Digital Photogrammetric Inversion: Theory and Application to Surface Reconstruction

Abstract

Recovering descriptions of objects represented in digital image data is the first step of digital photogrammetry. It deals with the dual issues of representation and inversion. The form of the input, i.e., digital images, and the desired output, e.g., descriptions of image events like lines, regions, or visible surfaces, has to be specified before the algorithms that derive object descriptions from primary image data are developed. This process is typically inverse because the measurements are performed indirectly. Many problems, including how to create constraints linking desired parameters and data, how to recover the unknown parameters from the constraints, how to integrate a priori knowledge to deal with uncertainty, and how to evaluate the quality of the solutions, can occur during this inverse process. All of these are the objective of this paper. We look at the issue of inverse problems in object-reconstruction from image data generally. Based on the Maximum A Posteriori *(MAP) principle, we introduce a theoretical framework for the decision problem in the inverse process with uncertainty. As its application, we show how this framework leads naturally to recover visible surfaces from multiple images. We propose an algorithm for automatic generation of digital elevation models (DEM) and illustrate its performance with experimental results.*

Introduction

Photogrammetry, as a traditional contact-free surveying technique, is entering into a new stage of development. This is perceptible through the application of more and more techniques based on the processing of digital image data for photogrammetric purposes. In comparison with traditional analog and analytical methods, digital photogrammetry can be chiefly characterized as follows:

- The input, namely the original observation, is extended from geometric values, e.g., coordinates of image points, to physical measurements, e.g., intensities of image points.
- The output, namely the desired result, is concerned not only with geometric but also with physical properties of objects depicted in images.
- The realization of input-output transformation aims at full automation, contrary to the traditional approaches which rely essentially on the actions and intelligence of human operators for image measurement and interpretation.
- The application range is becoming wider, from topography to robot vision, for instance.

These extensions make digital photogrammetry more challenging and more complex and therefore require new formulations of the problems and new frameworks for problem solving.

Digital photogrammetry attempts to recover descriptions of geometrical and physical properties of objects automatically from primary image data. Generally speaking, this is a typical inverse problem as the desired information about the world is inferred from image data. The first step of this image inversion is recovering descriptions of image events including lines, regions, structures, and visible surfaces from numerical arrays of pixels. Here the main goal could be formulated in the form "given a mapping $X \to Y$ and data Y, recover a set of parameters X." To facilitate understanding, let us look at a simple example of digital photogrammetric relative orientation. Here some pairs of corresponding features in two stereo images and the optical flow equations describing the geometric relationship between these two images are given. Now, the main question is how to recover the unknown parameters in the optical flow equations, based on the given data which may be inexact or erroneous.

Inverse problems are basic issues, of which we are all aware through our own experience in research on many problems in digital photogrammetry, including feature extraction, image and boundary segmentation, object reconstruction, and image interpretation. These kinds of problems are of universal significance, and their solutions are not always straightforward. Actually, some of them belong to the most difficult tasks in computer applications. There are at least two reasons for this. First, an image does not provide enough information, by itself, to recover the underconstrained scene. Many factors, including surface material, atmospheric conditions, light source, ambient light, camera angle and characteristics, etc., are compounded in the image and contribute to a single measurement, say the intensity of a pixel. The various factors cannot be separated, as long as they are not measured. Another reason for the difficulty consists in the fact that so much information is lost during the imaging process that projects the three-dimensional (3D) world into two-dimensional (2D) images with a total reduction of dimensions. This would be true even if there were no stochastic component in a digital image, because it is corrupted by both discrete spatial sampling and intensity quantization.

In this paper, we are concerned with inverse problems in digital photogrammetry. We first try to classify inverse problems into groups and discuss their solvability. After that, we focus our attention on the problem of "uncertainty" during image inversion which is also known as an "ill-posed" problem (Hadamard, 1923). Based on the *Maximum A Posteriori* (MAP) principle, we introduce a framework for the integration of *a priori* knowledge to solve the decision problem in the inverse process with uncertainty. To demonstrate the approach, we formulate the problem of surface reconstruction from image data within this framework. This leads to a novel algorithm for automatic generation of digital elevation models (DEM) from multiple digital images. We then demon-

Yong-Jian Zheng
Institute for Photogrammetry and Remote Sensing, University Karlsruhe, Englerstrasse 7, D-7500 Karlsruhe 1, Germany

strate briefly the implementation and illustrate its performance with experimental results. The main contributions of this paper are (1) formulating problems in digital photogrammetric object reconstruction in a more general way, (2) providing a theoretical framework for problem solving, and (3) applying this framework for the concrete problem of recovering the geometric representation of surfaces from image data.

Inverse Problems

Let **S** represent a physical system (for example, the Earth's surface, or an image event like a line or a region). Assume that we are able to define a set of model parameters which describe **S** to some extent. These parameters may not all be directly measurable. We can operationally define a set of some observable parameters $\mathcal{Y}$ whose actual values hopefully are relatable to a set of the model parameters **X**. To solve the *forward problem* is to predict the values of the observable parameters $Y \in \mathcal{Y}$ given arbitrary values of the model parameters $X \in X$. To solve the *inverse problem* (the identification problem) is to infer the values of the model parameters **X** from given observed values of the observable parameters **Y** (Tarantola, 1987). Mathematically, the inverse problem can be described as follows. Given a mapping f from set **X** into set $\mathcal{Y}$, i.e., $f : X \rightarrow \mathcal{Y}$, the solution of the inverse problem consists in the interpretation of data $Y \in \mathcal{Y}$ in order to recover the original image $X \in X$.

Obviously, the solvability of the inverse problem is connected strongly with the characteristics of the mapping f (Table 1). A mapping f is called *surjective* if every element $Y \in$ **Y** has at least one original image $X = f^{-1}(Y) \in X$. A mapping f is called *injective* if an element $Y \in \mathcal{Y}$ has only one original image $X = f^{-1}(Y) \in X$. A mapping f is *bijective* if it is both injective and surjective. So a surjective mapping ensures the existence of its inverse mapping and an injective mapping ensures the uniqueness of its inverse mapping.

Let us now consider a linear mapping $A : X \rightarrow \mathcal{Y}$. The inverse problem is to identify **X** from the "data" **Y** and the equation

$$Y = AX. \tag{1}$$

The solvability of this inverse problem could be discussed as follows:

(1) If **A** is bijective and A^{-1} is stable, one can easily get a unique solution $X = A^{-1}Y$.

(2) If **A** is injective but not surjective, the inverse problem is *overdetermined*. One can, however, get a unique pseudo solution through minimizing the norm of the residual $\|V\| = \|Y - AX\|$.

(3) If **A** is not injective, the inverse problem is *underdetermined* and there is a unique pseudo solution **X** $= A^+Y$, where A^+ is the so called Moore-Penrose Inversion which, unfortunately, is only stable when the domain of **A** is closed in **Y**.

