A New Algorithm for Surface Determination Based on Wavelets and its Practical Application

Jaan-Rong Tsay

Abstract

A new wavelets-based algorithm, FAST Vision (facets stereo vision), is presented for an automatic and simultaneous determination of an object surface and its ortho image. Two families of orthogonal and C^1-continuous (continuously differentiable) object gray-value models, called "S-D-model" and "S-model," respectively, were developed from the basic concept of multiresolution spaces. Both models establish two families of very simple gradient operators and enable FAST Vision to do a very high resolution representation of an object surface and a fast solution of a very large system of normal equations. Test results using digitized aerial images at a scale of 1:4000 show that FAST Vision is capable of a fast, highly resolved, reliable, and precise determination of an object surface in large windows and with rigorous error computations. The very high resolution of 2 by 2 pixels per height facet (0.12 by 0.12 m² in these tests) was obtained with the S-model in practical tests. The precision of the determined object surface was ± 0.02 to 0.06 m, i.e., 0.2 to 0.6 pixels or 0.03 to 0.1 ‰ of the flying height above ground, when compared with the control data measured by an operator on a Wild AC3 analytical stereoplotter. These figures correspond well with the natural roughness of the Earth's surface in the chosen test area.

Introduction to FAST Vision

FAST Vision (facets stereo vision) is a method for performing digital terrain reconstruction and ortho image computation, where digital stereo images are the basic observations. Its basic concepts lie in the inverse process of image formation, called "image inversion." The first draft of FAST Vision was proposed publicly in 1987 (Wrobel, 1987) and, since then, new or extended models and algorithms for an operational and multifunctional FAST Vision have been presented (Wrobel, 1987; Weisensee, 1992; Kempa, 1995; Tsay, 1996).

In the process of image formation, an object signal $G(X,Y)$ is projected from object space into image space and an image signal $\hat{G}'(x',y')$ is then formed in the image plane. On the other hand, in image inversion, the image signal $\hat{G}'(x',y')$ is backwards projected from image space into object space. Both processes can be generally described mathematically as follows:

$$\text{image formation: } \hat{G}'(x', y') \xleftarrow{(\mathbf{T}')^{-1}} G(X, Y) \qquad (1)$$

$$\text{image inversion: } \hat{G}'(x', y') \xrightarrow{\mathbf{T}'} G(X, Y) \qquad (2)$$

where $\hat{G}'(x',y')$ is the gray value of an image point with image coordinates (x',y'), $G(X,Y)$ is the gray value of an object point with horizontal object coordinates (X,Y) corresponding to $\hat{G}'(x',y')$, and $\mathbf{T}'$, $(\mathbf{T}')^{-1}$ are the transformation functions from $\hat{G}'(x', y')$ to $G(X, Y)$ and its inverse.

The relationship between the object gray value $G(X, Y)$ of a SURFEL (surface element) and the image gray values $\hat{G}'(x',y')$, $\hat{G}''(x'',y'')$, ... , of the corresponding pixels in stereo images B', B'', ... , respectively, is then given by the following *correspondence condition*:

$$\mathbf{T}'\left(\hat{G}'\left(x', y'\right)\right) = \mathbf{T}''\left(\hat{G}''\left(x'', y''\right)\right) = \ldots = G(X, Y) \qquad (3)$$

with the transformation functions $\mathbf{T}'$, $\mathbf{T}''$, ... , of images B', B'', ... , etc. The functions $\mathbf{T}'$, $\mathbf{T}''$, ... , can be strictly defined by complicated reflectance functions (e.g., see Weisensee (1992)), but the related computational costs would be very high. Instead of that, a very simple linear function has been used to approximate the functions $\mathbf{T}'$, $\mathbf{T}''$, ... , etc. This approximation is acceptable only in a small region, e.g., a computation window or even a height facet, and it has yielded great success.

In the following, one assumes that the interior and exterior orientation data of all stereo images B', B'', ... , are known. An image gray value $\hat{G}'(x',y')$ of a pixel i in the first image B' is linearly transformed and projected to the position X_i^0, Y_i^0 on an approximate object height surface. The corresponding object gray-value function is then expanded using a Taylor series in which the higher order terms have been omitted. The basic equation of FAST Vision is then derived (Wrobel, 1987): i.e.,

$$\begin{aligned} & g_0' + g_1' \cdot \left\{G'(x', y') + v_{G'}(x', y')\right\} \\ & = G(X_i, Y_i) \approx G^0(X_i^0, Y_i^0) \\ & + \frac{\partial G^0(X_i^0, Y_i^0)}{\partial X} dX_i + \frac{\partial G^0(X_i^0, Y_i^0)}{\partial Y} dY_i \\ & + dG^0(X, Y)_i \end{aligned} \qquad (4)$$

where g_1' and g_0' are the multiplication and addition parameter of the linear approximation function of the transformation function $\mathbf{T}'$ for image B'; $G'(x', y')$ and $v_{G'}(x', y')$ is the observation of the image gray value $\hat{G}'(x', y')$ of a pixel i and its random observation error, i.e.; $G'(x', y') + v_{G'}(x', y') = \hat{G}'(x', y')$; dX_i and dY_i are the corrections of the approximate values X_i^0, Y_i^0 of the horizontal object coordinates X_i, Y_i of the corre-

J.-R. Tsay is with the Department of Surveying Engineering, National Cheng Kung University, University Road 1, 70101 Tainan, Taiwan, R.O.C. (tsayjr@mail.ncku.edu.tw)

Photogrammetric Engineering & Remote Sensing, Vol. 64, No. 12, December 1998, pp. 1179–1188.

0099-1112/98/6412–1179$3.00/0
© 1998 American Society for Photogrammetry and Remote Sensing

sponding SURFEL i for $\hat{G}'(x', y')$; and $dG^0(X, Y)_i$ is the correction of the approximate object gray value $G^0(X_i^0, Y_i^0)$ of the SURFEL i that is caused by the corrections of the approximate values of the object gray-value model parameters.

In image inversion, the image signal $\hat{G}'(x', y')$ of a pixel i moves from image space along its perspective ray to object space, where the position and orientation of the ray in space is fixed. In the case of central perspective projection, one gets Equation 5 after differentiating the *collinearity equation* with known interior and exterior orientation data (Schwidefsky and Ackermann, 1976): i.e.,

$$dX_i = \frac{X_i^0 - X_0'}{Z_i^0 - Z_0'}\,dZ_i \qquad dY_i = \frac{Y_i^0 - Y_0'}{Z_i^0 - Z_0'}\,dZ_i \tag{5}$$

where X_0', Y_0', Z_0' are the object coordinates of the perspective center of image B'; and (X_i^0, Y_i^0, Z_i^0), (dX_i, dY_i, dZ_i) are approximate values of the object coordinates (X_i, Y_i, Z_i) of SURFEL i and their corrections.

By inserting Equation 5 into Equation 4, one obtains

$$\begin{aligned} & g_0' + g_1' \cdot \left\{G'(x', y') + v_{G'}(x', y')\right\} \\ & \approx G^0(X_i^0, Y_i^0) + dG^0(X, Y)_i \\ & + \left\{\frac{\partial G^0(X_i^0, Y_i^0)}{\partial X} \cdot \frac{X_i^0 - X_0'}{Z_i^0 - Z_0'} + \frac{\partial G^0(X_i^0, Y_i^0)}{\partial Y} \cdot \frac{Y_i^0 - Y_0'}{Z_i^0 - Z_0'}\right\} \\ & \cdot dZ^0(X, Y)_i \end{aligned} \tag{6}$$

where $dZ^0(X, Y)_i$ equals dZ_i and is the correction of the approximate height value Z_i^0 of SURFEL i that is caused by the corrections of approximate height values on all grid points in the handled DTM (digital terrain model) and depends on SURFEL position.

Equation 6 shows that the image gray-value measurement $G'(x', y')$ is directly related to the geometric and radiometric object surface functions $Z(X,Y)$ and $G(X,Y)$. The corresponding Equations 4 through 6 for other stereo images B'', . . . , are also used in FAST Vision.

The Taylor-linearized Equation 6 for all pixels i have to be solved iteratively, where the difference between the approximate value of every unknown and its *true value* must be smaller than its *convergence radius*. A larger convergence radius can be obtained by enlarging the pixel size. Therefore, a multiresolution approach (the image pyramid method) is essential and very applicable for FAST Vision. For details, please see Kaiser *et al.* (1992a).

