Spatial Sharpening of Thematic Mapper Multispectral Data Using Multiresolution and Neural Network Techniques

By George Lemeshewsky

Open-File Report 95-551

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

SPATIAL SHARPENING OF THEMATIC MAPPER MULTISPECTRAL DATA USING MULTIRESOLUTION AND NEURAL NETWORK TECHNIQUES

George Lemeshewsky
U.S. Geological Survey
512 National Center
Reston, VA 22092
Open-File Report 95-551

ABSTRACT

A new approach for spatial sharpening of Landsat thematic mapper (TM) multispectral band images by a higher resolution SPOT panchromatic (P) image is described. Reported multiresolution image pyramid techniques were used to integrate the spatial detail of the P and TM images. An adaptive neural network (NN) technique was developed to improve sharpening when there are contrast reversals between respective P and TM edge boundaries. Based on an assumed correlation between image edges, derived at multiple resolutions, neural networks were trained to modify the P edge data before sharpening. Sharpening tests of one TM band for simulated contrast reversals show definite improvements; however, the sharpening of three TM bands was only partially successful. This study suggests changes in NN training with more representative data to improve sharpening.

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

There are many reported sharpening techniques for increasing the spatial resolution of a lower resolution multispectral image by incorporating spatial information from a higher resolution panchromatic image. For this study, the objective of spatial sharpening of Landsat thematic mapper (TM) multispectral (XS) images with SPOT panchromatic (P) band images is to enhance the tasks of visual interpretation and manual delineation of class boundaries for land use and land cover mapping.

Schowengerdt and Filiberti (1994) note that there are two general categories of sharpening algorithms: spatial domain and spectral space substitution techniques. They described an effective spatial domain sharpening technique with respect to radiometric fidelity developed from point spread function models of the respective sensors. A primary assumption is that panchromatic and multispectral data have radiometric correlation. After coregistration and interpolation, each sharpened XS-band pixel is essentially the product of the XS-pixel value and the ratio of the P-pixel value and low-pass filtered P-pixel value. The low-pass spatial domain filter was determined from the sensor point spread functions, relative image scales, and interpolation method. They noted 'bleeding' of the sharpened near-infrared (IR) band image due to the low correlation between the near-IR and P data, which contradicted the implied assumption that they are correlated. Their model implies that the data have positive correlation.

Another spatial domain sharpening technique is based on the transfer of high-pass filtered, high-pass edge data from P to the XS images. Chavez and others (1991) described three sharpening methods: (1) high-pass filter sharpening, (2) spectral substitution using intensity, hue, and saturation, and (3) spectral substitution using principal components. The high-pass filter method produced the lowest spectral distortion.

For the high-pass filter technique, there can be reduced sharpening when there are contrast reversals between corresponding edge boundaries of the P and XS images. Schowengerdt (1980) noted that multispectral image edges can reverse sign from band to band. He described sharpening experiments where edge data with reversed sign were added to the XS images. He also observed that edges correlate more consistently from band to band than does the spectral information. Without sign correction, Schowengerdt observed a fringing effect in the sharpened image. These effects are illustrated in the image profile sketches of figure 1.

Another spatial sharpening technique, described by Tom, Carlotto, and Scholten (1985) relies on the assumption that the TM multispectral band images are locally correlated. Using this correlation property, a least squares (LS) optimal estimate of the 120-m ground sample distance IR-band image was developed from the visible and IR bands. To sharpen the 120-m IR data they first filtered the other six 30-m TM bands to 120-m spatial resolution. At this reduced resolution, spatially localized LS estimates of the IR data were computed from all remaining TM bands. Then a prediction estimate of 30-m IR data (that is, sharpened) was calculated by using these local LS estimator coefficients; however, the original spatial resolution data were used. The latter was high-pass filtered and then added to the original IR image to produce sharpened IR data.

Tom (1986) also described application of the LS estimator procedure for SPOT P and SPOT XS sharpening. An LS estimate of the 20-m XS data was determined from 10-m P data blurred to 20-m spatial resolution. Using the same LS estimator coefficients with 10-m P data produced the predicted, sharpened 10-m resolution XS image.

A different LS technique for multispectral sharpening was described by Shen, Lindgren, and Payton (1994). They applied LS predictive sharpening when there was local correlation between an XS-band image and the higher resolution P image. However, if an XS band was not correlated with the P band, sharpened XS data were calculated as the product of the XS-band pixel value and the ratio of the P-pixel value and the local average of four P pixels. The sharpening process was repeated twice when the spatial resolution of P to XS images was four to one.

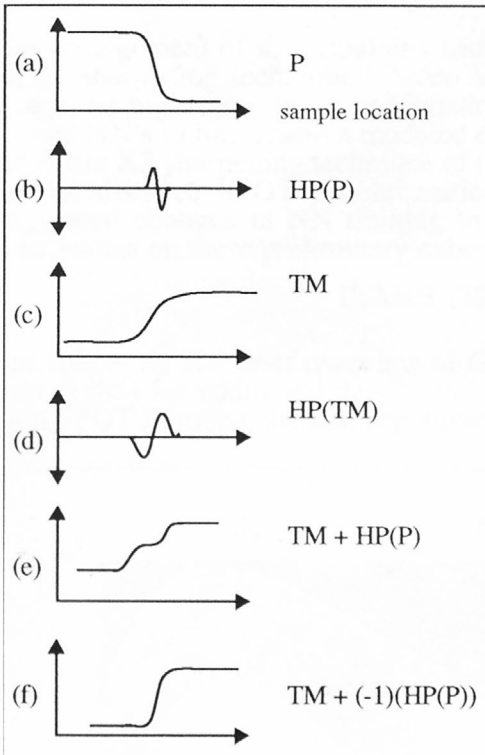


Figure 1. An illustration of TM sharpening by the addition of high-pass filtered P data.