So, it is quite clear that inverse problems are not always straightforward to solve by using traditional methods, i.e., least-squares estimation. In fact, some ambivalent non-injective inverse problems are practically not solvable through a numeric process, as even a few errors in **Y** can destroy the solution totally. Unfortunately, there are many problems in digital photogrammetry, as mentioned earlier, which are of a non-injective nature and whose solutions demand new inference techniques beyond traditional estimation methods.

Schematically, there are two reasons for the uncertainty in inverse problems: intrinsic lack of data, and observation uncertainties. With additional information, for example, some *a priori* assumptions regarding model parameters **X** or

TABLE 1. CHARACTERISTICS OF A MAPPING $f : X \rightarrow \mathcal{Y}$.

criterion	expression	supposition
MAP	$P(X \mid Y) \rightarrow \max$	
BE	$P(Y \mid X)P(X) \rightarrow \max$	1). $P(Y)$ = constant
ML	$P(Y \mid X) \rightarrow \max$	2). $P(X)$ = constant and 1)
LS	$V^{T}\Sigma^{-1}V \rightarrow \min$	3). $V \sim N(0,\Sigma)$ and 2)

an additional data set, many such problems can be reformulated into well-posed solvable problems. Now, the main question is how to integrate *a priori* knowledge to solve dubious inverse problems. One way is to restrict the space of admissible solutions by introducing suitable *a priori* knowledge. We need criteria to impose constraints on the solution space and a framework to integrate *a priori* knowledge in order to select a unique solution (the so called "best" solution) for given data. Intuitively, the best solution exists only in connection with criteria which are, of course, strongly task dependent.

The MAP Criterion

The criterion which we introduce in this section is the so-called *Maximum A Posteriori* (MAP) criterion which is based on probability theory (Geman and Geman, 1984). It selects as the best solution the model parameters **X** that maximize the conditional probability of **X** given the data **Y** : $P(X \mid Y)$, subject to the inverse problem (Equation 1). The MAP criterion leads to three important estimation methods, namely, the *Bayesian estimation* method (BE), the *maximum likelihood* method (ML), and the *least-squares* method (LS), which are widely used in data processing.

Using Bayes' theorem gives

$$P(X \mid Y) = \frac{P(Y \mid X)P(X)}{P(Y)}, \tag{2}$$

where $P(Y \mid X)$ is the conditional probability of getting data **Y** given the model parameters **X**, and $P(X)$ is the prior probability of **X**. The relation (Equation 2) shows how the "prior probability" $P(X)$ changes to the "posterior probability" $P(X \mid Y)$ as a result of acquiring new information **Y**. Intuitively, the MAP criterion will choose X that maximizes

$$P(Y \mid X)P(X), \tag{3}$$

if $P(Y)$ is constant. This is the principle of the Bayesian estimation. Furthermore, under the specification that the prior probabilities $P(X)$ are all the same, i.e., $P(X)$ is constant, the MAP criterion leads to the simpler maximum-likelihood principle of selecting that **X** which maximizes $P(Y \mid X)$. If the random variables to which the data **Y** refer are normally distributed, the maximum-likelihood estimation will give the same results as the least-squares estimation which has been widely used in different branches of science and engineering. If **V** is the vector of observational residuals, for which $E(V)$ = 0, and it is assumed that **V** is normally distributed, and Σ is the covariance matrix of the distribution, then we have

$$P(Y \mid X) = P(V) = C \cdot \exp\left[-\frac{1}{2}V^{T}\Sigma^{-1}V\right] \tag{4}$$

where C is constant. The least-squares criterion is to minimize $V^{T}\Sigma^{-1}V$ which is equivalent to yielding a maximum-likelihood estimation to maximize $P(Y \mid X)$.

So far we have discussed the MAP criterion and its progenies. Obviously, each criterion has its own supposition

(Table 2). The LS criterion, which is so widely used in photogrammetry as a general framework for problem solving, is only suitable for dealing with over-constrained inverse problems. For under-constrained inverse problems, the MAP criterion is more appropriate as it provides a flexible framework to integrate *a priori* knowledge to restrict the solution space, and one can take the probability behavior of both the data and the desired solutions into account. This is most important as we, then, have the chance to make full use of our *a priori* knowledge.

Restricting the Solution Space

The MAP criterion provides a general approach to handle the inverse problem in an uncertain environment. It gives a mechanism to restrict the solution space and to integrate *a priori* knowledge by specifying the appropriate prior probabilities P(X). However, the MAP criterion doesn't tell how to construct P(X). In this section we look at this issue.

The parameter set X, as mentioned above, represents a physical system and can be considered as a space. In principle, every point X in this space represents a possible solution. It can be easily imagined that not all points in the solution space are meaningful. Our job is to explore the solution space to find an appropriate point (solution). So, the first problem is how to measure the appropriateness of a solution and how to describe the solution space. A general way to do this is to define a possibility distribution of the solution space P(X)(Tarantola, 1987).

Constructing P(X) is strongly task dependent as we have to know the meaning of X. If we, for instance, know *a priori* that X can be described by using a Markov random field, then, according to the Hammersley-Clifford Theorem (Geman and Geman, 1984; Chou and Brown, 1988), P(X) can have the Gibbs form

$$P(X) = \frac{1}{C} \exp\left[-\frac{1}{T}\mathcal{E}(X) \right], \quad X \in X, \qquad (5)$$

where C is a normalizing constant, T is the "temperature" of the space that controls the flatness of the distribution of the configurations X, and $\mathcal{E}(X)$ is the energy of X which consists of the sum of the local potential

$$\mathcal{E}(X) = \sum_{x \in X} V_e(x). \qquad (6)$$

The relation (Equation 5) suggests that the point in X with a higher energy occurs is less likely to occur.

Now, let us come back to the inverse problem (Equation 1). According to the the least-squares criterion (Table 2), one can get a unique pseudo solution by minimizing $V^T\Sigma^{-1}V$ with respect to

$$V = AX - Y. \qquad (7)$$

This leads to solving the normal equation

$$(A^T\Sigma^{-1}A)X = A^T\Sigma^{-1}Y. \qquad (8)$$

Certainly, the normal matrix $N = A^T\Sigma^{-1}A$ is regular only if the problem (Equation 1) is overdetermined. This suggests that the least-squares criterion can only be used to deal with overdetermined inverse problems. For underdetermined inverse problems, the least-squares criterion can not lead to a satisfying solution, as it does not have a mechanism to restrict the solution space.

