(2)

AD-A238 821

CSDL-R-2303

CONCEPTS FOR THE USE OF THREE-DIMENSIONAL DEPTH MAPS IN UNDERWATER MINE CLASSIFICATION

by

M. Bello

January 1991

Sponsored by
Defense Advanced Research Projects Agency
Naval Technology Office
ARPA Order No. 6261-12
Issued by DARPA/CMO under Contract MDA972-88-C-0057

The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

Approved for public release; distribution unlimited.
OASD(PA) 91-3133 7/1/91

91-05844

DRAPER LABORATORY

The Charles Stark Draper Laboratory, Inc.
555 Technology Square, Cambridge, Massachusetts 02139-3563

91 043

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT
4. PERFORMING ORGANIZATION REPORT NUMBER(S) CSDL-R-2303	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Draper Laboratory 6b. OFFICE SYMBOL (*If applicable*)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (*City, State, and Zip Code*) 555 Technology Square, Cambridge, MA 02139-3563	7b. ADDRESS (*City, State, and Zip Code*)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Advanced Research Projects Agency 8b. OFFICE SYMBOL (*If applicable*) DARPA (TTO)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA 972-88-C-0057
8c. ADDRESS (*City, State, and Zip Code*) 1400 Wilson Blvd. Arlington, VA 22209-2308	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (*Include Security Classification*)
Concepts For The Use Of Three-Dimensional Depth Maps In Underwater Mine Detection

12. PERSONAL AUTHOR(S)
Dr. Martin G. Bello

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (*Year, Month, Day*)	15. PAGE COUNT
	FROM _____ TO _____	January 1991	48

16. SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (*Continue on reverse if necessary and identify by block number*)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (*Continue on reverse if necessary and identify by block number*)

Concepts are examined for the use of three dimensional depth map information for the construction of feature inputs to neural network classifiers for underwater mines. Features derived from both low and high resolution depth map construction techniques are considered. Feature calculation concepts are suggested based on both deterministic and stochastic modeling of the bottom surface.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (*Include Area Code*) / 22c OFFICE SYMBOL

000F/589/ReportDocPage

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

CSDL-R-2303

CONCEPTS FOR THE USE OF THREE-DIMENSIONAL DEPTH MAPS IN UNDERWATER MINE CLASSIFICATION

by
M. Bello

January 1991

A-1

The Charles Stark Draper Laboratory, Inc.
Cambridge, Massachusetts 02139-3563

1.0 Introduction

The following report is an examination of concepts for the utilization of three-dimensional depth map information for the construction of features to be employed by neural network based classifiers in the classification of underwater mines. The evaluation of the discriminatory power of a subset of the identified three dimensional features for mine detection/classification, in the context of actual or simulated data, will be reported separately. The context that forms the basis for the current study envisions the availability of coregistered side-scan sonar-derived intensity imagery and three-dimensional depth map information at the same resolution. Hence, the three-dimensional related features to be described are to be viewed as augmenting features computed from side-scan sonar imagery.

The discussion that follows focuses separately in Sections 2 and 3 on the construction of three-dimensional features that use low-resolution and high-resolution depth map construction concepts, respectively, and associated with the use of telesounding and swath-bathymetry techniques, respectively. Finally, Section 4 summarizes the study's conclusions.

2.0 Low-Resolution Three-Dimensional Depth Map Utilization Concepts

The discussion that follows will first review the telesounder concept for the construction of depth maps. Next, the combined use of side-scan sonar data and height map data in the estimation of underwater mine-size-related features will be described.

The geometry underlying the telesounder depth map construction concept is pictured in Figure 2-1. Two acoustic receivers mounted on a tow body are positioned vertically with a separation Δ, and with the lower receiver also functioning as a transmitter. It is assumed that the height of receiver #2 above some "reference" bottom level, h_R, is known, and that the sea-bottom terrain height above the "reference" bottom at a ground range x, is denoted by $h_b(x)$. In addition, let $r_{s1}(x)$, $r_{s2}(x)$, denote the slant range from receivers #1 and #2, to the point $(x,h_b(x))$, respectively.

The telesounder concept (Ref. [1]) consists in forming the signal z(t) as

$$z(t) = |r_1(t) + r_2(t)| \tag{2-1}$$

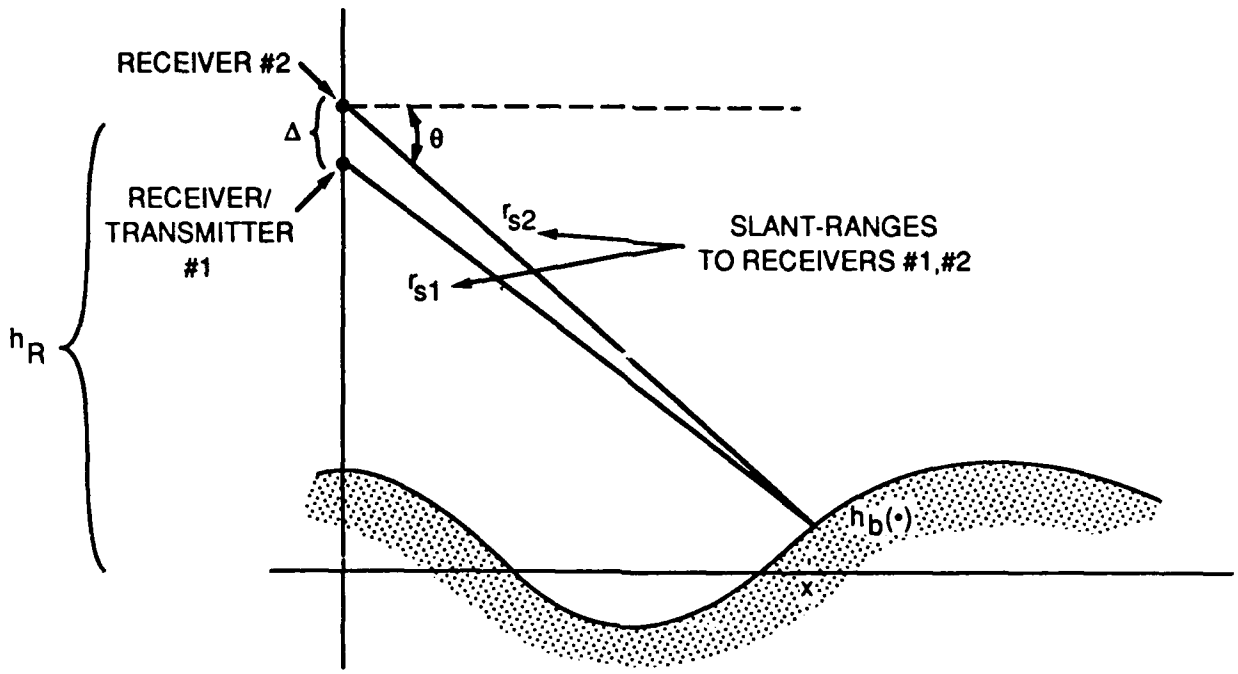

Figure 2-1. Telesounder depth map construction geometry.

where $r_1(\cdot)$, $r_2(\cdot)$ denote demodulated, complex signals associated with receivers #1 and #2, respectively. For the sake of simplicity, it is assumed here that reflections are only received from the specific bottom patch at $(x, h_b(x))$. Then, assuming a transmitted complex signal of the form

$$s(t) = U_T(t)\, e^{i\omega_o t} \tag{2-2}$$

$$U_T(t) \triangleq \begin{cases} 1 & 0 \le t \le T \\ 0 & \text{otherwise} \end{cases} \tag{2-3}$$

and assuming an idealized propagation process with no effect other than delay, it can be shown that over the time interval where the two received pulses interact, that the signal z(t) approximately includes a factor of the form $f(\theta\ (x))$ where

$$f(\theta) \triangleq \left|\cos\left(\frac{1}{2}\frac{\omega o\, \Delta}{v}\sin(\theta)\right)\right| \tag{2-4}$$

and v denotes the velocity of propagation of sound. Now, letting

$$\Delta = k\lambda \tag{2-5}$$

where λ denotes the acoustic wavelength and noting that

$$\omega_0 = \frac{2\pi v}{\lambda} \tag{2-6}$$

it can be seen that

$$f(\theta) = |\cos(\pi k \sin\theta)| \tag{2-7}$$

The conclusion to be obtained from the above discussion is that specific θ's corresponding to the maxima of (2-7) will be associated with specific maxima in the "interference" signal z(t). Hence, at those specific θ values, the depth may be deduced by the use of the approximate equation

$$h_b(x) \cong h_R - \Delta - r_{s1}(x)\sin\theta(x) \tag{2-8}$$

where $r_{s1}(x)$ is deduced from the time location of the appropriate z(•) maxima.

From (2-7) it can be shown that maxima of $f(\theta)$ occur at discrete θ_n's defined by

$$\theta_n = \sin^{-1}\left(\frac{n}{k}\right) n = 0 \ldots k \tag{2-9}$$

For illustrative purposes, assuring that the bottom in Figure 2-1 were flat, then the ground ranges associated with the $f(\theta)$ maxima would be specified by

$$x_n = h_R \cot(\theta_n) \tag{2-10}$$

Table 2-1 lists θ_n, x_n, $(x_{n-1} - x_n)$ values for a specific case in which

$$\begin{cases} h_R = 40\text{ m} \\ k = 10 \end{cases} \tag{2-11) / (2-12}$$

Table 2-1. θ values and ground ranges associated with interference signal maxima.

n	θ_n	x_n	$x_{n-1} - x_n$
0	0	∞	
1	0.10017 rad	397.98 m	
2	0.20136 rad	195.96 m	202.02 m
3	0.3469 rad	127.19 m	68.77 m
4	0.41152 rad	91.65 m	35.54 m
5	0.52360 rad	69.28 m	22.37 m
6	0.64350 rad	53.33 m	15.95 m
7	0.77540 rad	40.81 m	12.52 m
8	0.92730 rad	29.99 m	10.82 m
9	1.1198 rad	19.37 m	10.62 m
10	π/2 rad	0.0 m	19.37 m

Hence, for the specific case considered in Table 2-1, over the first 200 m of ground range, the separation between x_n values varies from 10.62 m up to 68.77 m.

In (Ref. [2]), a variety of cylindrical-shaped U.S. mines are discussed, and their dimensions are listed. The ranges of diameter and length dimensions are listed in Table 2-2. The above discussion of the telesounder-based depth estimation technique is sufficient to suggest that detailed three-dimensional surface and shape information characterizing individual mines is not likely to be obtainable from gridded, telesounder-derived depth maps. Hence, the discussion that follows suggests a way of employing low-resolution, telesounder-derived depth map information to estimate mine-size-related features.

Table 2-2. Cylindrical mine dimensions.

	Minimum	Maximum
Diameter	0.38 m	0.73 m
Length	1.8 m	4.1 m

When human sonar operators interpret side-scan sonar imagery, they look for highlight regions with associated shadows. The concept suggested here is to make use of side-scan sonar-derived slant-range and shadow extent information, combined

with approximate bottom map information, $\hat{h}_b(x)$, in order to estimate object diameter. In Figures 2-2 and 2-3, the geometry of the object diameter calculation problem for the case of an object on the bottom, or in the water column, respectively, is depicted. The notations r_s, ℓ_s will denote the slant range to the object "highlight" center and the slant-range extent of the shadow, respectively. In the case of Figure 2-3, g_s denotes the slant-range extent of the gap between the center of the "highlight" and the beginning of the shadow. The notations x_h, and $\{x_s, x_{su}, x_{s\ell}\}$ denote ground ranges associated with the "highlight" center and shadow edges, respectively. Finally, h_o and h_{wc} denote the diameter of the object and the water column height, respectively.

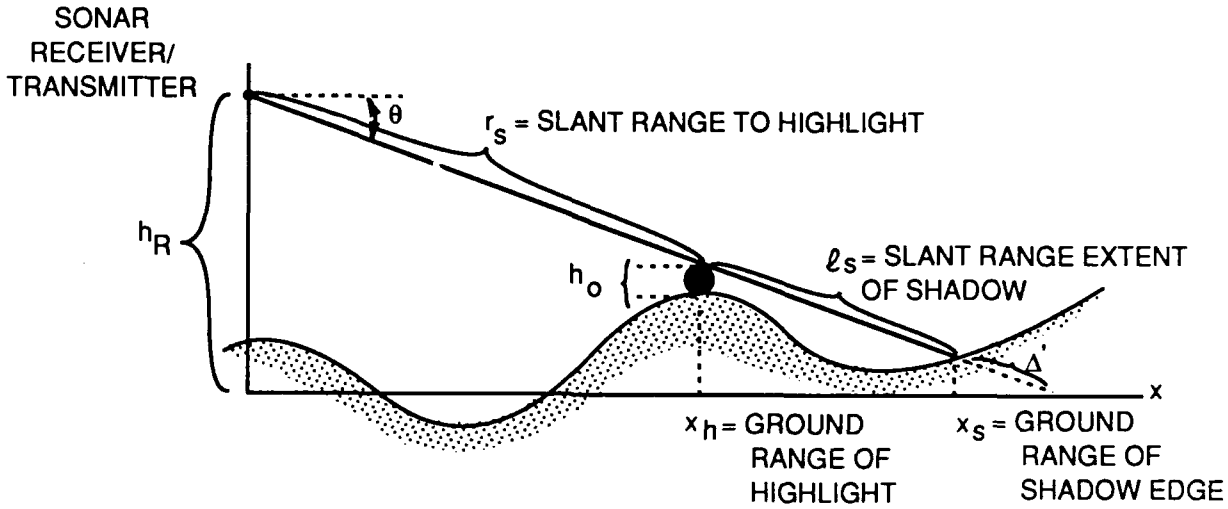

Figure 2-2. Geometry of object diameter calculation for case of object on the bottom.