For edge contrast reversal conditions between P (a) and TM (c) gray-scale image profiles, the edge profile sketches of (e) and (f) compare results of TM sharpening by the addition of high-pass (HP) filtered P image data , (HP(P)) without and with sign reversal. Without sign reversal, there is the degraded sharpening of (e). Reversing the sign of (HP(P)) gives the desired sharpened edge (f).

Adaptive, predictive sharpening using neural networks was described by Iverson and Lersch 1994) for SPOT P data resampled to 15 m and a 7-band TM image. Concepts of their sharpening echnique were described for Gaussian (*G*) and Laplacian (*L*) multiresolution image pyramids. Basically, the images were decomposed into *L* edge pyramid images at spatial resolutions equal to (or less than) 15 m. At reduced resolutions where there was high local spatial correation between the P and TM edge images, a NN was trained to predict TM edge data given P dge data. At higher resolution levels of the *L* image pyramid, the previously trained NN predicted the TM edge data given P edge data. The sharpened, predicted higher resolution TM mage was generated by reconstructing the TM image pyramid with predicted edge data.

This report describes development and preliminary results of a multiresolution image sharpening echnique that includes NN processing to improve TM multispectral band sharpening, specifically when there are contrast reversals at respective P and TM edges. It is a spatial domain, not spectral pace substitution technique. In contrast to NN based predictive sharpening, NN's were trained with simulated, correlated edge data at all resolution levels of a *L* image pyramid. That is, instead of the NN learning the SPOT P to TM correlation at reduced resolutions, then predicting multispectral data at higher levels, NN's were trained to learn a modeled correlation at all image pyramid levels. When trained, the NN's adaptively estimate or modify P-band edges to improve

sharpening. The two basic processing steps are (1) estimation of P edge data by trained NN's and (2) TM sharpening using the estimated P edge data. A reported multiresolution technique was used for the latter.

The arrangement of the remaining sections is as follows. Section II outlines the multiresolution image sharpening technique (Ogden and others, 1985) used in this study to transfer to the TM image the higher resolution information of the SPOT P image. Section III describes the use of trained NN's to implement a modeled correlation between P and TM edge data. Also described is the entire XS sharpening technique of this study. Section IV describes sharpening results for10-m sample distance SPOT panchromatic band image and TM 30-m image bands 5, 4, and 2. Suggested changes in NN training to improve sharpening are given here. Section V contains conclusions on these preliminary experiments.

II. MULTIRESOLUTION SHARPENING

The following is a brief overview of *G* and *L* image pyramids; see Ogden and others (1985) and Burt (1985) for additional details on image pyramids. Figure 2 shows *G* and *L* image pyramids for a SPOT P image; the numeric subscript denotes the pyramid level.

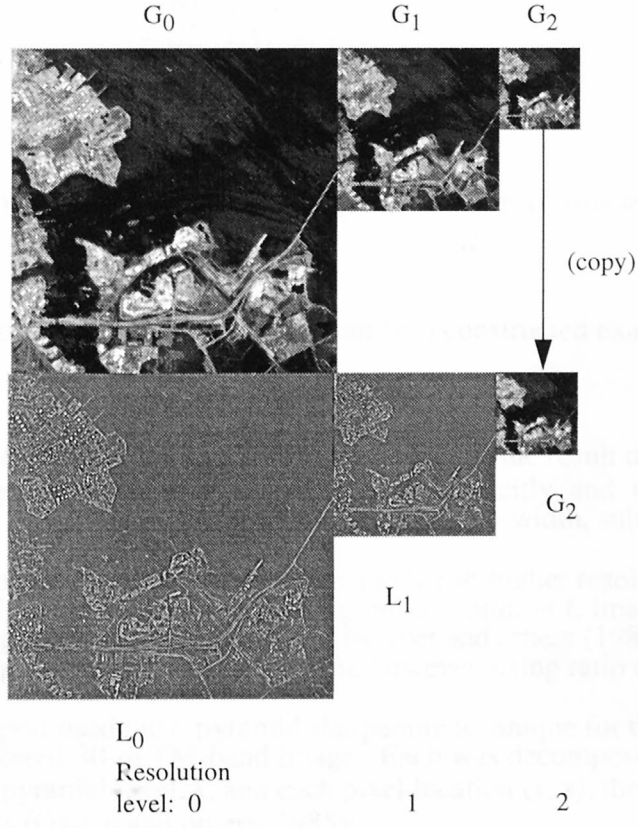


Figure 2. Example of Gaussian (*G*) and Laplacian (*L*) multiresolution image pyramids for SPOT panchromatic (P) data.

G_0 , G_1 , G_2, are a three resolution level (levels 0, 1, 2) *G* image pyramid for the full resolution, 10-m ground sample distance SPOT P image, G_0. Images L_0 and L_1 are pyramid derived *L* edge data from G_0 and *G*, respectively. To assure positive values for display, a gray-level value of 128 was added to the *L* images. The original P image, G_0 can be reconstructed exactly from L_0, L_1, and G_2; that is, the Laplacian-Gaussian pyramid decomposition of G_0.

The L images contain edge data derived at multiple resolutions. The images were computed as follows. Starting with the original image at full resolution level 0 (that is G_0), image G_1 is the result of low-pass filtering G_0 followed by subsampling by a factor of two. This process is called reduce, (RE). Each G image is reduced in size, in each dimension, relative to original image G_0 by 2^{-k}, where k is the image pyramid level.