Using the Bayesian estimation method (BE) (Table 2), we have the following optimizing problem (Equations 3, 4, and 5):

$$V^T\Sigma^{-1}V + \frac{2}{T}\mathcal{E}(X) \rightarrow \min, \qquad (9)$$

with respect to Equation 7. Intuitively, this criterion, in comparison with the least-squares criterion, is more powerful to deal with underdetermined inverse problems, as it gives not only a measure for the quality of the fitting , through the first term in Equation 9, but also a measure for the probability of the solution, through the second term in Equation 9. So we can integrate our *a priori* knowledge into the inverse process by designing the second term in Equation 9 appropriately. Of course, designing $\mathcal{E}(X)$ is a skill. One needs knowledge about the physical meaning of the solution and the internal coherence of unknown parameters.

Generally, the second term in Equation 9 can be designed to have the form

$$\frac{2}{T}\mathcal{E}(X) = E^T\Sigma_e^{-1}E, \qquad E = \Phi X - \Psi, \qquad (10)$$

where Φ is an operator, Ψ is a vector, and Σ_e is a matrix. They have to be determined using our *a priori* knowledge. If we, for instance, know *a priori* that the elements $x_i \in X$, $i = 1, ..., m$, should have values near $a_i, = 1, ..., m$, then we can construct

$$\Sigma_e = \frac{T}{2}\begin{pmatrix} \sigma_1^2 & 0 & ... & 0 \\ 0 & \sigma_2^2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \sigma_m^2 \end{pmatrix},$$

$$\Phi = \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & 1 \end{pmatrix}, \quad \Psi = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, \qquad (11)$$

where σ_i, $i = 1, ..., m$, denote the degree of the certainty of our *a priori* knowledge.

Let us now solve the inverse problem (Equation 1) again, but using the new criterion (Equation 9) which is equivalent to

$$V^T\Sigma^{-1}V + E^T\Sigma_e^{-1}E \rightarrow \min. \qquad (12)$$

This leads to the new normal equation

$$(A^T\Sigma^{-1}A + \Phi^T\Sigma_e^{-1}\Phi)X = A^T\Sigma^{-1}Y + \Phi^T\Sigma_e^{-1}\Psi. \qquad (13)$$

One can be certain that the new normal matrix $N = A^T\Sigma^{-1}A + \Phi^T\Sigma_e^{-1}\Phi$ is not singular even for underdetermined inverse problems, if Φ, Σ_e, and Ψ are all appropriately constructed.

Surface Reconstruction From Images

There are, as indicated above, many problems in digital photogrammetry which can be formulated as the inversion of image data. We have proposed general approaches for inverse problems which provide a sound theoretical basis but offer few practical computational methods for dealing with concrete tasks in digital photogrammetry. So, in this section, we go further into the application of the inverse problem theory to an elementary problem in photogrammetry, i.e., the computing of the representation of surfaces from multiple images, as a solid application example.

TABLE 2. THE MAP CRITERION AND ITS PROGENIES.

f	neither surjective nor injective	surjective noninjective	injective nonsurjective	bijective
X ↓ Y				
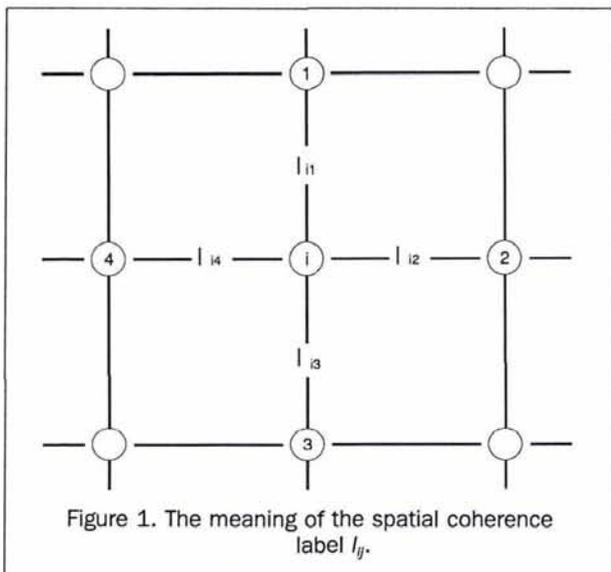
Figure 1. The meaning of the spatial coherence label l_{ij}.

Representation of Visible Surfaces

The role of a representation is to make certain information explicit at an appropriate point in the problem analysis because the abstract information must be expressed by concrete representation. Thus, the choice or design of a representation affects the success of analysis. The representation of object surfaces deals with strategies and techniques for describing their geometrical and physical properties in a way appropriate for numerical processing.

Let $\mathcal{S}$ be a set of parameters which describe the geometrical and physical properties of an object surface. The element

of $\mathcal{S}$ can be a concrete measure, e.g., elevation (depth), deformation, reflectivity, etc. Each element in $\mathcal{S}$ can be mapped onto the XY-plane in a 3D coordinate system and represented mathematically as $S = S(X,Y)$, $S \in \mathcal{S}$. For computational reasons, we would rather represent S by a grid of square 1×1 elements, where each element is centered at the coordinates (X_i, Y_i) of the ith element. Then, the object surface is described by $m \times n$ elements: $S_i = S(X_i, Y_i)$, $i \in \mathcal{I}$, where $\mathcal{I}$ can be thought of as a vector belonging to the set $(1, ..., m) \times (1, ..., n)$ which has totally $m \times n$ elements.

Sometimes we may also be interested in the spatial coherence (continuity) of S. So we introduce a label set **L** whose element l_{ij} represents the strength of the spatial coherence between two neighbors S_i and S_j (Figure 1). The label l_{ij} can be binary: $l_{ij} = 0$ for continuity between S_i and S_j, $l_{ij} = 1$ for discontinuity between S_i and S_j. l_{ij} can also take the value between 0 and 1, i.e., $l_{ij} \in [0,1]$, for continuously describing the coherence strength.

Forward Modeling for Constraints Generation

The purpose of forward modeling is to find constraints linking elements in $\mathcal{S}$ with observations, i.e., image densities (intensities), based on physical properties of the imaging. The well known relationship between the image density D of a photographic image and the exposure $H[lux.sec]$ is