Furthermore, the function $G(X,Y)$ represented by bilinear interpolation in each facet is on each facet border generally only C^0-, but not C^1-continuous, so that a point on the facet border generally has two derivative values $\frac{\partial G}{\partial X}$ (or $\frac{\partial G}{\partial Y}$). This ambiguous property is a drawback of bilinear interpolation. Other interpolation functions, such as higher order splines, have continuous derivatives, but they are not used here because the related computational costs are higher and the available window size is smaller.

Appropriate wavelets enable FAST Vision to have the following superior characteristics: (1) an eligible continuity degree of a represented object gray-value model $G(X,Y)$; (2) orthogonality between any two model parameters of $G(X,Y)$, which means also a high computation speed and independent parameters; (3) better stability at larger *support* than the one in bilinear interpolation; (4) a better representation accuracy because of the wonderful ability of wavelets: *multiresolution approximation*; and (5) a larger computation window, which means that more image information is available for surface determination. The reasons why wavelets are used in this particular application are then obvious.

Two families of wavelets-based, C^1-continuous and orthogonal *S-D-model* and *S-model* were developed to describe $G(X,Y)$. Both models were derived from the basic concepts of the so-called *multiresolution spaces* in the theory of wavelets, where the compactly supported Daubechies orthogonal wavelets (Daubechies, 1994) are utilized.

Fundamental Concept of Multiresolution Spaces

The concept of multiresolution spaces, denoted by $\mathbf{V}_j$, $\forall_j \in \mathbb{Z}$, has been given in many references with modern mathematical notation and concepts that are often unfamiliar to most readers. Therefore, this concept will be depicted briefly in a *user-friendly* manner using a very simple example of a multifrequency signal function f (see Figure 1).

The signal f is, e.g., a continuous image or height signal and can be decomposed into two components: a trend signal $\mathbf{T}$ and a detail signal $\mathbf{D}$, i.e., $f = \mathbf{T} + \mathbf{D}$. Similarly, this detail signal $\mathbf{D}$ can also be decomposed into two components: a low frequency component $\mathbf{D}_{low}$ and the rest term $\mathbf{D}_{high}$, i.e., $\mathbf{D} = \mathbf{D}_{low} + \mathbf{D}_{high}$.

The summation of the trend signal $\mathbf{T}$ and the detail signal $\mathbf{D}_{low}$, denoted by $\mathbf{T}' = \mathbf{T} + \mathbf{D}_{low}$, is usually a better approximation of f than $\mathbf{T}$. $\mathbf{T}'$ thus presents a finer resolution approximation than does $\mathbf{T}$. The change from $\mathbf{T}$ to $\mathbf{T}'$ results in a resolution refinement, e.g., from a coarser resolution space $\mathbf{V}_j$ to a finer resolution space $\mathbf{V}_{j+1}$. The relationship between both resolution spaces $\mathbf{V}_j$ and $\mathbf{V}_{j+1}$ is represented graphically in Figure 2, where $\mathbf{V}_{j+1}$ is decomposed into two orthogonal components $\mathbf{V}_j$ and $\mathbf{O}_j$. This decomposition is denoted mathematically by $\mathbf{V}_j \oplus \mathbf{O}_j = \mathbf{V}_{j+1}$, where $\mathbf{O}_j$ is called the "detail space" and is an orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j+1}$. The detail space is often named the "wavelet space" or "difference space." The resolution space is also named the "scaling space" or "approximation space."

In these three spaces $\mathbf{V}_j$, $\mathbf{O}_j$, and $\mathbf{V}_{j+1}$, $j \in \mathbb{Z}$, there exist three groups of orthogonal basis functions ϕ_{jk}, ψ_{jk}, and $\phi_{j+1,k}$, $\forall k \in \mathbb{Z}$, respectively. One can use these basis functions to extract the signals $\mathbf{T}$, $\mathbf{D}_{low}$, and $\mathbf{T}'$ from the original input signal f, respectively. If the index j becomes larger, the approximation space $\mathbf{V}_j$ offers more signal information about f, and $\mathbf{T}$ is a better approximation of f with a finer resolution. That presents us the basic concept of the so-called MRA (multiresolution analysis).

The space $\mathbf{V}_{j+1}$ can be decomposed into the spaces $\mathbf{V}_j$ and $\mathbf{O}_j$. Conversely, one can use these two spaces $\mathbf{V}_j$ and $\mathbf{O}_j$ to reconstruct the space $\mathbf{V}_{j+1}$. Similarly, the signal $\mathbf{T}'$ in the space $\mathbf{V}_{j+1}$ can be decomposed into the signals $\mathbf{T}$ and $\mathbf{D}_{low}$ in the spaces $\mathbf{V}_j$ and $\mathbf{O}_j$, respectively. One can also reconstruct the

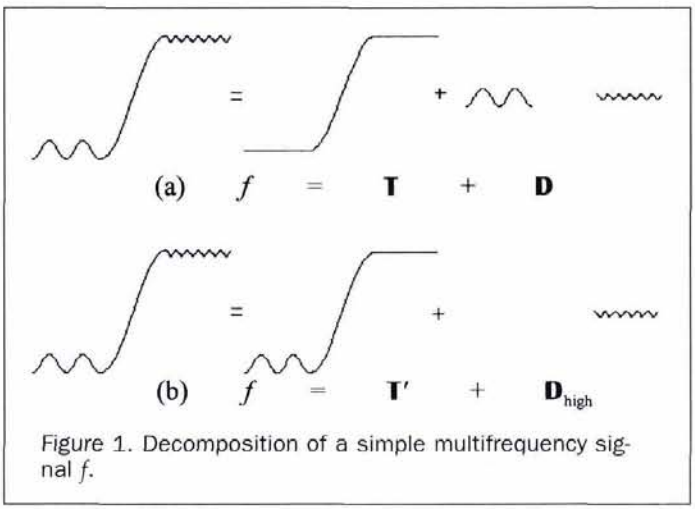

Figure 1. Decomposition of a simple multifrequency signal f.

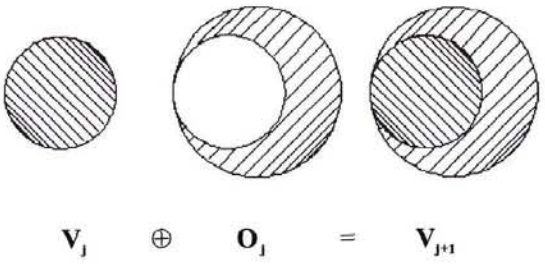

Figure 2. Graphical representation of the equation $\mathbf{V}_j \oplus \mathbf{O}_j = \mathbf{V}_{j+1}$.

signal $\mathbf{T}'$ using the signals $\mathbf{T}$ and $\mathbf{D}_{low}$ (Figure 3). In the case of discrete raster image data, both mean image decomposition and image reconstruction.

In practice, discrete raster image data with a specified resolution give FAST Vision the information about real ortho image function of object surface also in a specified resolution space, e.g., $\mathbf{V}_{j+1}$. These data represent a discrete approximation of the ortho image function in space $\mathbf{V}_{j+1}$. This approximation can be represented using the orthogonal basis functions in the spaces $\mathbf{V}_j$ and $\mathbf{O}_j$. From that idea, a new interpolation model, called the S-D-model, is derived.

If the approximation of the ortho image function in the space $\mathbf{V}_{j+1}$ is a relatively smooth function, its detail component in the space $\mathbf{O}_j$ will be close or equal to zero. In this case, the approximation in the space $\mathbf{V}_{j+1}$ is almost or completely the same as its smooth component in the space $\mathbf{V}_j$ that can be represented using the orthogonal basis functions in the space $\mathbf{V}_j$. From this idea, one gets a new approximation model called the S-model.

In other words, the S-D-model is suitable for image data which are rich in texture and lacking in image noise. However, the S-model is appropriate for a relatively smooth image signal with noise.

A New Interpolation Model: The S-D-Model

The S-D-model describes a signal transformation between the spaces $\mathbf{V}_{j+1}$ and $(\mathbf{V}_j, \mathbf{O}_j)$ (Figures 2 and 3). It is a new interpolation model among some given grid points. If reference data $\mathbf{G}_i$ on grid points with coordinates $\mathbf{X}_i$, $i = 1\ (1)\ 2n$, $n \in \mathbb{N}$, are given, their smooth and detail components $\mathbf{S}_k$, $\mathbf{D}_k$, $k = 1\ (1)\ n$, can be determined as follows (Tsay, 1996, pp. 53–55):

$$\mathbf{S}_k = \sum_{m=0}^{2N-1} \mathbf{h}_m \cdot \mathbf{G}_{2k+m-1} \tag{7}$$

$$\mathbf{D}_k = \sum_{m=0}^{2N-1} (-1)^m \mathbf{h}_{(2N-1)-m} \cdot \mathbf{G}_{2k+m-1} \tag{8}$$

where $\mathbf{h}_m$, $m = 0\ (1)\ 2N-1$, are the known filter coefficients (low-pass filter) for the compactly supported Daubechies orthogonal wavelets (see Daubechies (1994)); N is the order of Daubechies-wavelets; and $N \geq 3$ is used here to present a C^1-continuous function.

