INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Multi-scale surface reconstruction in the object space

Doorn, Bradley Dean, Ph.D.
The Ohio State University, 1991
MULTI-SCALE SURFACE RECONSTRUCTION IN THE OBJECT SPACE

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Bradley D. Doorn, B.S., M.S.

The Ohio State University

1991

Dissertation Committee:
Professor Anton F. Schenk
Professor Kurt Novak
Professor Burkhard Schaffrin
Professor Dean C. Merchant

Approved by

Advisor
Department of Geodetic Science & Surveying
ACKNOWLEDGMENTS

I would like to express my deep appreciation to the United States Army for funding my doctoral study. Furthermore, I would like to thank the U. S. Army Engineer Topographic Laboratories, Fort Belvoir, VA. for providing financial support, technical support, and general advice toward my dissertation research.

I am extremely grateful to my advisor, Professor Anton F. Schenk, for his patience, guidance, and sincere interest in my research. Thanks is also extended to Professor Dean C. Merchant who assisted and encouraged me to pursue an doctoral degree. Finally, I would like to thank Professors Kurt Novak and Burkhard Schaffrin for their efforts and comments while serving on my reading committee.

Finally, and most importantly, I would like to express my indebtedness to my wife, Kristen, who patiently endured the long hours and stress that accompanies doctoral research. Also, gratitude goes out to my daughter, Hannah, and my son, Robert, for whom I could not fail.
VITA

23 September 1962 .........................Born – Mitchell, South Dakota

1985 .................................................B. S., Geological Engineering, S. D. School of Mines and Technology, Rapid City, South Dakota

1987 ..................................................M. S., Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio

1987-1988 ..........................................U. S. Army Engineer Officer Basic Course; Platoon leader and company executive officer - 30th Engineer Battalion (Topographic)

1988-Present .................................Graduate student assigned to U. S. Army Student Detachment with duty at The Ohio State University

FIELDS OF STUDY

Major Field: Geodetic Science and Surveying
# Table of Contents

ACKNOWLEDGMENTS ................................................................. ii  
VITA .................................................................................... iii  
LIST OF FIGURES .................................................................... vii  
LIST OF TABLES ....................................................................... ix  

## CHAPTER

<table>
<thead>
<tr>
<th>I</th>
<th>INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>General</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Scope and Objectives of the Research</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Organization of the Report</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>II</th>
<th>STATUS OF AUTOMATIC STEREOMATCHING</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Definition of the Stereo Matching Problem</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Stereo Model Refinements</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Concepts</td>
<td>11</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Techniques</td>
<td>14</td>
</tr>
<tr>
<td>2.3</td>
<td>Similarity Measures</td>
<td>21</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Concepts</td>
<td>21</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Techniques</td>
<td>22</td>
</tr>
<tr>
<td>2.4</td>
<td>Constraints</td>
<td>25</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Concepts</td>
<td>25</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Techniques</td>
<td>27</td>
</tr>
<tr>
<td>2.5</td>
<td>General Problems/General Solutions</td>
<td>33</td>
</tr>
</tbody>
</table>
A Results of Implementation Comparison (SYNTHETIC MODEL) .... 97

B Results of Implementation Comparison (OSU MODEL) ....... 106
## List of Figures

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Scale space/pyramid representation</td>
</tr>
<tr>
<td>2</td>
<td>Warping Images</td>
</tr>
<tr>
<td>3</td>
<td>Multi-scale concept</td>
</tr>
<tr>
<td>4</td>
<td>Traditional, image space strategy</td>
</tr>
<tr>
<td>5</td>
<td>Proposed, object space strategy</td>
</tr>
<tr>
<td>6</td>
<td>Relationship between the image space and object space within the proposed strategy</td>
</tr>
<tr>
<td>7</td>
<td>Image/Object space relief displacement relationship at occlusions</td>
</tr>
<tr>
<td>8</td>
<td>Occlusion detection in the image space via the object space DEM</td>
</tr>
<tr>
<td>9</td>
<td>Window/search area relationship for cross-correlation similarity measure</td>
</tr>
<tr>
<td>10</td>
<td>Reliability factor of cross-correlation similarity measure</td>
</tr>
<tr>
<td>11</td>
<td>Conjugate ray intersection error vector</td>
</tr>
<tr>
<td>12</td>
<td>Bilinear interpolation resampling technique for densifying conjugate locations in the image space strategy</td>
</tr>
<tr>
<td>13</td>
<td>Interpolation from an irregular grid raw surface</td>
</tr>
<tr>
<td>14</td>
<td>Synthetic stereo images and object space model</td>
</tr>
<tr>
<td>15</td>
<td>OSU stereo images</td>
</tr>
<tr>
<td>Page</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>16</td>
<td>Gray value maps of correlation coefficients (scale levels 0-2) (SYN. MODEL)</td>
</tr>
<tr>
<td>17</td>
<td>Gray value maps of correlation coefficients (scale levels 3-4) (SYN. MODEL)</td>
</tr>
<tr>
<td>18</td>
<td>Manually detected breaklines on DEMs (SYN. MODEL)</td>
</tr>
<tr>
<td>19</td>
<td>Occlusions (SYN. MODEL)</td>
</tr>
<tr>
<td>20</td>
<td>Refined DEMs from scale levels 0-2 (SYN. MODEL)</td>
</tr>
<tr>
<td>21</td>
<td>Refined DEMs from scale levels 3-4 and the true DEM (SYN. MODEL)</td>
</tr>
<tr>
<td>22</td>
<td>DEM error magnitudes from scale levels 0-1 and the initial, approximate DEM (SYN. MODEL)</td>
</tr>
<tr>
<td>23</td>
<td>DEM error magnitudes from scale levels 2-4 (SYN. MODEL)</td>
</tr>
<tr>
<td>24</td>
<td>Gray value maps of correlation coefficients (scale levels 0-2) (OSU MODEL)</td>
</tr>
<tr>
<td>25</td>
<td>Gray value maps of correlation coefficients (scale levels 3-4) (OSU MODEL)</td>
</tr>
<tr>
<td>26</td>
<td>High error vector distances from ray intersection (OSU MODEL)</td>
</tr>
<tr>
<td>27</td>
<td>Breaklines detected from LoG operator over DEMs (OSU MODEL)</td>
</tr>
<tr>
<td>28</td>
<td>Occlusions (OSU MODEL)</td>
</tr>
<tr>
<td>29</td>
<td>Omitted match locations (OSU MODEL)</td>
</tr>
<tr>
<td>30</td>
<td>Refined DEMs from scale levels 0-2 (OSU MODEL)</td>
</tr>
<tr>
<td>31</td>
<td>Refined DEMs from scale levels 3-4 and idealized DEM (OSU MODEL)</td>
</tr>
<tr>
<td>32</td>
<td>Raw match point planimetric distributions</td>
</tr>
</tbody>
</table>

viii
# List of Tables

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Matching technique parameters (SYN. MODEL)</td>
</tr>
<tr>
<td>2</td>
<td>Similarity measure quality and error comparison (SYN. MODEL)</td>
</tr>
<tr>
<td>3</td>
<td>Matching technique parameters (OSU MODEL)</td>
</tr>
<tr>
<td>4</td>
<td>Precision comparison with manually derived object space coordinates (OSU MODEL)</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

1.1 General

Automatically measuring and interpreting the physical world or object space from remote sensor data or image space is one of the most challenging problems in many engineering and physical science fields. The tasks of measurement and interpretation, essentially, define photogrammetry as stated in the Manual of Photogrammetry [87]:

...photogrammetry is the art, science and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring, and interpreting photographic images and patterns of electromagnetic radiant energy and other phenomena...

More specifically, surface reconstruction is a component of photogrammetry comprising the determination of three-dimensional coordinates in the object space from sensor information. Automatic surface reconstruction, however, has primarily developed as a sub-problem of Computer Vision in the field of artificial intelligence. Recent advancements in computer technology, e.g., digital sensors, advanced microprocessors, array processors, and the decrease in software and hardware costs, have catalyzed efforts in the photogrammetric sub-field of Digital Photogrammetry and
other disciplines. For instance, automatic surface reconstruction can be applied in most geo-sciences (e.g., forestry, environmental sciences, polar science, geography, and geology), engineering (e.g., surveying, parts inspection, quality control, computer aided design, autonomous navigation, and object tracking), and medical fields (e.g., anatomical movement analysis, micro-biology, and disease or defect diagnoses).

Automatic surface reconstruction has been addressed from basically three types of digital imagery. Active sensors directly represent range information as sensor data which simplifies the problem to that of only interpretation. Monocular imagery utilize the shading or texture information of passive, reflectance image gray values to measure and interpret the object space. Stereo images are perspective images of the same object space where the geometric disparity (parallax) between the conjugate locations directly relates to the geometry of the object space. It is assumed that the digital imagery has been directly obtained from digital sensors or indirectly from digitized photographs. Stereo images are the most popular imagery type because of its potential of robust, quantifiable measurements and its wide application base.

The primary task of automatic surface reconstruction using stereo images is the detection of conjugate locations (stereo matching) in dissimilar images. Automatic stereo matching has, in fact, been isolated as an independent task in most digital photogrammetry and computer vision research. However, numerous problems have limited its success. Hardware limitations restrict the flexibility and speed of computations (see discussions in e.g. [31, 42]). The other primary problem is the complexity of the object space. This complexity amplifies the difference between the stereo images
and may be categorized as geometric complexities e.g., discontinuities and occlusions, and radiometric complexities e.g., image noise and variable object space reflectance and illumination properties. These object space complexities cause stereo matching techniques to find ambiguous or inaccurate conjugate locations.

In effect, successful automatic surface reconstruction over a wide variety of object space surfaces must take into consideration object space complexities before or during stereo matching. This entails expanding the stereo matching problem to a associative set of matching, interpreting, and interpolating tasks. Therefore, while the terms stereo matching and surface reconstruction are used interchangeably in this research and most literature, this research is committed to addressing stereo matching as a sub-problem of surface reconstruction problem.

1.2 Scope and Objectives of the Research

In order to state the objectives of this research it is first necessary to define four terms used throughout this document.

- **concept**... a general hypothesis explaining the premise behind a problem solving strategy.

- **strategy**... a plan or tactic encompassing multiple tasks that when integrated result in a common goal.

- **task**... an intermediate function or purpose that produces a specific output, e.g., matching or interpolation.
technique... an implementation of a task or tasks with a specific algorithm or computation, e.g., correlation matching or moving average interpolation.

The first objective of this research is to present the rationale behind the concept of a proposed multi-scale, object space surface reconstruction strategy. The rationale is provided from a review of the current automatic stereo matching shortcomings. The second objective is to formalize the strategy to encompass a wide variety of matching and contiguous surface reconstruction tasks. The third objective is to implement the strategy and compare it to a more traditional strategy of surface reconstruction. Note that the emphasis of this research is not on improving specific techniques, but to improve the strategy that may encompass and integrate numerous techniques. The implementation comparison strives to use generic techniques that will legitimately contrast the effects of the strategies and not the techniques. The last objective is to analyze the implementation results and judge the rigor of the proposed strategy. This will then project what types of techniques would enhance the effectiveness of the strategy.

1.3 Organization of the Report

Chapter II defines the problem of stereo matching and the current concepts and techniques that are used to approach the problem. Specifically, the techniques developed in Digital Photogrammetry and Computer Vision are presented concurrently within the definition of the stereo matching problem. The result is a general look at where and why current techniques fail and succeed. From this analysis Chapter III
presents and rationalizes two hypotheses (multi-scale and object space) that will improve the performance of the current matching techniques and integrate them into surface reconstruction. Chapter IV describes a proposed strategy and its tasks that will encompass the hypotheses described in Chapter III. This strategy is compared to a traditional strategy of stereo matching developed from techniques described in Chapter II. Chapter V compares an implementation of the proposed and traditional strategies using fundamental and proven techniques. Chapter VI provides conclusions and comments on the proposed strategy and the directions of future research.
CHAPTER II

STATUS OF AUTOMATIC STEREO MATCHING

The status of automatic stereo matching is an analysis that objectively defines the stereo matching problem, the tasks that are required to solve the problem, and the current concepts and techniques used to perform the tasks. The result of this analysis will be the theoretical framework needed to deduce hypotheses for improving automatic stereo matching and, thus, surface reconstruction.