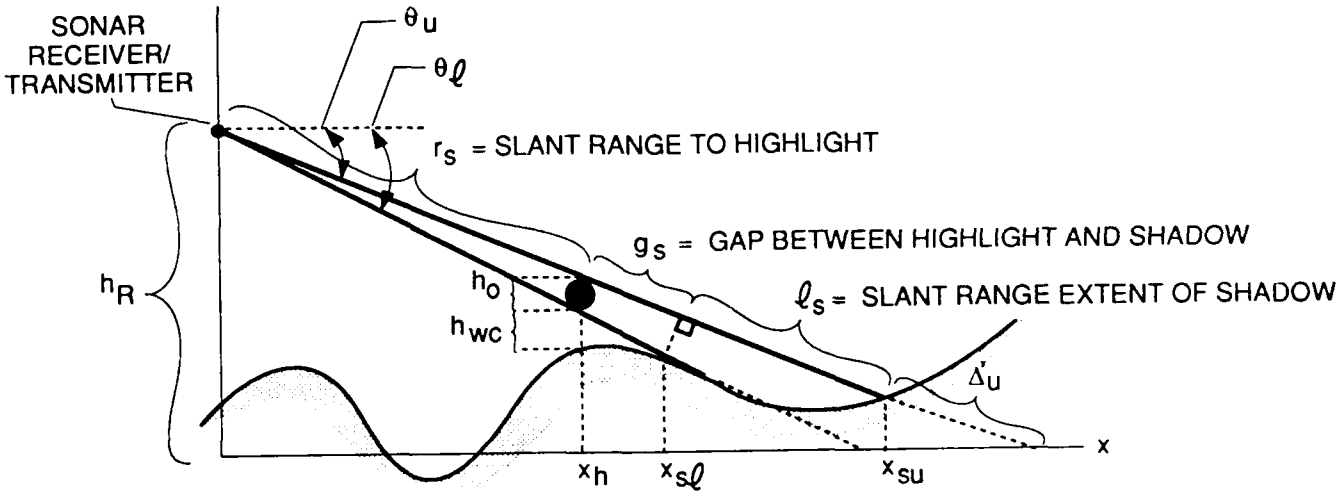

Figure 2-3. Geometry of object diameter calculation for case of object in the water column.

The relationships that allow the calculation of h_o, h_{wc} are derived from equating the ratios of corresponding sides of similar triangles. In the case of an object on the bottom,

$$\frac{h_R}{(r_s + \ell_s + \Delta')} = \frac{(h_o + h_b(x_h))}{(\ell_s + \Delta')} \tag{2-13}$$

where

$$\Delta' = \frac{h_b(x_s)}{\sin(\theta)} \tag{2-14}$$

and

$$\theta = \cos^{-1}\left(\frac{x_s - x_h}{\ell_s}\right) \tag{2-15}$$

and therefore implying that

$$h_o = \frac{(\ell_s + \Delta')}{(r_s + \ell_s + \Delta')} h_R - h_b(x_h) \tag{2-16}$$

For an object in the water column depicted in Figure 2-3, two sets of similar triangles yield the relations

$$\frac{h_R}{(r_s + g_s + \Delta'_\ell)} = \frac{(h_{wc} + h_b(x_h))}{(g_s + \Delta'_\ell)} \tag{2-17}$$

$$\frac{h_R}{(r_s + g_s + \ell_s + \Delta'_u)} = \frac{(h_o + h_{wc} + h_b(x_h))}{(g_s + \ell_s + \Delta'_u)} \tag{2-18}$$

where

$$\Delta'_u = \frac{h_b(x_{su})}{\sin(\theta_u)} \tag{2-19}$$

$$\Delta'_\ell = \frac{h_b(x_{s\ell})}{\sin(\theta_\ell)} \tag{2-20}$$

and

$$\theta_u = \cos^{-1}\left(\frac{x_{su} - x_h}{g_s + \ell_s}\right) \tag{2-21}$$

$$\theta_\ell = \cos^{-1}\left(\frac{x_\ell - x_h}{g_s}\right) \tag{2-22}$$

Relations (2-17) - (2-12) therefore imply that

$$h_{wc} = \frac{(g_s + \Delta'_\ell)}{(r_s + g_s + \Delta'_\ell)} h_R - h_b(x_h) \tag{2-23}$$

$$h_o = \frac{(g_s + \ell_s + \Delta'_u)}{(r_s + g_s + \ell_s + \Delta'_u)} h_R - h_{wc} - h_b(x_h) \tag{2-24}$$

The above discussion only addresses the geometric relationships that could allow the construction of h_o estimates, $\hat{h}_o$. We next consider how some of the specific quantities required in the above relations could be obtained in practice. By computing an empirical histogram of side-scan sonar intensity image values over a given strip, it is possible to determine thresholds $\tau_{\varepsilon\ell}$, $\tau_{\varepsilon h}$ so that ε percent of intensity values lie below and above these respective values; then by employing $\tau_{\varepsilon\ell}$, $\tau_{\varepsilon h}$, it is possible to form candidate shadow and highlight pixel sets. In each case, these sets will consist of a finite number of connected components whose number can be reduced by a size constraint, i.e., requiring each connected component to have over m pixels. Next, by introducing further geometric constraints on the relative location of highlight blobs and shadow blobs, a candidate set of pairs (highlight blob, shadow blob) can be determined. Finally, by making use of our knowledge of the sets (highlight blob, shadow blob) and counting pixels in the cross-range direction corresponding to the center of each highlight blob, and knowing the slant-range resolution per pixel, it should be possible to determine r_s, g_s, ℓ_s values.

The remaining quantities required to support h_o calculation consist in the ground ranges (x_h, x_s) or (x_h, $x_{s\ell}$, x_{su}), and corresponding bottom height values, $h_b(\bullet)$. It is assumed here that approximate bottom height information, $\hat{h}_b(\bullet)$, is available at the same resolution of the side-scan sonar imagery, through the application of some gridding algorithm to discrete, scattered telesounder-derived height measurements. In addition, note that ground-range information is assumed to be obtainable through the

use of geometric corrections that make use of $\hat{h}_b(\bullet)$. For example, consider a reflector at a height h_a above the bottom and at a slant range r_{sr}; then, the associated estimated ground range, $\hat{x}_r$, can be viewed as the solution of the equation

$$\left[h_R - h_a - \hat{h}_b(x)\right]^2 + x^2 = r_{sr}^2 \tag{2-25}$$

In the application of (2-25), $h_o = 0$ for bottom points associated with shadow endpoints identified in Figures 2-2 and 2-3. In the case of an object in the water column, as depicted in Figure 2-3, h_a could be set to zero initially, and the calculations repeated in an iterated manner, using better estimates for h_a as they are obtained.

The above discussion has identified a concept for using telesounder-derived, low-resolution, three-dimensional depth map information to estimate object diameter, h_o. The quantities (r_s, g_s, ℓ_s) required in the above analysis might be expected to be accurate to within a few slant-range resolution cells. The accuracy of ground range and $\hat{h}_b(\bullet)$ values are coupled through (2-25). In addition, the accuracy of $\hat{h}_b(\bullet)$ values will directly affect the accuracy of h_o estimates through (2-14) and (2-16) or (2-19), (2-20), (2-23), respectively. In turn, the accuracy of $\hat{h}_b(\bullet)$ values will depend on

- The accuracy of discrete telesounder-derived height measurements.
- The interpolation error implicit in the adopted depth map gridding algorithm.

In the final analysis, the utility of the above approach may be best evaluated through attempting to implement it using actual data.

3.0 High-Resolution Three-Dimensional Depth Map Utilization Concepts

The discussion that follows will first review the swath-bathymetry concept for constructing high-resolution depth maps, as outlined by Denbigh in Ref. [1]. Based on Denbigh's analysis, figures for the required signal-to-interference ratios in received signals will be obtained in order to achieve desired depth accuracies. Next, in Sections 3.1 and 3.2, feature calculations based on stochastic and deterministic depth surface modeling approaches, respectively, will be considered.

The telesounder-based depth measurement concept was based on the addition of received demodulated signals $r_1(\bullet)$, $r_2(\bullet)$ obtained from two receivers as depicted in Figure 2-1. The factor in (2-4) that weights the magnitude of the sum signal z(t), is a result of the phase difference between the two received signals. In the swath-bathymetry depth measurement approach, the phase difference between the two received signals is estimated directly from the individual signals and employed to

allow depth measurements at each slant-range resolution cell. For the sake of consistency with Denbigh's discussion and analysis, the geometry of the two receivers and the associated depth calculation is depicted in Figure 3-1.

Hence, from Figure 3-1, the depth below receiver #1, $h_D(x)$, may be calculated as

$$h_D(x) = r_{s1}(x) \sin(\beta + \alpha(x)) \tag{3-1}$$

where $\alpha(x)$ is related to the phase difference between the two received signals, $\phi(x)$, by

$$\phi(x) = \frac{2\pi\Delta}{\lambda} \sin \alpha(x) \tag{3-2}$$

In summary, $\phi(x)$ and $r_{s1}(x)$ are estimated directly from the received demodulated signals $r_1(\bullet)$, $r_2(\bullet)$; $\phi(x)$ is then used to calculate $\alpha(x)$ from (3-2); and then $\alpha(x)$, $r_{s1}(x)$ are used to form $h_D(x)$ from (3-1).

In Ref. [1], Denbigh investigates the question of determining the minimal signal-to-interference ratio necessary to guarantee some desired fractional depth accuracy. If Δh_D, $\Delta\phi$ denote depth and phase errors, respectively, then based on a linearized analysis, Denbigh shows that

$$\frac{\Delta h_D}{h_D(x)} = \Delta\phi \frac{1}{\left(\frac{2\pi\Delta}{\lambda}\right) \cos(\psi(x) - \beta) \tan(\psi(x))} \tag{3-3}$$

Based on a phase estimation error analysis Denbigh obtains the following empirically fitted result for the standard derivation $\sigma_{\Delta\phi}$ associated with the phase estimation error

$$\sigma_{\Delta\phi} = \frac{2.5}{\rho^{.43}} \tag{3-4}$$

where ρ denotes the signal-to-interference power ratio in the received signal. Hence, if

$$|\Delta\phi| \cong \sigma_{\Delta\phi} \tag{3-5}$$

Figure 3-1. Denbigh's swath bathymetry depth map construction geometry.

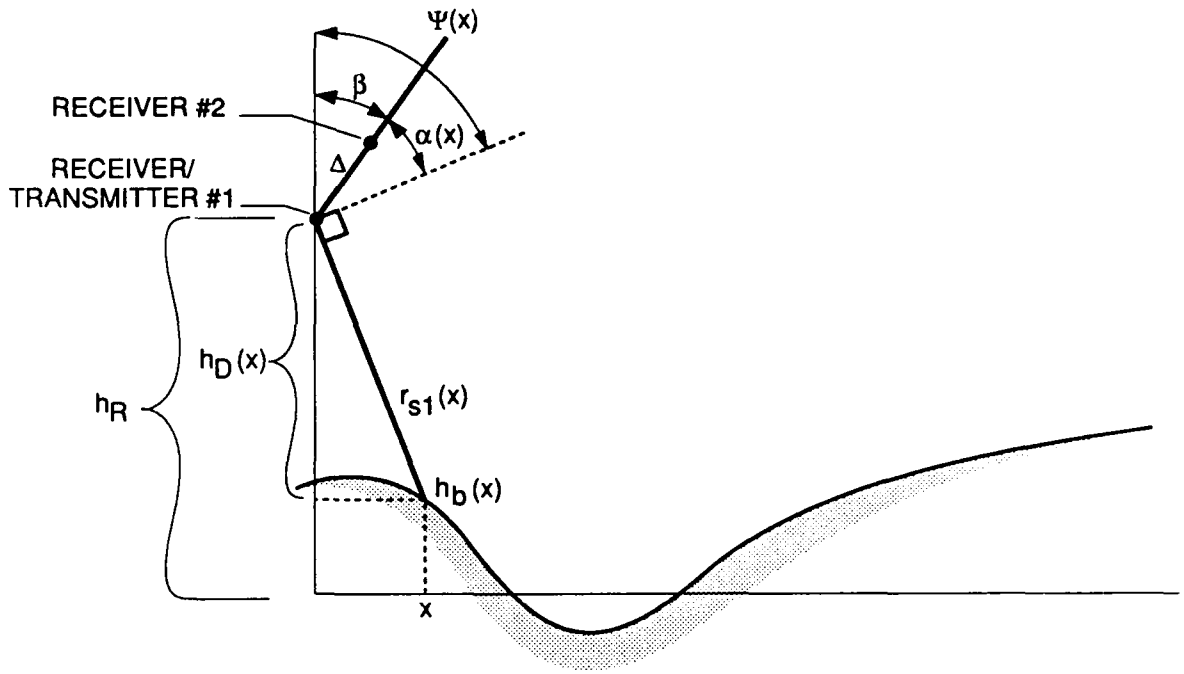

Figure 3-1.

then in order that

$$\left|\frac{\Delta h_D}{h_D(x)}\right| \cong \varepsilon \tag{3-6}$$

we must require that

$$\rho \cong \left|\frac{2.5}{\varepsilon\left(\frac{2\pi\Delta}{\lambda}\right)\cos(\psi(x)-\beta)\tan\psi(x)}\right|^{1/.43} \tag{3-7}$$

We next apply (3-6) and (3-7) to a specific case. From Table 2-2, the minimal diameter of a cylindrical-shaped mine was $\cong$ 0.4 m, and hence, we let

$$\Delta h_D = (0.1)\times(0.4\text{m}) = 0.04\text{ m} \tag{3-8}$$

then if $h_D \cong h_R$, as defined by (2-11), then

$$\varepsilon \cong 10^{-3} \tag{3-9}$$

Letting

$$\beta = 0 \tag{3-10}$$

$$\psi(x) = \frac{\pi}{4} \tag{3-11}$$

$$\Delta = 10\lambda \tag{3-12}$$

and employing (3-7) and (3-9) results in a value for ρ of

$$(\rho)_{dB} = 40.7 \tag{3-13}$$

The result (3-13) illustrates the stringent signal-to-interference ratios required in order to obtain depth measurements sufficiently accurate to yield shape information concerning cylindrical objects with diameters as small as $\cong$ 0.4 m. Reference [1] discusses techniques for reducing signal-to-interference ratios. One approach is to

average phase measurements over some collection of slant-range and along-track side-scan sonar image cells in order to reduce $\sigma_{\Delta\phi}$ values. By averaging N phase values, a reduction of $1/\sqrt{N}$ may be achieved in $\sigma_{\Delta\phi}$, and hence, for example, by letting

$$N = 100 \tag{3-14}$$

the figure in (3-13) becomes

$$(\rho)_{dB} = 17.4 \text{ dB} \tag{3-15}$$

Unfortunately, averaging phase information results in a loss of slant-range and along-track resolution in the resulting depth map. In the above case, to ensure a minimum of 5 resolution cells associated with a 0.4-m diameter cylindrical cross section requires a slant-range resolution of ≅8 cm or 3.1 in.