$$D(x,y) = \gamma \log H + D_0 \qquad (14)$$

for normal exposure, where x,y are image coordinates of an image point, $\gamma(\approx 1)$ is the gradation, and D_0 is a constant. Usually, γ depends on the developer, the development time and temperature, and the photographic material. The exposure H depends, first of all, on the reflection properties of surfaces (Figure 2). Many natural terrain features roughly approximate diffuse reflectors. A Lambertian surface is a per-
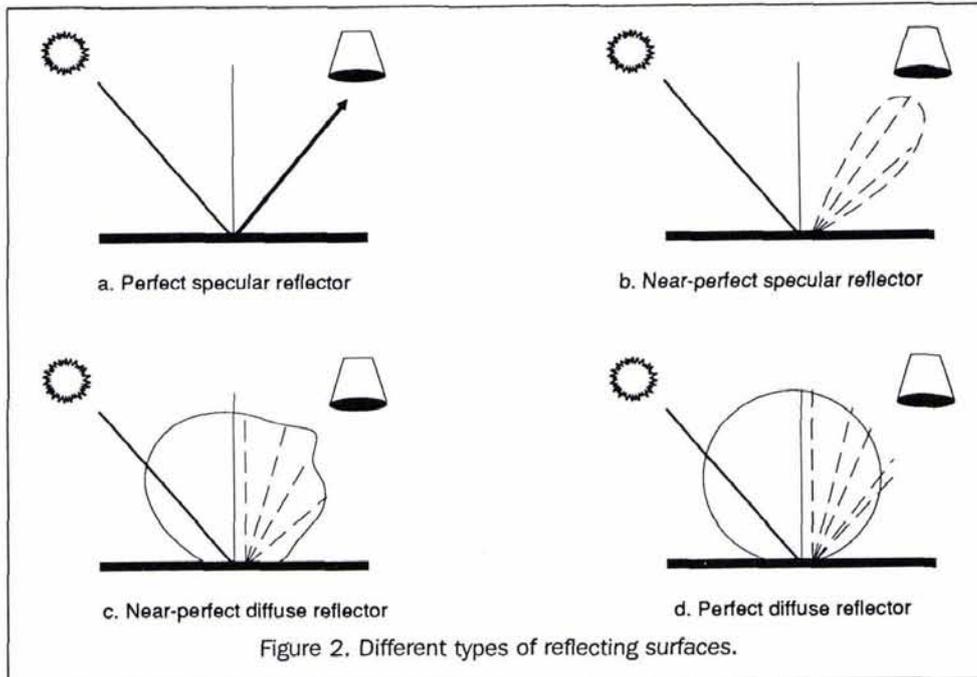
Figure 2. Different types of reflecting surfaces.

fect diffuse reflector with the property that the radiance $L[cd.m^{-2}]$ is constant for any incident angle.

The exposure of an emulsion (Slater, 1983) can be written as

$$H = \frac{t\pi \cos^4\alpha}{4N^2}(L_g \cdot \tau_a + L_s)\tau_0$$

$$= \frac{t\pi c^4}{4N^2(c^2 + x^2 + y^2)^2}(L_g \cdot \tau_a + L_s)\tau_0, \qquad (15)$$

where t [sec] and α[rad] are exposure time and field angle of a pixel, c is calibrated focal length, N is relative aperture of the lens, $(L_g \cdot \tau_a)$, L_s are luminance [$cd \cdot m^{-2}$] from the surface element and of skylight intercepted by the lens; and τ_a, τ_0 are transmittance of atmosphere and optics.

Substituting Equation 15 in Equation 14 gives

$$D(x,y) = \gamma \log \frac{t\pi c^4}{4N^2(c^2 + x^2 + y^2)^2}(L_g \cdot \tau_a + L_s)\tau_0 + D_0$$

$$= f(\gamma,t,c,x,y,N,L_g,L_s,\tau_a,\tau_0,D_0), \qquad (16)$$

which is a basic formula for digital photogrammetric inversion as it explains the physical meaning of observations, i.e., image intensities. It can be seen that there is uncertainty during the inversion of model parameters from Equation 16 even if there were no stochastic components to the observations D. Thus, Equation 16 is practically useless as long as no information about the physical calibration of the whole imaging process is provided. Rewriting Equation 16 gives

$$D(x,y)/\gamma + 2 \cdot \log(c^2 + x^2 + y^2) + \eta + \psi(x,y) = 0, \quad (17)$$

where η can be considered as a constant for all pixels in the same image; but ψ is a local parameter which changes from pixel to pixel. The physical meaning of ψ in Equation 17 is the logarithm of the luminance intercepted by the lens for a pixel.

The image coordinates x and y in Equation 17 are further functions of the object coordinates of the surface element, ac-

cording to the well known projection equation:

$$\begin{pmatrix} x \\ y \\ -c \end{pmatrix} = \frac{1}{m}\mathbf{R}^T \left[\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} - \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} \right], \qquad (18)$$

where m is a scale factor, $\mathbf{R}$ is the rotation matrix containing three rotation angles (ϕ,ω,κ), (X,Y,Z) are the corresponding ground coordinates of the image point (x,y), and $\Omega = (\phi,\omega,\kappa, X_0,Y_0,Z_0)$ are camera orientation parameters. In addition, the intensity $D(x,y)$ of the image point (x,y) has to be re-sampled from the neighboring digitized pixels by using, for example, a bilinear interpolation

$$\mathbf{D} = \mathbf{G}^T\mathbf{L}, \qquad (19)$$

where $\mathbf{L}$ denotes a set of intensities of neighboring pixels and $\mathbf{G}$ denotes a corresponding coefficient vector, respectively. Thus, for the ground surface point (X,Y,Z), the left side of Equation 17 is a function of many parameters: i.e.,

$$f(X,Y,Z,\Omega,\mathbf{L},\gamma,\eta,\psi) = 0. \qquad (20)$$

Let us now look at the problem of surface reconstruction from multiple images. To facilitate the analysis, we discuss here only the solution of recovering the surface profile illustrated in Figure 3, which is represented by using K discrete profile points: Z_i, $i \in \mathcal{I} = [1, ..., K]$, from J images: M_j, $j \in \mathcal{J} = [1, ..., J]$, which are taken from different views and depict the same surface (Figure 3). In addition, we also suppose that the orientation parameters of these images are all known.

For the ith profile point, we can write J constraints like Equation 17. For all K profile points we can write totally $J \times K$ constraints like Equation 17. Supposing a Lambertian surface and $\gamma \simeq 1$, these constraints can be further simplified (Zheng, 1990): i.e.,

$$D_1(x_{i1},y_{i1}) + p_j D_j(x_{ij},y_{ij}) + 2 \log \left(\frac{c_1^2 + x_{i1}^2 + y_{i1}^2}{c_j^2 + x_{ij}^2 + y_{ij}^2} \right) + q_j = 0,$$
$$i \in [1, ..., K], j \in [2, ..., J], \qquad (21)$$
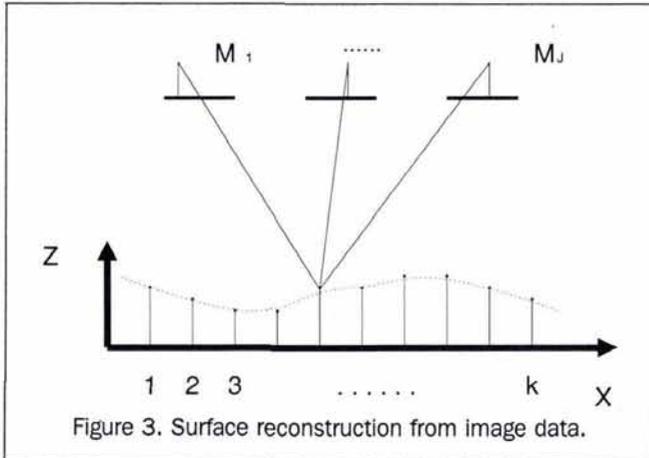
Figure 3. Surface reconstruction from image data.

where $p_j = -\gamma_1/\gamma_j$, $q_j = \eta_1 - \eta_j$, and we have a set of new parameters p_j and q_j, $j \in [2,...,J]$, which describe approximately the radiometric relationship of the image M_1 with the other images M_j, $j \in [2,...,J]$. Our *a priori* knowledge about p_j and q_j is that p_j should be near 1 and q_j should be near 0.

