, or in this case, the L_{PAN} and L_{IFSAR} image pairs. That is, generalization allows the NN to approximate the correct output when given input samples similar to, but not identical to, training set samples (Wasserman, 1993).

A separate NN was trained for each L image level (0, 1), in contrast to training one NN with data from two resolution levels. Additional study is needed to determine if, because of generalization, a single NN would be sufficient.

These preliminary tests were made with one relatively small, single hidden layer of five neurons, NN configuration (50:5:1) because previous edge pattern mapping experiments (not reported here) showed 5 to 10 hidden-layer neurons were appropriate. No other sizes were tested at this time. Additional training should be made with more and less hidden-layer neurons to establish an NN that performs best with training and test data. Two simple, common methods for determining the number of hidden neurons (and (or) hidden layers) are to (1) begin with a large number of hidden-layer neurons and reduce their number until performance degrades, or (2) begin with a small number of neurons and increase their number until performance stops improving.

Next, untrained 50:5:1 networks were trained in increments of a fixed number of training samples and the network weights (which are modified during training) saved after each increment of training (NeuralWare, 1993). Backpropagation training with momentum was used and weights were updated after each training sample presentation. Learning rate and momentum values were, respectively, (a) output layer, 0.0750, 0.0075, and (b) hidden layer, 0.1000, 0.0100 (NeuralWare, 1993). In these very preliminary tests, the training increment size was approximately, instead of exactly, equal to the number of training samples.

Network performance on the test data set (that is, RMS error) was computed after each incremental training cycle to give a measure of generalization. As incremental training progressed, each NN (level 0, 1) had a local minimum error to the test data; these NN's were retained for this study. With further training, this error increased and then decreased a small amount. Although this may be due to local minimum conditions, the lesser trained NN's were used in this study. This error variation may be due to the unequal training increment size and training file size, and thus not all training samples are presented during each training increment. Further testing is needed to resolve this.

It is likely that the NN's trained quickly, that is, after about twopresentations of the training file samples, because the training files contained many redundant training samples. Network performance on training and test data is summarized in tables 2 and 3; the RMS error is for the NN output range 0 to 1. NN's used in this study are denoted with "*." The fact that test set error is less than training set error is possibly because the test file contained many more input-output patterns that have lower RMS error; that is, input-output samples from non-edge regions.

Table 2: Neural network, level 0 performance versus training; 25,200 training and test samples

NN no.	Cumul. no. of samples, thousands	RMS training error	RMS test error
1	25	0.0300	0.0271
2	50	0.0276	0.0255
3	75	0.0266	0.0248
4	100	0.0261	0.0246
5	125	0.0259	0.0247
6	150	0.0254	0.0246
7	175	0.0254	0.0245

Table 3: Neural network, level 1 performance versus training; 6,930 training and test samples

NN no.	Cumul. no. of samples, thousands	RMS training error	RMS test error
1	7	0.0342	0.0295
2	14	0.0324	0.0284
3	21	0.0313	0.0277
4	28	0.0306	0.0264
5	35	0.0299	0.0265
6	42	0.0297	0.0263
7	49	0.0289	0.0253

VI. TEST RESULTS

The following examples were made with a SPOT pan image acquired in 1993, and a C-band (IFSAR) image acquired in June 1993, by the JPL TOPSAR system aboard a DC-8 aircraft (Giglio and Carlisle, 1995). The pan image was a georeferenced SPOT "Digital Ortho-Image" from the SPOTView sampler (SPOTView, 1995). It was geometrically coregistered (with cubic interpolation) to the IFSAR image. In these examples, the IFSAR image has been converted, and compressed, from 32-bit real data to 256 gray level, logarithm of intensity (magnitude) by the

following log_{10} (intensity) to gray level mapping: -3.5 to 0 and 0.5 to 255 (Decision-Science Applications, Inc., 1995, unpub. data). A linear, gray-level stretch was applied to all the following images to enhance the illustrations.

Figure 3 shows IFSAR images before (a) and after (b) noise reduction by soft threshold modification (threshold = 10.0) of only the L_0 image. This arbitrary threshold appears to give reasonable noise reduction without loss of image detail; however, more quantitative methods for its selection are needed for this application.