In summary, the above discussion suggests that slant-range resolutions of 8 cm and signal-to-interference ratios of as high as 40.7 dB may be necessary to adequately characterize cylindrically-shaped mines as small as 0.4 m in diameter.

3.1 Feature Calculation Based on Stochastic Depth Surface Modeling

The discussion that follows adopts the assumption that the two-dimensional terrain bottom height, $h_b(x,y)$, pictured as the two-dimensional extension of the terrain in Figure 3-1 and its discrete grid counterpart, $h_b(k,l)$, is modeled as a random field. In this context, a bottom mine will appear as an "anomalous" patch of terrain whose statistical characteristics are inconsistent or unpredictable relative to the surrounding "background" terrain. Hence, the feature calculation approaches described here are motivated by the literature on texture measures (Refs. [3], [4], [5]) as well as random field-based approaches to image processing (Ref. [6]).

The statistics of a stationary random field are characterized in the spatial domain by its mean value and correlation function, or in the spectral domain by its power spectral density function. Hence, typical choices for features are related to empirical correlation function or power spectral densities. Empirical autocorrelation function values, $\hat{c}(\Delta_x, \Delta_y)$, computed based on field values $h_b(k, l)$ associated with a specified two-dimensional, processing "estimation" window, R_E, are defined as follows:

Let empirical mean and standard deviation estimates, $\hat{m}, \hat{\sigma}^2$ be defined by

$$\hat{m} = \frac{1}{|R_E|} \sum_{(k,l) \in R_E} h_b(k,l) \tag{3-16}$$

$$\hat{\sigma}^2 = \frac{1}{|R_E|} \sum_{(k,l) \in R_E} (h_b(k,l) - \hat{m})^2 \tag{3-17}$$

where $|R_E|$ denotes the number of elements in R_E, then

$$\hat{c}(\Delta_x, \Delta_y) \triangleq \frac{1}{N(\Delta_x, \Delta_y)\hat{\sigma}^2} \sum_{\left\{\begin{matrix}(k,l) \in R_E \\ \& (k+\Delta_x, l+\Delta_y) \in R_E\end{matrix}\right\}} (h_b(k+\Delta_x, l+\Delta_y) - \hat{m})(h_b(k,l) - \hat{m}) \tag{3-18}$$

where $N(\Delta_x, \Delta_y)$ denotes the number of terms in the summation defined by (3-18). A vector of two-dimensional correlation coefficient features can be obtained by varying $(\Delta_x, \Delta_y) \in R_c^{(n)}$, where $R_c^{(n)}$ denotes a region in the 2D correlation plane of the form depicted in Figure 3-2. Due to the symmetry of the autocorrelation function through the origin, knowledge of $\hat{c}(\Delta_x, \Delta_y)$ for $(\Delta_x, \Delta_y) \in R_c^{(n)}$ implies knowledge of autocorrelation function estimates over the region $[-n, n] \times [-n, n]$.

Next, the construction of power spectral density related features is described. Assuming for simplicity that $R_E = [0, N_E-1] \times [0, N_E-1]$, the discrete Fourier transform $H_b(u,v)$ associated with the zero-mean field, $h_b(\bullet, \bullet) - \hat{m}$ is defined by

$$H_b(u,v) = \frac{1}{N_E^2} \sum_{k,l=0}^{N_E-1} [h_b(k,l) - \hat{m}]\, e^{-i2\pi(ku+lv)} \qquad u,v, \in [0, N_E-1] \tag{3-19}$$

Standard normalized texture features based on ring and wedge-shaped samples of the discrete Fourier power spectrum are defined as follows

$$\phi_{r_1, r_2} \triangleq \frac{1}{\hat{\sigma}^2} \sum_{\left\{\begin{matrix} r_1^2 \le u^2 + v^2 < r_2^2 \\ u, v \in [0, N_E-1] \end{matrix}\right\}} |H_b(u,v)|^2 \tag{3-20}$$

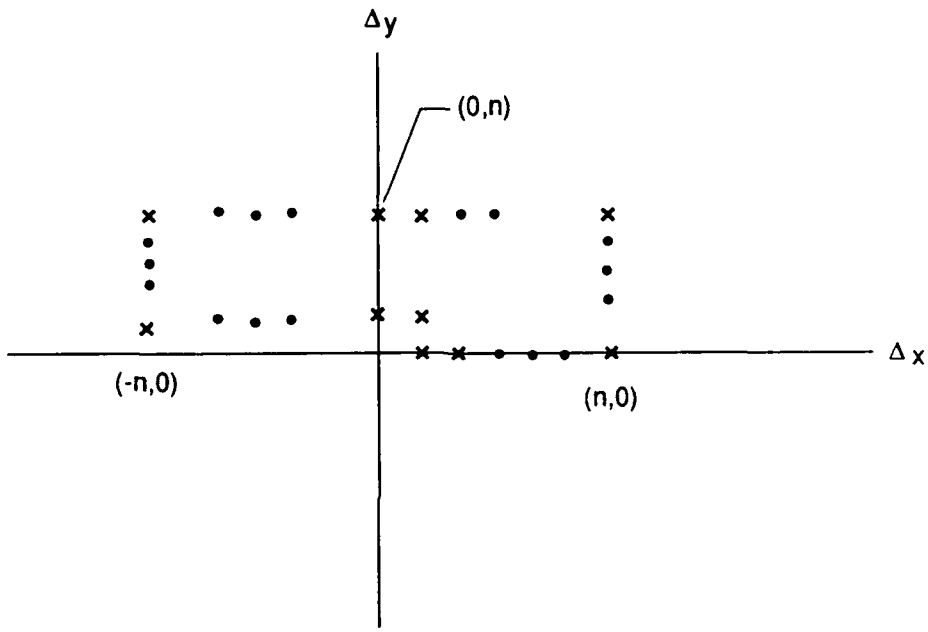

Figure 3-2. R_C region in 2D correlation plane.

$$\psi_{\theta_1,\theta_2} \triangleq \frac{1}{\hat{\sigma}^2} \sum |H_b(u,v)|^2$$

$$\left\{ \begin{array}{c} \theta_1 \le \tan^{-1}(v/u) < \theta_2 \\ u, v \in [0, N_E - 1] \end{array} \right\} \tag{3-21}$$

By choosing a discrete collection of (r_1, r_2), (θ_1, θ_2) pairs, a vector of power spectral density related features may be defined.

The features defined by (3-18), (3-20) and (3-21), together with $\hat{\sigma}^2$, use nonparametric techniques to characterize the statistics in the spatial and spectral domains, respectively, of the zero mean random field determined by $[h_b(\bullet,\bullet) - \hat{m}]$. An alternative approach for constructing statistical features would be to adopt the assumption that $h_b(\bullet, \bullet)$ can be modeled by a parametric, 2D autoregressive (ar) model of the form

$$h_b(k,l) - \sum_{(n,m) \in S^{(n_a)}} \alpha_{n,m} h_b(k+n, l+m) = w_{k,l} + \mu_w \tag{3-22}$$

where $S^{(n_a)}$ denotes a 2D region pictured in Figure 3-3, $w_{k,l}$ denotes a zero-mean spatially independent random field with variance

$$E\left\{ w_{k,l}^2 \right\} = \sigma_w^2 \tag{3-23}$$

and μ_w denotes a nonzero mean component to the prediction error process defined by $[\mu_w + w_{k,l}]$. In the above context, estimated 2D model parameters $\left\{\hat{\sigma}_w^2, \hat{\alpha}_{n,m} (n,m) \in S^{(n_a)}\right\}$ could be identified as features that determine the spatial or spectral properties of the zero-mean random field $[h_b(\bullet,\bullet) - \hat{m}]$. Estimates of $\left\{\hat{\alpha}_{n,m}, \hat{\mu}_w\right\}$ are determined through the solution of a linear least-squares problem defined by optimizing Φ_{LS}, where

$$\Phi_{LS} = \sum \left(h_b(k,l) - \sum_{(n,m) \in S^{(n_a)}} \alpha_{n,m} h_b(k+n, l+m) - \mu_w \right)^2$$

$$\left\{ \begin{array}{l} (k,l) \in R_E \\ \text{so that } (k+n, l+m) \in R_E \\ \forall\, (n,m) \in S^{(n_a)} \end{array} \right\} \tag{3-24}$$

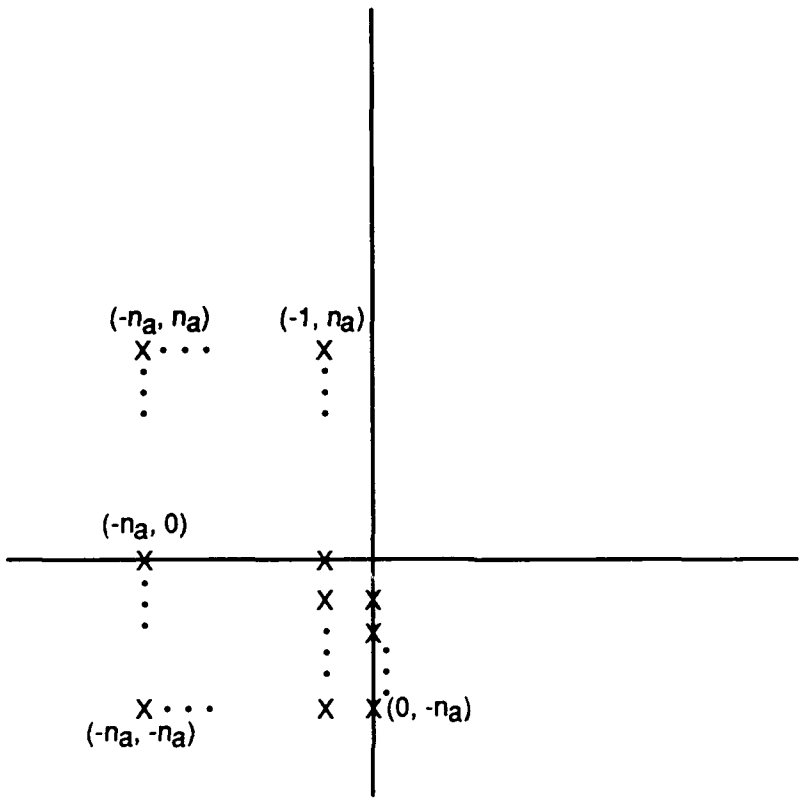

Figure 3-3. Support, $S^{(n_a)}$, of 2D autoregressive model.

If m_{LS} denotes the number of terms in the summation defined by, then

$$\hat{\sigma}_w^2 = \frac{1}{m_{LS}} \Phi_{LS} \Big|_{\{\hat{\alpha}_{n,m}, \hat{\mu}_w\}} \tag{3-25}$$

Refs. [7] and [8] discuss solution techniques for linear least-squares problems. The performance functional Φ_{LS} can be expressed

$$\Phi_{LS} = ||Ax - b||^2 \tag{3-26}$$

for an appropriately defined m_{LS} x n_{LS} - dimensional matrix, A, m_{LS} - dimensional vector b, and n_{LS} - dimensional vector x defined as

$$x \triangleq \begin{pmatrix} \alpha_{n,m} (n, m) \in S^{(n_a)} \\ \mu_w \end{pmatrix} \tag{3-27}$$

when n_{LS} is expressed as

$$n_{LS} = 2 n_a^2 + 2n_a + 1 \tag{3-28}$$

The most computationally inexpensive technique for optimizing Φ_{LS} involves the solution of the normal equations

$$[A' A] x = A' b \tag{3-29}$$

and requires a total of

$$(n_{LS}^2/2)[m_{LS} + (n_{LS}/3)] \tag{3-30}$$

floating-point operations using the Cholesky decomposition to solve (3-29), by factorizing $A'A$ into a product of upper and lower triangular factors, and solving a sequence of triangular systems.