Image L_0 at level 0 is generated by first up-sampling (by a factor of two) image G_1 followed by interpolation; subtracting this result from G_0 gives L_0. Up-sampling followed by interpolation is called expand, (EX). Similarly L_1 is computed by applying RE and EX to G_1 followed by subtraction from G_1.

Low-pass filtering and interpolation can be performed by convolution with Gaussian-like filters; however, other filters are possible. This study used the separable interpolation kernel (1, 4, 6, 4, 1) (Burt, 1985). A 5-by 5-element low-pass filter kernel (Wilson and Bhalerao, 1992) with closer to ideal low-pass response was used for the low-pass filtering; benefits, if any, were not evaluated.

G and L image pyramid generation is as follows (Ogden and others, 1985):

$$G_1 = RE[\, G_0 \,] \tag{1}$$
$$L_0 = G_0 - EX[\, G_1 \,] \tag{2}$$
$$G_2 = RE[\, G_1 \,] \tag{3}$$
$$L_1 = G_1 - EX[\, G_2] = G_1 - EX[\, RE[\, G_1 \,]]. \tag{4}$$

Rearranging equation 2 shows that G_0 can be reconstructed exactly from L_0 and G_1:

$$G_0 = L_0 + EX[\, G_1]. \tag{5}$$

Also from equations 4 and 5, G_0 can be reconstructed exactly from

$$G_0 = L_0 + EX[\, L_1 + EX[\, G_2 \,]]. \tag{6}$$

L_0 is the result of high-pass filtering G_0. L_1 is the result of band-pass filtering G_0. The band-pass and high-pass filtering are performed implicitly and the filters obtained heuristically (Lim, 1990). Also the G-L pyramid produces octave-width, subband decompositions.

The sharpening of one image with another of higher resolution was based on a technique reported by Ogden and others (1985) using multiresolution L image pyramids. A similar sharpening and merging technique was described by Toet and others (1989) and Toet (1992) for multiscale image merging and contrast enhancement, however, using ratio of low-pass instead of L pyramids.

This report used the L pyramid sharpening technique for the images pairs: 10-m SPOT P and each coregistered 30-m TM-band image. Each was decomposed into G and L image pyramids and for each L pyramid level, k, and each pixel location (x, y), the sharpened L_{TM} (that is, L'_{TM}) was computed as (Ogden and others, 1985):

$$\begin{aligned} &\text{if } |\, L_P\,(x, y)\,| \geq |\, L_{TM}(x, y)\,| \text{ then} \\ &\qquad L'_{TM}(x, y) = L_P(x, y) \\ &\text{else} \\ &\qquad L'_{TM} = L_{TM}(x, y) \end{aligned} \tag{7}$$

and where L_P is the L pyramid image derived from the P data. Reconstructing the TM image from L'_{TM} and highest pyramid level $G_{2,TM}$ by equation 6 produces the sharpened image. Note that

when there are no edges in the P image (that is, L_P is zero), the TM image will be recovered exactly.

When there are contrast reversals at respective edge boundaries of the P and TM images, the signs of L_P and L_{TM} will differ and thus the sharpening by equation 7 can be ineffective. NN preprocessing of L_P data was used to improve sharpening.

III. NEURAL NETWORK PROCESSING

NN or artificial NN is any computing architecture that consists of massively parallel interconnection of simple "neural" processors (Lau and Widrow, 1990). Often NN's are computer simulations with the "neural" processor or artificial neuron modeled as a weighted summation of the inputs to the neuron followed by a nonlinear transfer function.

This study used a multilayer, fully connected, feedforward NN. Starting with the input layer of neurons, the NN consists of consecutive layers of neurons. From input-to output-layer, each neuron output is connected to all inputs of the next layer. Neuron input values are scaled or weighted by the network weights.

In this image sharpening application the purpose of the NN is to implement a mapping of input edge patterns (L image edges) to desired output edge-pixel value. This mapping is 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 based on the error between desired and actual output such that the NN output approximates the desired output. This is called backpropagation training with weight modifications by the generalized rule (Rumelhart and others, 1986).

NN's were trained to produce a mapping of input edge patterns to desired output using training samples taken from a representative subregion of the image. Because of their generalization property (similar inputs produce similar outputs), the trained NN's function as adaptive estimators for the remaining image, or other similar images. 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).

The underlying premise was that for contrast reversals, TM sharpening by equation 7 would improve if the sign of L_P edge data was modified. That is, the sign of L_P edges should be reversed when L_P , L_{TM} edges have opposite signs. This was done by training NN's to estimate either L_P or $(-1)(L_P)$, based, however, on the local spatial correlation between the edges of P and TM images.

It was assumed that at reduced resolution levels where L_P and L_{TM} have higher local spatial correlation, a NN could be trained to estimate L_P with appropriate sign given L_P and L_{TM}. However, at higher resolution levels where there is lower P to TM spatial correlation, it would be more difficult to train a NN to accurately estimate L_P from L_{TM} and L_P . This problem was solved by training a second NN to estimate L_P with appropriate sign, given L_P and the NN estimate of L_P from the previous or lower resolution level.

Figure 3 roughly indicates the desired input-to output-relationship between L_P and estimated L_P implemented by training NN_1 with simulated, local spatially correlated data. As shown in figure 3(i), at pyramid level 1 the desired NN_1 output, after training is either L_P or $-L_P$.

