Nonlinear Mean-Square Estimation with Applications in Remote Sensing

Mark J. Carlotto (mjcarlotto@ma.psrw.com)
Pacific-Sierra Research Corporation, 1400 Key Blvd., Suite 700, Arlington, VA 22209

Abstract

An approach to image modeling based on nonlinear mean-square estimation that does not assume a functional form for the model is described. The relationship between input and output images is represented in the form of a lookup table that can be efficiently computed from, and applied to images. Three applications are presented to illustrate the utility of the technique in remote sensing. The first illustrates how the method can be used to estimate the values of physical parameters from imagery. Specifically we estimate the topographic component (i.e., the variation in brightness caused by the shape of the surface) from multispectral imagery. The second application is a nonlinear change detection algorithm which predicts one image as a nonlinear function of another. In cases where the frequency of change is large (e.g., due to atmospheric and environmental differences), the algorithm is shown to be superior in performance to linear change detection. In the last application, a technique for removing wavelength-dependent space-varying haze from multispectral imagery is presented. The technique uses the IR bands, which are not affected significantly by haze, to predict the visible bands. Results show a significant reduction in haze over the area considered. Additional application areas are also discussed.

Key words: Nonlinear Estimation, Multispectral Imagery, Change Detection, Topographic Extraction, Atmospheric Correction

1. Introduction

Many remote sensing applications are concerned with the determination and exploitation of relationships between a sensed image and physical properties of the scene. Remote sensing models provide a means for understanding the relation between the image and scene (Strahler and Woodcock 1986). Physically-based models relate properties of the scene to the sensed data in the form of a functional relation. An example is the exponential relationship between image radiance and water depth used in shallow-water bathymetry. In this case the model is invertible and thus can be used to infer water depth from image radiance. However, in many cases the model is not invertible, or the underlying physical mechanisms are too complex to represent by a simple function.

Image-based models provide an alternative approach in which empirical relationships are inferred from image-scene measurements (training data). When the training set is small it is necessary to assume a functional form for the image model where the parameters of the function are adjusted to fit the training data. The function effectively extends the image-scene measurements in the training set so that they can be used to infer the value of scene properties from image values not in the training set. Unfortunately many scene parameters are not related to the sensed image in a simple way. For example, image brightness depends on albedo and the topography. Unless the albedo is known, topography cannot be computed from brightness in an obvious manner.

This paper describes an approach to image modeling based on nonlinear mean-square estimation that does not assume a functional form for the model (Section 2). The relationship between input and output images is represented in the form of a lookup table that can be efficiently computed from, and applied to images. The method can be used to derive an empirical relationship between an image and an existing low resolution database in order to extend and spatially enhance the database using the image (Section 3). The method can a

lso be used to model two or more images acquired at different times by one another for change detection (S ection 4), and to model time-coincident images acquired at different wavelengths and by different sensors f or image restoration, atmospheric correction, and sensor simulation (Section 5).

2. Non-Linear Mean Square Estimation

Let x and y be two random variables and $p(x,y)$ be the joint density function. Our goal is to estimate y b y a function f of x such that the mean-square error (MSE)

$$E\left[y - f(x)\right]^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[y - f(x)\right]^2 p(x,y)\,dx\,dy \tag{1}$$

is minimized. As shown in Papoulis (1965) the MSE is minimized when f is the conditional expected valu e

$$f(x) = E\left[y \mid x\right] = \int_{-\infty}^{\infty} y p(y \mid x)\,dy \tag{2}$$

where $p(y \mid x) = p(x,y) / p(x)$. When x and y are integers in the range $0 \le x, y < L$, $p(x,y)$ is an $L \times L$ a rray. f is a table containing L entries where the l-th entry has the value $f[x = l]$ and is given by

$$f(x = l) = \sum_{y=0}^{L-1} y p(y \mid x = l) = \sum_{y=0}^{L-1} y p(x = l, y) \Big/ \sum_{y=0}^{L-1} p(x = l, y) \tag{3}$$

The method extends directly to vectors. Let $\mathbf{x} = [x(1) \cdots x(M)]$ and $\mathbf{y} = [y(1) \cdots y(N)]$ be M- and N-dimens ional vectors and $p(\mathbf{x}, \mathbf{y})$ be the joint density. When the elements $x(m)$ of $\mathbf{x}$ and $y(n)$ of $\mathbf{y}$ are in the rang e $0 \le x(m), y(n) < L$, $p(\mathbf{x}, \mathbf{y})$ is an array containing L^{M+N} elements. Storage and associated computational r equirements can be reduced by the following steps:

1) Estimate each $y(n)$ by a separate function f_n of $\mathbf{x}$. This reduces the storage requirements for the joint density from L^{M+N} to NL^{M+1}.
2) Since the joint density tends to become sparse as the number of dimensions M increases, store t he elements in a hash table.