Then, a new interpolation function $G(X)$, called the "S-D-model," is defined as

$$G(X) = \sum_k \{h(x + 1 - 2k) \cdot S_k + q(x + 1 - 2k) \cdot D_k\} \tag{9}$$

where

$X_1 \leq X \leq X_{2n}$, $\forall\ X \in \mathbb{R}$;

$x = \dfrac{X - X_1}{\Delta X_G} + 1$, $x \in \mathbb{R}$;

X_i is the coordinate of the i-th grid point, $i = 1\ (1)\ 2n$ (i.e., the object coordinate of the i-th pixel of ortho image in FAST Vision);

$\Delta X_G = X_{i+1} - X_i$, $i = 1\ (1)\ 2n-1$;

$h(x + 1 - 2k) = 2^{-1/2} \int_{-\infty}^{+\infty} \phi\left(\frac{u}{2}\right) \cdot \phi(u - (x + 1 - 2k))du$;

$q(x + 1 - 2k) = 2^{-1/2} \int_{-\infty}^{+\infty} \psi\left(\frac{u}{2} - (N-1)\right) \cdot \phi(u - (x + 1 - 2k))du$; and

ϕ and ψ are the Daubechies orthogonal scaling and wavelet functions derived from a minimum phase filter (Daubechies, 1994).

Equation 9 defines a C^1-continuous curve that passes through all reference points G_i, $i = 1\ (1)\ 2n$.

Similarly, for the two-dimensional case, if reference data $G_{i,l}$ on grid points with coordinates (X_i, Y_l), $i = 1\ (1)\ 2n$, $l = 1\ (1)\ 2m$, are given, one can use, e.g., the following two-dimensional wavelet decomposition to determine their smooth components S_{jk} and detail components D^X_{jk}, D^Y_{jk}, and D^{XY}_{jk} in the X-, Y-, and XY-directions, respectively, with $j = 1\ (1)\ n$, $k = 1\ (1)\ m$: i.e.,

$$S_{jk} = \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} h_p h_q G_{2j+p-1,2k+q-1};$$

$$D^Y_{jk} = \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} (-1)^q h_p h_{(2N-1)-q} G_{2j+p-1,2k+q-1}$$

$$D^X_{jk} = \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} (-1)^p h_q h_{(2N-1)-p} G_{2j+p-1,2k+q-1};$$

$$D^{XY}_{jk} = \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} (-1)^{p+q} h_{(2N-1)-p} h_{(2N-1)-q} G_{2j+p-1,2k+q-1}$$

$$\forall_{p,q} \in \mathbb{N} \tag{10}$$

Then, the two-dimensional S-D-model $G(X,Y)$ is defined by

$$G(X,Y) = \sum_j \sum_k \{h_x h_y \cdot S_{jk} + q_x h_y \cdot D^X_{jk} + h_x q_y \cdot D^Y_{jk} + q_x q_y \cdot D^{XY}_{jk}\} \tag{11}$$

where

$h_x = h(x + 1 - 2j)$; $q_x = q(x + 1 - 2j)$;

$h_y = h(y + 1 - 2k)$; and $q_y = q(y + 1 - 2k)$.

Equation 11 calculates a C^1-continuous surface that passes through all reference points $G_{i,j}$, $i = 1\ (1)\ 2n$, $l = 1\ (1)\ 2m$.

New Symmetrical Gradient Operators

Equation 6 shows clearly that the object gray-value gradients $\frac{\partial G}{\partial X}$ and $\frac{\partial G}{\partial Y}$ are of great importance for FAST Vision and that they must be computed. For that, one obtains the following gradient function by differentiating Equation 9 for X: i.e.,

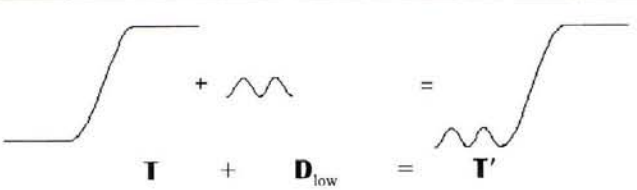

Figure 3. Signal decomposition $\mathbf{T}' \rightarrow \mathbf{T} + \mathbf{D}_{low}$; Signal reconstruction $\mathbf{T} + \mathbf{D}_{low} \rightarrow \mathbf{T}'$.

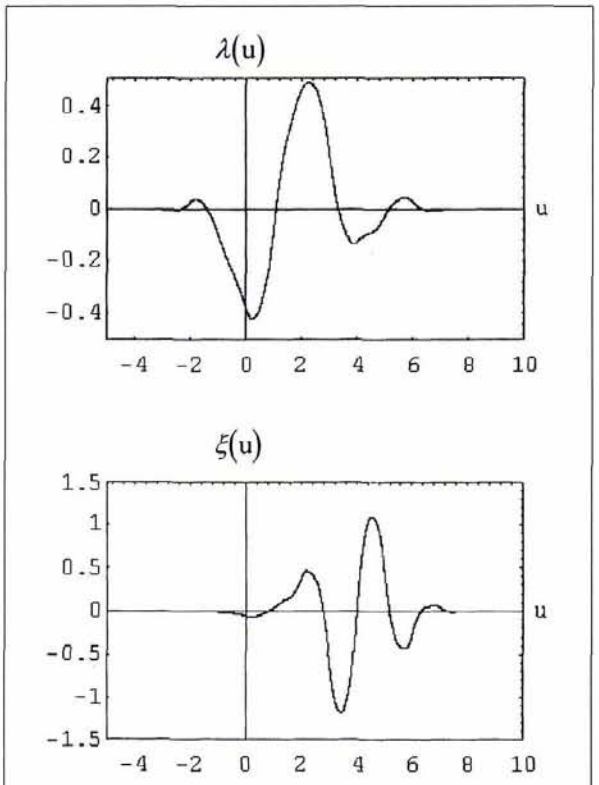

Figure 4. The function graphs $\lambda(u)$,$\xi(u)$ with $N = 3$ and $\Delta u = 2^{-10}$.

$$\frac{\partial G}{\partial X} = \frac{-\sqrt{2}}{\Delta X_G} \sum_k \{\lambda(u) \cdot S_k + \xi(u) \cdot D_k\} \tag{12}$$

where

$$u = x + 1 - 2k,$$

$$\lambda(u) = 2^{-1} \int_{-\infty}^{+\infty} \phi\left(\frac{t+u}{2}\right) \cdot \phi'(t)dt, \text{ and}$$

$$\xi(u) = 2^{-1} \int_{-\infty}^{+\infty} \psi\left(\frac{t+u}{2} - (N-1)\right) \cdot \phi'(t)dt.$$

There is no closed-form analytic formula for the functions ϕ and ψ, so that the usual conventional procedures to compute their function graphs, derivatives, integrals, etc., do not apply. Nevertheless, one can use some algorithms to compute them with arbitrarily high precision. For example, both functions λ and ξ can be calculated as follows:

$$\lambda(u) = \lim_{\Delta t \to 0} \sum_{m=-\infty}^{+\infty} 2^{-1} \phi\left(\frac{t_m + u}{2}\right) \cdot \phi'(t_m)\Delta t \tag{13}$$

$$\xi(u) = \lim_{\Delta t \to 0} \sum_{m=-\infty}^{+\infty} 2^{-1} \psi\left(\frac{t_m + u}{2} - (N-1)\right) \cdot \phi'(t_m)\Delta t \tag{14}$$

where $\Delta t = t_{m+1} - t_m$, $\forall m \in \mathbb{Z}$.

It is obvious that the scaling function ϕ, the wavelet function ψ, and the derivative ϕ' must be determined before one can compute the functions λ and ξ. For that purpose, several methods are available to compute the scaling function ϕ, e.g., the so-called "cascade algorithm" (Daubechies and Lagarias, 1991; Daubechies, 1994), Strang's method (Strang, 1989; Chui, 1992), the inverse Fourier Transform (Daubechies, 1994; Kaiser, 1994), and the method of cumulants (Kaiser, 1994). In fact, Strang's method is the best one because it is simple, quick, and accurate (Chui, 1992).