It is not difficult to see that in this example the left side of Equation 21 is a function of many parameters, including the elevation of each profile point (see Equations 19 and 18): i.e.,

$$f(L_{i1}, L_{ij}, Z_i, p_j, q_j) = 0, \quad i \in [1, ..., K], j \in [2,...,J].$$

These $K \times (J-1)$ constraints can be further denoted as

$$F(L, Z, Q) = 0 \qquad (22)$$

with

$$L = (L_{11},...,L_{K1}, L_{12},...,L_{K2},...,L_{1J},...,L_{KJ})^T$$
$$Z = (Z_1,...,Z_K)^T,$$
$$Q = (p_2, q_2,..., p_J, q_J)^T.$$

Linearization of these constraints gives

$$UV + A\Delta Z + B\Delta Q + W = 0 \qquad (23)$$

with

$$U = \frac{\partial F}{\partial L}, \; A = \frac{\partial F}{\partial Z}, \; B = \frac{\partial F}{\partial Z}, \; W = F(L_0, Z_0, Q_0),$$

where V denotes the residual vector of L; L_0 denotes the approximate values of L, Z_0 and ΔZ denote the vector of approximate elevations of profile points in $\mathcal{I}$ and their corrections, and Q_0 and ΔQ denote approximate values of the unknown parameters in Q and their corrections, respectively. It is clear that Equation 23 is strongly underdetermined as V, ΔZ, and ΔQ are all unknown. The total number of unknowns is much larger than that of the constraints, and one could generally hypothesize an infinite number of different solutions that would meet Equation 23. So, we have to use criteria to restrict the space of acceptable solutions and to find a unique solution which will be a "best" one to interpret the image data.

The MAP Based Method for Surface Reconstruction

According to Table 2, the MAP criterion would choose $Z = Z_0 + \Delta Z$ and $Q = Q_0 + \Delta Q$ in Equation 23 such that the conditional probability $P(Z, Q \mid L)$ is maximized, which is equivalent to maximizing $P(L \mid Z, Q)P(Z, Q)$, if $P(L)$ is constant.

As mentioned above, the conditional probability $P(L \mid Z, Q)$ can be simply assumed as the probability that the observational residuals were produced by a normally distributed random variable (Equation 4). The problem is, now, how to exploit our *a priori* knowledge about Z and Q to constitute their probabilities, i.e., $P(Z, Q)$. If Z and Q are statistically independent, we have $P(Z, Q) = P(Z) \cdot P(Q)$.

To construct $P(Z)$ and $P(Q)$, one has to know the meaning of Z and Q. The vector Z, for instance, represents some elevations of discrete surface profile points. So, our *a priori* assumption about Z is the spatial coherence of its elements. This suggests that the local potential of the element $Z_i \in Z$, $i \in \mathcal{I}$ can be written as

$$V_a(Z_i) = \sum_{k \in N_i} \left[l_{ik} \frac{Z_i - Z_k}{\sigma_Z} \right]^2 \qquad (24)$$
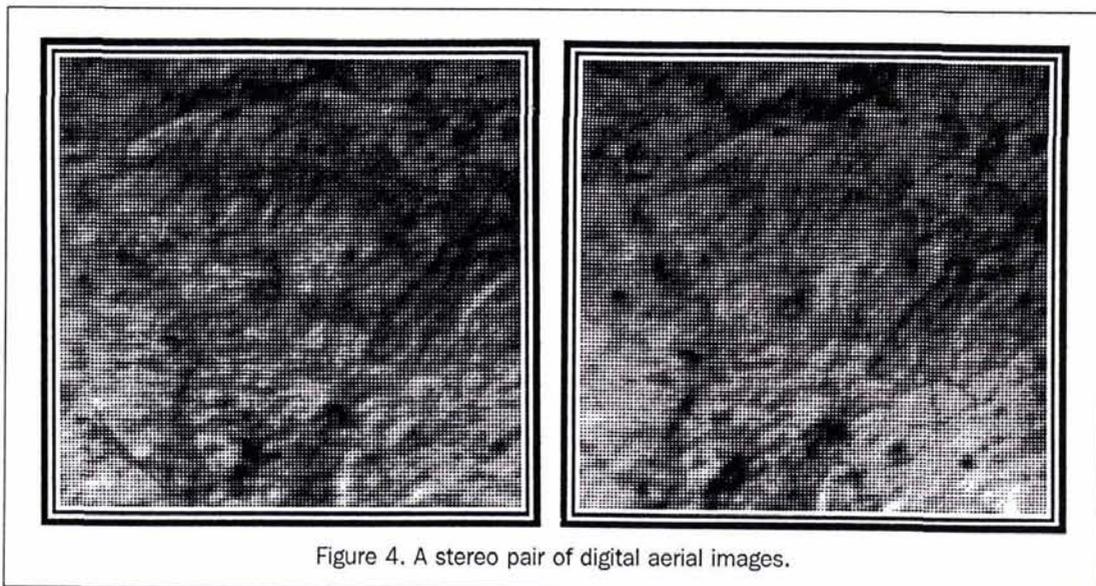
Figure 4. A stereo pair of digital aerial images.

where N_i denotes the set of totally connected subgraphs (cliques) with respect to the element Z_i, $l_{ik} \in [0,1]$ is the connection strength between Z_i and Z_k, and σ_Z is a normalizing constant. According to the Hammersley-Clifford theorem, the energy of Z can be computed

$$\mathcal{E}(Z) = \sum_{i \in I} \sum_{k \in N_i} \left[l_{ik} \frac{Z_i - Z_k}{\sigma_Z} \right]^2 = E_Z^T \Sigma_Z^{-1} E_Z \quad (25)$$

with

$$E_Z = \Phi_Z Z - \Psi_Z, \quad (26)$$

where

$$\Phi_Z = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{pmatrix}, \quad Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_K \end{pmatrix},$$

$$\Psi_Z = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \Sigma_Z = \begin{pmatrix} \frac{\sigma_Z^2}{l_{12}} & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & \frac{\sigma_Z^2}{l_{(K-1)K}} \end{pmatrix} \quad (27)$$