The first step in effect reduces the problem to that of estimating a scalar y by a function f of a vector $\mathbf{x}$. If there are K non-zero elements in the joint density $p(\mathbf{x}, y)$, f is a table containing K entries where the k-t h entry has the value $f[\mathbf{x} = \mathbf{x}_k]$ and is given by

$$f(\mathbf{x} = \mathbf{x}_k) = \sum_{y=0}^{L-1} y p(\mathbf{x} = \mathbf{x}_k, y) \Big/ \sum_{y=0}^{L-1} p(\mathbf{x} = \mathbf{x}_k, y) \tag{4}$$

The actual calculation of the table is actually simpler than it appears. Let $\boldsymbol{X} = \{\mathbf{x}(i,j)\}$ and $Y = \{y(i,j)\}$ den ote a pair of images where $(i, j) \in T$. We refer to T as the training set. $\boldsymbol{X}$ consists of M input images and Y is the output image; i.e., $y(i,j)$ is the output value associated with the input vector $\mathbf{x}(i, j)$ at pixel (i, j). L

et $S_k \subset T$ denote the set of pixels where $\mathbf{x}(i, j) = \mathbf{x}_k$ and N_k be the number of pixels in S_k. Eq. 4 can be wr itten

$$f(\mathbf{x} = \mathbf{x}_k) = \frac{1}{N_k} \sum_{(i,j) \in S_k} y(i, j) \tag{5}$$

The table f can be thus be computed by summing the output values that occur for each unique combinatio n of input values and dividing by the number of pixels in the input with that value.

It is noted that when x or y represent symbols (i.e., thematic data such as land cover), the table f is obtain ed by assigning the output value with the largest posterior probability which, according to Bayes rule, is

$$f(x = l) = \arg\max_y \{p(y \mid x = l)\} = \arg\max_y \{p(x = l, y) / p(x = l)\} \tag{6}$$

where $p(x = l, y)$ is sometimes written as the product of a conditional probability times a prior probability, $p(x = l, y) = p(x = l \mid y)\, p(y)$.

3. Topographic Extraction from Imagery

Techniques which estimate topography from a single image (shape-from-shading, or photoclinometry) con vert shading information into slope that is, in turn, integrated to form an elevation surface. If we assume that atmospheric effects have been removed, the brightness in a multispectral image at pixel (i, j) in band n is given by

$$x(i, j, n) = a(i, j, n) t(i, j) \tag{7}$$

where $a(i, j, n)$ is the spectral albedo and $t(i, j)$ is the brightness due to the topography (topographic comp onent). It is usually assumed that the topographic component does not depend on wavelength. When the al bedo of the surface is constant over the area, brightness is related to slope by the reflectance map (Horn 19 77). Except in limited situations (e.g., planetary surfaces, terrestrial deserts and snow), surfaces are covered by different materials so the albedo must also be treated as a variable. In a previous approach (Eliason et al 1981), multispectral imagery is clustered into regions with similar spectral properties using band ratios. Un der certain conditions, the average value of a band over each region can be used as an estimate of the albedo of the region. The topographic component can then be estimated by dividing a band by its albedo. The perf ormance of this approach depends critically on the clustering.

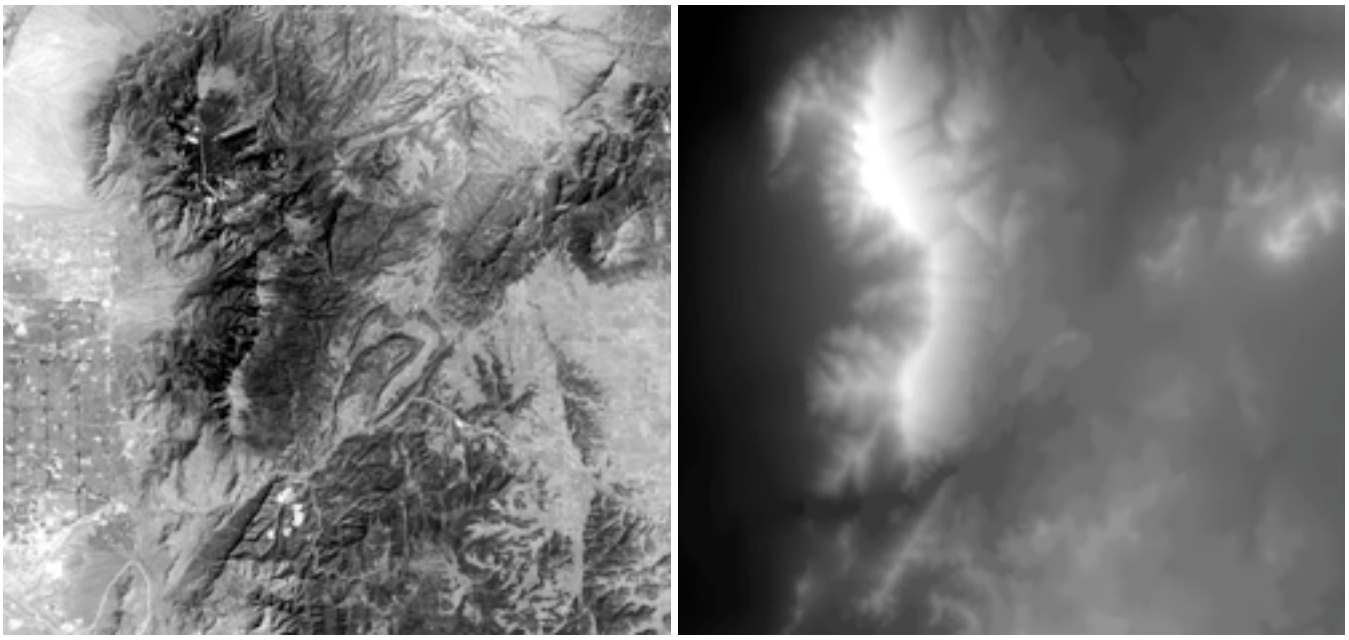
We estimate the topographic component directly from multispectral imagery without clustering (Carlotto 19 96). It is assumed that the surface reflectance function is known, and that elevation data is available over a tr aining area that is representative of the land cover and topography of the full area of interest. The topograph ic component computed from elevation data $e(i, j)$ over the training area T