Once the scaling function ϕ is determined, the corresponding wavelet function ψ can be computed using the so-called *dilation equation* (Strang, 1989; Chui, 1992; Daubechies, 1994).

The derivative function of first order ϕ' exists and is a continuous function only in the case of $N \geq 3$. In the appendix, an algorithm is presented to compute the derivatives of the k-th order, i.e., $\phi^{(k)}$. One can use that algorithm with $k = 1$ to compute the function ϕ'.

Now, both functions λ and ξ can be computed using Equations 13 and 14. For example, Figure 4 shows the function graphs $\lambda(u)$ and $\xi(u)$ with $N = 3$ and $\Delta u = 2^{-10}$.

On a grid point $X = X_i$, $\forall i \in \mathbb{N}$, Equation 12 is identical to

$$\left(\frac{\partial G}{\partial X}\right)_i = \left.\frac{\partial G}{\partial X}\right|_{x=x_i} = \frac{-\sqrt{2}}{\Delta X_G} \sum_k \{\lambda(n) \cdot S_k + \xi(n) \cdot D_k\} \tag{15}$$

where

$n = 2 - 2N\,(2)\,4N - 4$, if i is an odd number,
$n = 3 - 2N\,(2)\,4N - 3$, if i is an even number, and
$k = (i + 1 - n)/2$.

Inserting Equations 7 and 8 into Equation 15, one obtains

$$\left(\frac{\partial G}{\partial X}\right)_i = \frac{1}{\Delta X_G} \sum_m G_{i+m} \cdot \beta_m \tag{16}$$

where

$m = 4 - 4N\,(1)\,4N - 3$, if i is an odd number,
$m = 3 - 4N\,(1)\,4N - 4$, if i is an even number, and
$\beta_m = -\sqrt{2}\sum_n \{\lambda(n) \cdot h_{m+n} + \xi(n) \cdot (-1)^{m+n} \cdot h_{(2N-1)-(m+n)}\}$.

For example, Table 1 shows the elements β_m with $N = 3$ and $m = -8\,(1)\,9$ or $m = -9\,(1)\,8$ for odd or even index i, respectively. It is evidently a new symmetrical gradient operator. Each of its elements has the same values in both cases (odd or even index i). On the other hand, there are only eight non-zero elements in this gradient operator. The other elements are equal to zero because of the characteristics of Daubechies-wavelets and their filter coefficients h_m, e.g., $\sum_m h_m \cdot h_{m+2n} = \delta_{0,n}$, where $\delta_{0,n}$ is the so-called Kronecker delta.

TABLE 1. A NEW GRADIENT OPERATOR DERIVED FROM THE S-D-MODEL WITH N=3 (SEE EQUATION 16)

	i is odd	i is even
m	β_m	β_m
−9	---	0
−8	0	0
−7	0	0
−6	0	0
−5	0	0
−4	−0.000342	−0.000342
−3	−0.014612	−0.014612
−2	0.145205	0.145205
−1	−0.745205	−0.745205
0	0	0
1	0.745206	0.745206
2	−0.145206	−0.145206
3	0.014612	0.014612
4	0.000342	0.000342
5	0	0
6	0	0
7	0	0
8	0	0
9	0	---

Similarly, one can get the derivative functions $\frac{\partial G}{\partial X}, \frac{\partial G}{\partial Y}$ of the two-dimensional S-D-model $G(X,Y)$ by differentiating Equation 11 for X and Y, respectively. Then, one gets the gradients on grid points (X_i, Y_l), $\forall i,l \in \mathbb{N}$: i.e.,

$$\left(\frac{\partial G}{\partial X_i}\right) = \frac{-\sqrt{2}}{\Delta X_G} \sum_l \sum_n \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} \{\lambda_u h_p + (-1)^p \xi_u h_{(2N-1)-p}\} \cdot \{h_n h_q + (-1)^{n+q} h_{(2N-1)-n} h_{(2N-1)-q}\} \cdot G_{2j+p-1,2k+q-1} \quad (17)$$

$$\left(\frac{\partial G}{\partial Y_l}\right) = \frac{-\sqrt{2}}{\Delta Y_G} \sum_m \sum_k \sum_{p=0}^{2N-1} \sum_{q=0}^{2N-1} \{\lambda_u h_q + (-1)^q \xi_u h_{(2N-1)-q}\} \cdot \{h_m h_p + (-1)^{m+p} h_{(2N-1)-m} h_{(2N-1)-p}\} \cdot G_{2j+p-1,2k+q-1}$$

The characteristic $\sum_m h_m \cdot h_{m+2n} = \delta_{0,n}$ results in many elements in Equation 17 being equal to zero. Therefore, Equation 17 reduces to

$$\left(\frac{\partial G}{\partial X_i}\right) = \frac{1}{\Delta X_G} \sum_n G_{j+n,l} \cdot \beta_n \text{ and } \left(\frac{\partial G}{\partial Y_l}\right) = \frac{1}{\Delta Y_G} \sum_n G_{i,l+n} \cdot \beta_n \quad (18)$$

The above-mentioned S-D-model is a wavelets-based new interpolation model among all given grid points (G_i or $G_{i,l}$). It yields a new family of symmetrical gradient operators and thus provides an alternative for interpolation and gradient computation. In the following, we present another available model for FAST Vision that is a new approximation model among all given grid points. It is called the S-model.

A New Approximation Model: The S-Model

One assumes that some grid point data G_i, $\forall i \in \mathbb{N}$ are given. They represent a discrete approximation of a real continuous signal function $G(X)$ with a certain resolution, e.g., in a resolution space $\mathbf{V}_{j+1}$. The approximation in $\mathbf{V}_{j+1}$ can be simply represented by its smooth component in $\mathbf{V}_j$ if its detail component function in $\mathbf{O}_j$ is close or equal to zero. The basic idea of the S-model was drawn thereupon.

The S-model $G(X)$ describes a best approximation $A_jG(X)$ of $G(X)$ in space $\mathbf{V}_j$, where the "best" is defined by a quadratic minimal norm. For the detailed derivation, please see Tsay (1996, pp. 85–86). The final formula for the one-dimensional S-model is defined as follows:

$$G(X) = A_jG(X) = \sum_k S_k \cdot \phi\left(\frac{1}{2} \cdot \left(\frac{X - X_1}{\Delta X_G} + 1\right) - k + \frac{\tau + J}{2}\right) \quad (19)$$

where S_k, $\forall k \in \mathbb{N}$ are computed by Equation 7 using the given grid point data G_i, $\forall i \in \mathbb{N}$; ϕ is the Daubechies scaling function with the support $[0, 2N - 1]$; X_1 is the coordinate of the first grid point G_1; ΔX_G is the interval of the given grid points G_i, $\forall i \in \mathbb{N}$; $\tau = \frac{1}{\sqrt{2}} \sum_{n=0}^{2N-1} h_n \cdot n$; and $h_J = \text{MAX}(h_n, \forall n)$, e.g., $J = 1$ for $N = 3$ or $N = 4$.

The two-dimensional S-model $G(X,Y)$ is defined by the product of two one-dimensional S-models $G(X)$ and $G(Y)$: i.e.,

$$G(X,Y) = G(X)\,G(Y) = \sum_j \sum_k S_{jk} \cdot \phi\left(\frac{1}{2} \cdot \left(\frac{X - X_1}{\Delta X_G} + 1\right) - j + \frac{\tau + J}{2}\right) \cdot \phi\left(\frac{1}{2} \cdot \left(\frac{Y - Y_1}{\Delta Y_G} + 1\right) - k + \frac{\tau + J}{2}\right) \quad (20)$$

The scaling function ϕ is C^1-continuous only when its order N is larger than 2. One chooses $N = 3$ to keep the computation costs as small as possible and also to have a C^1-continuous function $G(X)$ or $G(X,Y)$. Both models (Equations 19 and 20) with $N = 3$ have a support that contains 5 or 5 by 5 S-grid point data S_k or S_{jk}, respectively. However, their discrete models with $X = X_i$ and $Y = Y_l$, that are used in surface determination by means of the indirect method of FAST Vision, only need 3 or 3 by 3 S-grid point data, because the function values $\phi((\tau + n)/2)$ are exactly equal to zero for $n > 5$, i.e., for $n = 6$ (1) 9. Therefore, the cost to compute the function values $G(X_i)$ or $G(X_j, Y_l)$ becomes minimal.