Similarly, because we know a priori that p should be near 1 and q should be near 0, so the energy of Q is

$$\mathcal{E}(Q) = E_Q^T \Sigma_Q^{-1} E_Q, \quad (28)$$

with

$$E_Q = \Phi_Q Q - \Psi_Q, \quad (29)$$

where

$$\Phi_Q = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} p_2 \\ q_2 \\ \vdots \\ p_J \\ q_J \end{pmatrix},$$

$$\Psi_Q = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \Sigma_Q = \begin{pmatrix} \sigma_p^2 & 0 & \cdots & 0 & 0 \\ 0 & \sigma_q^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_p^2 & 0 \\ 0 & 0 & \cdots & 0 & \sigma_q^2 \end{pmatrix} \quad (30)$$

and σ_p and σ_q are constants encoding the reliability of our a priori knowledge about p and q.

Considering Equations 3, 4, 5, 25, and 28, the MAP based surface reconstruction is to solve the optimizing problem

$$\frac{1}{2} V^T \Sigma^{-1} V + \frac{1}{T} E_Z^T \Sigma_Z^{-1} E_Z + \frac{1}{T} E_Q^T \Sigma_Q^{-1} E_Q \rightarrow \min, \quad (31)$$

subject to

$$UV + A\Delta Z + B\Delta Q + W = 0,$$
$$E_Z - \Phi_Z \Delta Z + (\Psi_Z - \Phi_Z Z_0) = 0, \text{ and}$$
$$E_Q - \Phi_Q \Delta Q + (\Psi_Q - \Phi_Q Q_0) = 0.$$

Surely, the surface profile Z which is reconstructed in this way depends on $\Gamma = (\Sigma, \Phi_Z, \Psi_Z, \Sigma_Z, \Phi_Q \Psi_Q, \Sigma_Q)$ and they should be determined using a priori knowledge, before the inversion process takes place. The quality of Z is sometimes not satisfying if our knowledge is not good enough to ensure an appropriate determination of Γ. So, a very interesting question is how to enlarge our knowledge and how to adapt Γ during the inversion in order to improve the quality of Z. Information processing systems that improve their performance or enlarge their knowledge bases are said to "learn." This ability would clearly have value in digital image inversion. Using parameter estimation technique, we can, for instance, adapt Σ, Σ_Z, and Σ_Q in Equation 31 iteratively during the inversion so that the result is robust against image noises with different properties, against surface discontinuities, and against different reliabilities of our a priori knowledge (Zheng and Hahn, 1990).

Experimental Results

To demonstrate the feasibility of the approach for surface reconstruction, an algorithm was implemented (Zheng, 1990); it can be understood as an iterative optimization procedure which consists of the following steps:

(1) For each of J images, create a pyramid by recursive methods using a low pass filter such as a Gaussian filter to remove gray level noise and the high-frequency components in the images. The filtering of level $l+1$ is done using the corresponding image at level l: For each pixel at level $l+1$, its value is, for instance, equal to the sum of four pixels at level l divided by 4. The locations of these four pixels are such that each is centered at a quadrant of a square of $s \times s$, where s is equal to $2l$ at level l.

(2) Assume $Q_0 = (1,0,...,1,0)^T$ and Z_0 which denotes an approximate DEM and can be for instance a plain represented by a grid of $m \times n$ surface points with the same heights.

(3) Assume the matrices Σ, Φ_Z, Ψ_Z, Σ_Z, Φ_Q, Ψ_Q, Σ_Q using the a priori knowledge (Equations 27 and 30).

(4) Determine L_0 using Z_0 and the images at the top level of the pyramids (Equation 18).

(5) Determine the coefficient matrices U, A, B, and the vector of constants W in Equation 23 using L_0, Z_0, and Q_0 (Equations 19 and 21).

(6) Solve the optimizing problem (Equation 31) to adapt Z_0 and Q_0.

(7) Estimate and adapt the matrices Σ, Φ_Z, Ψ_Z, Σ_Z, Φ_Q, Ψ_Q, Σ_Q using parameter estimation technique.

(8) Return to the Step 4 and determine L_0 again, but using the images at a lower level of the pyramids and the new Z_0 and Q_0.

This iterative optimization procedure is finished after the base level of the pyramids is reached. At this level the surface is reconstructed from the original images.

Figure 4 shows a stereo pair of digital aerial images. They represent a piece of steep and rough wilderness with rock-debris. Each of them has 240 by 240 pixels. The image scale is about 1:10,000. This image material was also used to test the feature based and least-squares matching algorithms and is regarded as the hardest one within three selected projects (Hahn and Förstner, 1988).

Figure 5 shows the automatically generated DEM and its a posteriori accuracy. It contains 30 by 30 lattice points with 1 by 1 m² lattice size. All heights of the same lattice points (900 points) was also manually measured on the Planicomp C100 analytical measuring device as reference (Figure 6). The precision of the manual measurements is about 0.22m (≈ 0.14 ‰ of the flying height). Figure 7 illustrates the difference between the automatically and the manually measured DEM. This difference can be characterized by its mean (bias) and its standard derivation against the bias. Taking the a posteriori accuracy of the automatically measured DEM (Figure 5) into account, the results are

MEAN DIFF: -0.313m (bias),
SDEV: 0.207m (≈ 0.13‰ of the flying height),

Figure 5. The automatically derived DEM and its *a posteriori* accuracy.

Figure 6. The manually measured DEM.

where the precision of the automatically derived DEM is about the same as observed by the operator.

Finally, we look at the image pair "House" (Figure 8), which is one of the 12 image pairs for the test on image

Figure 7. The difference between the automatically and manually measured DEM.

matching of Working Group III/4 of the International Society for Photogrammetry and Remote Sensing (Gülch, 1988). This image pair has been classified by the test organizer into the group of high complexity for image matching, as it contains almost all problems, including discontinuities, occlusions, shadows, and corruptions. Each image has a size of 240 by 240 pixels and the image scale is about 1:3,000. In Figure 9, we show the computed surface field by means of a perspective view and a contour map. Using this example, we want to demonstrate that the approach presented in this paper for surface reconstruction is able to deal with such problems as discontinuities by iteratively estimating and adapting the matrices Σ_z in Equation 31. Of course, the result shown in Figure 9 is not the final one, but it provides important guidance for finding physical discontinuities of the surface. In this case, the final surface should be reconstructed by integrating the semantic information obtained from image analysis as our *a priori* knowledge. The approach presented in this paper could be used as a theoretical framework for integrating different kinds of *a priori* knowledge.