$$y(i, j) = h[e(i, j)], \; (i, j) \in T \tag{8}$$

where the function h depends on the surface reflectance function and the sensor geometry. Instead of usin g elevation data, we would like to estimate the topographic component directly from an image:

$$\hat{y}(i, j) = f[\mathbf{x}(i, j)], \; (i, j) \in T \tag{9}$$

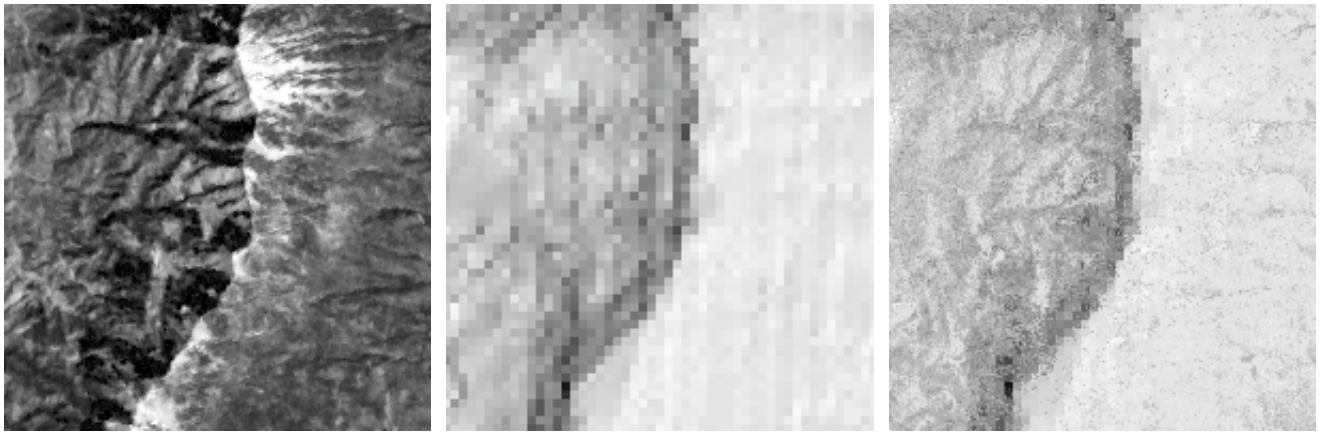
where the function f minimizes the mean-squared error between Eqs. 8 and 9. This function is given by th e condition expected value $f[\mathbf{x}] = E[y \mid \mathbf{x}]$.



a) Landsat TM (first principal component) b) USGS Digital Elevation Model

Figure 1 Full study area -- Albuquerque and Sandia Mountains

The topographic extraction method has been evaluated using Landsat TM imagery and USGS digital elevati on models (DEMs) over a region 1361 x 1285 pixels in size in New Mexico that includes the eastern porti on of Albuquerque and the Sandia Mountains (Figure 1a). This area is particularly challenging due to the d iversity of land cover categories. The DEM (Figure 1b) was resampled to 25 meters and registered to the i magery.



a) First principal component b) Topography from DEM c) Topography from Landsat TM

Figure 2 Training area

In order to reduce the complexity of the algorithm and the size of the lookup table, we use the first three La ndsat TM principal components to predict the topographic component. Figure 2a shows the first principal c omponent image over a 200 x 200 pixel training area in the Sandia Mountains. Figure 2b is the topographic

component computed from the DEM using a Lambertian reflectance map. Within the training area, 14,538 unique combinations of the three principal component values were found. Figure 2c is the topographic component estimated from the three principal component images over the training area. Within the training area, the method is able to spatially enhance the topographic component, revealing subtle detail not visible in the lower resolution DEM.