While the estimated ar model parameters $\{\hat{\alpha}_{n,m}, \hat{\sigma}_w^2\}$ can be employed as features, an additional option employed in random field-based approaches to object

detection (Ref. [6]) involves the construction of normalized, prediction error residuals defined by

$$r_{k,l} \triangleq \left[h_b(k,l) - \sum_{(n,m)\in S^{(n_s)}} \hat{\alpha}_{n,m}\, h_b(k+n, l+m) - \hat{\mu}_w \right] / \hat{\sigma}_w^2 \tag{3-31}$$

When the random field model defined by (3-22) holds, the normalized prediction error residuals defined by (3-31) will approximately represent a spatially independent random field. In addition, large $r_{k,l}$ values will be representative of areas of the field which are atypical. The above discussion suggests that correlation coefficient, and mean square power estimates constructed using $r_{k,l}$ values over processing sub-windows within R_E, will be useful features.

The above described, stochastic model based approaches to feature construction were based on viewing $h_b(\bullet, \bullet)$ as a continuous valued random field. The next texture measure related features described here are defined using grey-level images, i.e., images that assume a discrete set of values. Such grey-level images arise naturally in an imaging setting, but may be formed by quantizing a continuous valued image into a discrete set of ranges, or bins. It is assumed here that the image values $[h_b(k, l) - \hat{m}]$ are mapped into the values $0 \ldots n_G-1$ by an unspecified quantization function $G_Q[\bullet]$, in order to obtain the discrete field, $f(k, l)$, i.e.

$$f(k,l) \triangleq G_Q\left[h_b(k,l) - \hat{m}\right] \tag{3-32}$$

The first grey-level statistic-related texture measures to be defined are computed from empirical joint grey-level distribution estimates. Let Δ denote a two-dimensional discrete shift vector defined by

$$\Delta = (\Delta_x, \Delta_y) \tag{3-33}$$

and let

$$P_\Delta(i,j) \triangleq \Pr\left\{ (f(k,l) = i) \,\&\, (f(k+\Delta_x, l+\Delta_y) = j) \right\} \tag{3-34}$$

$$\psi_\Delta(i,j) = \left[P_\Delta(i,j) + P_\Delta(j,i)\right]/2 \tag{3-35}$$

If $\hat{\psi}_\Delta(i,j)$ denotes an empirically computed estimate of $\psi_\Delta(i,j)$, then Ref. [3] defines the following spatial grey-level dependence (SGLD) related features:

(1) Energy:

$$E_\Delta \triangleq \sum (\hat{\psi}_\Delta(i,j))^2 \tag{3-36}$$

(2) Entropy:

$$H_\Delta \triangleq \sum -\hat{\psi}_\Delta(i,j) \ln[\hat{\psi}_\Delta(i,j)] \tag{3-37}$$

(3) Correlation:

$$C_\Delta \triangleq \sum [(i-\hat{\mu}_x)(j-\hat{\mu}_y)\hat{\psi}_\Delta(i,j)/(\hat{\sigma}_x\hat{\sigma}_y)] \tag{3-38}$$

(4) Local Homogeneity:

$$L_\Delta \triangleq \sum \frac{1}{1+(i-j)^2}\hat{\psi}_\Delta(i,j) \tag{3-39}$$

(5) Inertia:

$$I_\Delta = \sum (i-j)^2 \hat{\psi}_\Delta(i,j) \tag{3-40}$$

with $\{\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x, \hat{\sigma}_y\}$ defined by

$$\hat{\mu}_x \triangleq \sum i\,\hat{\psi}_\Delta(i,j) \tag{3-41}$$

$$\hat{\mu}_y \triangleq \sum j\,\hat{\psi}_\Delta(i,j) \tag{3-42}$$

$$\hat{\sigma}_x^2 \triangleq \sum (i-\hat{\mu}_x)^2 \hat{\psi}_\Delta(i,j) \tag{3-43}$$

$$\hat{\sigma}_y^2 \triangleq \sum (j-\hat{\mu}_y)^2 \hat{\psi}_\Delta(i,j) \tag{3-44}$$

The features $\{E_\Delta, H_\Delta, C_\Delta, L_\Delta, I_\Delta\}$ should be computed for a collection of 2D shifts over a region of the form defined by Figure 3-2, due to the symmetry of $\psi_\Delta(\bullet,\bullet)$ through the origin in Δ.

A second set of grey-level statistic-related texture measures are computed from the empirical distributions of the absolute value of grey-level differences. Let $\eta_\Delta(i)$ be defined by

$$\eta_\Delta(i) = \Pr\left\{\left|f(k, \ell) - f(k + \Delta_x, \ell + \Delta_y)\right| = i\right\} \tag{3-45}$$

and let $\hat{\eta}_\Delta(i)$ denote a corresponding empirical estimate. Then Ref. [3] defines the following grey-level difference (GLD)-related features:

(1) Contrast:

$$CON_\Delta \triangleq \sum i^2 \hat{\eta}_\Delta(i) \tag{3-46}$$

(2) Angular Second Moment:

$$ASM_\Delta \triangleq \sum (\hat{\eta}_\Delta(i))^2 \tag{3-47}$$

(3) Entropy:

$$ENT_\Delta \triangleq -\sum \hat{\eta}_\Delta(i) \ln[\hat{\eta}_\Delta(i)] \tag{3-48}$$

(4) Mean:

$$MEAN_\Delta \triangleq \sum i \hat{\eta}_\Delta(i) \tag{3-49}$$

(5) Inverse Difference Moment:

$$IDM_\Delta \triangleq \sum \frac{\hat{\eta}_\Delta(i)}{(i^2 + 1)} \tag{3-50}$$

Note from the definition of $\eta_\Delta(i)$ that the features computed will be dependent on the sign of the direction, Δ, and hence $\{CON_\Delta, ASM_\Delta, ENT_\Delta, MEAN_\Delta, IDM_\Delta\}$ would typically be computed for a collection of Δ's over a neighborhood of the origin, from which (0, 0) was excluded.

A third and final set of grey-level statistic-related texture measures are computed by counting grey-level "runs" of various lengths. A "run" of grey levels consists in a collection of adjacent pixels, along some direction θ, having the same grey level value. Let $\xi_\theta(i, j)$ denote the number of runs of length j, associated with

direction θ and grey level value i, contained within the processing window R_E. Then Ref. [3] defines the following grey-level run length (GLRL)-related features:

(1) Short-Run Emphasis:

$$RF_{1,\theta} \triangleq \frac{1}{T_R} \sum \frac{\xi_\theta(i,j)}{j^2} \tag{3-51}$$

(2) Long-Run Emphasis:

$$RF_{2,\theta} \triangleq \frac{1}{T_R} \sum j^2 \, \xi_\theta(i,j) \tag{3-52}$$

(3) Grey-Level Distribution:

$$RF_{3,\theta} \triangleq \frac{1}{T_R} \sum_i \left\{ \sum_j \xi_\theta(i,j) \right\}^2 \tag{3-53}$$

(4) Run-Length Distribution:

$$RF_{4,\theta} \triangleq \frac{1}{T_R} \sum_j \left\{ \sum_i \xi_\theta(i,j) \right\}^2 \tag{3-54}$$

(5) Run Percentages:

$$RF_{5,\theta} \triangleq \frac{1}{|R_E|} \sum \xi_\theta(i,j) \tag{3-55}$$

where

$$T_R \triangleq \sum \xi_\theta(i,j) \tag{3-56}$$

The features $\{RF_{1,\theta}, RF_{2,\theta}\ RF_{3,\theta}, RF_{4,\theta}\ RF_{5,\theta}\}$ are typically computed for directions $\theta = 0°, 45°, 90°, 135°$.

Research in Refs. [3], [4] has investigated the use of SGLD, GLD, GLRL features in discriminating between textures in aerial terrain photo and synthetic fields generated using Markov chain models. These results suggest that SGLD and GLD-derived features were more effective in separating different texture classes than GLRL features. Hence, of the grey-level texture measures discussed above, the SGLD and GLD features may be the most promising.

The above discussion has centered on traditional spatial and spectral texture measures. Recent mathematical research under the heading of fractal geometry has resulted in new geometric concepts for describing and quantifying the degree of irregularity of natural objects. Hence, an additional feature of interest may be estimates of the fractal dimension of the 3D terrain surface over a processing window, R_E. Fractal dimension features have already been successfully employed in Ref. [9] to discriminate natural from man-made objects in images. The brief discussion that follows outlines the use of Mandelbrot measures in fractal dimension estimation as described in Ref. [10]. Let a 3D point set $v_{k,l}$ be defined by

$$v_{k,l} \triangleq \begin{pmatrix} x_{k,l} \\ y_{k,l} \\ h_b(k,l) \end{pmatrix} (k,l) \in R_E \tag{3-57}$$

Then, by centering spheres of radius, L, at each of the $v_{k,l}$ and counting the number of 3D terrain points that are enclosed, the "Mandelbrot" probability distribution P(m,L) may be estimated, where

$$P(m, L) \triangleq \text{ the probability that a sphere of diameter L encloses m-points} \tag{3-58}$$

Now, by computing M(L) defined as

$$M(L) \triangleq \sum m\, P(m, L) \tag{3-59}$$

and repeating the above procedure for a sequence of L values, the fractal dimension estimate may be obtained as the slope, D, of the graph of ln(M(L)) versus ln(L), i.e., through solving a linear least-squares problem involving the minimization of

$$\sum_i [\ln(M(L_i)) - D \ln(L_i)]^2 \tag{3-60}$$

Next, the present subsection on stochastic depth surface model-based feature construction techniques is concluded by sketching the potential use of wavelet transform-based features. Recall that through (3-9) - (3-21), 2D Fourier spectra related features were defined. The basis functions associated with Fourier expansions, i.e., complex exponentials, are localized in their frequency content, but not

in terms of their spatial duration. Wavelet signal representation techniques (see Ref. [11]) allow for expansion in terms of basis functions that are both orthogonal and localized in terms of both spatial duration and frequency content.

In the discussion that follows, the wavelet signal representation of a 2D continuous function g(x,y) is briefly described, in order to motivate the construction of wavelet transform related features. Two effectively finite duration 1D functions that play a key role in wavelet based signal representatives are termed the "scaling function", $\phi(\bullet)$, and "wavelet function", $\psi(\bullet)$. The following orthonormal 2D functions are defined, and associated with 2D basis functions underlying the wavelet signal representation:

$$H_{n,m}^{(0,j)}(x, y) = 2^j \phi(2^j (x - 2^{-j} n)) \phi((2^j (y - 2^{-j} m)) \tag{3-61}$$

$$H_{n,m}^{(1,j)}(x, y) = 2^j \phi(2^j (x - 2^{-j} n)) \psi((2^j (y - 2^{-j} m)) \tag{3-62}$$

$$H_{n,m}^{(2,j)}(x, y) = 2^j \psi(2^j (x - 2^{-j} n)) \phi((2^j (y - 2^{-j} m)) \tag{3-63}$$

$$H_{n,m}^{(3,j)}(x, y) = 2^j \psi(2^j (x - 2^{-j} n)) \psi((2^j (y - 2^{-j} m)) \tag{3-64}$$

The index "j" in the above notation denotes the j-th resolution, while the indices (n, m) serve to determine the regions over which the basis functions are effectively non-zero. Using the notations $H^{(0,j)}\{\bullet\}$, $H^{(1,j)}\{\bullet\}$, $H^{(2,j)}\{\bullet\}$, $H^{(3,j)}\{\bullet\}$ to denote projection operators mapping a given function $g(\bullet, \bullet)$ onto the linear spaces spanned by the associated orthonormal basis functions, then the key practical result underlying the use of wavelet expansions can be expressed as

$$H^{(0,0)}\{g\} = H^{(0,-J)}\{g\} + \sum_{j=-J}^{1} \left[H^{(1,j)}\{g\} + H^{(2,j)}\{g\} + H^{(3,j)}\{g\}\right] \tag{3-65}$$

Due to the special properties of $\phi(\bullet)$, $\psi(\bullet)$, each term on the right-hand side of (3-65) is mutually orthogonal and is itself represented in terms of orthogonal basis functions. The $H^{(0,-J)}\{g\}$ term represents a coarse, low frequency approximation to the function,

while the $H^{(1,j)}\{g\}$, $H^{(2,j)}\{g\}$, $H^{(3,j)}\{g\}$ terms represent higher spatial frequency effects that appear at progressively finer resolutions, and correspond to effects resembling vertical "edges," horizontal "edges," and "corners," respectively.

When the above 2D signal representation framework is applied to the representation of discrete images as in Refs. [11] and [12], the coefficients associated with the expansion of each of the projections on either side of (3-65) can be associated with "images." These coefficients can then be used as features for the detection of "anomalous" objects, which appear at some specified resolution. Hence, one feature of potential interest would be ms power estimates associated with sub-windows of the "images" associated with each of the $H^{(1,j)}\{\bullet\}$, $H^{(2,j)}\{\bullet\}$, $H^{(3,j)}\{\bullet\}$ terms.

3.2 Feature Calculation Based on Deterministic Depth Surface Modeling

The preceding discussion adopted a statistical perspective on the representation of the terrain bottom height $h_b(x,y)$, and the construction of features for mine classification. The discussion that follows will adopt a deterministic surface model-based perspective, similar to that adopted in the 3D object recognition and machine/computer vision literature. A comprehensive review of 3D object recognition approaches by Besl and Jain (Ref. [13]) suggests that object recognition is best solved by recognizing the individual surface regions associated with the depth map functions for specific object types. Hence, the structure of 3D object recognition algorithms often consists of

(1) The fitting of surface representations and extraction of geometric features.

(2) The construction of a symbolic representation of the collection of surfaces defining a depth map, using the features extracted in (1).

(3) The "matching" of derived symbolic representations against analogous symbolic representations computed for specific object types.

Our discussion will focus on step (1), with the view that the geometric features identified will become inputs to neural-network-based classifiers for mine recognition.