New Asymmetrical Gradient Operators

The necessary gradient function for FAST Vision is determined by differentiating Equation 19 for X. In the case of $N = 3$, its function value on a grid point $X = X_i$ is equal to

$$\left(\frac{\partial G}{\partial X_i}\right) = \frac{1}{\Delta X_G} \sum_m a_m \cdot S_{\frac{i-m}{2}} \quad (21)$$

where

$m = -1$ (2) 7 if i is an odd number;
$m = 0$ (2) 8 if i is an even number; and
$a_m = \frac{1}{2} \phi'\left(\frac{\tau + m + 1}{2}\right)$ (see Table 2).

Similarly, the gradient functions $\partial G/\partial X$ and $\partial G/\partial Y$ of the two-dimensional S-model are determined, and their function values at a grid point with $X = X_i$ and $Y = Y_l$ are computed using Equation 22 for the case of $N = 3$: i.e.,

$$\left(\frac{\partial G}{\partial X_i}\right) = \frac{\sqrt{2}}{\Delta X_G} \sum_m \sum_q \left(a_m h_{q+1} \cdot S_{\frac{i-m}{2}, \frac{l-q}{2}}\right) \text{ and } \left(\frac{\partial G}{\partial Y_l}\right) = \frac{\sqrt{2}}{\Delta Y_G} \sum_p \sum_n \left(h_{p+1} a_n \cdot S_{\frac{i-p}{2}, \frac{l-n}{2}}\right) \quad (22)$$

where

$m = -1$ (2) 7	and	$p = -1$ (2) 3	if	i is odd;
$m = 0$ (2) 8	and	$p = 0$ (2) 4	if	i is even;
$n = -1$ (2) 7	and	$q = -1$ (2) 3	if	l is odd; and
$n = 0$ (2) 8	and	$q = 0$ (2) 4	if	l is even.

In the S-model, all S-values S_k or S_{jk} are discrete grid point data, where their grid size is two times that of the grid size of their original grid point data G_i or G_{jl}. That means that each S-grid contains 2 or 2 by 2 G-grids (G-SURFELS), respectively.

Furthermore, the S-D- or S-model is used here to describe an object ortho image function in the indirect method of FAST Vision, where all observations of image gray values are located on predefined grid points. In this case, all parameters of the S-D- or S-model, i.e., (S_k, D_k or S_{jk}, D^X_{jk}, D^Y_{jk}, D^{XY}_{jk}) or (S_k or S_{jk}), $\forall j, k \in \mathbb{N}$, are orthogonal to each other, so that the solution and/or inversion of a very large system of normal equations becomes simpler than the one in any non-orthogonal model, e.g., bilinear interpolation.

Normal Equations

The normal equations in FAST Vision are derived by the well-known least-squares adjustment (Mikhail and Ackermann, 1976) and are defined by

$$\begin{bmatrix} N_{GG} & N_{GZ} & N_{Gg'} \\ N^T_{GZ} & N_{ZZ} & N_{Zg'} \\ N^T_{Gg'} & N^T_{Zg'} & N_{g'g'} \end{bmatrix} \cdot \begin{bmatrix} X_G \\ X_Z \\ X_{g'} \end{bmatrix} = \begin{bmatrix} b_G \\ b_Z \\ b_{g'} \end{bmatrix}$$

$$\mathbf{N} \qquad \mathbf{X} = \mathbf{b} \quad (23)$$

where X_G are unknown corrections of the approximate values of object gray-value model parameters; X_Z are unknown corrections of the approximate values of object height model

TABLE 2. A NEW GRADIENT OPERATOR DERIVED FROM THE S-MODEL WITH $N=3$

i is odd		i is even	
m	a_m	m	a_m
−1	0.70894	0	0.76885
1	−1.08080	2	−1.01089
3	0.53443	4	0.21525
5	−0.16220	6	0.02679
7	−0.00036	8	0.00000

parameters; X_g are unknown corrections of the approximate values of multiplication and addition parameters dg'_1, dg'_0 of the linear transformation function (see Equation 4); and N_{GG}, N_{GZ}, N_{Gg}, . . . , are sub-matrices of the normal equation matrix **N**.

The characteristic of Daubechies filter coefficients h_m, i.e., $\sum_m h_m \cdot h_{m+2n} = \delta_{0,n}$, results in the simplest form of sub-matrix $\mathbf{N}_{GG}$: i.e.,

$$\mathbf{N}_{GG} = \lambda \cdot \mathbf{E} \tag{24}$$

where

$\lambda = \sum_{b=1}^{B} 1/(g'^0_1)^2_b$ for the S-D-model or $\lambda = 4\sum_{b=1}^{B} 1/(g'^0_1)^2_b$ for the S-model;
B is the number of the used (stereo) images, $B \in \mathbb{N}$, $b \in \mathbb{N}$;
$(g'^0_1)_b$ is the approximate value of the parameter g'_1 for the b-th image; and
E is a unit matrix.

Compared with a band matrix $\mathbf{N}_{GG}$ established by any non-orthogonal (e.g., bilinear) object gray-value model, this sub-matrix $\mathbf{N}_{GG}$ needs the very least core memory to store the unique element λ. Therefore, the economized core memory is available for surface determination in a larger window. That is a wonderful property for FAST Vision.

A New Algorithm for Object Surface Determination Using the Method of FAST Vision

Basic Idea

In general, there exist many local image regions with homogeneous gray values in real images, e.g., large-scale aerial images. This kind of real image is, of course, not ideal for image matching or object surface determination. Object surface determination using these kinds of real images should be done in a possibly larger window in order to enlarge the chance to find in that window the sufficient and necessary image information for a successful and accurate matching. Therefore, the orthogonal S- and S-D-model are presented to enable FAST Vision to do an object surface determination in a possibly larger window.

Furthermore, a sufficiently accurate initial DTM plays an important role for a quick FAST Vision. In principle, the image pyramid method is the simplest one to get a good initial DTM. At each image pyramid level, the window size should be chosen as large as possible. This gives two advantages: (1) each computed DTM at the upper pyramid level could be used as the initial DTM for object surface determination in more windows at the lower level, so that the total computation cost/time for a very large area can become cheaper/shorter; and (2) FAST Vision can be more reliable and accurate at every pyramid level, because more image features with large gradients of image gray values would generally appear in a larger window.

In addition, a rigorous adjustment is not necessary for the computations of FAST Vision at the upper pyramid levels. Instead of that, a group adjustment is very suitable because it needs less core memory and computation time. With that, a sufficiently accurate initial DTM for object surface determination at the lower pyramid levels can be determined as quickly as possible.

At the lowest pyramid level, the approximate values of all parameters of FAST Vision can be corrected in the initial iterations by a group adjustment. These parameters can then be exactly determined by a rigorous adjustment, if a rigorous computation of the covariance matrix for all parameters is required.

This algorithm is outlined in Figure 5.

Algorithm

A minimal number of grid points is necessary, because the determination of border points in a window is worse than the determination of inner points. On the other hand, the window size cannot be too large, because it would create a very large system of normal equations and would result in high computation costs. Therefore, suitable algorithms are indispensable from the beginning, e.g., the so-called scan technique, for object surface determination in a large project area.

For the present, different scan techniques are applicable, e.g., the one proposed in Kaiser *et al.* (1992b). One wants to use here the relatively simple and practically suitable scan technique proposed by Kempa (1995). In that method, a constant number of grid points per scan window is applied which is equal to the number of grid points used in a window in the bottom level of the image pyramid. The overlap between any two neighboring windows is so defined that the computed heights on all grid points of a certain number of grid lines on the border of a window is not used as a part of

Define the project area and the number of image pyramid levels.

Compute the image pyramid data.

Define approximation of DTM: horizontal plane or better data if available.

Compute a better DTM in project area by the image pyramid method:

From the top to the bottom level:

- Define each window as large as possible.
- For the upper levels:
  - Group adjustment.
- For the bottom level:
  - *Group adjustment in the beginning iterations.*
  - Strict adjustment in the following iterations.
  - Output the computed DTM and ortho image

Figure 5. A new algorithm for object surface determination in a large project area using the method of FAST Vision.

the final entire DTM in a given project area. Only the other computed heights are available. Mean values of available heights on the common points, which are located in two neighboring windows, are used as final results.