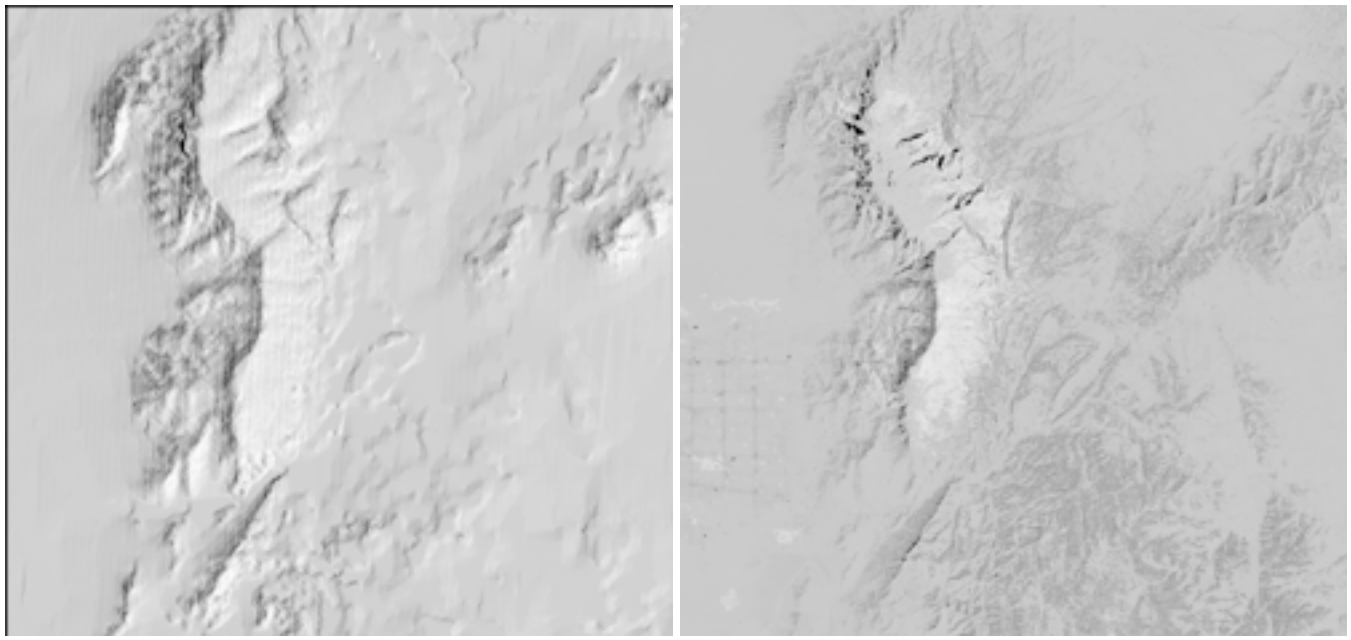
We then used the lookup table to the estimate the topographic component from the Landsat imagery outside of the training area T; i.e.,

$$\hat{y}(i,j) = f\left[\mathbf{x}(i,j)\right], (i,j) \notin T \tag{10}$$

A limitation of the basic method described in Section 2 is that the table only contains entries for input values in the training area. Values encountered that are not in the training area must be interpolated. The simplest interpolation technique is to assign the output value associated with the nearest entry in the table, $f(\mathbf{x} = \mathbf{x}') = f(\mathbf{x} = \mathbf{x}_{k'})$ where

$$k' = \arg\min_k \, d(\mathbf{x}', \mathbf{x}_k) \tag{11}$$

and $d(\mathbf{x}', \mathbf{x}_k) = (\mathbf{x}' - \mathbf{x}_k)^T (\mathbf{x}' - \mathbf{x}_k)$ is the Euclidean distance. To avoid searching the entire table, in our implementation, when a value not in the table is encountered, a search for values in a region of radius $\rho < \sqrt{d(\mathbf{x}', \mathbf{x}_k)}$ around the value is performed. If a value is not found the output is assigned a special value for "unknown".



a) Topography from DEM b) Topography from Landsat TM

Figure 3 Results comparison over full study area

Figure 3a shows the topographic component derived from the DEM over the full study area. Figure 3b is the topographic component estimated from the imagery using the lookup table derived over the training area. The full study area contains 133,213 unique combinations of the three Landsat TM principal component values. Within the training area only about 10% of the spectral diversity of the full image is represented. Almost 90% of the values shown in Figure 3b have thus been interpolated.

4. Change Detection

The objective of change detection is to identify features in one image that cannot be accounted for by features in another image over the same area. In simple image subtraction, the measure of change is based on the difference between corresponding pixel values. Adaptive techniques (Therrien et al 1986, Lee et al 1986, Carlotto et al 1992) use information over larger areas (e.g., within a sliding window, regions of similar material type, or the entire image) to model and predict the two images from each other. The difference between the actual and predicted image is used as a measure of change. For two images $x(i, j)$ and $y(i, j)$ we seek an estimate of y based on x, $\hat{y} = f(x)$, that minimizes the forward prediction error

$$E\left[\varepsilon_F(i,j)\right]^2 = E\left[y(i,j) - \hat{y}(i,j)\right]^2. \tag{12}$$

Similarly, in the other direction we seek an estimate of x based on y, $\hat{x} = g(y)$, that minimizes the backward prediction error

$$E\left[\varepsilon_B(i,j)\right]^2 = E\left[x(i,j) - \hat{x}(i,j)\right]^2. \tag{13}$$

When the functions f and g are linear, the parameters of the linear models

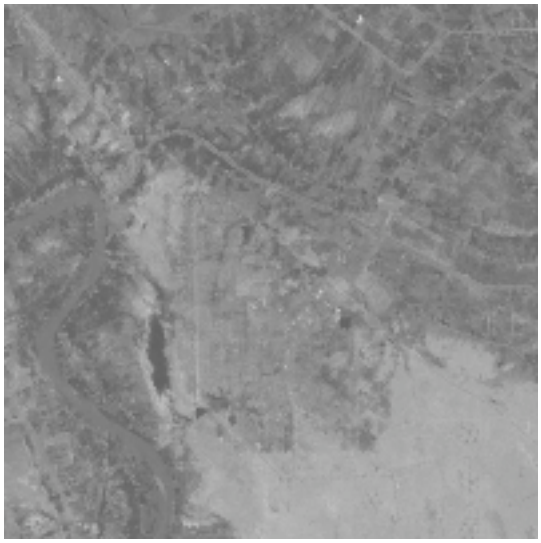
$$\begin{aligned} f(x) &= ax + b \\ g(y) &= cy + d \end{aligned} \tag{14}$$

are obtained using linear regression. In the linear model, the gains and offsets adjust for differences in solar angle, sensor gain, atmospheric transmission, scattering, and other global differences between the two images.