Automatic surface reconstruction has been driven by two major fields of research: Digital Photogrammetry and Computer Vision. Unfortunately, until recently these fields have worked independently. To analyze the full scope of automatic surface reconstruction it is necessary to first distinguish and relate the efforts made in both fields.

2.1 Definition of the Stereo Matching Problem

Digital Photogrammetry has recently developed as a legitimate sub-field of photogrammetry [79]. Classical photogrammetric research has been traditionally tied to physics, optics, and numerical analysis. The primary goal has always been to make quantifiable and highly accurate point measurements which will lead to accurate de-
scriptions of the object space. A qualitative interpretation of images has been left to
the human operator. It is understandable that when approached with the problem of
automatic surface reconstruction the emphasis is on analytically modelling the imag­
ing process and precisely defining conjugate points from the gray values (area-based
matching). Therefore, support for digital photogrammetric techniques has rested on
the photogrammetrist's expert knowledge of the image to object space model and the
statistical analyses of the results obtained from area-based techniques.

Computer vision has developed from the artificial intelligence field. Artificial intel­
ligence research is traditionally tied to the cognitive sciences and information process­
ing. The primary goal has always been to manage large and complex knowledge bases
and derive meaningful, qualitative descriptions of the object space. Therefore, the
problem of automatic stereo matching has generally been approached from the tasks
of object recognition, representation, and correspondence (feature-based matching).

Much support for the use of features is founded in biological evidence of the early
human vision system. Briefly, psychophysical evidence shows that light sensitive
photoreceptors (rods and cones) in the retina amass in clusters of variable sizes called
receptive fields. Each receptive field is connected to a single retinal ganglion cell
which conveys information to the visual cortex. The signal fed to the ganglion cell
is a combination of brightness signals from the photoreceptors. The response of the
ganglion cell is a function of the change in brightness (edges) over the area covered
by the receptive field. The size of the receptive field determines the acuity (scale
or resolution) of the edge. The result is an edge map of multiple edge features at
variable scale levels [46, 60]. These edge maps from both eyes are then matched at
higher levels of the human vision system. This evidence has been the driving force
behind developments of a computational theory of human vision [61, 58, 28, 49]. In
turn, this computational theory has been the basis for most computer vision research
in stereo matching. An introspective explanation of the human visual process for
non-experts of the neuro-sciences is given by Hubel [44].

It is hoped that by reviewing the current research in both fields that clues may
be derived that will help determine an improved strategy for stereo matching. See
Dhond and Aggarwal [19] for a detailed review of stereo matching from a computer
vision perspective and Doorn [20] for a review from a photogrammetric perspective.

In order to analyze this collection of research from different perspectives it is useful
to step back and analyze stereo matching as the process of making observations that
refer to a theoretical abstraction of the physical world called a model. In this sense,

- The *model* consists of perspective stereo images which are formed by a geometric
  model (perspective and relief properties) and a radiometric model (reflectance
  and illumination properties) of the object space.

- The unknowns or *parameters* of this model are the exterior orientation val-
  ues of the stereo images and the conjugate locations of the same object space
  event. Assuming the exterior orientation parameters are known, the conjugate
  locations define the geometric model.

- The *observations* are similarity measures between stereo images which determine
  the degree of resemblance between the gray values (or gray value properties) of
However, by definition stereo matching is a non-trivial inversion problem in which the three-dimensional object space surface cannot be uniquely defined from dissimilar images of that surface. Put in another way, the observations lack sufficient information to overcome image noise and an undefined radiometric model to uniquely and completely find a solution (conjugate locations) which would define the geometric model. Therefore, stereo matching is an ill-posed problem.

Currently, the ill-posed problem of stereo matching is addressed by assuming or deriving knowledge about the coincident object space, e.g., surface shape and locations of discontinuities. These assumptions are then applied in two forms: stereo model refinements and restrictions of the solution space (constraints).

- **Stereo model refinement** reduces the effects of noise and the effects of the geometric and radiometric model between the stereo images. This manifests in locally redefining the gray values or expanding (making more explicit) the gray value properties from each image. A single, local refinement of the model which comprises a single similarity measurement is called a primitive.

- **Constraints on the solution space** reduce the possibility of multiple ambiguous solutions. Constraints occur as approximate conjugate locations and a permitted spatial deviation (threshold) from the approximate.

Therefore, the accuracy of the assumptions has a significant effect on the matching process. Note that model refinements and constraints may be derived from the
same assumptions. Together, hopefully, the similarity measure or matching criteria (observations) can resolve ambiguities to find accurate, precise, and reliable conjugate locations (parameters). In this context accuracy will signify a general closeness to the 'truth', precision will define an absolute deviation of conjugate locations from an independent measurement [76], and reliability will define the uncertainty of the results.

To analyze the current status of automatic stereo matching, concepts and techniques are discussed in three sections: Stereo Model Refinements, Similarity Measures, and Constraints. Interestingly, the goals of the techniques in these sections correspond to the three measures for stereo matching described by Barnard and Thompson [9]: distinctness, similarity, and consistency.

Each section will address the shortcomings and benefits of current techniques which will provide information on where and how improvements should be made, Section 2.5. For organizational purposes all techniques are sub-divided into the two major matching vehicles of area-based and feature-based techniques. Furthermore, multiple scale level (scale space) techniques of deriving and applying knowledge from different resolutions of the images are treated separately in the Model Refinement and Constraints sections. Scale space representations are currently used by most notable stereo matching techniques in area-based matching and feature-based matching. These techniques are also termed hierarchical, pyramid, coarse-to-fine, and multi-resolution techniques.
2.2 Stereo Model Refinements

2.2.1 Concepts

Stereo Model refinement redefines the original gray values as primitives representations. A primitive is a local description of the gray values which is used in a similarity measurement. The goal is to reduce the geometric and radiometric distortions between the images and detect primitives that relate to distinct object space events. These primitives are derived from assumptions about the object space. The types of primitives are determined by what and how much assumptions are utilized.

A fundamental assumption in most matching techniques is that the epipolar geometry is known. Epipolar geometry assumes that the images have been relatively oriented, transformed, and resampled to a common coordinate system. Therefore, the geometric distortions between the stereo images are restricted to the x-axis or image rows [82]. Primitives are now extracted from these epipolar images.

In the general case no object space information is assumed; therefore, the only option is to directly use the image gray values. Gray value primitives are represented in pixel format that cover an area of the object space; thus described as area-based.

Some general knowledge may be assumed such that the model may be expanded to make object space events more explicit. For instance, geometric discontinuities, reflectance changes, and illumination changes, are represented in the images as abrupt gray level changes or edges. This relationship is the basis of most feature-based primitives.

Both area-based and feature-based primitives have one common obstacle: noise.
Therefore, prior to feature extraction smoothing is needed to remove high spatial frequency variations that corrupt the primitive description. However, there is no standard to select the amount of smoothing that is necessary. Too much smoothing eliminates important object space information and too little smoothing leaves too much information for the primitives to discern. Therefore, the effects of smoothing in relation to the properties of the image must be more accurately defined.

Marr [58], and Koenderink and Van Doorn [49], expound on the nature of smoothing by explaining that in most cases image information encompasses a wide range of spatial frequencies. No one operation can optimally address information at all spatial frequencies simultaneously. In natural images spatial frequency will be locally dependent upon variations in the object space (see above) and globally dependent upon the scale (resolution) of the image. For instance, in one image a tree does not exist at the same scale as its leaves, nor that of the forest which it is a part. Also, a scale of a building imaged from a low flying airplane does not, most likely, exist at the same scale as the same building imaged from a satellite. Since the object space is unknown, spatial frequency diversity caused by object space events cannot be addressed without making unrealistic assumptions. Therefore, an intuitive solution is to deal separately with a range of different global scale levels. This solution is also supported from biological evidence of variable sizes of receptive fields which define the scale or acuity of the information that is transmitted to the visual cortex, see Section 2.1.

A multiple scale level (multiple resolution) image representation of the same ob-
ject space consists of a sequence of fixed, separate images represented at successively coarser scale levels (reduced resolutions). Each scale level is determined from a smoothing operator of a particular size. The basic assertion is that information in the coarser scale levels is implicit and roughly localized in the finer scale levels. Therefore, by first matching the more distinguishable, low frequency primitives in the coarse scales, less distinguishable high frequency primitives in the finer scale levels can be isolated.

Witkin [89] expanded the multiple scale level structure by describing a sequence of continuous scale levels, termed the scale space. The motivation comes from an assumption that an edge at some spatial location over two or more successive scale levels is the result of a particular object space event (spatial coincidence assumption [58]). Therefore, tracing edges through a continuum of scale levels results in a qualitative description or fingerprint [92] of the image structure. This continuous scale level representation is an attempt to manage the problem of scale by directly eliminating the scale dependency of primitive descriptions at single scale levels. Unfortunately, these descriptions will be very complex, especially in two-dimensions and a practical application of this concept in stereo matching has yet to be produced. Even so, the continuous scale space domain provides the stimulus for future model refinement research which can determinatively and completely describe object space events. Therefore, in this document scale space will refer to any primitive description that spans multiple scale levels.
2.2.2 Techniques

Area-based primitives

For the sake of definition, gray value primitives can be seen as being extracted from a \(1 \times 1\) operator of unit value. A singular gray value primitive is a small amount of information for any meaningful interpretation. Therefore, a window of multiple gray values is selected as a primitive, expanding the matching model. This has the effect of increasing the area and amount of information of the primitive as related to the object space. The window size and sampling pattern are usually chosen arbitrarily or empirically (see e.g. [75]) or from feature operators which determine the distinctness of the gray values, see Section 2.2.2.

The gray value primitives may be further expanded by redefining the gray values of one image relative to its stereo mate. Geometrical reshaping removes geometric differences by transforming gray value locations by affine transformation parameters\(^1\). The parameters are defined from the parallax values within the primitive. Resampling is necessary to reestablish the format of the image. Reshaping may be performed locally for each primitive [66, 1] or globally for an entire image [65, 72]. The radiometry of the gray value primitives is neglected or normalized which assumes a constant gray value difference between conjugate primitives.

\(^1\)An affine transformation is the most widely used; however, other transformations may be used.
Feature-based primitives

Feature-based techniques seek to identify and localize a particular image property. There are two primary types of low level, feature-based operators: interest operators and edge operators. Each of these operators evaluate local gradients of the gray values as a discrete form of a differentiation process. Interest operators describe the statistical nature of a point from its neighboring gray values and edge operators specifically detect local changes in the gray value gradients.

Interest operators detect point features by evaluating the gray values surrounding a point. Moravec [64] discusses a directional variance operator that selects points with the largest minimum variance of neighboring gray values in four directions. Förstner [25] uses an error ellipse, the axes of which are the elements of the covariance matrix, to measure the gradient distribution surrounding the point.

Interest operators are statistical measures of a neighborhood of gray values and provide little explicit information about the object space. Instead, these operators are useful as statistical cues that augment the descriptions of gray value or feature primitives before measuring similarity. For instance, Hannah [36] uses a maximum-variance interest operator to identify gray value primitives that have sufficient and distinct gray value information to find a unique match by cross-correlation. As a matter of organization, while interest operators are technically point-feature operators, they are used primarily as area-based techniques and rarely as stand-alone primitives.

Edges are, by far, the most widely used feature primitives. Edge operators detect local changes in gray values and may be classified into two types: first derivative and
second derivative operators.

First derivative operators detect edges as the local maxima of the first derivative, which is approximated by the gradient of the gray values. A straightforward implementation is to convolve the image with a series of operators that respond to gradients in different directions. Two such operators are the Roberts and Sobel operators. Canny [12] developed a more sophisticated operator which optimizes the performance based on three criteria: good detection, good localization, and one response to each edge. These operators directly result in edge locations along with their magnitudes (value of maxima), and orientation. The drawback of these techniques is that they utilize the gradient in only one direction at a time. Therefore, a set of passes must be made with different differentiation orientations to find the actual edge.

Second derivative operators detect edges as the zero crossings of the second derivative. The second derivative is most commonly approximated by the Laplacian of the gray values. The Laplacian operator is orientation-independent which means a single operation will define the edges. The Laplacian operator forms a convolution surface where the edge elements (zero crossings) are contours of that surface; providing boundaries are included. The magnitude and orientation must be determined separately from the zero crossings.