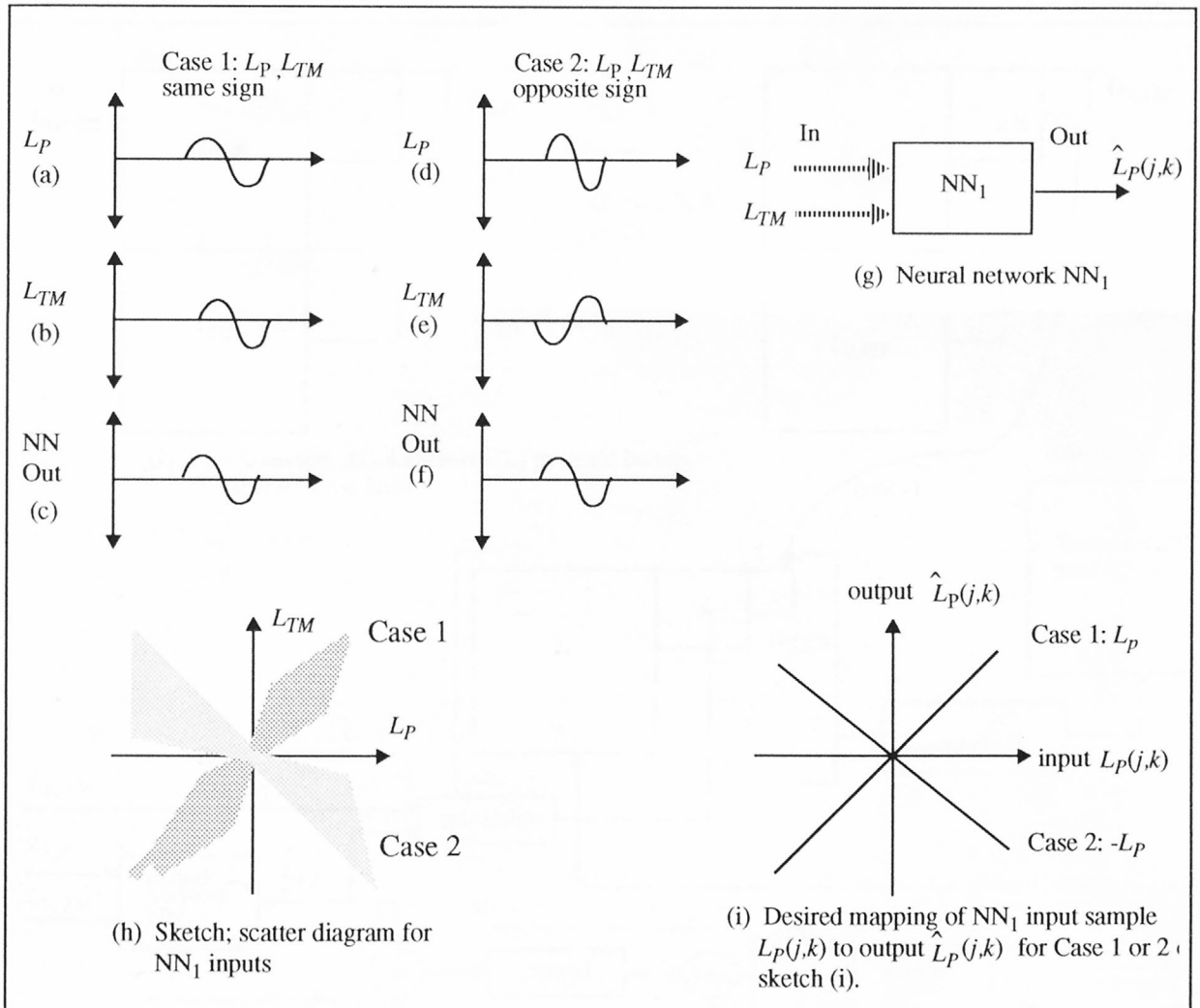


Figure 3. Illustration of neural network NN_1's input-to desired-output for two cases of P, TM edge conditions. Figures (a) to (f) and (h) show same and reverse sign P, TM edge data. Figure (i) shows the desired output of NN_1 for one of the two inputs, L_p. Depending on the spatial patterns of L_P and L_{TM}, the output is either an estimate of L_p (case 1) L_p with reversed sign (case 2).

Figure 4 shows the entire sharpening process used in this report. Only two L level images, L_0 and L_1 were used to facilitate initial tests. They have sample distances of 10 m and 20 m, respectively. Thus the TM image is reconstructed from the 40-m image $G_{2,TM}$, L'_1, and L'_0 where L'_1, L'_0 are the result of equation 7 applied to: (1) NN estimate of $L_{1,P}$ and $L_{1,TM}$ and (2) NN estimate of $L_{0,P}$ and $L_{0,TM}$, respectively.