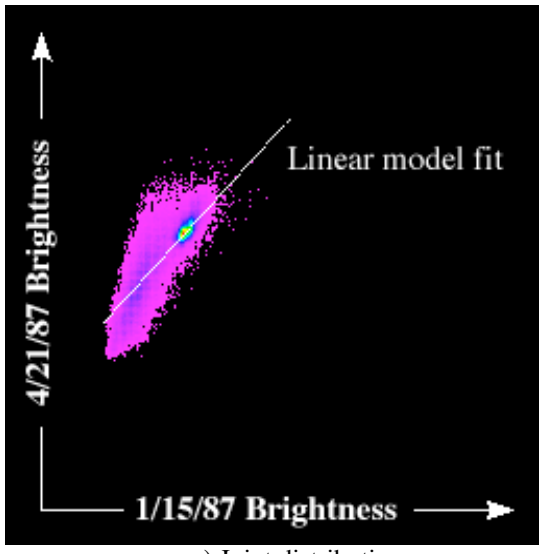
Figure 4 is an example of linear change detection. The area shown is a 400 x 400 pixel region from a Landsat TM image (band 1) in central Iraq. The first image was acquired on 1/15/87 just before the start of construction of an industrial facility near the center of the image. The second image was acquired on 4/21/87 during the early stages of construction. The January image (a) is darker than the April image (b) because of differences in solar angle. The high correlation (low scatter) in the joint density (c) implies that there are relatively few changes between the two images. The linear model adapts easily to the illumination difference and is able to detect changes in vegetation (lower left), water level (pond just left of center), and within the industrial facility (center) as shown in (d).

Consider now a second image pair (Figure 5) where there are significant differences between the two images. The two images were taken about a month apart in 3/30/88 and 5/1/88. Except for the clouds, they are relatively similar in content. The clouds in (b) significantly increase the scatter in the joint density (c). They effectively pull the model away from the hypothetical regression line assuming there were no clouds (dotted line) to the actual regression line with clouds (solid line). The backward prediction error image is shown in (d). Changes in the upper left are clouds and surface features obscured by clouds. A rectangular box is drawn around the facility under construction near the center of the image. The bright region outlined in the lower right is a spurious change caused by the model attempting to adapt to the clouds. Although the linear model works well when the amount of change is relatively small, when there are large changes, e.g., due to clouds or significant seasonal variations, the linear model, in attempting to adapt to these changes as if they were part of the background, will generate false alarms thus reducing its effectiveness.

a) 1/15/87 image



b) 4/21/87 image



c) Joint distribution

d) Forward prediction error image

Figure 4 Example of linear change detection where there are relatively few changes

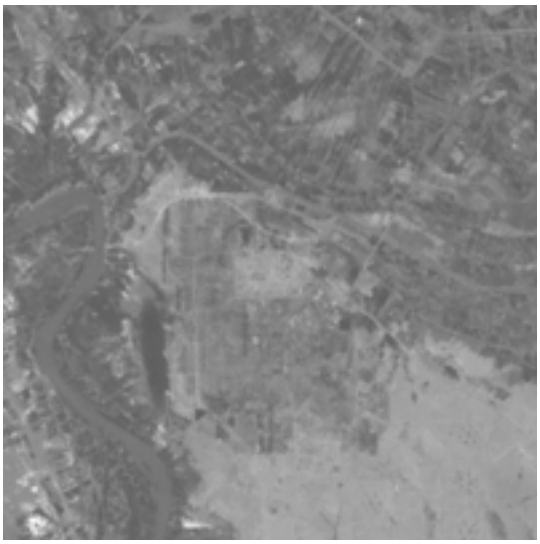
This motivates the development of a model that can handle cases where there may be large changes in the b ackground. In particular, we allow f and g to be nonlinear functions of image brightness

$$f(x) = \sum_{y} yp(y \mid x) \tag{15a}$$

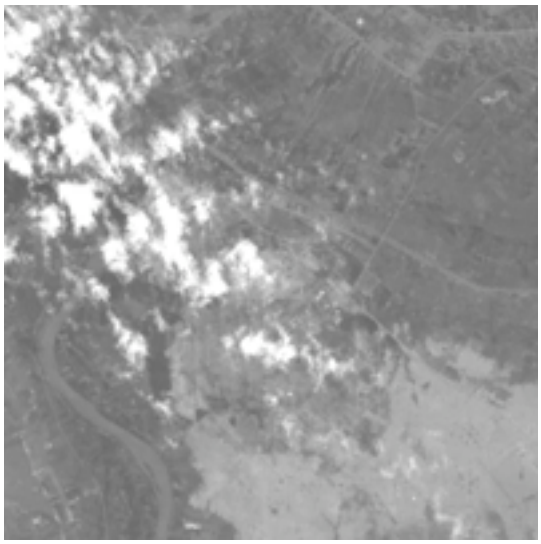
$$g(y) = \sum_{x} xp(x \mid y) \tag{15b}$$