Conclusion

In many practical problem solving situations such as digital image analysis and interpretation, the available information is incomplete or inexact and is inadequate to support the desired sorts of logical inferences. Probabilistic reasoning methods allow us to use uncertain or probabilistic knowledge and information in ways that take the uncertainty into account and help us accumulate evidence for hypotheses in a fair way. They are appropriate tools for making "just" decisions and they provide techniques that help to minimize risk in making decisions.

In this work, we concentrate our attention on the problem of digital photogrammetric object reconstruction and formulate it generally as a problem of image inversion and decision-making in an uncertain environment. Based on MAP criterion, we have introduced a theoretical framework for
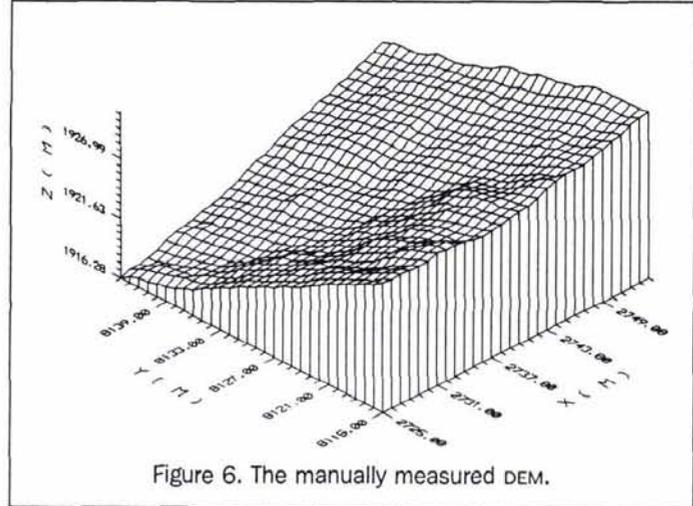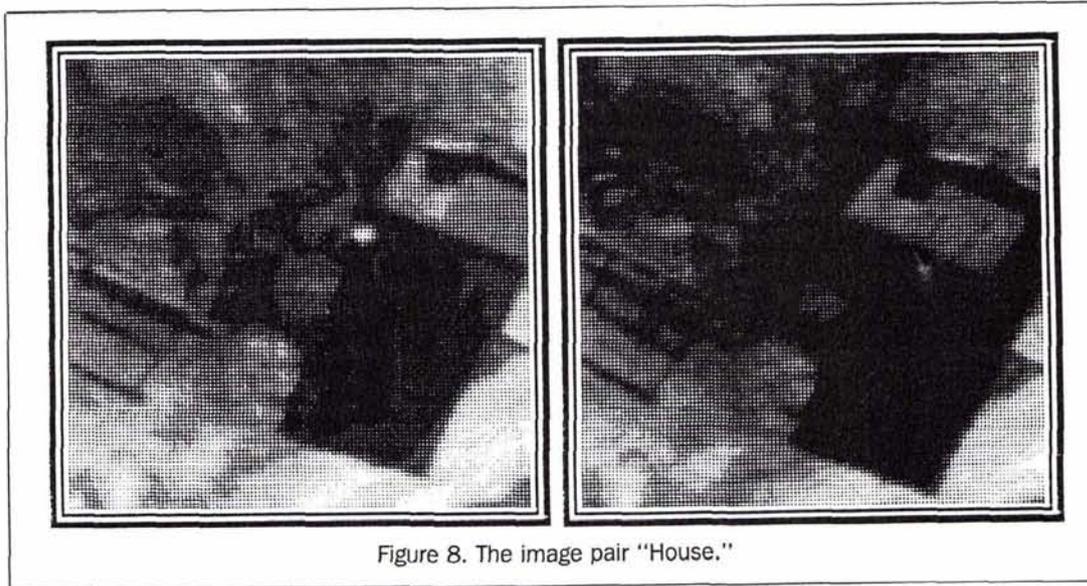
Figure 8. The image pair "House."
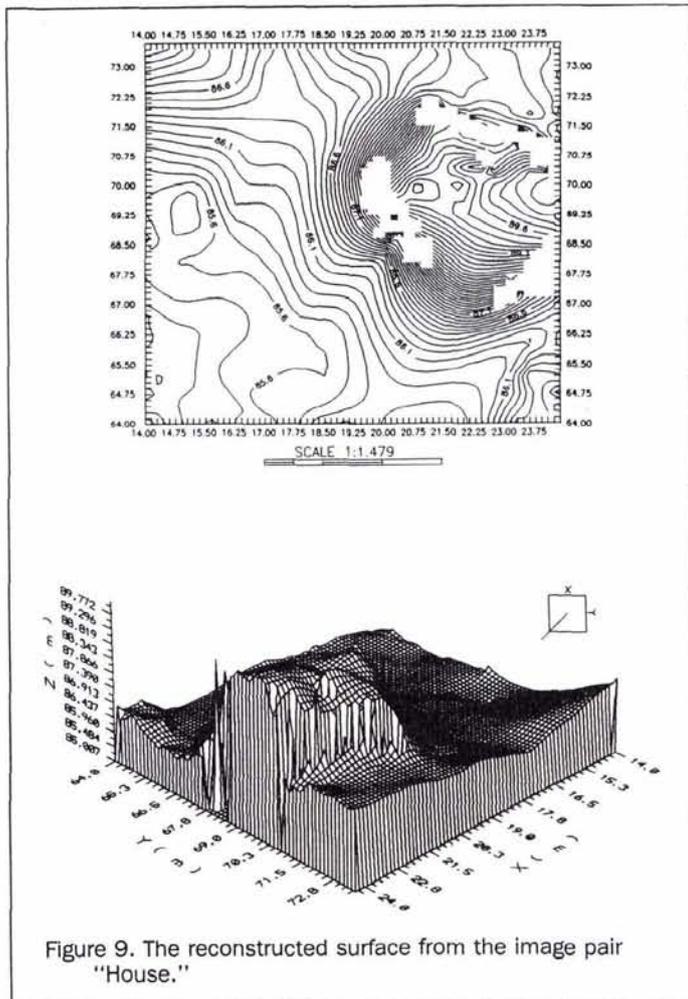
Figure 9. The reconstructed surface from the image pair "House."

dealing with *so* called ill-posed inverse problems. We have shown how surface reconstruction from images can be solved under this theoretical basis as an application example. Of *course*, there are many problems in digital photogrammetry which can also be solved using the framework given here. So, among the goals of future work will be (1) the investigation of the bias/variance dilemma during the inversion, (2) the introduction of a learning mechanism to improve and adapt our *a priori* knowledge during the inverse process, (3) the application of neural network technology to develop parallel algorithms for solving the optimizing problems mentioned above, and (4) the extension of the application range of the approach.