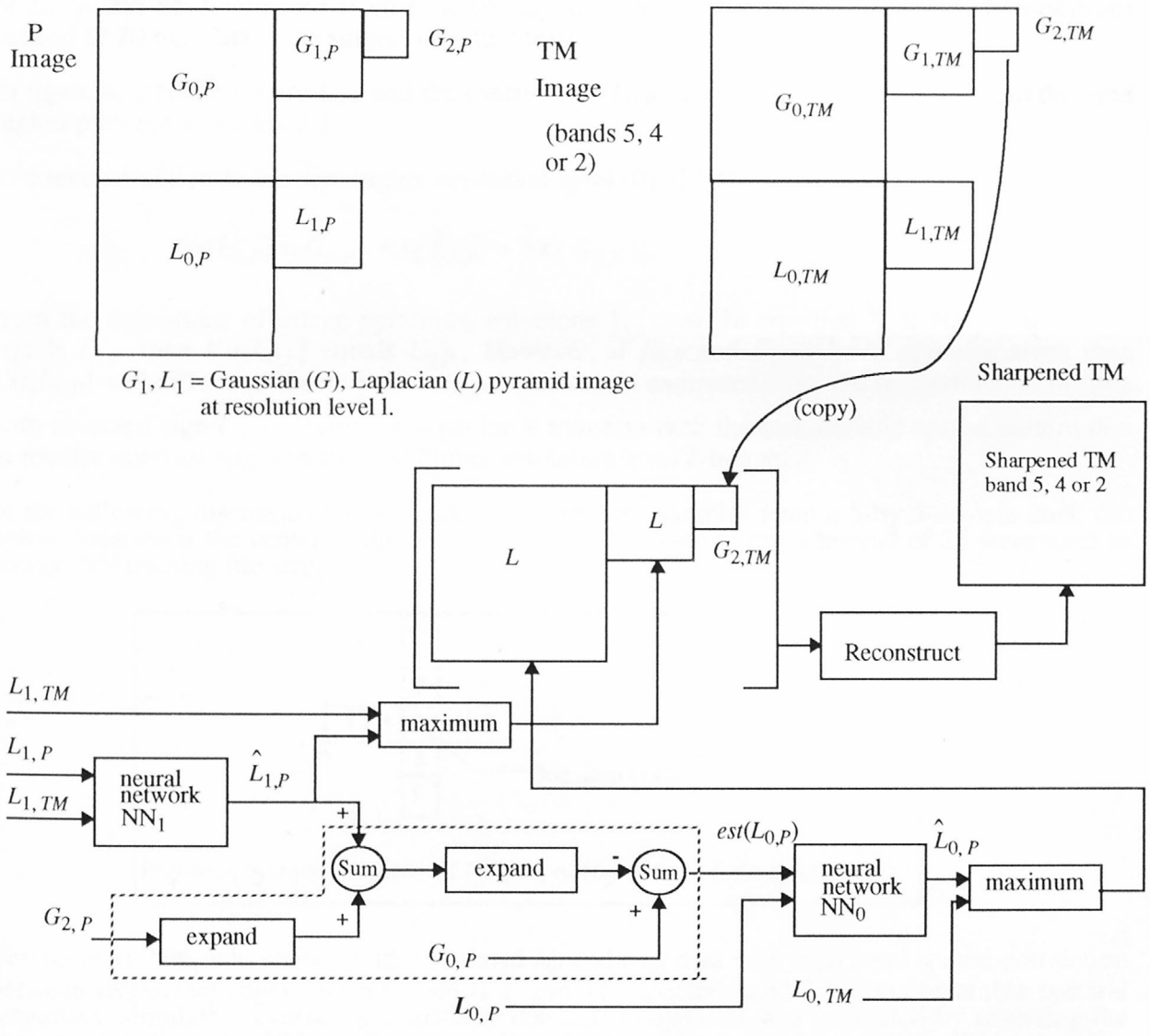


Figure 4. TM spatial sharpening by higher resolution SPOT panchromatic image P, with NN processing for edge contrast reversals.

Shown is the entire sharpening process for each TM band. First, coregistered pairs of P and TM bands 5, 4, or 2 are decomposed into three-level G / L pyramids. Sharpened TM image bands are reconstructed from the maximum, per equation 7, of L_{TM} , NN estimate of L_P , and $G_{2,TM}$. Each NN was trained to estimate or modify L_P depending upon edge conditions between P and TM band 2 data. Each NN input line represents nine samples; the output is one sample. Shown enclosed by ·-·-·-·-·- is the process for reconstructing an estimate (*est*) of $L_{0,P}$ given $G_{2,P}$, NN_1 output, and $G_{0,P}$; this follows from the definition of G and L image pyramids. Expand is the combination of up-sampling followed by interpolation, as described in section II.

An alternate approach to the multiresolution decomposition of figure 4 is to coregister and interpolate the 30-m TM images with a 7.5-m panchromatic image, and then decompose each to L_0, L_1, L_2 of 7.5-, 15-, 30-m sample distances. Then the initial, presumably more accurate, estimate of L_P would be determined from $L_{2,P}$ and $L_{2,TM}$ at 30 m (that is, the original TM resolution) instead of 20 m. This is the subject of future tests.

In figure 4, NN_0 inputs are $L_{0,P}$ and the estimate for $L_{0,P}$ ($Est[L_{0,P}]$), reconstructed from the next higher pyramid level, level 1.

The reconstruction to the next higher resolution level (0) of NN_1 output is

$$Est[L_{0,P}] = G_{0,P} - EX[\,\hat{L}_{1,P} + EX[\,G_{2,P}\,]\,] \tag{8}$$

from the definitions of image pyramids, equations 1, . . . ,6. In equation 8, if NN_1 output $\hat{L}_{1,P}$ equals $L_{1,P}$ then $Est[L_{0,P}]$ equals $L_{0,P}$. However, if $L_{1,P}$ and $L_{1,TM}$ have opposite signs then $Est[L_{0,P}]$ will differ significantly from $L_{0,P}$ because it is computed from the reconstruction of $G_{2,P}$ with reversed sign $\hat{L}_{1,P}$. Basically, equation 8 transfers both the location and spatial pattern due to reverse contrast edges to the next higher resolution level L image.

In the following discussion the network inputs are nine samples from a 5-by 5-sample area; the output location is the center of the 5-by 5-area (fig. 5). Nine samples instead of 25 were used to reduce NN training file size.

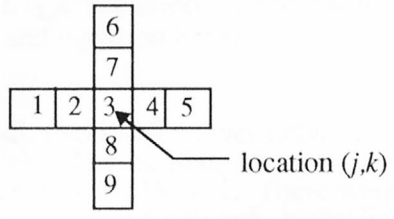


Figure 5. Sample locations of NN_1 input (L_p or L_{TM}) for output $L_P(j,k)$.