Figure 6 shows how the nonlinear change detection technique can better able adapt to changes caused featu res such as clouds and shadows that affect limited portions of the joint probability space. The joint density i s shown in (a) where the dotted line is a plot of the backward predictor function (Eq. 15b). The change in th e rectangular box in the center of Figure 6b is a better indication of the actual change caused by the facility

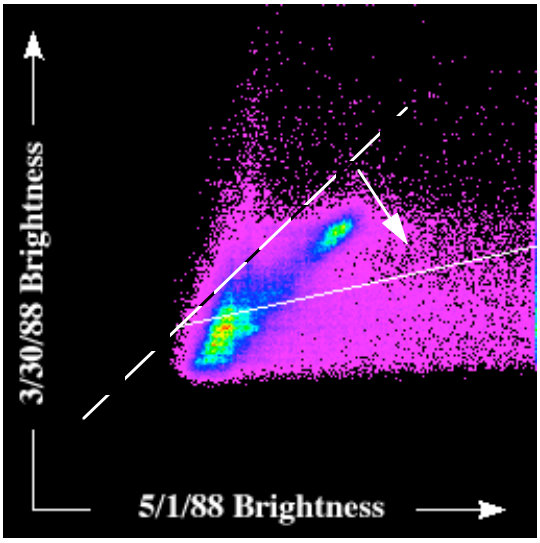
construction than that shown in Figure 5d and the large region of change outlined in the lower right has bee n largely eliminated.



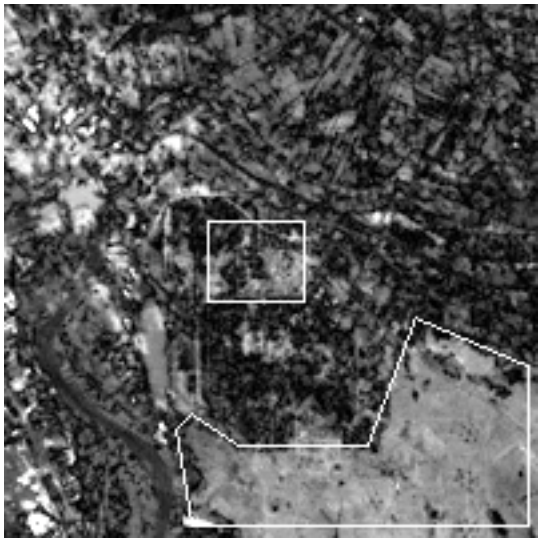
a) 3/30/88 image



b) 5/1/88 image

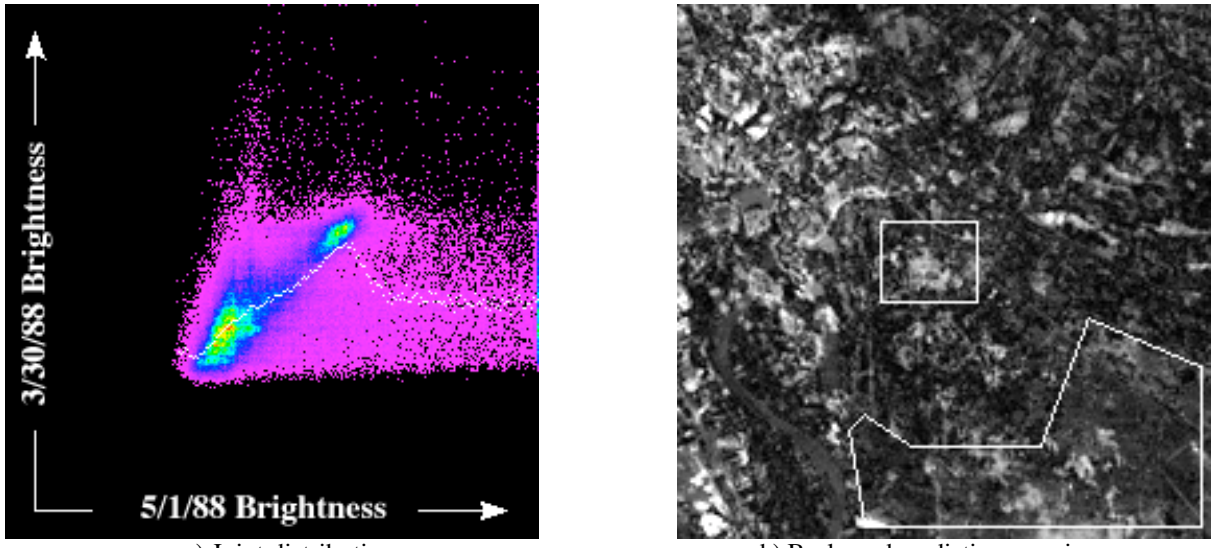


c) Joint distribution



d) Backward prediction error image

Figure 5 Example of linear change detection where there are many changes



a) Joint distribution b) Backward prediction error image

Figure 6 Example of nonlinear change detection

5. Haze Equalization

The two previous sections have used the nonlinear estimation technique to derive relations between multispectral imagery and terrain data, and between two images acquired at different times. This section uses the technique to predict one set of multispectral bands from another in order to reduce space-varying wavelength-dependent scatter caused by smoke and haze. The approach is based on using information from the infrared (IR) bands which are not affected by haze to predict the haze-free values of the visible bands. Assume that the atmosphere can be modeled as a horizontally homogeneous medium, the earth is a Lambertian reflector, and atmospheric properties which vary exponentially with altitude can be assumed to be constant over the scene (Sjoberg and Horn 1983). When the contribution of the ambient illumination (skylight) can be neglected and scattering is constant over the image, the brightness