The SPOT pan and the previous noise-reduced (fig. 3(b)) IFSAR images are shown in figure 4 for comparison with the later merging examples of figures 5 and 6. Figure 5 compares the *maximum amplitude* criterion merged pan-IFSAR image and the IFSAR image.

Figure 6 compares the *neural network* merged pan-IFSAR image with the IFSAR image. Although merging results (figures 5, 6) for the two methods of combining edge information are visually similar, the NN merged image has improved edge boundaries with less noise. Future work will include a quantitative evaluation.

VI. CONCLUSION

A new NN-based technique for merging SPOT pan and IFSAR images was developed. Existing multiresolution analysis techniques were used to extract spatial edge features from the image pair, and an NN was trained to combine multiresolution edge patterns on the basis of a logical OR-like function. A composite image was reconstructed from the NN combined edge features. These image results were compared with those from an existing multiresolution merging technique that combines edge information on the basis a maximum amplitude criterion. Simple visual comparisons of the reconstructed, merged IFSAR images from these two techniques show many similarities; however, the NN merging produced improved, relatively noise-free edge boundaries. Also described were results of a simple test of an IFSAR noise-reduction process that used multresolution image pyramids. This process was loosely based on a reported soft-thresholding of multiresolution derived wavelet coefficients. Although these preliminary results are promising, further quantitative evaluation is needed.

REFERENCES

Burt, P.J., 1985, Smart sensing within a pyramid vision machine: Proceedings IEEE, v. 76, no. 8, p. 1,006-1,015.

Burt, P.J., and Adelson, E.H., 1983, The Laplacian pyramid as a compact image code: IEEE Transactions on Communications, v. 31, no. 4, p. 532-540.

Giglio, D., and Carlisle, R., 1995, Tpoographic mapping using interferometric synthetic aperature radar, in Algorithms for Synthetic Aperature Radar Imagery II, Dominick A. Giglio, Editor: Proc. SPIE 2487, p. 381-392.

Lang, M., Guo, H., Odegard, J.E., and Burrus, C.S., 1995, Nonlinear processing of a shift invariant DWT for noise reduction, in Wavelet Applications II, Harold H. Szu, Editor: Proc. SPIE 2491, p. 640-651.

Lau, C.G., and Widrow, B., 1990, Scanning the Issue - Neural Networks, I: Theory and Modeling: Proceedings of the IEEE, v. 17, no. 9, p. 1,411-1,413.

Lim, J.S., 1990, Pyramid coding, section 10.3.5 *of* Two-dimensional signal and image processing: Englewood Cliffs, N. J., Prentice Hall, p. 632-640.

NeuralWare, Inc., 1993, *Neural Computing*: Pittsburgh, Pa. NeuralWare.

Ogden, J.M., Adelson, E.H., Bergen, J.R., and Burt, P.J., 1985, Pyramid-based computer graphics: RCA Engineer, v. 30, no. 5, p. 4-15.

Pao, Y., 1989, *Adaptive Pattern Recognition and Neural Networks*: Reading, Mass., Addison-Wesley.

Richards, J.A., 1993, Sources and characteristics of remote sensing image data, chapt. 1 *of* Richards, J.A., Remote sensing digital image analysis: Berlin, Springer-Verlag, p. 1-37.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986, Learning internal representations by error propagation, chap. 8 *of* Rumelhart, D.E., and McClelland, J.L., Parallel distributed processing 1: Foundations: Cambridge, Mass., MIT Press, p. 318-362.

SPOT Image Corporation, 1995, *SPOTView* data sampler CD-ROM: Reston, Va., SPOT Image Corp.

Wasserman, P.D., 1993, Neural engineering, chap. 11 *of* Advanced methods in neural computing: N.Y., VanNostrand-Reinhold, p. 214-244.