The development that follows will first motivate the selection of curvature and surface normal-related features by appealing to the differential geometry of surfaces. Next, specific generic surface fitting/representation approaches will be reviewed that can support the calculation of the identified geometric features. Then, noting that the above described approaches employ general surface representation techniques, while mines frequently have spherical or cylindrical surfaces; we will finally consider the use

of specialized quadric surface fitting techniques for the calculation of shape-related features.

We next define some concepts and notation from the differential geometry of surfaces (Ref. [14]). In general, a parametric 3D surface is represented by a vector valued function of two variables, $\mathbf{f}(\mathbf{u})$, where

$$\mathbf{f}(\mathbf{u}) \triangleq \begin{pmatrix} f_1(\mathbf{u}) \\ f_2(\mathbf{u}) \\ f_3(\mathbf{u}) \end{pmatrix} \tag{3-66}$$

with $f_1(\mathbf{u})$, $f_2(\mathbf{u})$, $f_3(\mathbf{u})$ representing x, y, z coordinate functions, respectively and

$$\mathbf{u} \triangleq \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{3-67}$$

In the case of a Monge or graph surface, (3-66) is specialized so that

$$f_1(\mathbf{u}) = u_1 \tag{3-68}$$

$$f_2(\mathbf{u}) = u_2 \tag{3-69}$$

If the differential of a vector valued mapping, $\mathbf{f}(\mathbf{u})$, denoted by $d\mathbf{f}$, is defined as

$$d\mathbf{f} \triangleq \mathbf{f}_{u_1} du_1 + \mathbf{f}_{u_2} du_2 \tag{3-70}$$

then the first and second fundamental forms are defined by

$$I(\mathbf{u}, d\mathbf{u}) \triangleq -d\mathbf{f} \bullet d\mathbf{f} = d\mathbf{u}' \, G \, d\mathbf{u} \tag{3-71}$$

$$II(\mathbf{u}, d\mathbf{u}) \triangleq -d\mathbf{f} \bullet d\mathbf{n} = d\mathbf{u}' \, B \, d\mathbf{u} \tag{3-72}$$

where $\mathbf{n}$ denotes the surface normal field defined by

$$\mathbf{n}(\mathbf{u}) = \frac{\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}}{\left|\left|\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}\right|\right|} \tag{3-73}$$

and G, B denote first and second fundamental form matrices that are functions of **u**. The normal curvature, k_n, denotes the projection onto the normal **n**(**u**) of the curvature vector associated with a curve on the 3D surface, passing through **f**(**u**), and is expressed by

$$k_n = \frac{\mathrm{II}(\mathbf{u}, \mathbf{du})}{\mathrm{I}(\mathbf{u}, \mathbf{du})} \tag{3-74}$$

The extreme values associated with k_n, denoted by k_1, k_2, are termed principal curvatures. The Gaussian and mean curvatures, K and H, respectively, are defined by

$$K = k_1 k_2 \tag{3-75}$$

$$H = \frac{1}{2}(k_1 + k_2) \tag{3-76}$$

and can in addition be expressed as functions of first and second fundamental form matrices as

$$K = \det\left[G^{-1} B\right] \tag{3-77}$$

$$H = \frac{1}{2} \mathrm{tr}\left[G^{-1} B\right] \tag{3-78}$$

Besl and Jain in Ref. [15] summarize the properties that make H,K attractive features for use in a 3D object recognition context:

(1) H and K are invariant to changes of variables in the **u** parameters, with the only exception that sign changes in H may occur if the sense of the surface normal **n**(**u**) changes.

(2) H and K are invariant to translations and rotations of the 3D surface

(3) The signs of H, K indicate the local shape of the surface.

In Refs. [15] and [16] Besl and Jain indicate the following eightfold classification of local surface shape based on the signs of H, K:

(1) $H < 0, K > 0 \rightarrow$ peaked surface

(2) $H < 0, K = 0 \rightarrow$ ridge surface

(3) $H < 0, K < 0 \rightarrow$ saddle ridge surface

(4) $H = 0, K = 0 \rightarrow$ flat surface

(5) $H = 0,\ K < 0 \rightarrow$ minimal surface

(6) $H > 0, K < 0 \rightarrow$ saddle valley surface

(7) $H > 0, K = 0$ valley surface

(8) $H > 0, K > 0 \rightarrow$ pit surface

The eight categories mentioned above are depicted in Figure 3-4.

Yang and Kak in Ref. [17] advocate the use of H,K histograms in order to discriminate between 3D objects. These authors distinguish between the following six categories of curvature histograms:

(1) Positive only with peak: curvature values nearly all positive with a well-defined peak value in the curvature histogram.

(2) Positive only without peak: curvature is positive, but continuously varying, so that no well-defined peak occurs in the curvature histogram.

(3) Negative only with peak: same as (1) except involving negative values.

(4) Negative only without peak: same as (2) except involving negative values.

(5) Peak at zero: curvature histogram has a single peak at the bin corresponding to zero curvature.

(6) Positive and negative without peak: corresponds to a surface with continuously varying curvature.

As specific examples, a spherical object has a K-histogram in category (1), and an H-histogram in category (3), while a cylindrical object has a K-histogram in category (5) and an H-histogram in category (3).

While specific approaches for calculating H, K values will be made explicit in the discussion that follows, the above background is sufficient to suggest the following curvature related features for use in neural network based mine classification. It is assumed here that the quantities designated below are obtained by the calculation of $H(\mathbf{u})$, $K(\mathbf{u})$, values over a discrete collection of $\mathbf{u}$ values, which are associated with a discrete grid defined over a specified processing, or estimation window, R_E. The proposed list of curvature related features includes

(1) Minimum and maximum H, K values.

(2) Average H, K values.

(3) The variances of H, K values.

(4) H, K histogram values.

Next, more explicit results than (3-75) - (3-78), obtained from Ref. [16], are stated for the calculation of H, K. If a_i i=1, 2, 3 is defined by

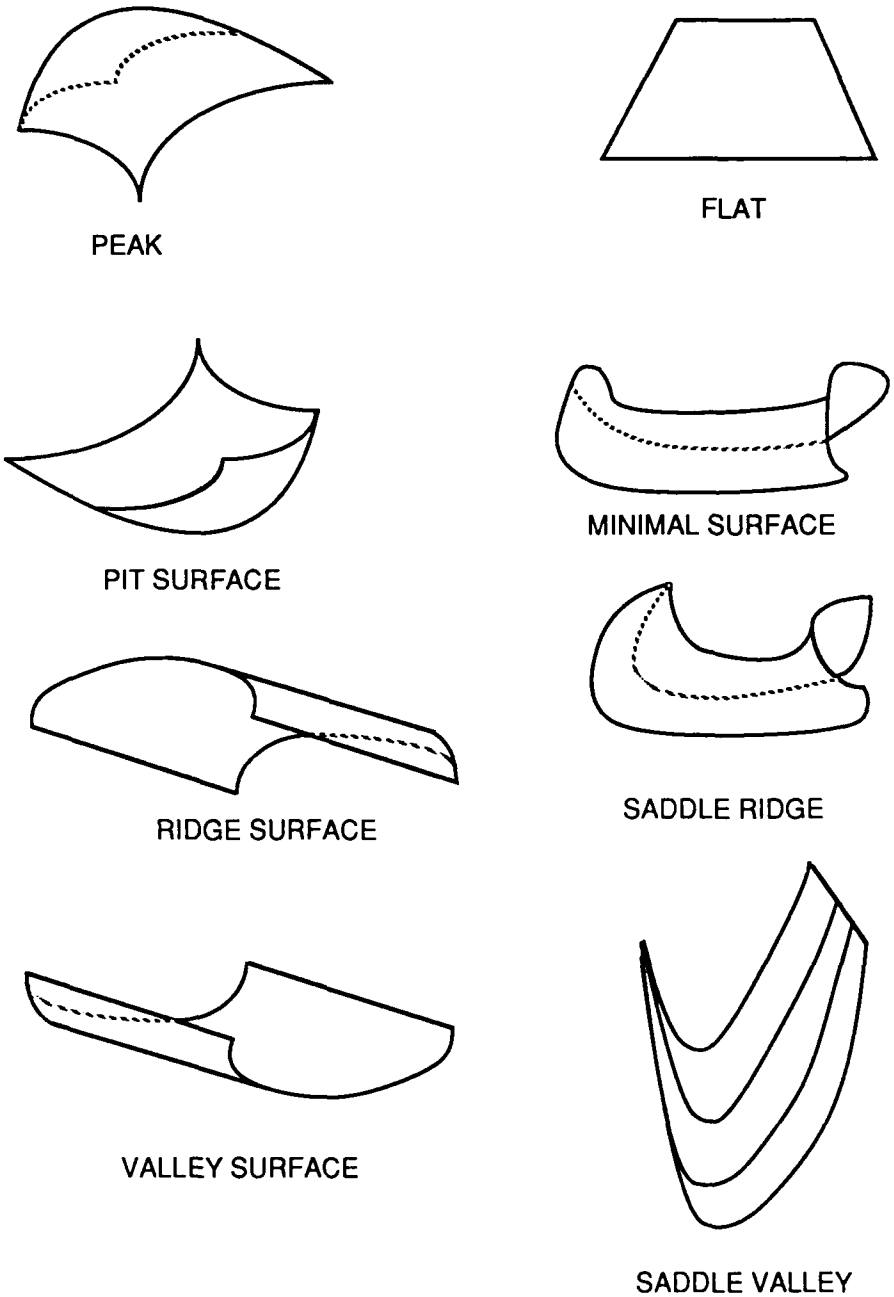

Figure 3-4. Local surface categories based on H,K signs.

$$a_1 \triangleq [\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}] \bullet \mathbf{f}_{u_1 u_1} \tag{3-79}$$

$$a_2 \triangleq [\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}] \bullet \mathbf{f}_{u_1 u_2} \tag{3-80}$$

$$a_3 \triangleq [\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}] \bullet \mathbf{f}_{u_2 u_2} \tag{3-81}$$

then

$$H = \frac{a_1 \, ||\mathbf{f}_{u_2}||^2 - 2\, a_2 [\mathbf{f}_{u_1} \bullet \mathbf{f}_{u_2}] + a_3 \, ||\mathbf{f}_{u_1}||^2}{2 \; ||\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}||^3} \tag{3-82}$$

$$K = \frac{(a_1 \, a_3 - a_2^2)}{||\mathbf{f}_{u_1} \times \mathbf{f}_{u_2}||^4} \tag{3-83}$$

The above general results simplify in the case of a Monge or graph surface ((3-68) and (3-69)) to

$$H = \frac{1}{2}\left[\frac{\left(f_{3,u_1u_1} + f_{3,u_2u_2} + f_{3,u_1u_1} f_{3,u_2}^2 + f_{3,u_2u_2} f_{3,u_1}^2 - 2 f_{3,u_1} f_{3,u_2} f_{3,u_1u_2}\right)}{\left(1 + f_{3,u_1}^2 + f_{3,u_2}^2\right)^{3/2}}\right] \tag{3-84}$$

$$K = \frac{\left(f_{3,u_1u_1} f_{3,u_2u_2} - f_{3,u_1u_2}^2\right)}{\left(1 + f_{3,u_1}^2 + f_{3,u_2}^2\right)^{3/2}} \tag{3-85}$$

The important conclusion to be obtained from (3-82) - (3-85) is that the calculation of H, K requires the calculation of second order derivatives of the coordinate functions, $f_i(\mathbf{u})$.

The above discussion has centered on the definition of curvature-related quantities for use as features in 3D object recognition. Other research in the 3D object recognition literature has considered the use of surface normal orientation related information. The unit vector surface normal field, $\mathbf{n}(\mathbf{u})$, defined by (3-73) can be characterized by its associated polar and azimuthal spherical coordinates, denoted here by $\phi(\mathbf{u})$, $\theta(\mathbf{u})$, respectively. Sethi and Jayaramamurthy (Ref. [18]), in the case of graph or Monge surfaces, define characteristic contains as curves C_ρ satisfying

$$\mathcal{C}_\rho \triangleq \{\mathbf{u} : \mathbf{n}(\mathbf{u}) \bullet \mathbf{n}_{ref} = \rho\} \tag{3-86}$$

The vector $\mathbf{n}_{ref}$ denotes a fixed, reference direction unit vector. As specific examples, note that the $\mathcal{C}_\rho$ contours associated with a sphere are a family of circles or ellipses, while for a cylinder, the analogous contours are a collection of parallel lines. A Hough transform based approach is suggested in Ref. [18] for the characterization of $\mathcal{C}_\rho$ contours. Letting the function $g_{r,\varphi}(\mathbf{u})$ be defined as

$$g_{r,\varphi}(\mathbf{u}) \triangleq r - u_1 \cos\varphi - u_2 \sin\varphi \tag{3-87}$$

letting (r_i, φ_j) denote a discrete collection of (r, φ) values covering alternative lines of practical interest, then the Hough Transform "image" $H(r_i, \varphi_j)$ is determined (see Ref. [19]) by initializing $H(r_i, \varphi_j)$ to zero and then incrementing by 1 for each discrete $\mathbf{u}_o \in \mathcal{C}_\rho$ satisfying

$$|g_{r_i,\varphi_j}(\mathbf{u}_0)| < \varepsilon \tag{3-88}$$

where ε is selected based on the resolution of the (r,φ) discretization. In the above context, features of potential use as inputs to a neural-network classifier include

(1) $H(r_i, \varphi_j)$ "image" values.

(2) (r, φ) moments computed viewing $H(r_i, \varphi_j)$ as a discrete mass distribution, e.g., such as the (r, φ) center of mass, and second moments of (r, φ) computed about the r, φ mass centers.