Edge operators are, by definition, noise enhancers; therefore, it is extremely important to smooth the images prior to detection. The Laplacian of the Gaussian (LoG) operator integrates the Laplacian operator with the Gaussian smoothing op-
erator [59]. The Gaussian is discussed in the next section. The LoG operator was first suggested from analysis of the human vision system, Section 2.1, that shows the center-surround organization of photoreceptors in a receptive field approximately mimics the LoG. Some other advantages are:

- Scale level representations can be determined by simply controlling the size of the Gaussian. This makes the LoG an extremely popular edge operator for scale space representations.

- Zero crossings are discrete forms of closed contours and, therefore, simplify the tasks of sorting and tying scattered edge elements to form more descriptive edge segments.

As stated above edge element descriptions can be expanded by defining characteristics of a series of connected edge elements, an edge segment, with shape and extent characteristics. For instance, Schenk and Hoffman [83] describe the use of edge segments for relative orientation, Schenk [85] describe edges as piece-wise linear segments in the $\psi - s$ parameter space. Medioni and Nevatia [63], Lim and Binford [52], and Perlant and McKeown [70] describe rectilinear edge segments assuming these types of edge segments are directly related to man-made structures, e.g. buildings.

$^2$Actually, the human vision system more closely mimics the difference of two Gaussians (DoG) at different scale levels
Scale space representations

A scale level image representation is derived by smoothing the gray values to a certain resolution. The size of the smoothing operator determines the resolution and, therefore, the amount of information (spatial frequency variation) that is to be represented. The tasks of producing scale space primitives are as follows: select the scale levels needed, smooth the image with the appropriate operator sizes, and extract the primitives (see previous sections).

Optimal scale level selection remains an open question. The most widely used design is a resolution reduction of $2^i$ where $i$ equals integer values (0, 1, 2, 3,...) see Figure 1. For instance, a square image size of 4096 will be decomposed into scale levels of size (4096, 2048, 1024, 512,...). The grid interval equals the amount of information reduction from the original image.

The smoothing technique requires proper selection of a smoothing operator. Marr and Hildreth [59] outlined the two considerations for selecting a smoothing operator. First, the operator must reduce the range of scales which the image represents, i.e., it must be localized in the frequency domain. Second, the operator must also be localized in the spatial domain (see also [49]).

Common smoothing operators are equal-weighted, moving average operators in the spatial domain and ideal low pass operators in the frequency domain (see discussions in [27, 6]). These operators are adequate for most rudimentary image enhancement tasks, e.g. noise reduction. They do not, however, adequately satisfy both considerations of spatial and frequency localization. For instance, a simple equal-weighted,
Figure 1: Scale space/pyramid representation with corresponding grid interval, scale level, and resolution parameters.
moving average operator is localized in the spatial domain; however, it creates sidelobes in the frequency domain (not bandlimited). An ideal low pass operator is bandlimited in the frequency domain; however, it creates sidelobes in the spatial domain (not spatially localized). In fact, there is no operator that strictly satisfies both of these considerations.

There is one operator that does optimize the conflicting relationship between spatial and frequency localization: the Gaussian operator. The reason is that a Gaussian in the spatial domain transforms to a Gaussian in the frequency domain. Therefore, in most images the Gaussian operator prevents spurious detail from being generated as the resolution diminishes. Stated in another way coarse scale level images contain no information not implicitly present in the finer scale levels. This is essential to reduce ambiguity when tracking between scale levels. These and other beneficial properties (and limitations) of the Gaussian operator in the scale space have been rigorously determined using the LoG edge operator in [89, 5, 48, 92, 45, 10]. Applications have also extended to tracking corners of edge segments [73, 38].

It is important to note that while the Gaussian operator is, overwhelmingly, the most popular smoothing operator, recently other operators have shown favorable properties in the scale space. For instance, morphological operators [14], a modified Bessel operator [53], and a Gabor operator [78] have recently been used. Also, adaptive smoothing operators are currently being developed that vary the amount of smoothing spatially within one representation [77, 71, 80]. Similar in premise to morphological operators, adaptive smoothing adjusts the operator size according to
local gray value information. Therefore, adaptive smoothing attempts to represent
different spatial frequencies which are caused by local object space events in one scale
representation, see Section 2.2.1.

After smoothing has been performed the primitives are extracted from the coarsest
scale level and similarities are measured, Section 2.3. The results (conjugate locations)
are used as assumptions for redefining the model and selecting the match primitives
at the next finer scale level.

A scale space primitive description is used in almost all relevant stereo matching
algorithms. For example, feature primitives in the scale space are used in [28, 90, 2,
8, 38, 88, 68] and gray level primitives in [36, 72, 50, 7, 65].

2.3 Similarity Measures

2.3.1 Concepts

Similarity measures are simply tools to measure the difference between individual
primitives in one image with primitives in its stereo mate. As discussed in Section 2.1
constraints are needed to restrict the location of conjugate primitives. This section
discusses only similarity measurements assuming that proper constraints will be im-
plemented, see Section 2.4.

The basis of the similarity measures is that conjugate primitives with the least
difference relate to the same object space event. The goal, as with any measuring tool,
is to provide accurate and reliable results. The type of measurement (observations)
depends on the type of primitive used to describe the model and how it is represented.
Consequently, the benefits and limitations of the definition of the model, Section 2.2,
directly affects the similarity measure (observations).

In area-based matching the primitives are represented as discrete signals or functions that cohesively and completely (except for occlusions) cover an area of the object space. Therefore, similarity is a straightforward measurement of the differences between gray values.

In contrast, feature-based primitives are distinct, qualitative information that represent certain object space events. Therefore, feature primitives are represented as a structural description that symbolizes the primitives and, most importantly, the spatial interrelationships among primitives in each image. The similarity measure of a single feature primitive attempts to correspond the feature primitives between stereo images at the expense of distorting the spatial interrelationships between neighboring primitives. The cost of distorting the interrelationships is, essentially, a constraint on the possible locations of conjugate primitives, see Section 2.4.

2.3.2 Techniques

Area-based measures

Two techniques are used to determine greatest similarity between stereo gray value primitives: maximizing the correlation coefficient with cross-correlation or minimizing the squared gray value differences by a least squares adjustment.

Cross-correlation selects a gray value primitive (window of gray values) in one image and correlates it directly on its stereo mate (see e.g. [66, 36]). The conjugate location in the stereo mate is chosen as the center of the gray value primitive at the
point of the highest correlation coefficient. The correlation coefficient is a measure of linear dependency between two primitives and is strongly affected by the gray value variation within each primitive. Therefore, homogeneous or repetitive texture, e.g., a sand dune, a parking lot, or an orchard, will result in high correlation coefficients for multiple locations. This problem is addressed by increasing the spatial size of the primitive or by only matching primitives with high variances selected from interest operators, Section 2.2.2.

Least squares matching (LSM) envelops the use of adjustment theory applied to gray level primitives. LSM was first introduced by Ackermann [1] and Förstner [24].

In its basic form LSM selects a gray value primitive from both images. The gray value differences are the observations. The geometric model is usually approximated by an affine transformation of gray value locations in one stereo mate. LSM must assume a direct and mathematical relationship between the observations and the model. In linearized form, the observation equations state that the gray value gradient in one or both images directly affects the transformation parameters. Therefore, after approximate conjugate primitives (the estimate) are selected the gray value primitive in the stereo mate is redefined (reshaped) from the updated transformation parameters. The process is iterated until the gray value differences are minimized and the conjugate locations are selected as the center points of the final gray value primitives (windows).

The major drawback of LSM is that the types of primitives needed are restrictive in many situations. For example, an assumption in LSM is that the gray value
gradient and the geometric model are directly related. This forms somewhat of a photometric model (see [43, 29, 30]) of the object space where gray values are highly correlated to the object space geometry and, therefore, correlated to each other. If this is the case highly correlated observations will also result in a statistical analysis that is suspect. If, on the other hand, the gray values are somewhat random (high variance), the linear dependency between the model and observations will be corrupt. Gruen and Baltsavias [32] suggests data snooping techniques to analyze the quality of the similarity measures.

LSM has the potential of obtaining highly precise results, up to .05 pixels [50]. Furthermore, LSM has the favorable quality of being able to integrate multiple models (refinements and constraints) into a unified adjustment. For instance, Helava [41], Wrobel [91], and Ebner [22] introduced a conceptual framework of a LSM technique that combines the similarity measure with a global model of surface reconstruction. The global model integrates an enhanced radiometric model and geometric model by matching gray value primitives from more than two images that are projected to object space facets. Unknown or incomplete models reduce the practical viability of this technique. However, recently, Heipke [39] implemented this concept and demonstrated the potential and comprehensive nature of a global approach.

Feature-based measures

The most widely matched feature primitive is the simple edge element. Matching these primitives is, at first glance, very simplistic. Either the edge element coincides
with an edge element in its stereo mate or not. This may be performed by simply correlating a primitive over primitives in its stereo mate. This, in fact, is the similarity measure used in relaxation [58, 51], regularization [88], and simulated annealing [8] techniques. Obviously, a simple on/off correlation of primitives will produce numerous ambiguous matches. The emphasis of these techniques, however, is not in the similarity measure. Instead, the similarity measure is weighted against spatial constraints to distinguish between ambiguous similarity measures, see Section 2.4.2.

If the primitives are edge segments or more advanced edge depictions, shape information (e.g., orientation, extent, curvature) may be used to distinguish edge element conjugate locations. For example, Medioni and Nevatia [63] and Perlant and McKeown [70] use linear edge segments, Boyer [11] matches structural descriptions of the edge segments, and Schenk [85] matches entire edge contours (entities) using the $\psi - s$ domain of the zero crossings of the LoG operator. However, more advanced descriptions of the primitives require a more advanced data structure (e.g., chain codes, quad trees, and semantic nets) and search techniques (e.g. consistent labelling [37]).

2.4 Constraints

2.4.1 Concepts

Match constraints are restrictions placed upon conjugate primitive locations. Therefore, constraints are assumptions about the geometry of the object space. One exception is the assumption of epipolar geometry, Section 2.2.1. Epipolar geometry restricts the geometric distortions between the stereo images to the x-axis; therefore, it constrains the space of possible matches to the x-axis or single rows in the stereo
mate.

All other constraints consist of two components: selecting approximate conjugate locations or starting points and selecting the space of possible conjugate locations. The starting conjugate locations are the initial approximations of the geometry of the object space. The space of possible conjugate locations, also called a search area, is the relief range that the geometry might deviate from the initial approximation. Both components may be determined from three sources.

- Independent knowledge, e.g., human interaction or a previously determined elevation model.

- Restricting the parallax difference between neighboring primitives (smoothness constraint [9]).

- Derived from coarser scale levels in the scale space.

The sources are the same for both area-based and feature-based techniques. The techniques differ only in how they are implemented and how well the assumed geometric model fits the model definition, Section 2.2. For instance, edge primitives detect object space changes; therefore, it is reasonable to assume that object space geometry is smooth between the edges and smooth along edge segments (configural continuity constraint [62]). In contrast, gray value primitives have no object space information; therefore, the search areas are merely independent thresholds with no association to the model definition.
2.4.2 Techniques

Area-based constraints

Cross-correlation implements constraints by simply selecting a search area and centering it on the starting point in one image. The correlation coefficients are then determined only within that search area. LSM centers two identical windows on the starting points and then constrains the shift parameters of the model (affine transformation) within the adjustment.

Norvelle [66] also predicts starting points by projecting a continuous trend from previous conjugate locations. Rosenholm [75] introduces continuity constraints between neighboring primitives using LSM. This technique matches multiple primitives simultaneously while constraining the search areas between the primitives from the gray value windows' normals (see also [50]).

Gruen and Baltsavias [34] introduce the object space points into the LSM model by constraining the conjugate locations in the collinearity equations. Constraining the search area in the object space is also presented in the conceptual global matching technique, see Section 2.3.2.

Straight-forward implementations of area-based matching techniques require a small search area to resolve similarity measure ambiguities, depending on the type of object space. Therefore, the starting points must be highly accurate. For instance, Norvelle [66] empirically determined that a search area of seven pixels is necessary for reasonable results from cross-correlation. Heipke [39] estimated that a search area of as few as one to three pixels is needed in standard LSM. With added constraints
the minimal search area may increase to as many as six to ten pixels as cited by Gruen and Baltsavias [33] or five pixels as cited by Heipke [39]. Even so, the need for accurate approximations further restricts the practical usefulness of many area-based matching techniques to interactive systems [66, 40].