The whole project area is then computed by moving the scan window of a constant size. First of all, the project area and the number of image pyramid levels are defined. The image pyramid data are then computed by way of image filtering using the low-pass filter proposed in Meer *et al.* (1987). The regularization (Tsay, 1996, pp. 105–108) must be done to solve the ill-posed problem in object surface determination using real, generally non-ideal images.

In the above-mentioned algorithm, the indirect method of FAST Vision is used, where the S-D- or S-model with $N = 3$ is utilized as the C^1-continuous and orthogonal object gray-value model and the parametric cubic convolution function is applied as the weighting function of a C^1-continuous object height model (Tsay, 1996, pp. 21–30).

In this paper, two models are presented (S- and S-D), but they were not compared in our experiments. The experimental comparison may be done later. To verify the superior applicability of wavelets for object surface determination using this algorithm, only the S-model was tested and its test results are shown and analyzed briefly in the following section.

Test Results Using Large-Scale Aerial Images

The chosen test area (Figure 6) is located in a rural region in Walddorf-Haeslach in the southwest of Germany. It has a natural Earth surface with vegetation. The aerial photos used in the test had a mean scale of 1:4000 and were taken with a Zeiss RMK A 15/23 camera (15cm focal length, 23cm by 23cm photo format). Altogether, four photos were selected as test data. They were taken from two cross flight lines. The overlap between the two stereo photos in each flight line was about 60 percent. These four photos were digitized with a Zeiss PS1 photo scanner using a pixel size of 15 by 15 μm and 8 bits in gray scale. All four digitized images show that they really have weaker contrast than their original analog photos. Some images do have very weak texture, e.g., the image of window C shown in Figure 6.

Three test windows were chosen in that test area: window A with relatively good texture, window B with a medium quality of texture in a scarp region of a highway, and window C with weak texture. Thus, object surface determination in different kinds of windows could be tested and analyzed. The Earth's surface in the windows is covered essentially with grass before spring. Some individual small bushes exist in window B. In addition, the height differences in A and C windows and the B window are 0.8 m and 2.5 m, respectively. Compared with other small-scale (aerial) images or images in other regions, the digitized images of the selected test area do have relatively very weak texture and, therefore, give a proper touchstone for automatic surface determination using aerial images.

The image pyramid algorithm is utilized in all tests to get sufficiently accurate initial DTMs, where a horizontal plane is used as an initial DTM for each top level. The computed DTM on each level is then transferred to the next lower level. Generally speaking, the convergence radius of the FV (FAST Vision) method depends on the wavelength of the image signal and is equivalent to the pixel size, if the image signal is sampled at the Nyquist rate (Weisensee, 1992). Therefore, in order to get a successful and fast surface determination, four or five image pyramid levels were chosen for the A and C windows and the B window, respectively, which have a precision of initial heights of 1.5 pixels and 1.1 pixels for the fourth or fifth pyramid levels, respectively. This results in a precision of initial heights of 2‰ (1.2m) or 3‰ (1.8m) of the flight height.

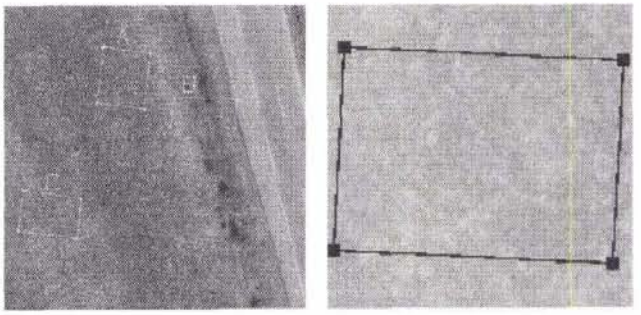

Figure 6. The test area (800 by 800 pixels, left) and window C (168 by 168 pixels, right) extracted from two digitized images with a 15- by 15-μm pixel size.

To check the accuracy of the DTM computed by FV, the DTM in each window with a height grid size of 0.5m by 0.5m was measured by an operator on a Wild AC3 analytical stereo plotter. The standard deviations $\sigma_{Z,OP}$ of those height measurements show clearly that window C does have the worst image texture among all three windows. The $\sigma_{Z,OP}$ values of the measured DTM in window C have a mean and maximum of 0.08m and 0.16m, respectively. In the following, only the analysis of the test results for window C are presented because the analyses for windows A and B (Tsay, 1996, pp. 122–133) are essentially the same as for window C.

The data below are given in Tables 3, 4, 5, and 6 to show the precision of every Z-value computed by FV, denoted by Z_{FV}, which are compared with the Z-values measured by an operator, denoted by Z_{OP}. In those tables, $\hat{\sigma}_0$ is the standard deviation of unit weight (units: gray value); $\bar{\sigma}_{Z,FV}$ and $\tilde{\sigma}_{Z,FV,max}$ are the mean and maximum of the standard deviations of the Z-values which are computed by FV; s_{FV}/s_{OP} is the *a posteriori* standard deviation for the Z-value $(Z_{FV})_i$ computed by FV and for the mean $(\bar{Z}_{OP})_i$ of the Z-values measured by operator on each grid point i, $i = 1$ (1) k, respectively, where k is the number of grid points; $\Delta Z \pm s_{\Delta Z}$ is the constant height difference between both DTM $(Z_{FV})_i$ and $(\bar{Z}_{OP})_i$, $i = 1$ (1) k, and its standard deviation; and $|dZ_r|_{max}$ is the maximum absolute value of the cleared height differences $(dZ_r)_i$, $i = 1$ (1) k, with $(dZ_r)_i = (Z_{FV})_i - (\bar{Z}_{OP})_i - \Delta Z$.

Table 3 shows the test results using two images, where each G- and Z-facet contains 2 by 2 and 4 by 4 pixels, respectively, denoted by P:G:Z = $1:2^2:4^2$. It shows that all computations converge in two or three iterations. From the upper to the lower level, the standard deviation $\hat{\sigma}_0$ becomes larger. The reason is obvious because the noise of image gray values is damped by the low-pass filter from the lower to the upper level. In addition, the computed standard deviations of heights by FV become smaller on lower levels with finer resolution. The *a posteriori* standard deviation s_{FV} of the Z-values computed by FV is smaller than or equal to 3 cm on the bottom level, which corresponds to 0.05‰ of the flight height or 0.3 pixel in image space. In those cases, each constant height difference ΔZ is not significant, i.e., $\Delta Z \sim 0$. The maximum absolute value of the height differences $|dZ_r|_{max}$ on the bottom level is equal to 13cm. Furthermore, 82.3 percent of the cleared height differences are located in the interval ±6 cm (0.1‰ of the flight height or 0.6 pixel). Figure 7 shows their distribution histogram.

Table 4 shows the test results using four images. Comparing it with Table 3 for two images, the precision using four images is better than the one using two images with regard to $\hat{\sigma}_0$ and the *a posteriori* standard deviation s_{FV} of the

TABLE 3. TEST RESULTS WITH THE FV METHOD USING TWO IMAGES, WINDOW C (WEAK TEXTURE)
Number of Pixels per Facet: P:G:Z = $1:2^2:4^2$
Facet size in first level of image pyramid: ΔX(P): ΔX(G): ΔX(Z) = 0.0625m : 0.125m : 0.25m

level of image pyramid	number of iterations	$\hat{\sigma}_0$ [gray value]	$\bar{\tilde{\sigma}}_{Z,FV}$ [m]	$\tilde{\sigma}_{Z,FV,max}$ [m]	s_{FV}/s_{OP} [m]	$\lvert dZ_i \rvert_{max}$ [m]	$\Delta Z \pm s_{\Delta Z}$ [m]
4	2	3.8	0.25	0.38	0.22/0.07	0.45	−0.04 ± 0.04
3	2	5.3	0.11	0.23	0.17/0.12	0.45	−0.03 ± 0.02
2	2	7.5	0.07	0.15	0.11/0.12	0.42	−0.02 ± 0.01
1.1	3	9.1	0.04	0.09	0.02/0.04	0.13	−0.00 ± 0.00
1							
1.2	3	10.3	0.05	0.10	0.03/0.04	0.13	−0.00 ± 0.00

Z-values computed by FV. On the first level, all *a posteriori* standard deviations s_{FV} are smaller than 3cm (0.05‰ of the flight height or 0.3 pixel in image space), where the constant height difference ΔZ is not significant.