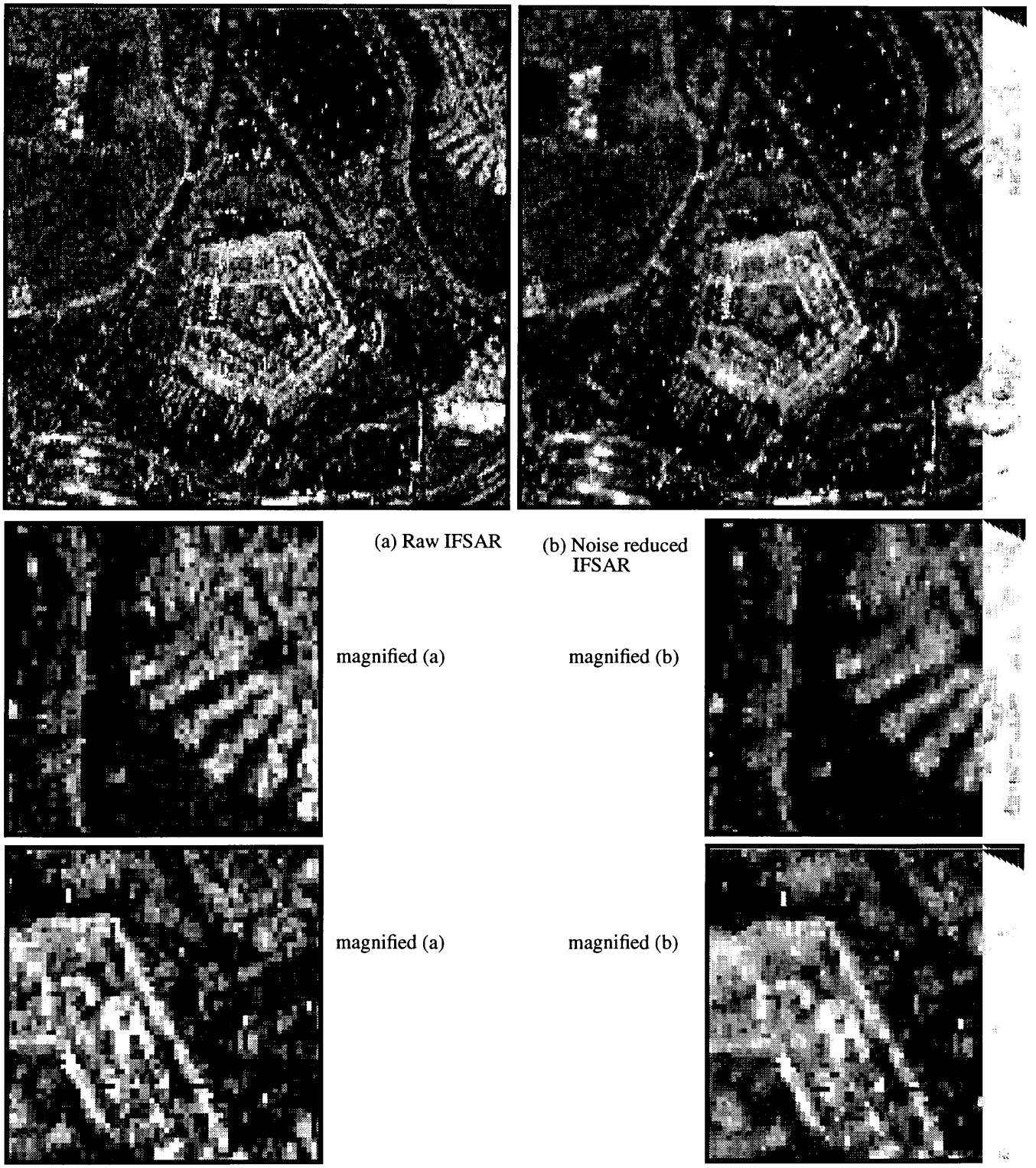

Figure 3. Noise-reduced IFSAR image.

(a) and (b) are the raw and noise-reduced IFSAR images, respectively. Image (b) is the result of soft threshold modification (threshold = 10.0) applied to the L_0 pyramid image of (a).