While the technique of Sethi and Jayaramamurthy (Ref. [18]) uses surface normal information to identify the type of characteristic contours $\mathcal{C}_\rho$ in the (u_1, u_2) plane, and hence to identify object type, an approach proposed by Herbert and Pence in Ref. [20] suggests the direct use of the Hough transform approach on calculated surface normal components in order to test for relationships satisfied for alternative surface types. As specific examples, consider the following Hough transform function, $g^{(PL)}_{\phi^*,\theta^*}(n_1, n_2, n_3)$, $g^{(CYL)}_{\phi^*,\theta^*}(n_1, n_2, n_3)$ which play the same role as $g_{r,\varphi}(\mathbf{u})$ in (3-87), in defining Hough transform "images" $H^{(PL)}(\phi_i^*, \theta_j^*)$, $H^{(CYL)}(\phi_i^*, \theta_j^*)$:

$$g^{(PL)}_{\phi^*,\theta^*}(n_1, n_2, n_3) = \sin\phi^* \cos\theta^* n_1 + \sin\phi^* \sin\theta^* n_2 + \cos\phi^* n_3 - 1 \tag{3-89}$$

$$g^{(CYL)}_{\phi^*,\theta^*}(n_1, n_2, n_3) = \sin\phi^* \cos\theta^* n_1 + \sin\phi^* \sin\theta^* n_2 + \cos\phi^* n_3 - 1 \tag{3-90}$$

Note, that (3-89) and (3-90) specify quantities that are exactly zero for surface normal components associated with planar and cylindrical surfaces, respectively, when (ϕ^*, θ^*) denotes the spherical coordinates of the planar surface normal or the axis of the cylinder, respectively. The above discussion suggests that Hough transform-based features, analogous to those described in the case of characteristic contours, C_ρ , but based on generalized Hough transforms computed using (3-89) and (3-90), may be of interest.

The approaches presented in Refs. [18] and [20] make use of $\mathbf{n}(\mathbf{u})$ vector field alone. Horn (Refs. [21] and [22]) suggests the use of derived image termed the "extended Gaussian image" (EGI), which combines both surface normal and curvature-related information. In the case of convex objects, there exists a unique correspondence between surface points and associated surface normal directions. Hence, for a convex object, the continuous EGI, which may be viewed as a surface mass density defined on the unit sphere, $G^{(CONV)}(\phi, \theta)$ is defined by

$$G^{(CONV)}(\phi(\mathbf{u}), \theta(\mathbf{u})) = \frac{1}{K(\mathbf{u})} \tag{3-91}$$

For nonconvex objects, with a finite or countable number of points having the same orientation, the continuous EGI is defined by

$$G(\phi, \theta) = \sum_{\{i:\ \phi(\mathbf{u}) = \phi,\ \theta(\mathbf{u}) = \theta\}} \frac{1}{|K(\mathbf{u}_i)|} \tag{3-92}$$

For objects with an uncountable number or points having the same orientation, the continuous EGI involves impulses defined on the surface of the unit sphere (see Ref. [21]) for a precise definition in this case). For our purposes, (3-92) is an adequate definition since our main interest is in a discrete version of the EGI, termed the DEGI.

The discrete EGI, or DEGI, is determined by

- A discrete set of cells σ_p $p = 1 \ldots M_S$ defined on the unit sphere.
- A discrete set of rectangular cells in the $\mathbf{u}$ parameter space, with associated center $\mathbf{u}_\ell$, and areas A_ℓ $\ell = 1 \ldots Mp$.

so that

$$G_{\sigma_p} = \sum_{\{\ell:\ (\phi(\mathbf{u}_\ell),\ \theta(\mathbf{u}_\ell)) \in \sigma_p\}} \frac{1}{|K(\mathbf{u}_\ell)|} \tag{3-93}$$

Alternatively, since the Gaussian curvature can be formulated in terms of area-related quantities, G_{σ_p} can also be approximated as

$$G_{\sigma_p} = \sum_{\{\ell:\ (\phi(\mathbf{u}_\ell),\theta(\mathbf{u}_\ell)) \in \sigma_p\}} A_\ell \quad \sqrt{G_{11}(\mathbf{u}_\ell)\, G_{22}(\mathbf{u}_\ell) - G_{12}^2(\mathbf{u}_\ell)} \tag{3-94}$$

where G (**u**) denotes the first fundamental form matrix defined through (3-71). Note, that for the case of a Monge or graph surface

$$\sqrt{G_{11}(\mathbf{u}_\ell)\, G_{22}(\mathbf{u}_\ell) - G_{12}^2(\mathbf{u}_\ell)} = \sqrt{1 + f_{3,\,u_1}^2 + f_{3,\,u_2}^2} \tag{3-95}$$

The use of the DEGI in a 3D object recognition context is complicated by the fact that, in general, it is viewpoint-dependent, and that typically only the visible portion of a surface is available for its computation. One way to derive features from the DEGI whose viewpoint dependence is lessened is to regard G_{σ_p} as defining a discrete mass distribution on the unit sphere and to compute an appropriate center of mass, with moments defined based on expected powers of spherical distances from the mass center. It is assumed that G_{σ_p} is viewed as a point mass associated with an appropriately selected center (ϕ_p, θ_p) of the spherical cell σ_p. Then, the spherical coordinates of a center of mass, (ϕ_{cm}, θ_{cm}) can be defined as

$$\phi_{cm} = \frac{\sum \phi_p\, G_{\sigma_p}}{\sum G_{\sigma_p}} \tag{3-96}$$

$$\theta_{cm} = \frac{\sum \theta_p\, G_{\sigma_p}}{\sum G_{\sigma_p}} \tag{3-97}$$

Hence, if $d_s(\phi_1, \theta_1, \phi_2, \theta_2)$ denotes the great circle arclength between the two points (ϕ_1, θ_1), (ϕ_2, θ_2) then expected powers of the spherical distance from the mass center (ϕ_{cm}, θ_{cm}) can be defined by

$$m_q = \frac{\sum d_s^q(\phi_p, \theta_p, \phi_{cm}, \theta_{cm})}{\sum G\sigma_p} \tag{3-98}$$

To conclude our discussion of the EGI and DEGI, moment features like that defined by (3-98) constitute an additional class of potential inputs for neural-network-based mine classifiers.

The above discussion, motivated by ideas from differential geometry and the 3D object recognition literature, has centered on the identification of curvature and surface normal-related features for use in mine classification. The discussion that follows will focus on a review of selected techniques for the fitting of surface representations in order to support the calculation of the above-described features. The majority of the various surface representation techniques available from the approximation theory and computer graphics literature divide into interpolatory approaches (denoted here by the abbreviation IA) in which the approximant function exactly reproduces the data at specified points, and least-squares-based approaches (denoted here by the abbreviation LSA), in which the approximant function solves a least squares optimization problem. In addition, techniques can be categorized based on their ability or inability to handle nonuniformly spaced data (denoted here by the abbreviations NUDA, UDA, respectively). The selected approaches described here, together with their categorizations based on the above defined abbreviations are as follows:

(1) Regression using discrete orthogonal polynomials (Ref. [23]) (LSA, UDA).

(2) Moving weighted least-squares-based surface representation (Ref. [24]) (LSA or IA, NUDA).

(3) Polynomial-spline-based surface representation (Ref. [24]) (IA, NUDA).

(4) Smoothing-spline-based surface representation (Refs. [25] and [26]) (LSA, NUDA).

Before reviewing surface representation approaches (1) - (4) defined above, it is worth noting that in the application of any of the techniques, fundamental decisions must be made concerning:

(i) The manner in which the 3D surface will be parameterized, i.e., through a general parametric representation for each coordinate function like

that defined in (3-66), or through the more specialized Monge surface representation implied by (3-68) and (3-69).

(ii) The manner in which surface representations will be employed to calculate curvature and surface normal related features.

In the case of (i), the use of the more general parametric surface representation implies the need to solve a surface representation problem for *each coordinate function*. Approaches for constructing a discrete set of 2D parameter vectors $\mathbf{u}_{i,j}$ to be associated with a discrete collection of 3D vector values $\mathbf{f}_{i,j}$ are described in Refs. [27] and [28]. In the case of (ii), there exist two alternative approaches:

(a) 3D vector values $\mathbf{f}_{i,j}$ over some collection of processing windows, each denoted by R_E, can be used to construct a single surface representation over each R_E, which can then be differentiated and employed to obtain curvature and surface normal related features at every enclosed pixel.

(b) By implicitly solving surface representation problems over windows, $R_{E,i,j}$ centered *about each pixel,* 2D convolutional filters may be determined for the *explicit* calculation of first and second derivative information necessary to support calculation of curvature and surface normal-related features. (This approach is only an option when the distribution of data points is uniform in the parameter space.)

The case (a) above enjoys the philosophical advantage of corresponding to the representation of a single well-defined approximating surface for each processing window, R_E. However, the speed with which convolutional filtering operations may be performed suggests a substantial computational advantage to (b).

The above discussion suggests that the adoption of a surface modeling approach based on a Monge representation and a feature calculation approach based on case (b) are favored from a computational standpoint. Hence, most research reported in the 3D object recognition literature makes these assumptions. However, the work reported in Ref. [27] has made use of the more general 3D surface parameterization, as well as a feature calculation approach based on case (a).

The discussion that follows will review surface representation approaches (1) - (4) mentioned above in the context of approximating a 2D function F(x,y), for which discrete values $F_{i,j}$ are available on the discrete grid (x_i, y_j) $i = 1 \ldots N_x$, $j = 1 \ldots N_y$. In the discussion of approach (1), regression using discrete orthogonal polynomials, to be consistent with Refs. [23] and [16], it is assumed that the (x_i, y_j) grid is uniform in both coordinates, and that

$$\begin{cases} N_x = 2M_x + 1 \\ N_y = 2M_y + 1 \end{cases} \qquad \begin{matrix} (3\text{-}99) \\ (3\text{-}100) \end{matrix}$$

Given (3-99) and (3-100), it is possible to define the shifted discrete coordinate parameters

$$\begin{cases} k = i - (M_x + 1) \\ \ell = j - (M_y + 1) \end{cases} \qquad \begin{matrix} (3\text{-}101) \\ (3\text{-}102) \end{matrix}$$

and the discrete data field $G_{k,\ell}$ defined by

$$G_{k,\ell} = F_{k+(M_x+1),\ell+(M_y+1)} \begin{cases} k = -M_x \ldots M_x \\ \ell = -M_y \ldots M_y \end{cases} \qquad (3\text{-}103)$$

Hence, the $(k, \ell) = (0, 0)$ point corresponds to the center of the originally defined discrete data field, $F_{i,j}$.

Now, let $\phi_{xq}(\bullet)$, $\phi_{yr}(\bullet)$ $q = 0 \ldots m_x$, $r = 0 \ldots m_y$ denote orthogonal polynomials associated with x, y coordinate directions, i.e.

$$\sum_{u=-M_x}^{M_x} \phi_{xq_1}(u)\,\phi_{xq_2}(u) = N_{xq_1}\,\delta_{q_1,q_2} \qquad (3\text{-}104)$$

$$\sum_{v=-M_y}^{M_y} \phi_{yr_1}(v)\,\phi_{yr_2}(v) = N_{yr_1}\,\delta_{r_1,r_2} \qquad (3\text{-}105)$$

Then the least-squares approximant $\hat{G}_{k,\ell}$ for the $G_{k,\ell}$ field, when represented in terms of the basis $[\phi_{xq}(\bullet)\phi_{yr}(\bullet)]$ functions, can be expressed as

$$\hat{G}_{k,\ell} = \sum_{q,r} a_{q,r}\,\phi_{xq}(k)\,\phi_{yr}(\ell) \qquad (3\text{-}106)$$

where

$$a_{q,r} = \frac{\sum_{k,\ell} G_{k,\ell}\, \phi_{xq}(k)\, \phi_{yr}(\ell)}{(N_{xq} \cdot N_{yr})} \tag{3-107}$$

If (k, ℓ) are now viewed as continuous parameters, (3-106) can be differentiated to provide expressions for first and second derivatives. In particular, by using the resulting expression for $k = 0$, $\ell = 0$, we obtain convolutional filter forms for calculating first and second derivatives mentioned in the above discussion. Ref [16] provides specific examples of $\phi_{xq}(\bullet)$, $\phi_{yr}(\bullet)$ for the case of discrete orthogonal polynomials that are no higher than second order, i.e., at most quadratic.