The network NN_1 was trained with simulated L_P and L_{TM} data with high local spatial correlation between respective edges. Because SPOT P and TM spectral band 2 have comparable spectral responses, simulated reverse sign, spatially correlated edge data was generated by reversing the gray-scale of the P (or TM band 2) image, then decomposing to L edge images. The scatter diagram sketch of figure 3(h) roughly depicts a measure of this correlation. L^R denotes the reversed sign edges.

Table 1 gives the three cases of input and output training data for NN_1.

Table 1: NN_1 training data

Case	input $_1$	input $_2$	NN_1 output
1	$L_{1,P}$	$L_{1,TM}$	$L_{1,P}$
2	$L_{1,P}$	$L^R_{1,TM2}$	$L^R_{1,P}$
3	$L_{1,P}$	0	$L_{1,P}$

When there are P but no TM band 2 edges, (that is, case 3), NN_1 is trained to output P edges. In these initial tests, networks were trained for only P and TM band 2 data. It was believed that one network, trained per table 1 and case 3 conditions, might be sufficient instead of separate networks trained for P and TM bands 4 or 5. This is discussed in the results section.

The effect of case 3 training on the sharpening process (that is, equation 7) is the transfer of P-band edge data to the TM image. However, instead of sharpening, only edge data are transferred. Subject to further study for case 3 edge conditions is the transfer of low-pass filtered data at the next lower resolution G pyramid image, $G_{2,P}$ to $G_{2,TM}$. This would transfer low-pass image data, not just P edges, to the TM image.

Because NN_1's desired output is $L_{1,P}$ or $L^R_{1,P}$ (fig. 3 (i)), there were two training cases for the second NN (that is, NN_0), as shown in table 2.

Table 2: NN_0 training data.

Case	input $_1$	input $_2$	NN_0 output
1	$L_{0,P}$	$L_{0,P}$	$L_{0,P}$
2	$L_{0,P}$	$Est[L_{0,P}]$	$L^R_{0,P}$

Simulated data for case 2, input 2 (that is, $Est[L_{0,P}]$) were simply calculated from equation 8 by substituting image $L_{1,P}$ for $\hat{L}_{1,P}$. An alternate approach is to calculate NN_0 input training data from NN_1's output estimate and equation 8. These data have more variation and should improve network generalization.

Fully connected, feed forward, two hidden layer networks with *tanh* nonlinear transfer functions were used (NeuralWare, 1993). The number of input, first hidden, second hidden, and output layer neurons of each network was 18, 18, 9, 1. There were nine samples per each input of tables 1 and 2 (fig. 5). No tests were made of network generalization versus network size. While one hidden layer would be sufficient (Cybenko, 1989), the two hidden layer network seemed to train more quickly.

Back-propagation training with the generalized delta rule (Rumelhart and others, 1986) was used. Also applied was momentum equal to 10 percent of the learning rate. The learning rate was decreased to a fixed value as training progressed. Basically, training was with a low learning rate (0.005) and a large, fixed number of sample presentations (800,000). Weight updates were made after each sample presentation.

Table 3 gives the root mean square (RMS) error between desired and actual NN output for test and training data. This error is for L edge data scaled from (-128 to 128) to (-1 to 1).

Table 3: NN test and training results

NN	Training error, RMS	Training samples	Test error, RMS	Test samples
NN_0	0.0417	4096	0.0312	4096
NN_1	0.0822	3456	0.0843	3456

The fact that test and training data were derived from different, nonoverlapping image areas may explain why NN_0's test error is less than the training error; this is subject to further study.

IV. IMAGE RESULTS

The following tests were made with a system level 1 corrected, 28.5-m (nominal 30-m) sample distance TM image. It was coregistered to the SPOT (© 1988, SPOT Image Corporation, Reston, Virginia) processing level 1A, 10-m P-band image and cubic interpolated to 10-m sample distance. SPOT P and TM XS scene acquisition dates were July 25 and July 5, 1988, respectively.

After NN training with P and TM band 2 data, TM sharpening tests were made using the previously described NN based sharpening process of figure 4. For comparison, results for sharpening by the maximum of L, 'max' equation 7 are also shown.

The first test was TM band 2 sharpening by P band for normal gray scale and reverse gray scale, P^R images. The results of the two sharpening methods were compared; table 4 summarizes these test conditions.

Table 4: Test conditions for sharpening test 1

TM band	P data	Sharpening method	Result as figure no.
2	P^R	NN	4, e
2	P	NN	4, f
2	P^R	'max'	4, d
2	P	'max'	4, c

For the conditions of simulated edge contrast reversals, figures 6 (d) and (e) show the sharpening without and with NN estimation/correction of P edges, respectively. This is the case where the NN predicts L edge data with reversed sign. The NN technique gives improved sharpening for reverse contrast conditions. The consequence of sharpening with reverse contrast P data and without NN correction is seen as a halo-like artifact near contrast boundaries; compare the top center area of images in figures 6 (d) and (e).

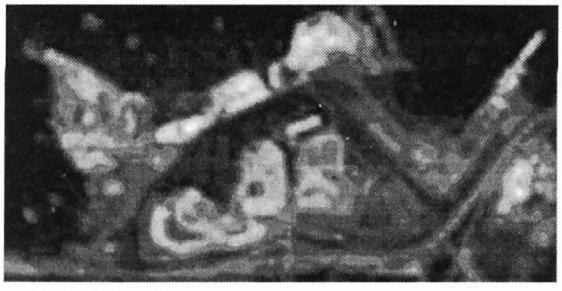
Figure 6(c) shows that when P, TM 2 edges have same sign (that is, no contrast reversals) 'max' sharpening (that is, no NN correction) of TM band 2 is effective.

Also shown in figure 6(f) is the result of P, TM2 sharpening (no contrast reversals) with NN processing. This is the condition where the NN predicts L edge data without sign reversal.