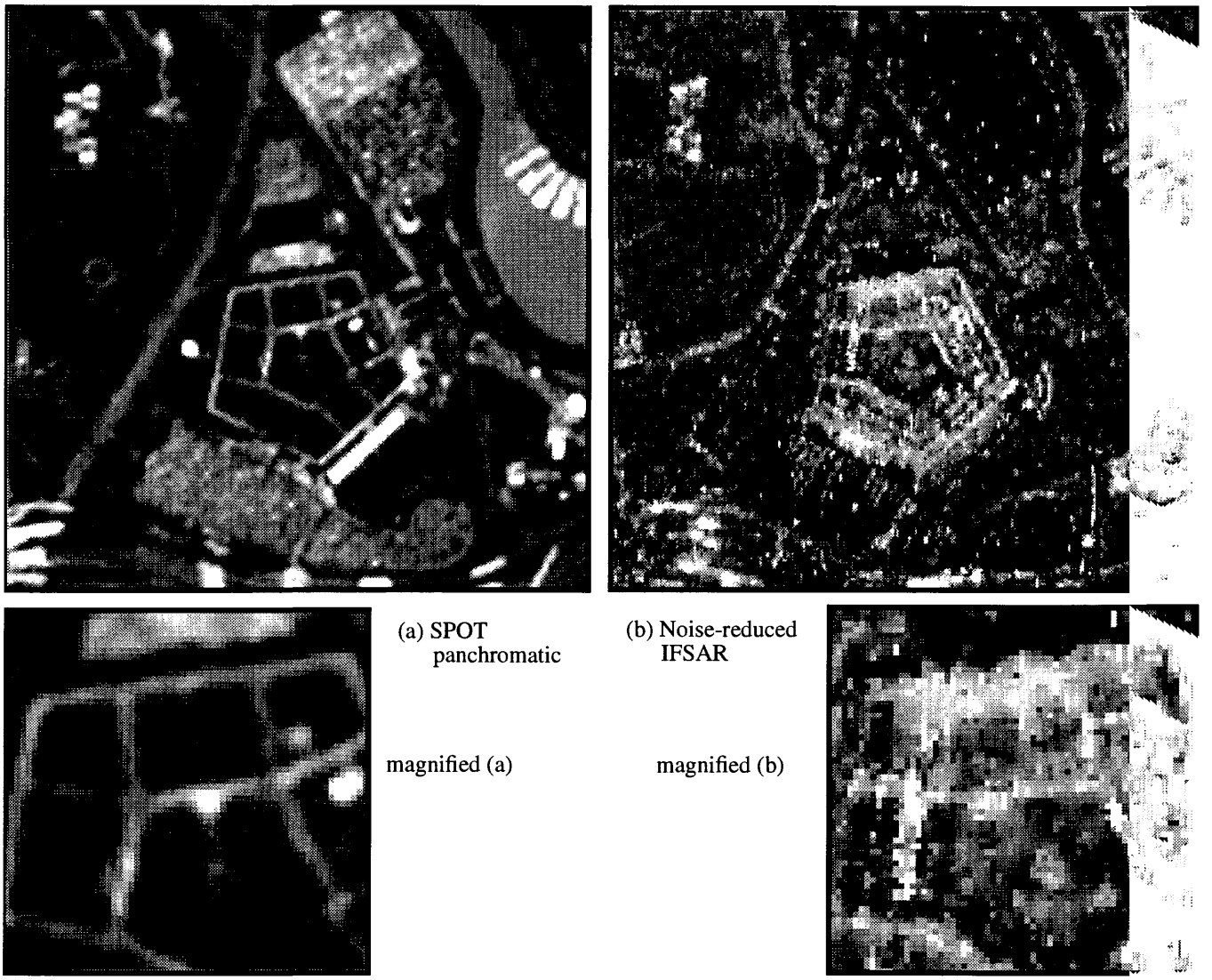

Figure 4. SPOT panchromatic and noise-reduced IFSAR image.

(a) and (b) are the SPOT panchromatic and noise-reduced IFSAR images, respectively used for the following me-g-ing tests.