Feature-based constraints

In its basic form feature-based constraints may be implemented alike to cross-correlation. The search area is centered on the starting point in one image and the possible conjugate primitives are restricted to that area (Marr-Poggio-Grimson algorithm discussed in [47]). The possible matches can then be further constrained by forcing the parallax of edge segments to vary continuously (configural continuity, Section 2.4.1).

More sophisticated algorithms have recently been developed for constraining edge elements under the topic of optimization. Relaxation techniques, regularization, simulated annealing, 2.3.2 and waveform analysis matching [56, 93] are all forms of optimization. The basic premise of these techniques is to distort the spatial relationships between neighboring primitives until a suitable correspondence occurs. The goal is to retain the basic structure of the model description (primitives) by placing a penalty on the similarity measure as the spatial relationships are distorted. For instance, there may be multiple similarity measures of equal value for a particular primitive. The conjugate location is chosen as the one that provides the least change in parallax (least resistance) between neighboring primitives.

This technique of cost/penalty functions may be more restrictive if more sophisti-
cated primitives are used. For instance, Medioni and Nevatia [63] penalize similarity measures that do not conform to rectilinear edge primitives describing the model of a building.

Scale space constraints

The question that remains when implementing constraints is how do we reasonably determine these constraints with the limited information usually known prior to matching. For instance, in an extreme case the base distance may be the only way of selecting starting points; therefore, a large search area is needed. Today, the solution is to measure similarity at a coarse scale level, which reduces the number of different possible similarity measures. The resulting crude and sparse conjugate locations may be used as starting points to the next finer scale level (tracking) and the search area can be reduced. The process is repeated until the finest scale level which will have highly refined starting points and, therefore, small search areas.

Most notable matching algorithms use some form of the scale space to reduce the amount of information needed prior to matching (see references, Section 2.2.2). Most area-based techniques use empirically derived thresholds on the search areas, where some feature-based techniques adjust the penalties placed upon the similarity measures through the scale space. The final conjugate locations then converge to the space of least resistance (cost). Recently, emphasis has been placed upon including the scale parameter (size of Gaussian operator) as a variable within the similarity measure [35, 90]. The latter techniques approach the problem of selecting constraints
for arbitrarily selected scale levels by making very small (theoretically continuous) steps through the scale space. The drawback is the large amounts of processing necessary for a large number of scale levels.

In summary, the current emphasis of automatic stereo matching research is on refining the model so that the similarity measures can reliably and accurately define conjugate locations without the need for restrictive constraints. While the scale space representation significantly improves the performance of both area-based and feature-based techniques, the conflict between finding a unique and complete solution that is also metrically accurate remains to be found. This can be seen in a summary of the disadvantages and advantages of area-based matching, feature-based matching, and scale space representations.

**ADVANTAGES:**

- **Area-based:**
  
  - dense object space description
  
  - potential of obtaining highly precise similarity measurements
  
  - well known statistical techniques can be used for measuring and analyzing similarity

- **Feature-based:**
  
  - explicit representation of object space events
  
  - the premise of smoothness constraints is justified from the explicit, structural description of object space events in feature primitives.
— less sensitive to radiometric and geometric distortions, therefore, more reliable - (Note that features are contaminated by radiometric and geometric distortions indirectly from the gray values they are detected.)

• Scale space:
  — the scale space provides a more distinct and complete representation of the image structure
  — the scale space reduces the need of unrealistically accurate assumptions

DISADVANTAGES:

• Area-based:
  — implicit representation of object space events
  — large gray value primitives and highly constrained multiple primitives provide more information but may over-generalize the object space as being a plane (flat plane without reshaping, tilted plane with reshaping) or a continuous surface (smoothness constraints); therefore, accurate descriptions result only on small scale images and/or images over slowly varying terrain
  — highly sensitive to noise, geometric distortions and radiometric variations
  — the premise of smoothness constraints is not justified from any object space information; therefore, constraints may over-generalize the object space (less accurate)
• Feature-based:

  - sparse object space description

  - features are qualitative descriptions requiring a more sophisticated (at least to photogrammetrists) representation

  - similarity measures usually result in less precise conjugate \(^{3}\) primitive locations

  - implementation of constraints may result in excessive computational burden, e.g. simulated annealing [90]

  - the high priority of constraints on feature-based primitives may still result in an over-generalized object space, especially over highly dynamic geometries, e.g. over urban terrain

• Scale space:

  - multiple scale levels significantly add to computational and storage requirements

  - there is no substantive criteria for selecting scale levels

  - there is no substantive criteria for selecting search areas for a particular scale level

\(^{3}\)While sub-pixel precision is obtainable, e.g. interpolating gradient slopes in edges, most computer vision research does not strive for geodetic surveying standards.
2.5 General Problems/General Solutions

The previous sections describe current concepts and techniques within the concep­
tual problem of stereo matching. From this analysis three general problems can be
surmised.

1. Assumptions needed for successful matching of a well-distributed and accurate
   surface description are either incomplete or unrealistic.

   A scale space representation reduces the need for unrealistically accurate assump­
tions. However, current techniques use solely the resulting conjugate locations to
refine the model and constrain the space of possible solutions. Is there more informa­
tion in the scale levels which can be used to more determinatively refine and constrain
the model?

2. Given the assumptions, geometric and radiometric distortions, especially in
   large scale images over urban terrain, have not been adequately modelled in
   both area-based and feature-based techniques.

   Currently, matching techniques seek to model the differences between images.
   Area-based techniques utilize an affine transformation for geometric differences and a
   constant or, at best, a linear transformation for radiometric differences. Feature-based
   techniques depend on features, e.g. edge elements, that are, hopefully, independent
   of geometric and radiometric distortions. However, these techniques form a model
   that is still tied to the original image space and is incapable of eliminating all of the
perspective, relief displacement and reflectance distortions. Instead of modelling the differences between the images, would this problem be addressed better by modelling the source of the distortions, the object space, directly from each image?

3. On a broader scope, implementation of image matching as part of the entire surface reconstruction process has not been fully adopted.

Currently, matching techniques have been addressed as independent tasks which satisfy only certain applications. This has resulted in a multitude of techniques each with different needs (assumptions) and different outcomes. For instance, the previous sections demonstrated the mutually exclusive nature of accuracy vs. reliability between and within area-based and feature-based techniques. However, there is no algorithmic function that can rigorously distinguish the needs of the application from the techniques available and the information at hand. In other words, the human operator cannot be replaced by an algorithm. Can a more global and conceptual approach to matching, and surface reconstruction as a whole, help formulate rules and a strategy that may begin the process toward a photogrammetric expert system? (see discussions in [81, 79])

This chapter has presented the theoretical framework and status of automatic stereo matching. Based on this chapter, Chapter III will present specific arguments supporting two hypotheses that will attempt to answer the questions posed in this section.
This chapter will present two hypotheses and their rationale for improving automatic surface reconstruction. These hypotheses were derived from past literature in Digital Photogrammetry, see Chapter II, with a notion to formalize a coherent strategy for surface reconstruction and not redefine the techniques used to perform surface reconstruction tasks.

3.1 Object Space Matching

3.1.1 Hypothesis 1

Matching stereo images that are first projected to the object space will improve similarity measure quality, enhance knowledge of the object space during matching, and optimally utilize information from prior surface reconstruction tasks.

3.1.2 Rationale and Significance

Chapter II explained that current matching techniques attempt to model and eliminate the geometric and radiometric differences before or simultaneously with (LSM, Section 2.3.2) measuring similarity. These techniques may be defined as operating in the image space environment. Therefore, an inherent property in all image space
environment techniques is that operations are performed on perspective projections of the ultimate outcome, the true surface. In fact, it is likely that if the approximate conjugate locations are the 'truth', the image space matching techniques would still have difficulty. Current matching techniques are simply unable to accurately and globally define the image-to-image relationship from the image space environment.

Since the goal of stereo matching is to remove the distortions resulting from images of the same object space, a logical environment for matching is the object space. Object space properties have been used in the past to constrain matching techniques [65, 34]. The actual matching, however, is always performed in the image space environment. Note that the global approach to LSM, Section 2.3.2, introduced the concept of simultaneously recovering geometric and radiometric distortions by matching object space representations (groundels) of multiple images in an all-in-one adjustment. While this concept is of notable interest, the global approach is too ambitious and inclusive for realistic application.

Using the same argument for object space representations as the LSM global approach, Schenk [84] introduced a more practical and open scheme for utilizing object space properties in stereo matching. In contrast to the image space environment techniques, the model is first expanded by projecting the image space gray values back to the object space (object space environment). As shown in Figure 2, the results are a three-dimensional location and two gray values from each stereo image for each surface grid (called groundels in the LSM global approach). If a true DEM is known, the result is two orthophotos; identical except for radiometric differences
and at occlusions. Obviously, only an initial, approximated DEM is known; therefore, the two object space representations will differ and are termed \textit{warped images}. The warped images' differences are projection differences analogous to the image space distortions, only to a much less degree. The object space environment will fully and globally compensate for geometric distortions from the approximate DEM.

The warped images and their coinciding surface coordinates most exemplify the object space from the known information. While unknown radiometric distortions are not yet addressed, the object space matching environment is open for any matching, interpretation, and image processing technique. Furthermore, the techniques are enhanced by providing a common, goal-oriented datum, the object space, for knowledge and computations to interrelate.

Performing stereo matching in the object space environment should improve surface reconstruction for the following reasons:

- The quality of the similarity measures should be enhanced by reducing, to the maximum extent, geometric distortions between the stereo images. A higher quality similarity measure will result in more reliable results. Also, better distortion removal will reduce the need for ambiguous constraints and, therefore, increase the robustness and flexibility of matching techniques.

- The three-dimensional, object space environment will create a fundamental base upon which all surface reconstruction tasks can operate and support each other. Therefore, known or derived knowledge of the object space surface can be used to directly support surface reconstruction tasks. For instance, absolute, three-
Figure 2: Object space representations (warped images) of the original images.
dimensional surface characteristics may be directly utilized as the matching model definitions or constraints and, furthermore, information from independent sources, e.g., other sensors or ground surveys, may be used directly to define these characteristics. In contrast, the image space environment reveals absolute three-dimensional surface information only after matching tasks have been performed.

• Creating the object space environment exploits knowledge from the prior surface reconstruction tasks; interior, relative, and absolute orientation. Assuming these tasks have been performed is consistent with most stereo matching applications. That is, common assumptions are that epipolar geometry is known and the image-to-object space transformation is known or performed separately.

3.2 Multi-Scale Analysis

3.2.1 Hypothesis 2

Matching warped images and interpreting the results at multiple scale levels (scale space) will reduce the need for extensive pre-matching knowledge or assumptions of the object space, reduce ambiguous conjugate locations, improve convergence to the true object space, and provide valuable information at each scale level which can be used to support surface reconstruction tasks at finer scale levels.

3.2.2 Rationale and Significance

Chapter II discussed the enhanced model description of the scale space. However, while scale space representations significantly improve the performance of stereo
matching techniques, problems remain. Lu and Jain [54] outlined two primary shortcomings of the scale space:

1. Arbitrary selection of scale levels,

2. Ineffective rules to combine information at different scale levels.

Koenderink [48] comments on the first problem stating explicitly,

If you have no a priori reasons to look for certain features, then you cannot decide on the "right scale" ... you must treat the image on all levels of resolution simultaneously.

Therefore, the continuous scale space, Section 2.2.1, is the only complete way of choosing intermediate scale levels; though it can be redundant and computationally expensive. Witkin [90] addresses the first problem applied to stereo matching by introducing the scale parameter into the matching function to track conjugate point locations through the continuous scale space. Hahn [35] includes the scale parameter in a LSM technique to match windows of different scale size. While a continuous scale space representation is the ideal solution, it is closed to any significant interpretation at individual scale levels.

The second problem may be effectively controlled by knowledge based reasoning through the scale space. The goal is to isolate certain events or objects from the image; not simply remove high frequency noise. Therefore, the behavior of the events or objects must be known to reason in the scale space. For instance, if the goal is
to isolate a certain edge, the behavior of that edge in the scale space will provide important clues for isolating that edge.