Tables 5 and 6 show the test results using the very fine resolution of 2 by 2 pixels per Z-facet and Z-facet = G-facet, denoted by P:G:Z = $1:2^2:2^2$, that is the finest resolution if the S-model is used as the object gray-value model. They show evidently that the finest resolution of 2 by 2 pixels per Z-facet is available for FV using the presented algorithm. In that case, all computations converge in two or three iterations. The standard deviation $\hat{\sigma}_0$ becomes larger from the upper to the lower level. That is the same as the case with the coarser resolution of 4 by 4 pixels per Z-facet.

In the case of four images, the *a posteriori* standard deviation s_{FV} of the Z-values computed by FV on the first level is equal to 3 to 4cm, which corresponds to 0.05 to 0.07‰ of the flight height or 0.3 to 0.4 pixel in image space. This is smaller than the mean value $\bar{\tilde{\sigma}}_{Z,FV}$ (8 to 9cm) of the standard deviations of the Z-values computed from the covariance matrix of FV. This indicates that the Z-values can be more precisely determined than their precision indicator $\tilde{\sigma}_{Z,FV}$. This is an expected, so-called bridging effect of the regularization.

The related constant height difference ΔZ between FV and the operator is not significant, i.e., $\Delta Z \sim 0$. The root-mean-square value RMS (dZ) of the height difference dZ ($Z_{FV} - Z_{OP}$) is 4 to 6cm, which corresponds to 0.07 to 0.1‰ of the flight height or 0.4 to 0.6 pixel in image space. This precision is limited by the following two factors: (1) the accuracy of the interior and exterior orientation of the images used in these tests is 0.2 to 0.3 pixel and (2) the roughness of the Earth's surface in the test area is obviously in the same order as the height precision of less than 10cm.

These tests show that FV is capable of a precise object surface determination using large-scale aerial images. In addition, the presented algorithm can do an object surface determination using the finest resolution of 2 by 2 pixels per Z-facet (0.12 by 0.12 m in these tests). That is much finer than the resolution that is used or available in other methods, e.g., 30 by 30 pixels per Z-facet in Krzystek (1991), 20 by 20 pixels per Z-facet in Heipke (1990) and Holm (1994), 16 by 16 pixels per Z-facet in Diehl and Heipke (1992), and 5 by 5 pixels per Z-facet in Diehl (1994).

Conclusion and Outlook

FAST Vision is a method for object surface determination using digital stereo images. In fact, it can integrate different kinds of information or data to perform simultaneously and automatically image matching, digital terrain reconstruction, and ortho image computation. At present, only the geometrical and radiometrical information has been taken into account. The potential application of semantic information for surface determination has not yet been exhausted.

In order to do a fast and accurate surface determination in a large window, this paper presents an algorithm and two families of wavelets-based orthogonal and C^1-continuous object gray-value models (S- and S-D). Test results using real aerial images with a large image scale of 1:4000 show that the S-model makes a finest resolution of 2 by 2 pixels per height facet available, corresponding to 0.12 by 0.12 m in object space. This resolution is much finer than the one used or available in other methods. Thus, an accurate surface representation is possible. In addition, the S-D-model enables FV to have a new potential to use a resolution of 1 by 1 pixel per height facet, but the related aspects have not yet been studied.

Moreover, the test results show that the presented algorithm can perform a fast, highly resolved, reliable, and accurate object surface determination in a rather large window. The precision of object surface determination is 0.2 to 0.6 pixels, which has been estimated by comparing the DTM computed by FV with the DTM measured by an operator on a Wild AC3 analytical stereo plotter. This precision is limited by the residual errors in the orientation data of the used images and from the roughness of the natural Earth surface in the test area. Therefore, FV can be utilized in many practical applications about surface and/or volume determination, e.g., computer vision, topography, land management, Earth volume computation, route design, and so on.

FV also can estimate the variances and covariances of the related surface parameters. It is a general defect of the other methods proposed in the area of computer vision (e.g., see Fua and Leclerc (1995)) that don't have this capability. Furthermore, FV can do surface determination using multi-images. The test results show that the precision of surface determination can be better if more stereo images are used.

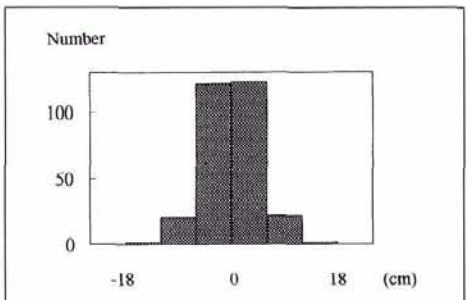

Figure 7. Histogram of the height differences dZ_i (= $(Z_{FV})_i - (Z_{OP})_i$ = $(Z_{FV})_i - \Delta Z$) on the bottom level in window C (weak texture) using two images and the facet size of P:G:Z = $1:2^2:4^2$.

TABLE 4. TEST RESULTS WITH THE FV METHOD USING FOUR IMAGES, WINDOW C (WEAK TEXTURE)
Number of Pixels per Facet: P:G:Z = $1:2^2:4^2$
Facet size in first level of image pyramid: ΔX(P): ΔX(G): ΔX(Z) = 0.0625m : 0.125m : 0.25m

level of image pyramid	number of iterations	$\hat{\sigma}_0$ [gray value]	$\bar{\tilde{\sigma}}_{Z,FV}$ [m]	$\tilde{\sigma}_{Z,FV,max}$ [m]	s_{FV}/s_{OP} [m]	$\lvert dZ_r \rvert_{max}$ [m]	$\Delta Z \pm s_{\Delta Z}$ [m]
4	2	3.7	0.21	0.35	0.21/0.08	0.45	−0.04 ± 0.04
3	2	4.9	0.09	0.19	0.16/0.14	0.45	−0.03 ± 0.02
2	2	6.6	0.06	0.12	0.10/0.13	0.44	−0.03 ± 0.01
1.1	3	8.8	0.04	0.09	0.02/0.04	0.18	0.01 ± 0.00
1							
1.2	3	9.7	0.04	0.09	0.03/0.05	0.18	−0.00 ± 0.00

TABLE 5. TEST RESULTS WITH THE FV METHOD USING TWO IMAGES, WINDOW C (WEAK TEXTURE) AND A HIGH RESOLUTION
Number of Pixels per Facet: P:G:Z = $1:2^2:2^2$
Facet size in first level of image pyramid: ΔX(P): ΔX(G): ΔX(Z) = 0.0625m : 0.125m : 0.125m

level of image pyramid	number of iterations	$\hat{\sigma}_0$ [gray value]	$\bar{\tilde{\sigma}}_{Z,FV}$ [m]	$\tilde{\sigma}_{Z,FV,max}$ [m]	s_{FV}/s_{OP} [m]	$\lvert dZ_r \rvert_{max}$ [m]	$\Delta Z \pm s_{\Delta Z}$ [m]	RMS(dZ) [m]	$\lvert dZ \rvert_{max}$ [m]
4	3	3.8	0.60	1.39	0.09/0.01	0.18	0.07 ± 0.01	0.11	0.25
3	2	4.6	0.20	0.52	0.07/0.03	0.17	0.07 ± 0.00	0.10	0.24
2	2	6.2	0.14	0.59	0.08/0.04	0.24	0.02 ± 0.01	0.09	0.27
1.1	2	7.4	0.08	0.49	0.03/0.03	0.13	0.08 ± 0.01	0.09	0.17
1.2	2	11.6	0.12	0.45	0.06/0.04	0.15	−0.06 ± 0.01	0.09	0.21
1									
1.3	2	10.6	0.10	0.66	0.05/0.04	0.16	−0.01 ± 0.01	0.07	0.16
1.4	3	10.4	0.10	0.33	0.05/0.04	0.17	0.02 ± 0.01	0.06	0.19

TABLE 6. TEST RESULTS WITH THE FV METHOD USING FOUR IMAGES, WINDOW C (WEAK TEXTURE) AND A HIGH RESOLUTION
Number of Pixels per Facet: P:G:Z = $1:2^2:2^2$
Facet size in first level of image pyramid: ΔX(P): ΔX(G): ΔX(Z) = 0.0625m : 0.125m : 0.125m

level of image pyramid	number of iterations	$\hat{\sigma}_0$ [gray value]	$\bar{\tilde{\sigma}}_{Z,FV}$ [m]	$\tilde{\sigma}_{Z,FV,max}$ [m]	s_{FV}/s_{OP} [m]	$\lvert dZ_r \rvert_{max}$ [m]	$\Delta Z \pm s_{\Delta Z}$ [m]	RMS(dZ) [m]	$\lvert dZ \rvert_{max}$ [m]
4	2	3.8	0.52	1.25	0.11/0.02	0.32	−0.02 ± 0.01	0.11	0.34
3	2	4.5	0.17	0.66	0.08/0.04	0.25	0.00 ± 0.01	0.09	0.25
2	2	5.9	0.11	0.46	0.06/0.04	0.24	−0.02 ± 0.01	0.07	0.26
1.1	3	8.5	0.09	0.41	0.04/0.04	0.13	−0.00 ± 0.01	0.05	0.13
1.2	2	10.1	0.09	0.31	0.04/0.04	0.12	−0.02 ± 0.01	0.06	0.14
1									
1.3	2	9.2	0.08	0.28	0.03/0.03	0.11	0.01 ± 0.01	0.05	0.12
1.4	3	10.3	0.09	0.31	0.03/0.03	0.13	−0.00 ± 0.01	0.04	0.12