References

Ackermann, F., and Y.-J. Zheng, 1990. Inverse and Ill-posed Problems in Photogrammetric Surface Reconstruction, *Intern. Archives of Photgrammetry and Remote Sensing*, Vol. 28, Part 3/2, pp. 22–47.

Chou, P. B., and C. M. Brown, 1988. Multimodal Reconstruction and Segmentation with Markov Random Fields and HCF Optimization, *Proc. Image Understanding Workshop*, San Mateo, California, pp. 214–221.

Geiger, D., and F. Girosi, 1991. Parallel and Deterministic Algorithms from MRF's: Surface Reconstruction, *IEEE PAMI*, Vol. 13, No. 5.

Geman, S., and D. Geman, 1984. Stochastic Relaxation, Gibbs Distribution and Bayesian Restoration of images, *IEEE PAMI*, Vol. 6, pp. 721–741.

Gülch, E., 1988. Results of Test on Image Matching of ISPRS WG III/4, *Intern. Archives of Photogrammetry and Remote Sensing*, Vol. 27.

Hadamard, J., 1923. *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, Connecticut.

Hahn, M., and W. Förstner, 1988. On the Applicability of a Feature Based and a Least Squares Matching Algorithm for DEM Aquisition, *Intern. Archives of Photogrammetry and Remote Sensing*, Vol. 27.

Poggio, T., and V. Torre, 1984. Ill-posed Problems and Regularization Analysis in Early Vision, *Proc. Image Understanding Workshop*, DARPA.

Slater, P. N., 1983. Photographic Systems for Remote Sensing, Manual of Remote Sensing, Second Edition, Vol. I, American Society for Photogrammetry and Remote Sensing, pp. 232–236.

Tarantola, A., 1987. *Inverse Problem Theory*, Elsevier, New York.

Zheng, Y.-J., 1990. *Inverse und schlecht gestellte Probleme in der photogrammetrischen Objekt-Rekonstruktion*, Inaugural-Dissertation, Fakultät für Vermessungswesen, Universität Stuttgart, Stuttgart.

————, 1992. Inductive Inference and Inverse Problems in Computer Vision, *Robust Computer Vision* (Förstner and Ruwiedel, Eds.), Wichmann, Karlsruhe, pp. 320–350.

Zheng, Y.-J., and M. Hahn, 1990. Surface Reconstruction from Images in the Presence of Discontinuities, Occlusions and Deformations, *Intern. Archives of Photgrammetry and Remote Sensing*, Vol. 28, Part 3/2, pp. 1121–1144.

FORTHCOMING ARTICLES

Kirk F. Adkins and *Carolyn J. Merry*, Accuracy Assessment of Elevation Data Sets Using the Global Positioning System.

Peter August, Joanne Michaud, Charles LaBash, and *Christopher Smith*, GPS for Environmental Applications: Accuracy and Precision of Locational Data.

David G. Barber, Mohammed E. Shokr, Richard A. Fernandes, Eric D. Soulis, Dean G. Flett, and *Ellsworth F. LeDrew*, A Comparison of Second-Order Classifiers for SAR Sea Ice Discrimination.

Pietro A. Brivio, Ilaria Doria, and *Eugenio Zilioli*, Aspects of Spatial Autocorrelation of Landsat TM Data for the Inventory of Waste-Disposal Sites in Rural Environments.

Russell G. Congalton and *Kass Green*, A Practical Look at the Sources of Confusion in Error Matrix Generation.

Liping Di and *Donald C. Rundquist*, A One-Step Algorithm for Correction and Calibration of AVHRR Level 1b Data.

Jean-Pierre Djamdji, Albert Bijaoui, and *Roger Maniere*, Geometrical Registration of Images: The Multiresolution Approach.

Peter Dreyer, Classification of Land Cover Using Optimized Neural Nets on SPOT Data.

John R. Dymond, An Improved Skidmore/Turner Classifier.

Maria Fiorella and *William J. Ripple*, Analysis of Conifer Forest Regeneration Using Landsat Thematic Mapper Data.

Giles M. Foody, Ordinal Level Classification of Sub-Pixel Tropical Forest Cover.

Tung Fung and *King-chong Chan*, Spatial Composition of Spectral Classes: A Structural Approach for Image Analysis of Heterogeneous Land-Use and Land-Cover Types.

Sucharita Gopal and *Curtis Woodcock*, Theory and Methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets.

Linda H. Graff and *E. Lynn Usery*, Automated Classification of Generic Terrain Features in Digital Elevation Models.

Edwin J. Green, William E. Strawderman, and *Teuvo M. Airola*, Assessing Classification Probabilities for Thematic Maps.

Olaf Hellwich and *Wolfgang Faig*, Graph-Based Feature Matching Using Descriptive and Relational Parameters.

B. R. Hunt, T. W. Ryan, and *E. A. Gifford*, Hough Transform Extraction of Cartographic Calibraion Marks from Aerial Photography.

Katarina Johnsson, Segment-Based Land-Use Classification from SPOT Satellite Data.

P. Manavalan, P. Sathyanath, and *G. L. Rajegowda*, Digital Image Analysis Techniques to Estimate Waterspread for Capacity Evaluations of Reservoirs.

Chris L. Lauver and *Jerry L. Whistler*, A Hierarchical Classification of Landsat TM Imagery to Identify Natural Grassland Areas and Rare Species Habitat.

Oğuz Müftüoğlu, A Data Reduction Approach Using the Collinearity Model from Non-Metric Photography.

M. Nazim, SOM Equations with Modifications to Estimate Longitude from Ascending Node, and Constants for the Clarke and Everest Spheroids.

J. Olaleye and *W. Faig*, An Evaluation of Three M-Processor Designs for a Digital Photogrammetric System.

Michael B. Smith and *Andrej Vidmar*, Data Set Derivation for GIS-Based Urban Hydrological Modeling.

Paul A. Sorensen and *David P. Lanter*, Two Algorithms for Determining Partial Visibility and Reducing Data Structure Induced Error in Viewshed Analysis.

I. Tannous and *B. Pikeroen*, Parametric Modeling of Spaceborne SAR Image Geometry Application: SEASAT/SPOT Image Registration.

J. A. Thomasson, C. W. Bennett, B. D. Jackson, and *M. P. Mailander*, Differentiating Bottomland Tree Species with Multispectral Videography.

E. Lynn Usery, Virtual Stereo Display Techniques for Three-Dimensional Geographic Data.

Erratum

The units of measure in Table 1 of the article, "Estimating Surface Reflectance and Albedo from Landsat-5 Thematic Mapper over Rugged Terrain," by Claude R. Duguay and Ellsworth F. Le Drew (*PE&RS*, May 1992, page 552) should read $mWcm^{-2}sr^{-1}\mu m^{-1}$, not $Wm^{-2}sr^{-1}\mu m^{-1}$.