How can knowledge based reasoning through the scale space be applied to surface reconstruction? The events or objects to be isolated are three-dimensional object space events. Analysis of the behavior of these surfaces requires a surface representation at each scale level. For instance, if geometric boundaries could be isolated as primitives, the behavior of these boundaries could be determined through the scale space. This behavior could then be formulated as a rule to help isolate three-dimensional object space events.

At this point it is important to emphasize that the type of stereo image matching, Section 2.2-2.4, is the basis for DEM sampling. The sampling pattern of matching dictates the interpolation complexity and accuracy of the DEM. Morphological sampling, composite sampling, and sampling by data compression capture those points which are most representative of the surface [55, 13, 69]. Progressive sampling starts at a coarse grid and densifies points as a function of the surface relief. Related work in terrain classification analyzes known DEMs based upon their scale characteristics [16, 21, 26, 4]. Therefore, successful isolation and representation of the object space depends not only on detecting optimal matches, but matching relevant locations of the object space. In principle sampling may be incorporated in scale space analysis by progressively selecting and densifying object space events if a three dimensional object space representation is available at each scale level.
The latter two arguments (isolating object space events and sampling), however, both rely on the ability to distinguish the localized spatial frequency tendencies of object space events, Section 2.2.1. In essence, the assumption is that scale level primitive representations correspond to a scaled (smoothed) geometry of the object space. In other words, they assume that smoothing gray values or radiometry directly effects the object space geometry. As stated in Chapter II this assumption is rarely true and is the basis of the ill-posed problem of stereo matching.

While a direct extension of the continuous scale space and scale space reasoning may not be feasible, the principle still applies. Analyzing task results at multiple scale levels in the object space will provide knowledge important for driving subsequent tasks and interpreting the final surface. This multi-scale concept in the object space is visualized in Figure 3. To distinguish this concept from past scale space applications, this concept will be called multi-scale analysis.

In summary, performing stereo matching in the object space environment and analyzing the results over multiple scale levels will improve surface reconstruction for the following reasons:

- Stereo matching over systematically decreasing scale levels has long been proven to reduce the need for pre-match assumptions and to decrease the number of false matches. Furthermore, convergence to the true object space will only be enhanced in the object space environment where warped images are refined to orthographic representations of the same object space.
Figure 3: Multi-scale concept using object space information.
• Multi-scale representations of the refined model (warped images) will reduce the presence of geometric distortions between the image representations that are still present in the scale space representations in the image space environment. Distortion removal before smoothing, in effect, is dynamic smoothing of the gray values as a function of the geometric model.

• Goal-oriented information from previous scale levels will provide knowledge that can be used in surface reconstruction tasks at subsequent scale levels. Also, interaction is not limited to single tasks. Results from matching, space resection, classification, and interpolation tasks can be used to constrain each other.

The hypotheses described and supported in this section must now be tested. Chapter IV will describe a strategy that will test the practical relevance of the hypotheses.
CHAPTER IV
MULTI-SCALE, OBJECT SPACE MATCHING STRATEGY

A strategy is a sequence of tasks that describe the steps necessary to obtain a desired result. Implementation occurs when techniques are assigned to perform these tasks, see Chapter V. The multi-scale, object space matching hypotheses described in Chapter III declare that improvements can be made by expanding traditional stereo matching strategies. Therefore, to begin testing the hypotheses a comparison of a traditional, image space strategy and the proposed strategy is presented.

4.1 Traditional, Image Space Strategy

Figure 4 illustrates the sequence of tasks performed in most scale space, stereo matching strategies. Refer to Chapter II for details on the scale space concept and related stereo matching techniques.

In brief review, a fundamental model refinement and constraint are derived from the initial assumptions of known relative orientation parameters forming the epipolar geometry. The scale space image representations are constructed from a smoothing operator of various sizes. The matching task envelops the extraction of primitives, similarity measures, and constraints performed on each scale level beginning at the
Initial assumptions

- Stereo images, exterior orientations, and approximate conjugate locations.

smooth images

scale space representations

match scale level(i) representation

- Corrected conjugate locations

finest scale level?

Yes

Final DEM

No

density to scale level(i) grid

- Raw object space surface

Interpolate to scale level(i) grid

Figure 4: Traditional, scale space strategy in the image space.
coarsest scale level \(i\). If \(scale \ level \_i\) is not at the finest (highest resolution), matching is repeated at the next finer \(scale \ level \_i\)+1. The refinements and starting conjugate locations are determined by simply densifying (interpolating) the resulting conjugate locations to \(scale \ level \_i\)+1. If \(scale \ level \_i\) is the finest, the conjugate locations are projected to the object space and subsequently interpolated to a DEM.

### 4.2 Proposed Strategy

Figure 5 illustrates the proposed sequence of tasks for a multi-scale, object space strategy. Notice that the initial assumptions and the ultimate goal are the same as the image space strategy. What differs is the environment in which stereo matching occurs.

In the proposed strategy the initial assumptions of epipolar geometry are expanded by using the absolute orientation parameters, Section 3.1. An approximate object space surface can be derived from the exterior orientation parameters by intersecting the known conjugate locations from the orientation procedure. Therefore, as few as the minimum 5-6 Von Gruber conjugate points from manual methods or as many as 100 conjugate points from automatic methods [85] may be used to define the initial object space surface. Whatever the number, an initial, approximate DEM can be interpolated. Note that absolute orientation must also be known in traditional strategies in order to ultimately derive the object space.

The next step is transforming the original stereo image gray values to orthogonal

---

1Conceptually, the relative orientation parameters may be used, where the conjugate locations are intersected in a model space instead of the object space.
Figure 5: Proposed, multi-scale strategy in the object space.
projections, warped images \(^2\), of the object space. Novak [67] describes and compares three digital techniques for indirectly transforming the image gray values to the object space: projective rectification, polynomial rectification, and differential rectification. Differential rectification is the most practical technique and produces the most accurate results for both aerial and satellite images when a DEM is available. Furthermore, each gray value of a warped image will be directly assigned a three-dimensional coordinate of the object space. This provides a more convenient format to relate the warped images to the object space.

The result of differential rectification (warping) is that each \((X, Y)\) location will have an elevation value and two gray values from each stereo image, Figure 6. The difference between the gray values is due to the error in the elevation at that location.

An attribute of rectification is the ability to detect and control occlusions in the stereo images. Occlusions are areas of the object space not visible from one or both of the images. For a central perspective image this occluded area occurs when the relief displacement along a radial line from the image nadir eclipses the object space surface further along the same radial.

Figure 7 demonstrates how the occluded area can be detected along a single radial line in ideal images. \(R_1, R_o\) and \(R_2\) are radial, object space points projected to radial, image points \(r_1, r_o\) and \(r_2\). The occluded point \(R_o\) can be detected from the relationship between its nearest, visible radial neighbor that is closer to the object space nadir, \(R_1\). If its projected radial distance, \(r_o\), is less than the projected radial distance, \(r_1\), \(R_o\) is occluded from \(R_1\) and is not visible in the image. The proceeding

\(^2\)Warped images are orthogonal projections of an approximate DEM, Section 3.1.2.
Figure 6: Relationship between the image space and object space within the proposed strategy.
Figure 7: Image/Object space relief displacement relationship at occlusions.

R2 \geq R1 \text{ and } r2 \geq r1 --- R2 \text{ is visible}

R_o \geq R1 \text{ and } r_o < r1 --- R_o \text{ is not visible}
points on the same object space radial are occluded until the image radial distance is greater than the last visible point radial distance, \( r_1 \).

The two warped images are now smoothed to a pre-selected scale level as in the image space strategy, Chapter II. The difference is that only one scale level is produced at a time. This is because the warped images will change from scale level to scale level and will not be tied to the original approximation of the object space.

The warped images can then be matched as in traditional matching tasks. The difference lies in how the refinements and constraints are derived. The primitives are directly extracted from the warped images with no need for reshaping since the warped images have globally and completely eliminated all known geometric distortions. Constraints on the conjugate locations can now directly utilize approximations of object space geometric properties. The starting conjugate locations are simplified to be the same \( X, Y \) location in each warped image. The search areas will restrict the possible conjugate locations along the epipolar plane defined by the projection centers \( L, R \), and DEM point \( D \), Figure 6. The size of the search area in the warped images can be directly derived from an approximated elevation deviation range. Schenk [82] derived orthographic position changes, \( p_Y, p_X \), in relation to elevation deviations, \( h \), the focal length of the image, \( f \), the position perpendicular to the base line connecting the projection centers, \( y_d \), and the stereo images' base distance, \( b \).

\[
p_y = \Delta h \times \frac{y_d}{f}
\]  \hspace{1cm} (4.1)
Notice that in contrast to the epipolar constraint in the original images, the search area may not restrict matches along single rows of gray values. While this creates a technical difficulty during implementation, the epipolar constraint is still valid and can be enforced as a new implementation technique.

The matching task will produce a matching vector\(^3\) between three-dimensional conjugate locations which signifies the geometric errors in the approximate \(DEM_i\), Figure 6. The three-dimensional conjugate locations \(D\) and \(PD\) are projected to their respective projection centers. If the conjugate locations were constrained to the epipolar plane then the two projection rays, \(LD\) and \(RPD\), will intersect. The ray intersection point, \(P\), will be a new \(X, Y, Z\) coordinate and collectively a irregular grid, raw object space surface.

The interpolation task computes a refined, approximate \(DEM_{i+1}\) from an irregular grid, raw object space surface. A grid format is needed to perform the differential rectification and is more convenient for DEM analysis, Section 3.2.

There are numerous interpolation techniques which are applicable for different levels of raw surface sparsity and complexity. The most common techniques used are weighted moving average operators and moving surface operators. Weighted moving average operators are brute force techniques which assume the grid points are the local maxima or minima. Moving surface operators overcome this limitation by fitting a

\[ px \approx \Delta h \times \frac{bx}{f} \]  

\((4.2)\)
plane or higher order function to the raw surface. The results of the above techniques will vary depending on the search pattern. The type of search pattern determines where and how many raw surface values are used for one grid point e.g., nearest points in any direction, a nearest point from each quadrant direction, or a nearest point in each octant direction.

The interpolation task is significantly enhanced by known geometric properties of the object space surface. These properties are determined from a surface analysis of the approximate \(DEM_i\) that results from the previous scale level or initial assumptions. Schut [86], Clarke [17], and Davis [18] all provide discussions on these and more advanced interpolation techniques.

In contrast, since traditional strategies remain tied to the original image space the conjugate locations resulting from the \(scale\ level_i\) can be directly utilized in \(scale\ level_{i+1}\) by densifying conjugate locations in each image. This has the benefit of retaining known conjugate locations for proceeding scale levels. Interpolating to a grid in the object space may not precisely retain known conjugate locations. For instance, the mere presence of occlusions indicates that the interpolation technique is 'predicting' the character of the object space which cannot be supported by similarity measures.

The new approximate \(DEM_{i+1}\) can now be analyzed (surface analysis) in the three-dimensional object space. This knowledge can be used to adjust matching parameters in the proceeding scale levels, e.g., sampling rates, window sizes, levels of constraint, and types of interpolation. For instance, an important object space event
are geometric discontinuities. The existence and location of geometric discontinuities provides a more explicit description of the $DEM_{i+1}$ which can be used to more accurately match and interpolate future DEMs. In depth discussions of DEM analysis are given in [17, 18, 3, 26].

$DEM_{i+1}$ is then used to rectify two new warped images and the tasks are repeated $i = i + 1$. If the finest scale level has been reached, the raw surface is interpolated to a refined $DEM_i$ grid interval and the matching is finalized.

In summary, this section described the strategies and tasks pertaining to a traditional, scale space matching in the image space and a proposed, multi-scale matching in the object space along with some common techniques for additional tasks in the latter strategy. A comparison reveals that the traditional strategy is tied to the image space and the direct results of the similarity measure. The actual object space is not disclosed until the matching is completed. However, in the proposed strategy the object space is immediately explicit for directing the tasks and interpreting the results.