$$y(i,j) = ar(i,j) + b \tag{16}$$

where r is the radiance of the surface, a represents the effects of atmospheric transmittance, sensor gain, and other multiplicative factors, and b represents the contribution of the path radiance. The brightness values at two or more pixels in the image together with measurements of the surface radiance at the corresponding locations on the ground can be used to solve for a and b. Alternatively, when ground truth is not available methods have been developed to estimate the path radiance component (e.g., Crane 1971, Potter and Mendlowitz 1975, and Crippen 1987). When the effects of scattering are not constant over the image, the image formation model becomes

$$y(i,j) = a(i,j)r(i,j) + b(i,j). \tag{17}$$

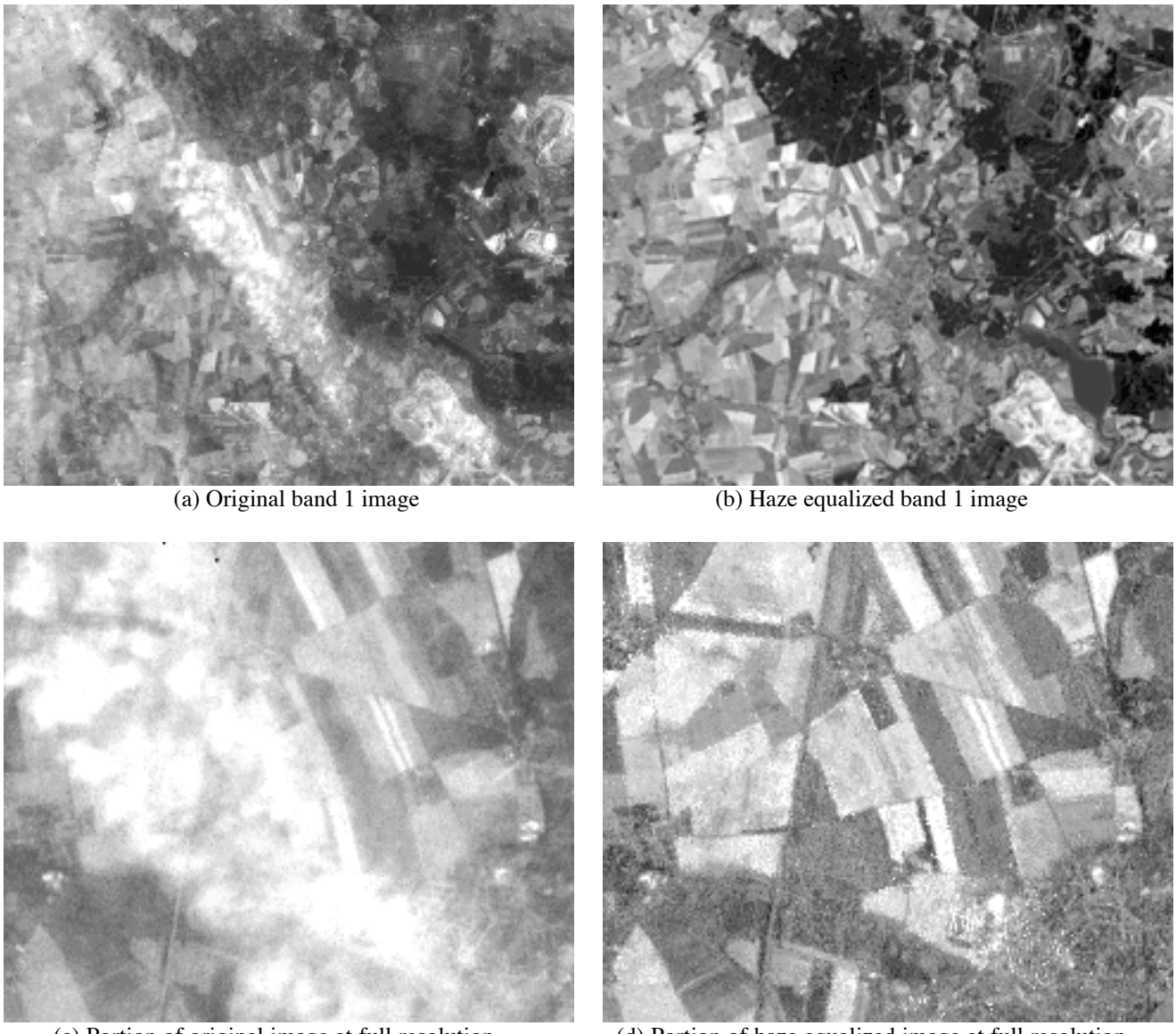
The dependence of scattering on wavelength is a function of the size of the particles relative to the wavelength of the radiation. In multispectral imagery, haze (and smoke) tends tend to affect the shorter wavelengths (i.e., the visible bands), often having negligible effect in the IR bands. Assume that Eq. 17 represents the effects of spatially-varying scattering in one of the visible bands. Let