The orthogonal polynomial regression based surface representation technique described above is very attractive in a computational sense, but is unable to handle the problem of scattered, unevenly distributed data. The moving weighted least-squares-based surface representation approach described next can handle nonuniformly distributed data, but at the price of requiring the solution to a general linear least-squares problem, and in some cases demanding the solution of a new least-squares problem for each point at which an approximant value is desired. Let $\mathbf{z}$ denote a vector representing a 2D point, and let $\mathbf{b}(\mathbf{z})$ denote a Q-dimensional vector valued function whose components correspond to the 2D basis functions to be employed in the formulation of the least-squares problem, e.g., $x^{p_1} x^{p_2}$ for $0 \leq p_1 + p_2 \leq 2$ specifies a case for which $Q = 7$. In addition, let $\mathbf{z}_m$ $m = 1 \ldots N_x N_y$ denote the collection of 2D (x_i, y_j) pairs and $\mathbf{F}$ denote the corresponding vector of associated function values made up of $F_{i,j}$'s. Then defining the matrices

$$B \triangleq \left(\mathbf{b}(\mathbf{z}_1) \vdots \ldots \vdots \mathbf{b}(\mathbf{z}_{N_x N_y})\right) \tag{3-108}$$

$$W(\mathbf{z}) \triangleq \begin{pmatrix} w(\|\mathbf{z} - \mathbf{z}_1\|) & & \\ & \ldots & \\ & & w(\|\mathbf{z} - \mathbf{z}_{N_x N_y}\|) \end{pmatrix} \tag{3-109}$$

where $w(\bullet)$ is some non-negative weighting function, the approximant function at $\mathbf{z}$, $\hat{F}(\mathbf{z})$ is expressed as

$$\hat{F}(\mathbf{z}) = \mathbf{b}'(\mathbf{z})\, \hat{a}(\mathbf{z}) \tag{3-110}$$

where $\hat{\mathbf{a}}(\mathbf{z})$ is the Q vector solution of the linear least-squares optimization problem expressed by

$$\inf_{\mathbf{a}} \left\{ \left\| W^{1/2}(z)\left[B'\,\mathbf{a} - \mathbf{F}\right] \right\|^2 \right\} \tag{3-111}$$

Recall that the solution of a generic linear least-squares optimization problem was discussed before in the context of 2D autoregressive model fitting, through (3-26) and (3-29). The type of surface approximant obtained in (3-110) varies with the selection of the weighting function w(•). In the case when w(•) is a non-negative constant, the approximant represents an ordinary least-squares surface fit, and $\hat{\mathbf{a}}(\mathbf{z})$ is not dependent on **z**. In the case when w(•) is a non-negative function, bounded at 0, e.g.,

$$w(\xi) = c_1\, e^{-\xi^2/c_2} \tag{3-112}$$

then the approximant represents a least-squares fit that weights data points according to their distance from the point of interest. In the case when w(•) is a non-negative function unbounded at 0, e.g.,

$$w(\xi) = \left[c_1\, e^{-\left\{\xi^2/c_2\right\}}\right] / \xi^2 \tag{3-113}$$

the approximant interpolates at the data point and represents a least-squares fit off the data points. Finally, it should be noted that the smoothness properties of the fitted surface $\hat{F}(\mathbf{z})$ varies with the smoothness properties of the $b_i(\mathbf{z})$'s and the choice of w(•). In the case when the $b_i(\mathbf{z})$'s are 2D polynomials and $w(\xi) = 1/\xi^{2k}$ for $k \geq 1$, $\hat{F}(z)$ has been shown to be infinitely continuously differentiable.

In general, the weighted moving least-squares-based surface representation approach described above requires the solution of a new least-squares problem for each new data point **z** at which $\hat{F}(\mathbf{z})$ is desired. In addition, the calculation of first and second spatial derivatives of $\hat{F}(\mathbf{z})$ requires the calculation of corresponding derivatives of $\hat{\mathbf{a}}(\mathbf{z})$, which in general must be calculated through the solution of linear systems of equations as discussed in Ref. [24].

Approach (2) as described above may be useful as a data gridding approach, i.e., an approach for obtaining a uniformly distributed data set from an irregular or scattered data set. In this case, the application of approach (2) might be followed by a

more computationally attractive interpolatory technique for the calculation of required surface derivatives.

Surface representation approaches (1) and (2) described above adopted a least-squares approximation formulation in a discrete and continuous space setting, respectively. The use of a least-squares approximation approach is motivated by the practical problem of fitting surfaces to noisy data. In situations where data noise effects are small, or the data has been smoothed (see Ref. [29]) or gridded using a least-squares approach like that described in approach (2) above, the use of an interpolatory approach like the polynomial spline surface representation approach (3) may be justified. In the discussion that follows, two approaches to the fitting of cubic polynomial splines are described, based on the use of cardinal splines and B-splines, respectively.

Both the cardinal spline and B-spline approaches are best introduced first in a 1D curve representation setting. The cardinal spline approach is based on representing a function of x, with associated discrete function values f_{xi} $i = 1 \ldots N_x$, by an approximant of the form $\hat{s}_{1D}(\rho)$, where

$$\hat{s}_{1D}(\rho) = \sum_{i=1}^{N_x} \Phi_{xi}(\rho) f_{xi} + \sum_{i=1}^{N_x} \Psi_{xi}(\rho) m_{xi} \tag{3-114}$$

where the m_{xi} denotes slope values, and the functions $\Phi_{xi}(\bullet)$, $\Psi_{xi}(\bullet)$ are piecewise cubic functions with support on $[x_{i-1}, x_{i+1}]$ and satisfying the following conditions:

$$\Phi_{xi_1}(x_{i_2}) = \delta_{i_1, i_2} \qquad \Phi'_{x_{i_1}}(x_{i2}) = 0 \tag{3-115}$$

$$\Psi_{x_{i_1}}(x_{i_2}) = 0 \qquad \Psi'_{x_{i_1}}(x_{i_2}) = \delta_{i_1, i_2} \tag{3-116}$$

Explicit expressions for $\Phi_{xi}(\bullet)$, $\Psi_{xi}(\bullet)$ are obtained in Ref. [24]. Letting $\mathbf{m}_x$, $\mathbf{f}_x$ denote vectors of m_{xi}, f_{xi} values respectively, then the 1D cardinal spline fitting approach involves the solution of a linear system of the form

$$A_x \mathbf{m}_x = [C_x \mathbf{f}_x] \tag{3-117}$$

in order tc guarantee the continuity of second derivatives of $\hat{s}_{1D}(\rho)$ at interval endpoints. The A_x, C_x matrices are tridiagonal matrices, entirely dependent on the

spacing of data points, and in part reflecting the form assumed for boundary conditions at the endpoints x_1, x_{N_x}. Note, that tridiagonal systems of linear equations of the form in (3-117) can be solved by specialized algorithms involving $O(N_x)$ operations ([8]). Letting O_{xi} denote the i-th column of the matrix O_x defined by

$$O_x = A_x^{-1} C_x \tag{3-118}$$

then $\hat{s}_{1D}(\rho)$ can in view of (3-114), (3-117), and (3-118) be expressed as

$$\hat{s}_{1D}(\rho) = \sum_{i=1}^{N_x} H_{xi}(\rho) f_{xi} \tag{3-119}$$

where

$$H_{xi}(\rho) \triangleq [\Psi_{x1}(\rho) \cdots \Psi_{xN_x}(\rho)] O_{xi} + \Phi_{xi}(\rho) \tag{3-120}$$

The $H_{xi}(\bullet)$'s are referred to as cardinal functions. By following through an analogous argument to the above for curve fitting along the y coordinate direction $H_{yj}(\bullet)$'s may defined for j =1 . . . N_y. Given $H_{xi}(\bullet)$'s and $H_{yj}(\bullet)$'s the following 2D spline approximant $\hat{s}_{2D}(\rho, \eta)$ may be defined for the originally defined 2D surface representation problem:

$$\hat{s}_{2D}(\rho, \eta) \triangleq \sum F_{i,j} H_{xi}(\rho) H_{yj}(\eta) \tag{3-121}$$

Note that (3-121) defines a surface representation with the property that derivatives up to second order in each variable are continuous. Hence, (3-121) may be differentiated in order to obtain expressions for desired first and second derivatives.

The cardinal spline approach is attractive since in cases when the cardinal functions and their derivatives can be precomputed, the operation of calculating spline approximant function and derivative values simply involves the evaluation of expressions of the form in (3-121). In the case when the distribution of data is uniform in both variables, convolution filter forms for calculating approximant function and derivative values at the center of a window can be obtained from (3-121).

The cardinal spline approach suffers from the disadvantage that the evaluation of expressions of the form in (3-121) in general requires the summation of all $N_x \cdot N_y$ terms. The B-spline approach avoids this pitfall by assuming finite support basis functions. The B-spline approach for representing a function of x is based on an approximant of the form $\hat{s}_{1DB}(\rho)$ where

$$\hat{s}_{1DB}(\rho) = \sum_{i=-2}^{N_x-1} \alpha_i B_{xi}(\rho) \tag{3-122}$$

where $B_{xi}(\bullet)$ denotes a piecewise cubic basis function, twice continuously differentiable with support on $[x_i, x_{i+4}]$. It should be noted that the representation in (3-122) requires the definition of knots, x_{-2}, x_{-1}, x_0 and x_{N_x+1}, x_{N_x+2}, x_{N_x+3} exterior to the approximant representation interval $[x_1, x_{N_x}]$. The $B_{xi}(\bullet)$ functions can be computed by recursions specified in Ref. [24].

Next, B-spline basis functions $B_{yj}(\rho)$ with support on $[Y_j, Y_{j+4}]$ for $j = -2 \ldots N_y-1$ may be defined and the B-spline approximant for our original surface representation problem may be defined as $\hat{s}_{2DB}(\rho, \eta)$, where

$$\hat{s}_{1DB}(\rho) = \sum_{i=-2}^{N_x-1} \sum_{j=-2}^{N_y-1} \alpha_{i,j} B_{xi}(\rho) B_{yj}(\eta) \tag{3-123}$$

The α_{ij} coefficients in the surface representation (3-123) must be determined by imposing a total of $(N_x+2) \cdot (N_y+2)$ linear constraints. These constraints can be defined in a variety of ways; two options include:

- Matching function values defined at (x, y) knots over the enlarged region $[x_0, x_{N_x+1}] \times [y_0, y_{N_y+1}]$
- Matching function values defined at (x, y) knots over $[x_1, x_{N_x}] \times [y_1, y_{N_y}]$, setting second partial derivatives with respect to x equal to zero along $x = x_1$, $x = x_{N_x}$ for $y \in [y_1, y_{N_y}]$, setting second partial derivatives with respect to y equal to zero along $y = y_1$, $y = y_{N_y}$ for $x \in [x_1, x_{N_y}]$, and finally setting the mixed second partial derivative with respect to x, and second partial derivative with respect to y equal to zero at the "corner" points: (x_1, y_1), (x_{N_x}, y_1), (x_1, y_{N_y}), $x_{N_x}, y_{N_y})$.

In either of the above cases, imposing the described constraints results in a system of linear equations of the form

$$L = \Omega_x \alpha \Omega_y' \tag{3-124}$$

where L contains the constraint function or derivative values, Ω_x, Ω_y are banded matrices that include appropriate function or derivative values associated with x-axis and y-axis B-spline basis functions, respectively, and α contains the $\alpha_{i,j}$ values. Hence, α can be obtained by solving two successive sets of linear systems indicated by the grouping of terms of the below equation:

$$\alpha = \left[\Omega_x^{-1} L\right]\left(\Omega_y'\right)^{-1} \tag{3-125}$$

Given the calculation of α in (3-125), the representation (3-123) can be differentiated in order to obtain expressions for first and second derivative values. Due to the fact that the B_{xi} (•)'s and B_{yj} (•)'s have supports spanning four intervals, the evaluation of the approximant $\hat{s}_{2DB}(\rho, \eta)$ at most requires the summation of 16 items. Finally, note that in cases when the (x_i, y_j) data grid is uniform, (3-123) can be used in conjunction with (3-125) in order to determine convolutional filter forms for evaluating first and second derivatives of the approximant surface, at the center of the *data window (see Ref. [17])*.

The use of an interpolatory spline-based surface representation approach can be criticized from the perspective of ignoring data uncertainty. Hence, in approach (4), the use of spline-based surface representations and least-squares optimization is combined. One example of this approach involving polynomial splines assumes a surface representation of the form in (3-123). Next, it is assumed that a finer "measurement" data grid is superimposed on the surface representation grid defined by the (x_i, y_j) pairs, which includes those points as well as others. In this context, the least squares surface fitting problem can be formulated (see Ref. [15]) as the problem of minimizing

$$\left\|\tilde{L} - \tilde{\Omega}_x \alpha \tilde{\Omega}_y'\right\|_F^2 \tag{3-126}$$

where $\|\cdot\|_F^2$ denotes the squared Frobrenius matrix norm corresponding to the sum of the squares of matrix elements, and $\tilde{L}, \tilde{\Omega}_x, \tilde{\Omega}_y$ have analogous interpretations to L, Ω_x, Ω_y in (3-124). The solution of the optimization over α in (3-126), denoted by $\hat{\alpha}$, can be expressed as

$$\hat{\alpha} = \left[(\tilde{\Omega}_x)^{\#} \tilde{L}\right] (\tilde{\Omega}_y')^{\#} \tag{3-127}$$

where # denotes the Moore-Penrose pseudo-inverse. The result in (3-127) can be obtained by solving two successive sets of linear least-squares problems indicated by the grouping of terms.

A final smoothing spline-based surface representation approach described here is the nonpolynomial spline under tension approach discussed by Cline in Ref. [26], and employed to support curvature related calculations in the 3D object recognition work of Vemuri, Mitche, and Aggarwal in Ref. [27]. Cline's approximant function $\hat{F}$ (ρ, η) is the solution of a variational problem defined minimizing

$$\int_{x_1}^{x_{N_x}} \int_{y_1}^{y_{N_y}} \left[\hat{F}_{xxyy}(\rho', \eta')\right]^2 d\rho' \, d\eta'$$

$$+ \sigma^2 \sum_{i=1}^{N_x-1} \sum_{j=1}^{N_y-1} \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \left[\hat{F}_{xy}(\rho', \eta') - \hat{F}_{xyd,\,i,\,j}\right]^2 d\rho' d\eta' \tag{3-128}$$

where

$$\hat{F}_{xyd,\,i,\,j} \triangleq \frac{\left(\hat{F}(x_{i+1}, y_{j+1}) - \hat{F}(x_i, y_{j+1}) - \hat{F}(x_{i+1}, y_j) + \hat{F}(x_i, y_j)\right)}{(x_{i+1} - x_i)(y_{j+1} - y_j)} \tag{3-129}$$

and subject to the constraint:

$$\sum_{i,j} \left\{ \frac{\hat{F}(x_i, y_j) - F_{i,j}}{\delta_{x_i} \delta_{y_j}} \right\}^2 \leq S \tag{3-130}$$

The S parameter controls the amount of "tension" associated with the resulting surface representation and the δx_i, δy_j are associated with representing the size of data uncertainties. In brief terms, the solution of the above variational problem involves

- The solution of linear equations for $\hat{F}_{xxyy}$ (•, •), $\hat{F}_{xx}$ (•, •), $\hat{F}_{yy}$(•, •) values at grid points, which are coupled with a single nonlinear equation related to the satisfaction of (3-130), and which then allows the construction of $\hat{F}$(•, •) values at grid points.