Most importantly, the proposed strategy integrates the tasks of image matching, interpolation, and surface analysis. The LSM global approach, see Section 2.3.2, suggests a simultaneous adjustment where these tasks are performed implicitly during the minimization to an optimal solution. In contrast, the proposed strategy integrates these tasks and their results iteratively by examining and modifying the similarity measure results at each scale level. Therefore, the image representations dynamically adjust to object space geometric and semantic information gained at each scale level.
The importance of the interpolation and surface analysis tasks cannot be over-emphasized. Simple techniques are used in this research, see Chapter V, to demonstrate how, otherwise unused, surface reconstruction tasks can be integrated into automatic stereo matching and improve the results. It is immediately apparent, even before experiments have been performed, that more sophisticated techniques must be investigated. Surface analysis and interpolation of DEMs over multiple scales during stereo matching is the topic of parallel research currently being performed in the Department of Geodetic Science and Surveying.
CHAPTER V

EXPERIMENTS AND RESULTS

The difference between the traditional strategy and the proposed strategy discussed in Chapter IV can be generalized as a difference between matching environments: image space and object space. This chapter will present an implementation comparison that emphasizes the effects of the matching environments, not the matching techniques. Therefore, the goal is to compare the two strategies by standardizing, as much as possible, the implementation techniques and test data.

Section 5.1 will describe and compare the techniques used for implementing the traditional, scale space strategy in the image space and the proposed, multi-scale strategy in the object space. Section 5.2 will describe the test data and Section 5.3 will analyze the results.

5.1 Techniques

The strategy tasks outlined in Figures 4-5 are implemented with techniques that facilitate a legitimate comparison between strategies, are well-known and tested, and are reasonably efficient.
5.1.1 Initial Assumptions (image space/object space)

The first task to implement is the selection of model refinements and constraints from the same initial assumptions. The assumptions are that the exterior orientation parameters are known \(^1\), a single elevation is known in the object space, and a maximum relief range from that elevation. It is presumed that the stereo images have been transformed and resampled into epipolar images from the exterior and interior orientation parameters.

In the image space strategy the initial starting conjugate locations are formatted to the grid of the left image and shifted by a constant parallax value in the right image. This parallax value coincides with the initial elevation approximate. In the object space strategy the initial elevation value is expanded to a DEM. This initial, approximate \(DEM_i\) covers the object space area in question.

5.1.2 Rectification (object space only)

The proposed strategy performs a differential rectification of both images to the approximate \(DEM_i\). Differential rectification projects a three-dimensional coordinate representing one object space grid of the approximate \(DEM_i\) to the image space. The projection equations are the collinearity equations defined by the exterior orientation coordinates of the image.

\[
x = x_p - f \frac{r_{11}(X - X_o) + r_{21}(Y - Y_o) + r_{31}(Z - Z_o)}{r_{13}(X - X_o) + r_{23}(Y - Y_o) + r_{33}(Z - Z_o)},
\]

\(^1\)Known exterior orientation parameters suggest that the interior, relative, and absolute orientations have been performed
\[
  y = y_p - f \frac{r_{12}(X - X_o) + r_{22}(Y - Y_o) + r_{32}(Z - Z_o)}{r_{13}(X - X_o) + r_{23}(Y - Y_o) + r_{33}(Z - Z_o)},
\]

where, \( x, y \) \ldots image space fiducial coordinates

\( X, Y, Z \) \ldots object space (DEM) coordinates

\( X_o, Y_o, Z_o \) \ldots camera coordinates

\( r \) \ldots camera rotation matrix

\( x_p, y_p, f \) \ldots interior orientation of camera

Note that a camera distortion model may be added to these equations to enhance accuracy. If necessary, the resulting fiducial coordinates are transformed to pixel locations using the parameters from the scanning instrument. Since the pixel location will not usually correspond to an integer value, a gray value is interpolated by resampling using bilinear interpolation. To reduce the effects of resampling and provide a consistent comparison with the image space strategy, the tessellation of the warped images and the original images are kept the same.

Occlusion detection is performed by mimicking the relief displacement precept of central perspective images, Section 4.2. An intermediate result of differential rectification is a collection of irregular, points in the image space which represent the projected locations of three-dimensional points from the DEM. Therefore, a simple radial line comparison is not applicable.

Instead, Figure 8 shows that the pixel locations and object space grid locations are first determined relative to the image and ground nadirs, respectively. The pixel
Figure 8: Occlusion detection in the image space via the object space DEM. DEM grid points projected to image pixel points.
and grid locations are represented as their center points. A pixel location, \((x_a, y_a)\), projected from an object space grid, \(A\), is compared with the pixel locations \((b, c, d)\) of the adjacent object space grids nearest the ground nadir \((B, C, D)\). If the pixel location satisfies the following condition its respective object space grid is occluded.

\[
(x_a < x_b \text{ and } y_a < y_b) \text{ or } (x_a < x_c \text{ and } y_a < y_c) \text{ or } (x_a < x_d \text{ and } y_a < y_d)
\]  

Note that if a portion of the object space grid is occluded the entire grid is flagged as an occlusion. The resulting occluded areas in both warped images are then omitted from the matching technique of the current scale level and saved for interpreting the results, Section 5.1.7.

5.1.3 Smoothing (image space/object space)

The scale levels are produced by convolving the images with a Gaussian smoothing operator, Section 2.2.2.

\[
G(R) = \frac{1}{\sigma} \exp \left( -\frac{R}{2\sigma^2} \right)
\]  

\[
w = 2\sqrt{2}\sigma
\]

where,  

- \(R\) \ldots radial distance to center of operator  
- \(\sigma\) \ldots space constant  
- \(w\) \ldots operator size
The size of the operator \( (w) \) relative to the Gaussian is chosen as the diameter of the excitatory region of the LoG function. Figure 1 shows how the scale levels are chosen prior to matching. In this implementation the scale levels represent \( 2^i \times 2^i \) resolutions of the images where \( i = (0,1,2,3,4) \). The grid interval \( (1,2,4,8,16) \) depicts the size of the smoothing operator \( (w) \) and, thus, the reduction factor of independent information relative to the original image. The image space strategy may create all scale level representations before measuring similarity. However, in the proposed strategy the smoothing is performed on the warped images after rectification at each scale level.

5.1.4 Matching (image space/object space)

The next task to be implemented is the matching task: primitive extraction, similarity measurement, and constraints. The primitives are gray value windows (area-based) where the size of the window and search area are predefined by human intervention and are the same for both strategies. The window sizes are chosen empirically. Those that resulted in the most accurate final DEM from a series of image space strategy implementations on the same data set were chosen for a comparison with the proposed strategy. The sampling rate is defined as the grid interval of the current scale level.

The initial search area is calculated by transforming the assumed relief range to parallax values in the image space strategy and from Equations 4.1-4.2 in the proposed strategy. This search area is used on the initial, coarsest scale level matching. The proceeding scale level search areas are sequentially reduced by a constant factor.
The similarity measure is area-based cross-correlation. As described in Chapter II, cross-correlation has numerous benefits and shortcomings. It is important to re-emphasize that the goal of this research is not to improve a particular matching technique, but to improve the matching strategy in which the techniques are implemented. In no way is cross-correlation advocated or opposed as a similarity measure from this research. Instead, cross-correlation has been chosen as a base test technique to examine the effects of matching in the object space. Therefore, cross-correlation has been chosen as a similarity measure for the following reasons.

- Correlation is a well known and tested technique in all scientific fields.

- Correlation is easy to implement and evaluate.

- Any area-based matching technique has extreme difficulty over complex, discontinuous geometry, Chapter II, which makes it an ideal technique to test the rigor of the matching strategy over such geometry.

The correlation technique used is developed from the stereo matching algorithms developed by Raye F. Norvelle at the U.S. Army Engineer Topographic Laboratories [66]. A window of gray values centered on a point in the left image is correlated with every possible position in a search area in the right image, see Figure 9. The position with the highest correlation coefficient from Equation 5.5 will define the match position in the right image. A function is fit to the maximum correlation value and it's surrounding values to interpolate the position to a sub-pixel value.
Figure 9: Window/search area relationship for cross-correlation similarity measure.

\[
\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R} = \frac{\sum_{k=1}^{n} (L_k - \bar{L})(R_k - \bar{R})}{\sqrt{\sum (L_k - \bar{L})^2 \sum (R_k - \bar{R})^2}}
\]

where, \(L, R\) \ldots image windows
\(\bar{L}, \bar{R}\) \ldots mean gray value of windows
\(n\) \ldots number of pixels in the window
\(\sigma_{LR}\) \ldots covariance of images' windows
\(\sigma_L, \sigma_R\) \ldots standard deviations of each image window

To further assist the image space strategy the right image gray value primitives are reshaped, Section 2.2.1. An affine transformation is used where the parameters are
The outcome of the similarity measure technique are conjugate locations, the correlation coefficient, and a reliability factor. The correlation coefficient may be termed the quality of the match and indicates the goodness of fit of the left image window to a patch in the right window. The reliability factor is the distinctness or uniqueness of the conjugate locations and is inversely proportional to the size of the angle, $\alpha_x$, see Figure 10. The angle $\alpha_x$ is formed from the tangents at two points surrounding the maximum correlation value$^2$. The reliability factor will be low in areas of repetitive texture or areas where the model is extremely inaccurate and all matches are poor quality. Each of these circumstances will result in near horizontal tangents. Conjugate locations in both strategies are omitted if the quality

---

$^2$Two angles are determined in the column and row directions. The largest angle is chosen as $\alpha_x$. 

Figure 10: Reliability factor of cross-correlation similarity measure.
Figure 11: Vertical, top-down, view of non-coplanar conjugate rays. Error vector indicates the shortest distance or normal between the rays.

or reliability of the similarity measure is low. These thresholds are determined as the thresholds that produce the most accurate results from a series of implementations of the image space strategy.

5.1.5 Ray intersection (object space only)

The resulting conjugate locations in the object space are three-dimensional coordinates in each image, $(D, PD)$, Figure 6. Rays (vectors) are then projected from the projection centers, $L, R$, of the images to their respective object space points, $D$, $PD$. The ray intersection position, $P$, of these two vectors is calculated giving an ad-
justed three-dimensional position which is the surface correction to the approximate $DEM_i$. If the two vectors do not intersect (non-coplanar), an error vector that is normal to each of the conjugate rays is calculated indicating the position where the rays are the closest, Figure 11. This normal vector is zero for rays that are coplanar. The mid-point of this error vector is the adjusted three-dimensional position. The result is an irregular grid, raw surface and a file of error vector distances. The error vector distances are used to omit raw surface points which deviate from the epipolar constraint in object space matching indicating wrong conjugate locations.

5.1.6 Interpolation (image space/object space)

*Image space* interpolation is performed to densify the conjugate locations from the current grid interval to the grid interval of the preceding scale level. The conjugate locations are interpolated separately for each image. The conjugate locations remain in grid format in the reference, left image. Omitted conjugate locations (see above) are simply replaced in the left image. In the right image they are linearly interpolated between present conjugate locations along the scan (epipolar) line (see e.g. [66]). Figure 12 then shows that parallax values are assigned to each grid point by subtracting the conjugate locations of the right image corresponding to the left image grid. The parallax values are then resampled by bilinear interpolation and redistributed as conjugate locations in the right image. Matching is now repeated at a finer scale level ($i = i + 1$).

In contrast, *object space* interpolation must first transform the irregular, raw surface to a refined $DEM_i$ at the current grid interval. A weighted moving average
Figure 12: Bilinear interpolation resampling technique for densifying conjugate locations in the image space strategy. Resampling is performed on the parallax values and redistributed as conjugate locations.
technique is used for its efficiency and flexibility. A grid point elevation is computed from the nearest surrounding elevations in each quadrant with weights inversely proportional to their radial distances from the grid point, Figure 13. The quadrant search pattern was selected so as not to be overly influenced by clusters of elevations. The size of the operator is selected as twice the size of the approximate current grid interval, e.g. for a grid interval of 8 the interpolation operator diameter is 16. This operator size was chosen as to not over-generalize the surface, at least as much as possible with a simple moving average technique, while assuring a usage of a reasonable number of elevations. If no elevation is found in one or more of the quadrants, the nearest elevation beyond the extent of the operator in that quadrant is used.

If a breakline is predicted in the approximate DEM$_i$ from the surface analysis in previous scale level (discussed shortly), the interpolation technique will not select any elevations opposite of the breakline. The result is a refined DEM$_i$ at the current scale level.