The second tests that compare sharpening with and without the NN estimation/correction of P edge data were for TM bands 5, 4, and 2 and are displayed as red, green and blue. Figure 7(a) and (b) are the raw SPOT P image and the coregistered and interpolated to 10-m sample distance TM image. Figure 7(c) is the TM image after sharpening by the NN process of figure 4. Figure 7(d) is the result of 'max' sharpening (that is, no NN processing). Figure 7(g) and (h) compare 'max' and NN sharpening and show the improved sharpening due to NN processing when there is a contrast reversal at a respective P and TM edge boundary. Lesser sharpening for other features such as roads may be the result of low correlation between edges of the P and TM bands 5, 4 data.

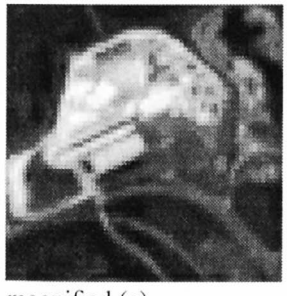
(a) SPOT panchromatic, P

(c) TM2 'max' sharpened by P

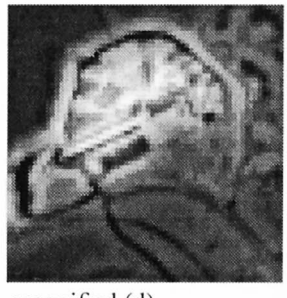


(b) TM band 2

(d) TM2 'max' sharpened by P^R



magnified (c)



magnified (d)

(e) TM2 'max' sharpened by NN modified P^R

magnified (e)

magnified (f)

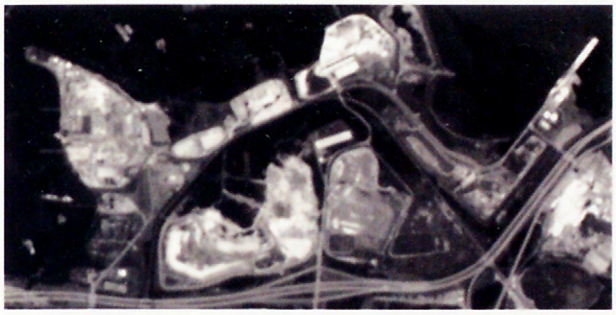
(f) TM2 'max' sharpened by NN modified P

Figure 6. Sharpening of TM band 2 with SPOT panchromatic, P or contrast reversed P data; with or without neural network correction for edge contrast reversals.

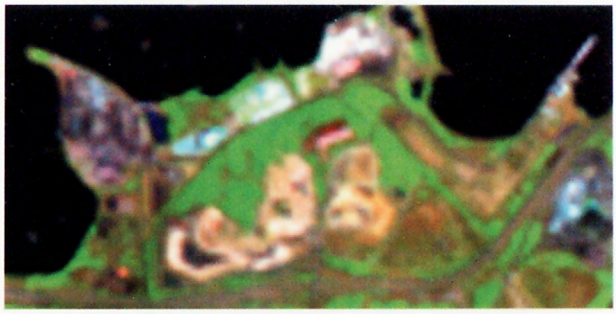
Images (d) and (e) compare sharpening for simulated contrast reversed P data (P^R), without and with NN processing. The NN process, result (e), gives improved sharpening for simulated contrast reversals. Without NN processing, (c) and (d) show, respectively, TM band 2 data sharpened by P and simulated, contrast reversed P, (P^R) using only the maximum of L, 'max' , equation 7. Results (e) and (f) show, respectively, TM 2 sharpened by P^R and P with NN processing for edge contrast reversals. Comparison of (e) and (f) indicates that the best NN sharpening was for the contrast reversal conditions of (e). NN training and test data were derived from the left third of (a) and (b).

Figure 7. Thematic Mapper (TM) bands 5, 4, 2 sharpened by SPOT panchromatic band data.

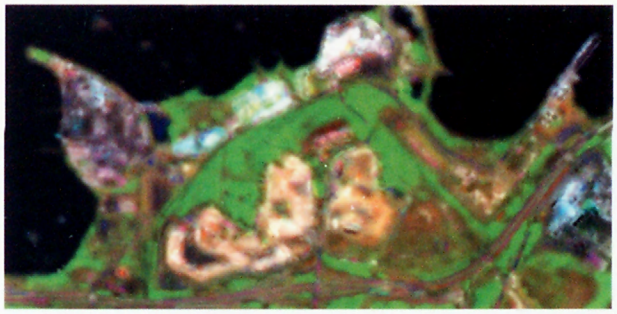
Figure 7(a), and (b) are the raw SPOT P image (10 m) and the coregistered and interpolated to 10-m sample distance TM image. Image (c) is the TM image after sharpening by the NN process of figure 4; image (d) is the result of 'max' sharpening (that is, no NN processing). Images (c) and (d) compare, respectively, TM bands 5, 4, 2 (image (b) as red, green, blue) sharpened by the SPOT P image (a), with and without NN processing. Images (g) and (h) compare results of these two sharpening techniques for contrast reversals at the vertical boundary separating the dark blue and green areas. At this boundary, there is a contrast reversal between the P data of (e) and TM band 4 (shown as green) in (f). Comparison of (g) (that is, NN sharpening) and (h) (that is, 'max' sharpening) shows the improvement due to the NN technique as the well-defined dark blue-to green-boundary and also the lack of distortion in the adjacent green area. This type of distortion, due to edge contrast reversals, was illustrated in figure 6.