There exists of course room for FV to be further improved, e.g., with appropriate algorithms/models to process/represent discontinuous surfaces, an optimal regularization, a strict reflectance model to describe more accurately the relationship between object and image gray values, geometric and radiometric calibration of the photo scanner or CCD camera, etc. One can expect that the precision of object surface determination can then be further improved after the above-mentioned improvements are carried out. On the other hand, the reconstruction of the Earth's surface is needed for topographic cartography. One should propose some solutions to extract the desired Earth surface from the reconstructed object surface that always includes anomalous surfaces, e.g., buildings and vegetation.

Appendix

An Algorithm to Compute the k-th Order Derivative $\phi^{(k)}$ of the Scaling Function

If the k-th order derivative function $\phi^{(k)}$ exists, one can use the dilation equation $\phi(x) = \sqrt{2}\sum_{n=0}^{2N-1} h_n \cdot \phi(2x - n)$ to derive a general condition equation for $\phi^{(k)}$ as follows:

$$\phi^{(k)}(x) = 2^k \sqrt{2}\sum_{n=0}^{2N-1} h_n \cdot \phi^{(k)}(2x - n),\ \forall x \in \mathbb{R}. \quad \text{(A1)}$$

On the other hand, there exists $\phi^{(k)}(x)=0$, $\forall x \in \mathbb{R}$, $x \le 0$ or $x \ge 2N - 1$. One can get $2N - 2$ linear equations using Equation A1 with $x = 1$ (1) $2N - 2$. They describe the conditions among the $2N - 2$ unknowns $\phi^{(k)}(1), \ldots, \phi^{(k)}(2N - 2)$. The rank of this system of linear equations is $2N - 3$, so that an additional independent condition among the unknowns $\phi^{(k)}(1), \ldots, \phi^{(k)}(2N - 2)$ must be found. For that, one can find the following condition from Dahmen and Micchelli (1990):

$$\sum_{j \in \mathbb{Z}} (-j)^r \phi^{(l)}(j) = l!\delta_{rl},\ 0 \le r \le l \quad \text{(A2)}$$

with

Kronecker delta δ_{rl}

$$= \begin{cases} 1 & ,r = l \\ 0 & ,r \ne l \end{cases} \quad \text{and} \quad l! = l(l-1)(l-2)\cdot\ldots\cdot 3\cdot 2\cdot 1.$$

For example, with $r = 1$ and $l = 1$, one can get

$$\sum_{j=1}^{2N-2} (-j)\phi'(j) = 1. \quad \text{(A3)}$$

Therefore, the unknowns $\phi^{(k)}(1), \ldots, \phi^{(k)}(2N-2)$ can then be determined by solving Equations A2 and A1, e.g., with $x = 1\ (1)\ 2N-3$. In the case of $k = 1$, one can get the derivative function values $\phi'(1), \ldots, \phi'(2N-2)$ after solving Equations A3 and A1, e.g., with $x = 1\ (1)\ 2N-3$. After that, one can compute all function values $\phi^{(k)}(x)$ on the so-called dilation points $x = \frac{n}{2^j}, \forall n,j \in \mathbb{Z}$ using Equation A1. If the resolution index j is large enough (e.g., $j > 9$), one can accurately compute $\phi^{(k)}(x)$ by linear interpolation between the corresponding two neighboring computation points.

References

Chui, C.K., 1992. *Wavelet Analysis and Its Applications, Vol. 2, Wavelets: A Tutorial in Theory and Applications*, Academic Press, Inc.

Dahmen, W., and C.A. Micchelli, 1990. Using the Refinement Equation for Evaluating Integrals of Wavelets, *Series A, School of Mathematics, Free University Berlin*, preprint No. A-9:1–14.

Daubechies, I., 1994. *Ten Lectures on Wavelets*, 3rd printing, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Daubechies, I., and J. Lagarias, 1991. Two-scale difference equations I. Existence and global regularity of solutions, *J. Math. Anal.*, 22: 1388–1410.

Diehl, H., 1994. Object based DTM generation by data of the MEOSS 3-line camera, *International Archives of Photogrammetry & Remote Sensing*, 30(Part 3/1):183–187.

Diehl, H., and C. Heipke, 1992. Surface reconstruction from data of digital line camera by means of object based image matching, *International Archives of Photogrammetry & Remote Sensing*, 29(Part B3):287–294.

Fua, P., and Y.G. Leclerc, 1995. Object-centered surface reconstruction: Combining multi-image stereo and shading, *International Journal of Computer Vision*, 16:35–56.

Heipke, C., 1990. Multiple Image Matching in Object Space, *International Archives of Photogrammetry & Remote Sensing*, 28(Part 3/2):294–302.

Holm, M., 1994. Global image matching and surface reconstruction in object space using three aerial images, *International Archives of Photogrammetry & Remote Sensing*, 30(Part 3/1):379–386.

Kaiser, B., M. Schmolla, and B.P. Wrobel, 1992a. Application of image pyramid for surface reconstruction with FAST Vision (=Facets Stereo Vision), *International Archives of Photogrammetry & Remote Sensing*, 29(Part B3):341–345.

Kaiser, B., J. Hausladen, J.R. Tsay, and B.P. Wrobel, 1992b. Application of FAST Vision for digital terrain model generation, *Proceedings, ISPRS Commission IV*, Washington, D.C., pp. 809–816.

Kaiser, G., 1994. *A Friendly Guide to Wavelets*, Birkerhaeuser, Boston/Basel/Berlin.

Kempa, M., 1995. *Hochaufgeloeste Oberflaechenbestimmung von Natursteinen und Orientierung von Bildern mit dem Facetten-Stereosehen*, Dissertation, TH Darmstadt, Institut für Photogrammetrie und Kartographie, Germany.

Krzystek, P., 1991. Fully automatic measurement of elevation models with Match-T, *43rd Photogrammetric Week*, Stuttgart, Germany, 15:203–214.

Meer, P., E.S. Baugher, and A. Rosenfeld, 1987. Frequency domain analysis and synthesis of image generating kernels, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):512–522.

Mikhail, E.M., and F. Ackermann, 1976. *Observations and Least Squares*, Thomas Y. Crowell Company, New York.

Schwidefsky, K., and F. Ackermann, 1976. *Photogrammetrie*, Second Edition, B.G. Teubner Verlag, Stuttgart, Germany.

Strang, G., 1989. Wavelets and dilation equations: A brief introduction, *Society for Industrial and Applied Mathematics, Review*, 31:614–627.

Tsay, J.-R., 1996. *Wavelets für das Facetten-Stereosehen*, Dissertation TH Darmstadt; Deutsche Geodätische Kommission, Series C, No. 454, Munich, Germany.

Weisensee, M., 1992. *Modelle und Algorithmen für das Facetten-Stereosehen*, Deutsche Geodätische Kommission, Series C, No. 374, Munich, Germany.

Wrobel, B.P., 1987. Digital image matching by facets using object space models, *SPIE*, 804:325–333.

(Received 26 September 1996; revised and accepted 24 February 1998; revised 16 March 1998)

REMINDER! If you received ASPRS BUCK$ vouchers for sponsoring a new member, please use them before December 31, 1998. After December 31, 1998, the vouchers will not be honored. Do not forget to write your name and your membership I.D. number on the back of the vouchers.

Each ASPRS BUCK$ voucher is worth $5 toward the purchase price of any publication in the ASPRS bookstore.

You can collect ASPRS BUCK$ and use them all at once or use them as you earn them!

How do you earn ASPRS BUCK$?
Every time you recruit a new ASPRS member have them complete the ASPRS membership application, then list your name as their sponsor.

All vouchers earned during 1998 must be used in 1998.