If the current scale level is not the finest, the refined DEM$_i$ is densified by bilinear interpolation to the grid interval of the proceeding scale level$_{i+1}$. Performing interpolation as a two step technique reduces the interpolation correlation effects and standardizes the densification technique with the image space strategy (see above).

Note that in image space matching DEMs are produced by space intersection using the collinearity equations for each scale level for comparison purposes. The irregular grid to grid interpolation technique is the same as in object space matching, without segmentation from breaklines.
Figure 13: Interpolation from an irregular grid raw surface for one grid point. Search pattern will not transverse breaklines.
5.1.7 Surface Analysis (object space only)

Interpolation and densification results in an approximate $DEM_{i+1}$, which is analyzed for relevant information to be used in proceeding scale levels. As stated in Chapter II geometric discontinuities are a major obstacle in stereo matching. This research does not confront the greater problem of surface analysis rigorously. Instead, surface analysis in this experiment is an initial attempt to determine what knowledge can be gained from scale level DEM representations and how can it be exploited in proceeding scale levels.

Therefore, in one data set geometric discontinuities are automatically derived and interactively detected from the geometry of $DEM_{i+1}$, occlusions, and low quality matches. In the second data set the discontinuities are detected as the zero crossings from a convolution of $DEM_{i+1}$ with a LoG operator (see LoG discussion in Section 2.2.2). Both techniques result are continuous edges which segment the DEM into low relief regions separated by breaklines. These breaklines are used to drive the interpolation technique in the proceeding scale level, Section 5.1.6. Matching in the object space is now repeated at the next finer scale level ($i = i + 1$).

Note that the surface analysis directly affects only the match results in the interpolation technique and not the matching technique parameters. This is an important point since almost any matching technique can perform successfully if the matching parameters, i.e. type and size of primitives, similarity measures, and constraints, are adjusted to the type of object space surface.
5.2 Test Data

Two data sets were selected to include complex object space geometric events, i.e. discontinuities and occlusions, that arise in large scale aerial images over urban terrain and cause the most difficulty to automatic stereo matching. The first data set is a synthetic stereo model. The synthetic model was constructed to reduce the effects of similarity measure errors on the results. Synthetic data also provides a ‘truth’ to compare results often not present in aerial image matching research.

5.2.1 Synthetic Stereo Images

The synthetic stereo images will emphasize problem areas of the geometric model commonly found in large scale aerial images over urban terrain, which include discontinuities, occlusions and planar to non-planar surfaces, Figure 14c. The techniques described in Section 5.1 standardize, as much as possible, the variables of the matching techniques between image space and object space strategies. The only major variable that will severely affect the results is the similarity measure (cross-correlation) errors. Section 2.3.2 described the ambiguous results that occur when the variance of the gray value window is low and in areas of repetitive texture. To reduce ambiguous results a radiometric model is constructed of random gray values. Therefore, the variance of each window will be large and consistent and the covariance between the two windows will be uniquely large only on the correct conjugate location. While this unrealistically simplifies the stereo matching problem, it does provide a more obvious comparison between matching strategies and a discerning examination of problems.
Figure 14: Stereo images produced from a scaled, random, radiometric model and an artificial geometric model. SYNTHETIC MODEL
caused by the geometric model.

The radiometric model is constructed by producing five separate images of random gray values. The UNIX random number generator, \textit{random}(seed), generated random numbers between 0 and 255 which are placed as gray values in a raster image format. Each image has a different \textit{seed} or starting point for the random number generator and a different resolution or scale level. These images are then added to one another and quantized to 256 gray values, Figure 14b. This produces a gray value image with highly variable spatial frequency responses over multiple bandwidths in the frequency domain, and thus multiple scale levels. Random gray values over multiple scale levels is needed so that when high frequency information is removed during multi-scale analysis there will remain some random information for the similarity measure to utilize.

The resulting gray values are assigned to the geometric model as an orthophoto. The geometric model and radiometric model are then densified by a factor of four. The geometric model and its assigned radiometric model are projected to two, predefined image planes. Occluded points are detected and eliminated as in Section 5.1.2. The result is an irregular grid of points in each image, each with an assigned gray value. A nearest neighbor, resampling technique is used to produce a grid format without correlation effects and is the reason for the over densification (factor of four) of the models.

The outcome is two stereo images with the object space models already defined, Figure 14a.
The equivalent characteristics of the synthetic data to actual aerial image data are stereo images that are vertical, epipolar and separated by a base of 1000 meters. Furthermore, an equivalent focal length of 0.1 meters and a flying height of 1000 meters coincides with an image scale of $1 : 10,000$. The pixels are 100 microns in size and approximately correspond to a one meter grid cell in the object space. The relief range in the object space is 40 meters. The match area is approximately $128 \times 128$ meters, equivalent to $128 \times 128$ grid cells or pixels.

### 5.2.2 OSU Campus Stereo Images

Stereo images of The Ohio State University were flown at an approximate height of 590 meters above ground level and are separated by a base of 343 meters, Figure 15. The orientation parameters are known so the images could be transformed into epipolar geometry. The focal length is 0.152 meters and the scale is $1 : 4,000$. The images have a pixel size of 120 microns that corresponds to approximately 0.48 meters on the object space surface. The match area is approximately $64 \times 64$ meters, equivalent to $128 \times 128$ grid cells or pixels. An idealized representation of the elevation model is determined manually from the actual photographs and is used as part of the error analysis.

### 5.3 Results

The results of the surface reconstruction techniques on the test data are discussed from numerical values or visual representations of the outcomes at each step. Explanations of the results are given in Sections 5.3.1-5.3.2.
Figure 15: Stereo images and manually derived, ideal DEM. OSU MODEL
Table 1: Matching technique parameters. Size units in pixels (image space) and surface grids (object space). SYNTHETIC MODEL

<table>
<thead>
<tr>
<th>scale levels (i)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>grid interval size</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>smoothing parameter $\sigma$</td>
<td>5.66</td>
<td>2.83</td>
<td>1.41</td>
<td>$\approx$1</td>
<td>n/a</td>
</tr>
<tr>
<td>reference window size (x/y)</td>
<td>25/25</td>
<td>17/17</td>
<td>13/13</td>
<td>11/11</td>
<td>9/9</td>
</tr>
<tr>
<td>search area size (x/y)</td>
<td>65/27</td>
<td>45/19</td>
<td>31/15</td>
<td>21/13</td>
<td>13/11</td>
</tr>
<tr>
<td>quality threshold (%) (lowest correlation coefficient accepted)</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>reliability threshold (%) (largest angle, $cz$, accepted (Figure 10))</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>coplanarity threshold (%) (largest error vector distance accepted - only object space strategy)</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

5.3.1 Synthetic Stereo Images

The initial, approximated $DEM_i$ is a flat surface at 125 meters with an approximated relief range of 20 meters. The predefined parameters are listed in Table 1. Notice that the smoothing parameter is chosen so that the size of the gaussian operator ($w$) equals the size of the grid interval. Refer to Section 5.1 for explanations on how the smoothing and matching parameters were selected. The results of the image space and object space matches through five scale levels are illustrated in Figures 16-23 in Appendix A.

Obviously, the quality of the matches is very high in both the image space and
Table 2: Similarity measure quality and error comparison between image space and object space strategies. SYNTHETIC MODEL

| SCALE LEVELS | IMAGE SPACE | | | OBJECT SPACE | | |
| | DEM Error (in meters) | Correlation Coefficient (Mean) | | DEM Error (in meters) | Correlation Coefficient (Mean) | |
| | Mean | St. Dev. | | Mean | St. Dev. | |
| 0 | 3.12 | 2.03 | 0.82 | 3.43 | 2.23 | 0.82 |
| 1 | 2.09 | 1.85 | 0.83 | 1.33 | 1.15 | 0.83 |
| 2 | 1.79 | 1.93 | 0.74 | 0.63 | 0.62 | 0.84 |
| 3 | 1.63 | 1.97 | 0.77 | 0.42 | 0.45 | 0.92 |
| 4 | 1.56 | 2.06 | 0.68 | 0.38 | 0.43 | 0.89 |

object space strategies due to optimal radiometric conditions. However, the object space matches have significantly better overall quality, Table 2. The utility of the reliability factor of the matches is insignificant as the synthetic radiometric surface eliminates uniqueness problems. The omitted grid cells due to thresholding correlation coefficient values, Figures 16-17, cluster near object space events. However, a relative comparison of the quality shows that the image space is less flexible to object space surfaces that deviate from a plane. In fact, image space matching also has difficulty in recovering steep planar surfaces, NW section of image, and almost eliminates an entire non-planar event, half-sphere in SW section of image. Object space matching generates much higher quality matches with lower quality only on thin outlines at discontinuities and, of course, the occluded region. The thin outlines of poor matches at discontinuities are not surprising due to the naive, linear interpolation technique used to transform irregular to regular DEM grids.
The thresholding of error vector distances in object space matching provides marginal information in this model. The reason is that the area matched is centered on the flight line of the images and that the DEM axes is parallel to the image space axes. Therefore, \( p_y \) (Equation 4.1) will be small and the resulting search area in the \( px \) direction (Equation 4.2) will follow very closely to the epipolar plane. Combined with the ideal radiometry for detecting correct conjugate locations, the resulting error vectors tend to be very small, less than one meter at all scale levels.

The segmentation technique for object space matching in this model is performed manually to observe and use realistic DEM information from previous scale levels. Figure 18 shows the breaklines which were selected from the DEMs, the occlusions (Figure 19), and low quality matches (Figures 16-17) at each scale level. Assuming these automatically derived cues indicate significant object space events, the object space was manually segmented from the derived data for use in the interpolation technique in the preceding scale level.

Object space matching provides a unique ability to detect occlusions geometrically and concurrently in the matching strategy. The occlusions appear because of wrong conjugate location selection or, more likely in this model, from the interpolation technique. As stated in Section 4.2 the interpolation technique predicts the character of the object space that may not be supported by or tied to the image space.

Even with a very rudimentary interpolation technique, the occluded areas in the right image (there are no occlusions actually or detected in the left image), Figure 19, show a very accurate depiction of the true occlusions.
This information is invaluable to surface acquisition and terrain classification. Combined with the other omitted points from Figures 16-23, an impressive isolation of the occlusion occurs.

The most important results of the matching techniques are the DEMs and differences from the 'truth' through the scale levels, Figures 20-23. Both strategies provided similar results in the initial scale levels. However, the image space strategy results stabilize at scale level 2 and diverge from the 'truth' in the final scale levels in certain areas, even though the initial scale level mean error is less than in the object space strategy. The results of the object space strategy show a steady convergence to the 'true' surface in both mean error and error dispersion. In Table 2 shows improved convergence of the mean DEM error and the dispersion of errors in the object space strategy.

The resulting DEMs and DEM errors confirm very nicely that warping the images to the object space at each scale level eliminates more distortions than a simple affine transformation between images. Furthermore, the warped images will improve their representation at each scale level, supplemented with the knowledge gained from the detected breaklines. In contrast, the image space strategy provided accurate results only in areas which it could model the object space sufficiently. Also, as shown in Figures 22-23, since lower quality similarity measures are accepted, wrong conjugate locations are extended through the scale levels. This causes the matching technique to essentially become 'lost' and verifies the need for smoothness constraints, Section 2.4.2, even in the most ideal of radiometry.
Table 3: Matching technique parameters. Size units in pixels (image space) and surface grids (object space). OSU MODEL

<table>
<thead>
<tr>
<th>scale levels (i)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>grid interval size</td>
<td>16</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>smoothing constant $\sigma$</td>
<td>5.66</td>
<td>2.83</td>
<td>1.41</td>
<td>$\approx$1</td>
<td>n/a</td>
</tr>
<tr>
<td>reference window size (x/y)</td>
<td>33/33</td>
<td>31/31</td>
<td>29/29</td>
<td>27/27</td>
<td>25/25</td>
</tr>
<tr>
<td>search area size (x/y)</td>
<td>85/37</td>
<td>71/35</td>
<td>59/31</td>
<td>47/29</td>
<td>35/27</td>
</tr>
<tr>
<td>quality threshold (%) (lowest correlation coefficient accepted)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>reliability threshold (%) (largest angle, $\alpha_x$, accepted (Figure 10))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>coplanarity threshold (%) (largest error vector distance accepted - only object space strategy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
</tr>
</tbody>
</table>

5.3.2 OSU Campus Stereo Images

The initial, approximated $DEM_i$ for the OSU campus images is a flat surface at 230 meters with an approximated relief range of 25 meters. The predefined matching technique parameters are listed in Table 3. Refer to Section 5.1 for discussions on how the smoothing and matching parameters were selected. Notice that the reference matching window size and search areas are much larger than in the previous model due to the less optimal radiometric conditions in the natural, OSU images. The results of the matching techniques are shown in Figures 24-32 in Appendix B.