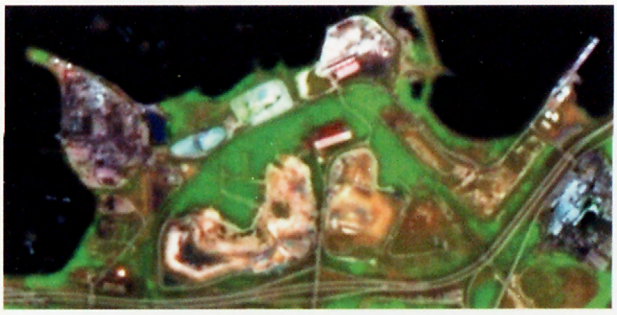
(a) SPOT PAN



(b) TM bands 5, 4, 2



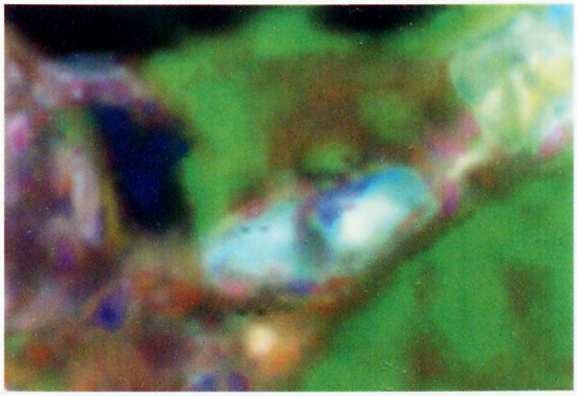
(c) NN sharpened TM



(d) 'max' sharpened TM

(e) magnified (a)

(f) magnified (b)



(g) magnified (c): NN sharpened

(h) magnified (d): 'max' sharpened

Figure 7. Thematic Mapper bands 5, 4, 2 sharpened by SPOT panchromatic band data.

Following are some possible causes and suggested changes to improve the NN sharpening process. Also, it is noted that some simplifications were made to produce these preliminary test results:

• NN training data were the highly correlated, simulated *L* edge data of similar amplitude, derived from P and TM 2 bands. Training data having differing amplitudes should improve network generalization.

• Sharpening tests with NN's trained for P and TM bands 5 or 4 data and reverse contrast data need to be conducted.

• Only 9 NN input samples per image were used for these initial tests; all 25 samples should be used to better represent the edge patterns.

As mentioned, to give more variation to NN_0's training data, these data should be derived from the output of NN_1.

V. CONCLUSION

A new approach to multispectral image sharpening, specifically for contrast reversals at respective P and TM edge boundaries consisted of a two step sharpening process: (1) adaptive modification, by trained NN's, of multiresolution derived P edge data to correct for contrast reversals and (2) sharpening by a reported multiresolution technique. NN's were trained for the conditions of correlated and uncorrelated P, TM edges. Preliminary results showed that the NN estimated (or sign corrected) edge data improved sharpening for simulated contrast reversals. Sharpening results for P and TM spectral bands 5, 4, 2 were partly successful. It is suggested that changes in NN training are needed to improve sharpening and that further tests are required.

REFERENCES

Burt, P.J., 1985, Smart sensing within a pyramid vision machine: Proceedings IEEE, v. 76, no. 8, p. 1006-1015.

Chavez, P.S., Sides, S.C., and Anderson, J.A., 1991, Comparison of three different methods to merge multiresolution and multispectral data: Photogrammetric Engineering & Remote Sensing, v. 57, no. 3, p. 295-303.

Cybenko, G., 1989, Approximation by superpositions of a sigmoidal function: Mathematics of Control, Signals, and Systems, v. 2, no. 4, p. 303-314.

Iverson, A.E., and Lersch, J.R., 1994, Adaptive image sharpening using multiresolution representations: Proceedings SPIE, v. 2231, p. 72-83.

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. 1411-1413.

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

NeuralWare, Inc., 1993, *Neural Computing*. Pittsburgh, PA, NeuralWare.

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

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986, Learning Internal Representations by Error Propagation, chap. 8 *of* Parallel Distributed Processing 1: Foundations, D.E. Rumelhart and J.L McClelland, MIT Press, Cambridge, MA, p. 318-362.

Schowengerdt, R.A., and Filiberti, D., 1994, Spatial frequency models for multispectral image sharpening: Proceedings SPIE, v. 2231, p. 84-90.

Schowengerdt, R.A., 1980, Reconstruction of multispectral image data using spatial frequency content: Photogrammetric Engineering and Remote Sensing, v. 46, no. 10, p.1325-1334.

Shen, S.S., Lindgren, J.E., and Payton, P.M., 1994, Panchromatic band sharpening of multispectral image data to improve machine exploitation accuracy: Proceedings SPIE, v. 2304, p. 124-131.

SPOT Image Corporation, 1988, *SPOT User's Handbook*: SPOT Image Corp., Reston, VA.

Toet, A., vanRuyven, L.J., and Valeton, J.M., 1989, Merging thermal and visual images by a contrast pyramid: Optical Engineering, v. 28, no. 7, p. 789-792.

Toet, A., 1992, Multiscale contrast enhancement with applications to image fusion: Optical Engineering, v. 31, no. 5, p. 1026-1031.

Tom, V.T., Carlotto, M.J., and Scholten, D.K., 1985, Spatial sharpening of Thematic Mapper data using a multiband approach: Optical Engineering, v. 24, no. 6, p. 1026-1029.

Tom, V.T.,1986, A synergistic approach for multispectral image restoration using reference imagery: Proceedings IGARS symposium, Zurich, p. 559-564.

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

Wilson, R., and Bhalerao, A.H., 1992, Kernel design for efficient multiresolution edge detection and orientation: IEEE Transactions on Pattern Analysis & Machine Intelligence, v. 14, no. 3, p. 384-390.