$$\hat{y}(i,j) = f[\mathbf{x}(i,j)] \tag{18}$$

be an estimate of the visible band as a function of the IR bands $\mathbf{x}(i, j) = [x(i, j, 1) \cdots x(i, j, M)]$. It can be shown that under fairly general conditions (Carlotto 1986)

$$\hat{y}(i, j) = ar(i, j) + b \tag{19}$$

The predictor thus equalizes the haze so that it is uniformly distributed over the image. This effectively reduces the more difficult problem of space-varying scattering (Eq. 17) to the simpler constant scattering problem (Eq. 16) which can be handled by existing methods.



(a) Original band 1 image

(b) Haze equalized band 1 image

(c) Portion of original image at full resolution

(d) Portion of haze equalized image at full resolution

Figure 7 Haze equalization results for Bitterfeld image

Figure 7a shows a Landsat TM image (band 1) acquired in September 1987 over Bitterfeld, Germany. The image is 960 x 808 pixels in size and has been reduced by 25% and stretched for visual presentation. Smoke and dust from a concentration of chemical manufacturing and coal mining obscures much of the area. Figure 7b shows the band 1 image predicted from the three TM reflective IR bands (TM bands 4, 5, and 7). Fi

gures 7c and d show a portion of the original and haze equalized images north of Bitterfeld at full resolutio n. The algorithm has been compared to an earlier algorithm for removing space-varying haze and found to better preserve the spatial detail within, and the spectral balance between corrected bands (Carlotto 1996).

6. Summary

An approach to image modeling based on nonlinear mean-square estimation that does not assume a functio nal form for the model was described. The relationship between input and output images is represented in t he form of a lookup table that is efficiently computed from, and applied to images. Three applications were presented to illustrate the utility of the technique in remote sensing. The first illustrated how the method can be used to estimate the values of physical parameters from imagery. Specifically we estimated the topograp hic component (i.e., the variation in brightness caused by the shape of the surface) from multispectral image ry. The second application was a nonlinear change detection algorithm which predicts one image as a nonli near function of another. In cases where the frequency of change is large (e.g., due to atmospheric and envi ronmental differences), the algorithm is shown to be superior in performance to linear change detection. In t he last application, a technique for removing wavelength-dependent space-varying haze from multispectral i magery was presented. The technique uses the IR bands, which are not affected significantly by haze, to pre dict the visible bands. Results showed a significant reduction in haze over the area considered.

Other applications of the technique are currently being explored. They include bathymetric extraction, imag e restoration (in which one band or sensor is used to correct for pixel errors in another band or sensor), and sensor morphing in which one sensor can be made to look like another.

References

M.J. Carlotto, "Reducing the effects of space-varying wavelength-dependent scattering in multispectral ima gery," *International Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, 1996.

M.J. Carlotto, "A new method for extracting topographic information from a single multispectral image," *In ternational Geoscience and Remote Sensing Symposium*, Lincoln, Nebraska, 1996.

M. J. Carlotto, M. B. Lazaroff and M. W. Brennan, "Multispectral image processing for environmental mo nitoring," *Proc. SPIE*, Vol. 1819, pp 113-124, November 1992.

R.B. Crane, "Preprocessing techniques to reduce atmospheric and sensor variability in multispectral scanne r data," *Proceedings 7-th International Symposium on Remote Sensing of Environment*, 1971.

R.E. Crippen, "The regression intersection method of adjusting image data for band ratioing," *Int. J. Remot e Sensing*, Vol. 8, No. 2, pp 137-155, 1987.

P. Eliason, L. Soderblom and P. Chavez, "Extraction of topographic and spectral albedo information from multispectral images, *Photogrammetric Engineering and Remote Sensing*, Vol. 48, No. 11, pp 1571-1579, 1981.

B.K.P. Horn, "Understanding image intensities," *Artificial Intelligence*, Vol. 8, pp 201-231, 1977.

B. G. Lee, V. T. Tom and M. J. Carlotto, "A signal-symbol approach to change detection," *Proc. AAAI*, Au gust 1986.

A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1965.

J.F. Potter and M.A. Mendlowitz, " On the determination of haze levels from Landsat data," *Proceedings 1 0-th International Symposium on Remote Sensing of Environment*, 1975.

R.W. Sjoberg and B.K.P. Horn, "Atmospheric effects in satellite imaging of mountainous terrain," *Applied Optics*, Vol. 22, No. 11, pp 1702-1716, 1983.

A.H. Strahler and C.E. Woodcock, "On the nature of models in remote sensing," *Remote Sensing of Enviro nment,* Vol. 20, pp 121-139, 1986.

C. W. Therrien, T. F. Quatieri and D. D. Dudgeon, "Statistical model-based algorithms for image analysis, " *Proc. IEEE*, Vol. 74, No. 4, April 1986.