The omitted points due to thresholding quality factors are depicted in Figures 24-
25. Once again, the reliability factors are insignificant and not of great use. Notice that the omitted points in the image space strategy are concentrated in one section, top of the building, and remain almost constant through the scale levels. As in the synthetic data image space matching loses object space events or portions of events in lower scale levels and will not recover. Object space matching has similarly concentrated poor matches in the lower scale levels. However, the omitted points converge to specific object space events, i.e. the occluded and discontinuous regions, that are highly distorted and will probably never match successfully with the simple techniques used.

In contrast to matching the synthetic data the error vector provides substantial information in the OSU model. Figure 26 presents the omitted grid points due to high error vector distances. The large error vectors are concentrated near the discontinuous and occluded regions. This match point analysis is much needed to independently discriminate match results when only simple quality and reliability measures are otherwise used, especially in urban terrain.

The segmentation technique on the OSU model uses a Laplacian of Gaussian (LoG) filter on the grid DEM at each scale level. A large filter size was chosen, 31 x 31 meters, to only segment large object space events. The results are zero-crossings that represent relief slope changes on the highly smoothed DEM. Figure 27 reveals the zero-crossings roughly outlining the building in the DEM.

The occluded areas for the left and right images of the OSU campus model are shown in Figure 28. The object space strategy detects the occlusions in both the left
Table 4: Error comparison with manually derived object space coordinates on original photographs. OSU MODEL

<table>
<thead>
<tr>
<th>NUMBER OF POINTS</th>
<th>IMAGE SPACE - mean DEM precision (meters)</th>
<th>OBJECT SPACE - mean DEM precision (meters)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (2304 pts)</td>
<td>3.48</td>
<td>1.88</td>
</tr>
<tr>
<td>Best 60% (1382 pts)</td>
<td>0.56</td>
<td>0.07</td>
</tr>
</tbody>
</table>

and right image. Even with the higher chance of ambiguous conjugate locations in this model, the interpolation technique predicted a crude, but characteristic outline of the building. When combined with the other omitted points from the low correlation coefficients and high error vectors, the highly distorted relief areas are evident, Figure 29.

The resulting DEMs are shown in Figures 30-31. The object space DEM's resolve the character of the building and stabilize near the 'truth' at the locations of accepted matches. The image space strategy could not resolve the form of the building and, thus, it resulted in a highly distorted DEM. In both strategies the occluded regions are, predictably, the most distorted.

This visual analysis provides a general glance at the caliber of the results in relation to an idealized approximate of the DEM that does not consider trees and other minor events of the surface. Interpolation bias also limits the ability of the strategies to acquire an accurate depiction of such complex terrain. A more precise and an ultimate
examination of the results is a comparison of the match locations with manually
determined conjugate points from the original photographs. A well distributed grid
of 48 x 48 raw, object space surface points (prior to interpolation) from each strategy,
Figure 32, were input into the Zeiss Planicomp analytical stereoplotter driving the
floating marks to the appropriate conjugate locations on the stereo diapositives. Note
that the match points of the image space strategy were intersected to the object space
surface to form the raw surface points for comparison. The operator corrected and
recorded the error in the elevation only.

Table 4 depicts the mean errors from this manual check of the precision of each
strategy's final scale level results. The image space match errors are almost twice as
large as the object space match errors. If the largest 40% of the errors are omitted
the mean error in the object space strategy reduces to 0.07 meters (equivalent to
0.087 pixels) or nearly \(\frac{1}{10,000}\)th of the flying height for over 1,300 points. The results
compare favorably to the suggested normal vertical accuracy results of \(\frac{1}{5,000}\)th of
the flying height for fully analytic aerotriangulation techniques [87]. These are very
encouraging results, especially considering the rudimentary techniques used for such
a distorted stereo model.
CHAPTER VI

Conclusions and Recommendations

6.1 General

The implementation techniques in Chapter V are admittedly and purposely elementary. Instead, an effort was made to examine the effectiveness of the proposed strategy that, ever so slightly, begins to address the automatic surface reconstruction from a conceptual perspective. Synthetic data was constructed to reduce matching errors and examine the responses due to complex geometric distortions in stereo images. The complicated urban terrain model was chosen to force the rudimentary tools to fail and see if the strategy could resolve a complex geometric model. In no instance is correlation alone suggested for matching large scale urban terrain.

However, the concept of representing image information in the object space and testing the relationships between these warped images and the elevation model has proved successful. The multi-scale representations provide the opportunities needed to exploit coarse descriptions and iteratively converge to an optimal solution, i.e. near equal object space representations from both stereo images. By iteratively matching, interpolating, and interpreting in the object space environment, the proposed strategy provides a foundation upon which multiple sensor data, multiple surface re-
construction techniques, and domain knowledge (quantitative or qualitative) may be combined during DEM acquisition.

DEM acquisition from the object space environment is a more reliable and robust strategy of acquiring DEMs when compared to an image space strategy and provides a more suitable environment for more advanced measuring and interpretation techniques. Essentially, the multi-scale, object space surface reconstruction concept and strategy provides a basis to realize the now widely concluded opinion that multiple surface reconstruction tasks must be used to interactively and heuristically build a complete description of the object space. Some more specific conclusions are described below.

**Advantages of multi-scale, object space matching:**

- *Similarity measurement quality is globally improved.*
- *Perspective and relief distortions are reduced.*
- *Convergence to a more accurate DEM is enhanced.*
- *Errors are concentrated (localized) at specific object space events by globally examining and modifying the results.*
- *Direct tracing of object space events through the scale space is enabled.*
- *The three-dimensional object space environment is conducive to integration and interaction with other surface reconstruction and image understanding tools.*
- *Concurrent, geometric occlusion detection is permitted.*
• DEMs from previous scale levels can be used as information cues for preceding scale level techniques.

• Sampling pattern and rate are chosen in the object space where DEM sampling techniques can be utilized. Notice the sampling patterns in Figure 32 for the final match points before interpolating to a grid. The image space pattern is referenced to the left image; therefore, the resulting sampling pattern is a function of the perspective projection of that photograph. The object space pattern is referenced to the grid of the DEM varying only with the ray intersection adjustment. This pattern is much more consistent and controllable.

Disadvantages of multi-scale object space matching:

• More tasks, i.e., rectification, interpolation, ray intersection, etc. make the procedure slower and more complicated.

• Performance is highly dependent on the interpolation technique

• Scan line epipolar constraint is lost.

6.2 Recommendations for Further Research

An obvious deduction from this study is that more advanced techniques should now be used in the object space procedure to take maximum advantage of coarse level DEMs and provide a more clever approach to detecting conjugate locations and interpreting the object space. It is recommended that more advanced techniques be studied for the following tasks.
• Surface Analysis - A more advanced analysis of the object space is needed to provide more useful and accurate information to be used by other tasks, e.g., three-dimensional object recognition (see e.g. [23]), Fourier analysis (see e.g. [16, 4], and combining of radiometric and geometric information.

• Interpolation - A more sophisticated interpolation technique is needed in order to retain the character of the irregular, raw surface, e.g., TIN model to grid lattice or polynomial adjustment within segmented regions using least squares adjustment [86]. The interpolation technique should fully utilize the object space information derived from the surface analysis.

• Matching - A more advanced matching (model refinement, similarity measure, and constraints, see Chapter II) technique is needed to supplement the warped images, eliminate false conjugate locations, and more effectively utilize knowledge from previous scale levels.

• Integration - Integration and management of data, e.g., by probabilistic means in Markov Random Fields [15], is needed to improve the efficiency of combining large amounts of derived information acquired in the proposed strategy.

• Control - A higher level of control is needed to effectively and intelligently react to myriad of knowledge and situations, e.g., an expert system (see e.g. [81]).

Finally, this research confirmed the hypotheses of multi-scale, object space surface reconstruction and tested its robustness when compared to a traditional, image
space strategy. This strategy should now be expanded with more advanced techniques (see above) and a wide variety of test data. If surface reconstruction is approached from a conceptual point of view (as in this research) the proposed strategy of multi-scale surface reconstruction in the object space should provide a sound basis for enhanced automation.
BIBLIOGRAPHY


Appendix A

Results of Implementation Comparison
(SYNTHETIC MODEL)

This appendix contains the graphic representations of the implementation results comparing the traditional, image space strategy and the proposed, object space strategy from the synthetic stereo images. The following is a list of the figures.

16. Gray value maps of correlation coefficients (scale levels 0-2)

17. Gray value maps of correlation coefficients (scale levels 3-4)

18. Manually detected breaklines on DEMs

19. Occlusions

20. Refined DEMs from scale levels 0-2

21. Refined DEMs from scale levels 3-4 and the true DEM

22. DEM error magnitudes from scale levels 0-1 and the initial, approximate DEM

23. DEM error magnitudes from scale levels 2-4
Figure 16: Gray value maps of correlation coefficients (black = 0 and white = 1) and the low 5% of the correlation coefficients. SYNTHETIC MODEL.
<table>
<thead>
<tr>
<th>Image Space</th>
<th>Scale Levels</th>
<th>Object Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(b)</td>
<td>(a)</td>
</tr>
</tbody>
</table>

Figure 17: Gray value maps of correlation coefficients (black = 0 and white = 1) and the low 5% of the correlation coefficients. SYNTHETIC MODEL
Figure 18: Manually detected breaklines on DEMs from each scale level. SYNTHETIC MODEL
Figure 19: Occlusions in right image detected from DEMs in the object space strategy and from the true DEM. SYNTHETIC MODEL
Figure 20: Refined DEMs from scale levels 0-2. SYNTHETIC MODEL
Figure 21: Refined DEMs from scale levels 3-4 and the true DEM. SYNTHETIC MODEL
Figure 22: DEM error magnitudes (absolute values) from scale levels 0-1 and the initial, approximate DEM. SYNTHETIC MODEL
Figure 23: DEM error magnitudes (absolute values) from scale levels 2-4. SYNTHETIC MODEL
Appendix B

Results of Implementation Comparison
(OSU MODEL)

This appendix contains the graphic representations of the implementation results comparing the traditional, image space strategy and the proposed, object space strategy from the OSU campus stereo images. The following is a list of the figures.

24. Gray value maps of correlation coefficients (scale levels 0-2)

25. Gray value maps of correlation coefficients (scale levels 3-4)

26. High error vector distances from ray intersection technique

27. Breaklines detected from LoG operator over DEMs

28. Occlusions

29. Omitted match locations

30. Refined DEMs from scale levels 0-2

31. Refined DEMs from scale levels 3-4 and idealized DEM

32. Raw match point planimetric distributions
<table>
<thead>
<tr>
<th>Image Space</th>
<th>Scale Levels</th>
<th>Object Space</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>(a)</td>
<td>(a)</td>
</tr>
<tr>
<td>(b)</td>
<td>(b)</td>
<td>(b)</td>
</tr>
</tbody>
</table>

Figure 24: Gray value maps of correlation coefficients from scale levels 0-2 (black = 0 and white = 1) and the low 15% of the correlation coefficients. OSU MODEL.
Figure 25: Gray value maps of correlation coefficients from scale levels 3-4 (black = 0 and white = 1) and the low 15% of the correlation coefficients. OSU MODEL
Figure 26: Highest 25% of error vector distances from ray intersection technique. OSU MODEL
Figure 27: Breaklines detected from LoG operator over DEMs. OSU MODEL
Figure 28: Occlusions in each and both images detected from DEMs in the object space strategy and from idealized DEM. OSU MODEL.
Figure 29: Omitted match locations determined from occlusions, low quality factors, and high error vector distances. OSU MODEL
Figure 30: Refined DEMs from scale levels 0-2. OSU MODEL
Figure 31: Refined DEMs from scale levels 3-4 and idealized DEM: OSU MODEL
Figure 32: Raw match point planimetric distributions prior to interpolation to the final DEM.