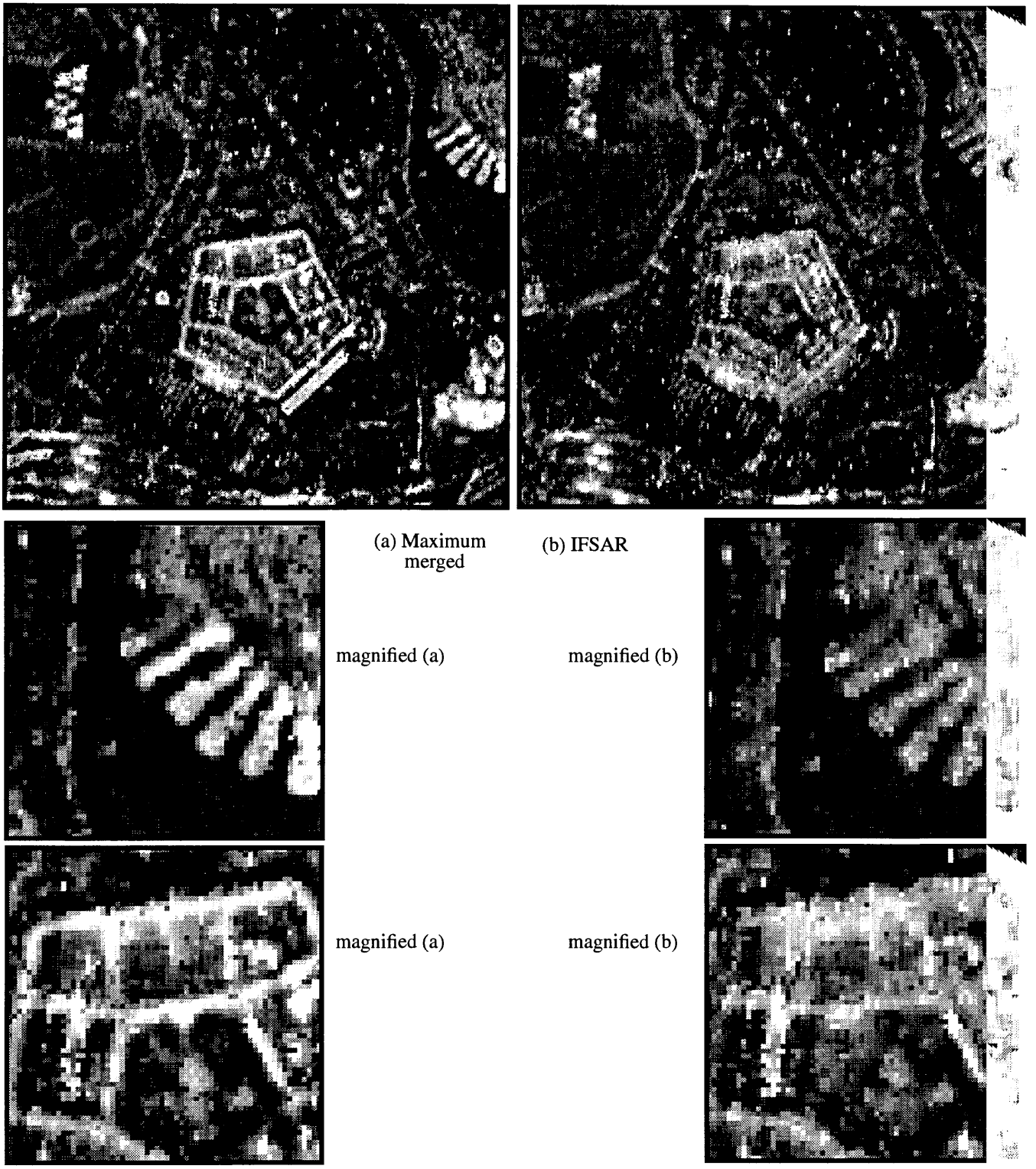

Figure 5. Maximum amplitude merged IFSAR images.

(a) is the result of maximum amplitude edge pattern merging of SPOT pan and noise-reduced IFSAR (b) images. A comparison between the maximum merged image (a) here with image (a) of the next figure 6 (neural network merged) shows the improved (with less noise) edge boundary structure for the NN method.