- The use of a 2D Hermite interpolation formula to represent $\hat{F}(\rho, \eta)$ values given the calculated $\hat{F}(\bullet, \bullet)$, $\hat{F}_{xxyy}(\bullet, \bullet)$, $\hat{F}_{xx}(\bullet, \bullet)$, $\hat{F}_{yy}(\bullet, \bullet)$ values at grid points.

The 2D Hermite interpolation-based surface representation mentioned above may then be differentiated to obtain first and second derivative values at any desired points.

The above discussion has reviewed four alternative approaches to surface representation, which can be employed to support curvature, and surface normal related feature calculation for mine classification. Approach (2), the moving weighted least-squares surface representation, may be most attractive for use in gridding non-uniform data sets, which can then be processed using the more computationally attractive orthogonal polynomial regression or polynomial spline-based surface representation techniques of approaches (1) or (3) in calculating desired first and second derivatives. Approach (4), while having the advantage of combining spine-based surface representation techniques and a least-squares approximation approach, appears to require very complex, computationally intensive algorithms.

The focus of the preceding development has been on the use of *general* surface representation techniques in order to calculate curvature and surface normal related features, as motivated by differential geometry and the 3D object recognition literature. In the present discussion, we recognize the fact that mines frequently have specific forms, i.e., cylindrical or spherical, and show how a more specialized quadric surface fitting technique introduced in Ref. [30] may be employed to calculate shape-related features for use in mine classification.

A quadric surface is represented by an implicit function of three spatial variables of the form

$$\mathbf{f}' \Omega \mathbf{f} + \mathbf{f} \bullet \mathbf{v} + d = 0 \qquad (3\text{-}131)$$

where **f** denotes a vector of x, y, z coordinate values, **v** denotes a constant vector, and d a scalar constant. In the case of a central quadric, i.e. an ellipsoid or hyperboloid, (3-131) can be simplified into the form

$$\mathbf{f}^{*\prime} \Omega \mathbf{f}^{*} + \gamma = 0 \qquad (3\text{-}132)$$

where

$$\mathbf{f}^* = \mathbf{f} + \mathbf{c} \tag{3-133}$$

and

$$\mathbf{c} = -\frac{1}{2}\Omega^{-1}\mathbf{v} \tag{3-134}$$

$$\gamma = \left[d - \frac{1}{4}\mathbf{v}'\Omega^{-1}\mathbf{v}\right] \tag{3-135}$$

Spheres and cylinders may be represented as special ellipsoidal surfaces. In the above context, the suggested shape-related features for use in mine classification are the eigenvalues of (Ω/γ), i.e. $\lambda_i\{\Omega/\gamma\}$ i = 1, 2, 3. These eigenvalues are invariant under 3D rotational transformations for the standardized central quadric surface form obtained by dividing both sides of (3-132) by γ.

At this point, the quadric surface fitting approach suggested in Ref. [30] is described. The fitting process is formulated in terms of ten parameters $a_1 \ldots a_{10}$ which are related to Ω, $\mathbf{v}$, and d as follows:

$$\Omega = \begin{pmatrix} a_1 & a_4/\sqrt{2} & a_5/\sqrt{2} \\ a_4/\sqrt{2} & a_2 & a_6/\sqrt{2} \\ a_5/\sqrt{2} & a_6/\sqrt{2} & a_3 \end{pmatrix} \tag{3-136}$$

$$\mathbf{v} = \begin{pmatrix} a_7 \\ a_8 \\ a_9 \end{pmatrix} \tag{3-137}$$

and

$$d = a_{10} \tag{3-138}$$

In addition, let

$$\mathbf{p} = \begin{pmatrix} a_1 \\ \vdots \\ a_6 \\ \hline a_7 \\ \vdots \\ a_{10} \end{pmatrix} = \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix} \tag{3-139}$$

and assuming a collection of 3D data points $\mathbf{f}_{i,j}$ the fitting problem is formulated as the optimization of E defined as

$$E \triangleq \sum \left[\mathbf{f}_{i,j}' \, \Omega \, \mathbf{f}_{i,j} + \mathbf{f}_{i,j} \bullet \mathbf{v} + d \right]^2 \tag{3-140}$$

over the $\mathbf{p}$ vector variable, subject to the constraint that

$$||\mathbf{p}_1|| = 1 \tag{3-141}$$

The condition (3-141) guarantees a nontrivial solution to the optimization over E.

To define the solution of the above constrained optimization problem, note that individual terms in the summation in (3-140) can be expressed as

$$\left[\mathbf{f}_{i,j} \, \Omega \, \mathbf{f}_{i,j} + \mathbf{f}_{i,j} \bullet \mathbf{v} + d \right]^2 = \mathbf{p}' \begin{pmatrix} B_{i,j} & C_{i,j} \\ C'_{i,j} & D_{i,j} \end{pmatrix} \mathbf{p} \tag{3-142}$$

for appropriately defined matrices $B_{i,j}$ (6×6), $C_{i,j}$ (6×4), $D_{i,j}$ (4×4). Next, defining

$$B_a = \sum B_{i,j} \tag{3-143}$$

$$C_a = \sum C_{i,j} \tag{3-144}$$

$$D_a = \sum D_{i,j} \tag{3-145}$$

then, from Ref. [30], the optimal $\mathbf{p}$ vector partitions, $\hat{\mathbf{p}}_1$, $\hat{\mathbf{p}}_2$ are defined as

$$\hat{\mathbf{p}}_1 = \text{eigenvector of } (B_a - C_a D_a^{-1} C_a') \text{ corresponding to the minimum eigenvalue, } (B_a - C_a D_a^{-1} C_a') \tag{3-146}$$

$$\hat{\mathbf{p}}_2 = D_a^{-1} C_a' \hat{\mathbf{p}}_1 \tag{3-147}$$

Finally, note that the minimum value of the optimization criterion E, E_{min}, can be expressed as

$$E_{min} = \lambda_{min}\left(B_a - C_a D_a^{-1} C_a'\right) \tag{3-148}$$

The above discussion defines an efficient procedure for the least-squares fitting of quadric surfaces that avoids the need for the use of nonlinear optimization techniques. Since E_{min} indicates the quality of the fit, it is suggested here that $\{E_{min}, \lambda_i\ (\Omega/\gamma)\ i = 1,2,3\}$ be the complete set of quadric surface derived features for use in mine classification.

4.0 Conclusions

The preceding development has considered a wide variety of features for use in mine classification, derived from 3D depth map information. These feature calculation concepts are summarized in Figure (4-1). In the case of low-resolution, telesounder-derived depth maps, the focus of our discussion was on the use of shadow information to calculate mine size-related features, motivated by the use of shadows by human interpreters of side-scan sonar data. In the case of high-resolution, swath-bathymetry derived depth maps, both stochastic and deterministic depth map model based feature calculation concepts were identified. The stochastic depth map model based feature calculation concepts were motivated by the literature on texture measures, and random field based approaches to image processing. Finally, the deterministic depth map model based feature calculation concepts were motivated by ideas from differential geometry and the 3D object recognition/computer vision literature.

Low-Resolution Depth Map Feature Calculation Concepts	(1) Mine Diameter Estimation Using Side-Scan Sonar and Depth Map Information
High Resolution Depth Map Feature Calculation Concepts: Stochastic Model-Based Concepts	(1) Autocorrelation Estimates (2) Spectral Density-Related Quantities (3) 2D Autoregressive Model Coefficients (4) Normalized Prediction Error Residuals Based on Fitted 2D Autoregressive Model (5) Spatial Grey-Level Difference Texture Measures (6) Absolute Grey-Level Difference Texture Measures (7) Grey-Level Run Length Texture Measures (8) Fractal Dimension Estimates (9) Wavelet Transform-Related Quantities
High-Resolution Depth Map Feature Calculation Concepts: Deterministic Model-Based Concepts	(1) Mean and Gaussian Curvature-Related Quantities (2) Hough-Transform-Related Quantities Derived From Surface Normal Information (3) Discrete Extended Gaussian Image Derived Quantities Obtained from Surface Normal and Curvature-Related Information (4) Quadric Surface Fit-Related Quantities

REFERENCES

[1] Denbigh, Phillip, "Swath Bathymetry: Principles of Operation and Analysis of Errors," *IEEE Journal of Ocean Engineering*, Vol. 14, No. 4, October 1989, pp. 289-298.

[2] NAVSEA Mine Familiarizer, Naval Mine Warfare Engineering Activity, Yorktown, Virginia, April 1985.

[3] Conners, Richard and Charles Harlow, "A Theoretical Comparison of Texture Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, No. 3, May 1980, pp. 204-222.

[4] Weszka, Joan et al., "A Comparative Study of Texture Measures for Terrain Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 6, No. 4, April 1976, pp. 269-285.

[5] Haralick, Robert et al., "Textural Features for Image Classification," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 3, No. 6, November 1973, pp. 610-621.

[6] Therrien, C. W. et al., "Statistical Model-Based Algorithms for Image Analysis," *Proceedings of the IEEE*, Vol. 74, No. 4, April 1986, pp. 532-551.

[7] Lawson, Charles and Richard Hanson, *Solving Least Squares Problems*, Prentice Hall Inc., Englewood Cliffs, New Jersey, c1974.

[8] Golub, Gener and Charles Van Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, Maryland, c1983.

[9] Stein, Michael, "Fractal Image Models and Object Detection," *SPIE*, Vol. 845, *Visual Communications and Image Processing* - II, 1987.

[10] Peitgen, Heniz and Dietmar Saupe (ed.), *The Science of Fractal Images*, Springer-Verlag, New York, c1988.

[11] Mallat, Stepnare, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," Vol. 11, No. 7, July 1989, pp. 674-693.

[12] Strang, Gilbert, "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review*, Vol. 31, No. 4, December 1989, pp. 614-627.

[13] Besl, P. J. and R. C. Jain, "Three-Dimensional Object Recognition," *ACM Computing Surveys*, Vol. 17, No. 1, March 1985, pp. 75-145.

[14] Lipschutz, Martin, *Theory and Problems of Differential Geometry*, McGraw Hill Book Co., New York, 1969.

[15] Besl, P. and R. Jain, "Intrinsic and Extrinsic Surface Characteristics," *Proceedings of the Computer Vision and Pattern Recognition Conference,* San Francisco, California, June 9-13, 1985, pp. 226-233.

[16] Besl, P. and R. Jain, *Surface Characterization for Three Dimensional Object Recognition*, Report RSD-TR-20-84 Electrical Engineering and Computer Science Department, University of Michigan, Ann Arbor, December 1984.

[17] Yang, H. and A. Kak, "Determination of the Identity, Position, and Orientation of the Topmost Object in a Pile," *Computer Vision, Graphics and Image Processing,* Vol. 36, 1986, pp. 229-255.

[18] Sethi, I. and S. Jayaramamurthy, "Surface Classification using Characteristic Contours," *Proceedings of the 7-th International Conference on Pattern Recognition,* Montreal, Canada July 30 - August 2 1984, IAPR and IEEE Press, New York, pp. 438-440.

[19] Ballard, Dana and Christopher Brown, *Computer Vision*, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1982.

[20] Herbert, M., and J. Ponce, "A New Method for Segmenting 3D Scenes into Primitives," *Proceedings of the 6-th International Conference on Pattern Recognition,* Munich, West Germany October 19-22, 1982, IAPR and IEEE Press, New York, pp. 836-838.

[21] Horn, B., "Extended Gaussian Images," *Proceedings of the IEEE*, Vol. 72, No. 12, 1984, pp. 1671-1686.

[22] Horn, B., *Robot Vision*, MIT Press & McGraw Hill Book Co., Cambridge, Mass. and New York, c1986.

[23] Haralick, R., "Digital Step Edges from Zero Crossing of Second Directional Derivatives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 1, January 1984, pp. 58-68.

[24] Lancaster, P. and K. Salkauskas, *Curve and Surface Fitting: An Introduction*, Academic Press, New York, c1986.

[25] Cox, M., "Data Approximation by Splines in One and Two Independent Variables," pp. 112-138 in A. Iserles and M. Powell (eds.), *The State of the Art in Numerical Analysis*, Clarendon Press, Oxford, c1987.

[26] Cline, A. K., *Surface Smoothing with Splines under Tension*, Report CNA-170 Department of Computer Science, University of Texas; Austin, Texas, January 1981.

[27] Vemuri, B. et al., "Curvature-Based Representation of Objects from Range Data," *Image and Vision Computing*, Vol. 4, No. 2, 1986, pp. 107-114.

[28] Farin, Gerald, *Curves and Surfaces for Computer Aided Geometric Design: A Practical Guide*, Academic Press, Boston, c1988.

[29] Fan, T. et al., "Segmented Descriptions of 3-D Surfaces," *IEEE Journal of Robotics and Automation*, Vol. 3, No. 6, December 1987, pp. 527-538.

[30] Faugeras, O., and M. Herbert, "The Representation, Recognition, and Locating of 3D Objects," *International Journal of Robotics Research*, Vol. 5, No. 3, 1986, pp. 27-52.