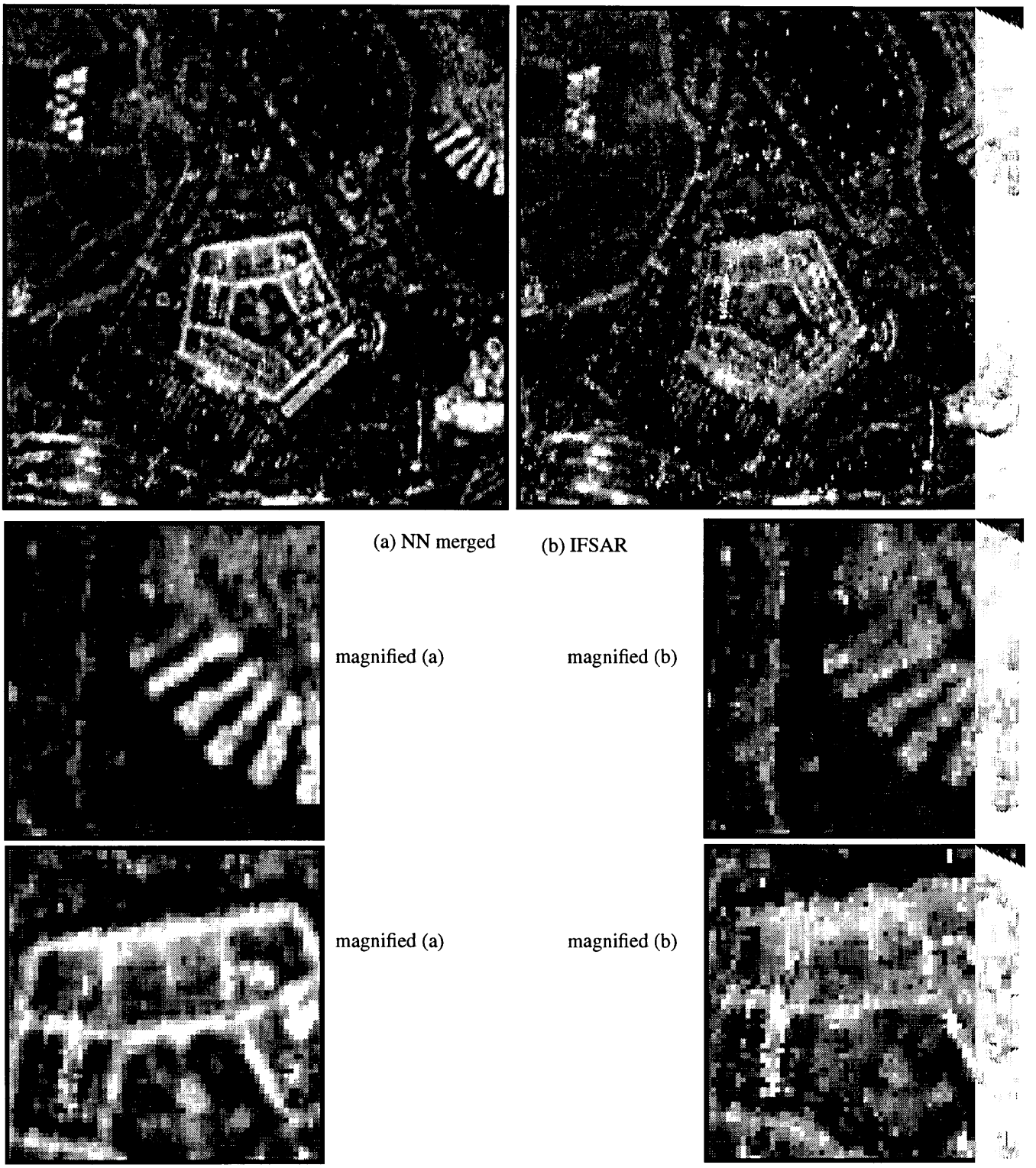

Figure 6. Neural network merged IFSAR images.

(a) is the result of neural network edge pattern merging of SPOT pan and noise-reduced IFSAR (b) images. A comparison between the NN merged image (a) here with image (a) of the previous figure 5 (maximum amplitude merged) shows improved (with less noise) edge boundary structure for the NN method.