Michael A. Crombie
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, VA 22060

Coordination of Stereo Image Registration and Pixel Classification

Correlation is an inadequate measure when there is less than, or more than, a moderate amount of image structure at and around points selected for image matching.

Abstract: *An inadequate concept of how corresponding points relate to one another on dissimilar images has a greater effect than do exposure geometry or data collection on registration problems in stereo photogrammetry.*

Conventional correlation, or one of its relatives, is the measure of similarity used in all automated stereo correlation systems. Correlation, a measure of the linear dependence between two sets of data, is an inadequate measure when there is less than, or more than, a moderate amount of image structure at and around points selected for image matching. The existence of structure should be recognized and utilized in an appropriate manner for image matching. Similarly, the absence of structure should be recognized, and the surrounding imagery should be used to complete matches where it is possible. The concurrent determination of what a pixel is, as well as where it is, can alleviate much of the registration problem. A variety of features, including point-density data, texture, and edges, as well as existing cartographic knowledge, can be combined and organized through rules in order to more completely describe a point. The overall throughput of the compilation process will be improved in both time and accuracy if those functions which tend to support one another are concurrently, rather than sequentially, performed. If the compilation process takes place in image space, then the image matching operation, as well as the other feature extraction operations, can be ordered by the data processing manager to best suit the function of the process.

Introduction

It is likely that the image matching component of most photographic feature extraction systems has been overemphasized. There is a widespread notion that the image matching component must occur first, after triangulation, and that it must be carried out quickly because other feature extraction processes cannot proceed until it is completed. Part of the problem may stem from the practice of developing the cartographic product directly in the preferred object-space rather than in an intermediate domain such as image-space. This notion is false, regardless of the reason. The correlation-first imperative can be dismissed immediately if the collection of image features, including x-parallax data, is carried out in image-space.

The gist of this paper is that, if the time line requirement of the match process is relaxed, then a more complete and accurate registration of the stereo pair can be effected. It is asserted that the allowable time for image matching can be increased significantly if the correlation process takes place in image-space. Furthermore, if digital or digitized pictures, rather than hardcopy pictures, are processed, then the entire feature extraction process can be developed and managed more effectively with state-of-the-art technology.

Digital Processing

Most of the observations presented in this paper were developed under Army funded technical base efforts in digital correlation and in

Photogrammetric Engineering and Remote Sensing,
Vol. 49, No. 4, April 1983, pp. 529-532.

0099-1112/83/4904-0529$02.25/0

digital feature extraction (USAETL, 1980). The operation of extracting x-parallax data in image space (Rosenberg *et al.*, 1974) does not require the imagery to be in softcopy format. It is recognized, however, that redesigning existing hardcopy processors may not be practical at this time. There are four basic ground rules being followed in the development of the feature extraction work. They are

- Feature extraction in image space;
- Rectified digital or digitized image input;
- Coordinated mensuration; and
- Rule based extraction methods.

The development of a system that operates in image-space will provide for better management of the feature extraction process whether the input is in digital format or in hard copy format. In the first place, if all of the feature extraction operations take place in image space, then the system should be capable of handling output from a variety of collection systems. The ordering of the various operations is not rigid, and this can be organized to best suit the function of the process which in turn need not be cast in concrete. For example, there is no need to go to the expense of creating an orthophoto with its attendant loss of resolution. In the image matching operation there is a need only to shape one of the images in only one direction if the images are rectified (Crombie, 1976; Norvelle, 1981). Operations such as image matching and feature extraction can be iterated if need be. Furthermore, features can be extracted from either or both images of the stereo model.

Pictures are not always collected in a manner that leads directly to a successful and economic processing exercise. Whether the material is collected in a rectified format, or reformatted after collection, it is assumed that the feature extraction system has available reformatted digital-image pairs with most of the y-parallax removed. The primary reason for this requirement is to simplify the image matching operation. Note that this requirement does not rule out the possibility that the feature extraction system can perform the rectification function. Because the bulk of the match process will be performed by conventional correlation methods which utilize, essentially, low-frequency information (up to 10 line pairs per millimetre) (Crombie and Rand, 1977; Helava, 1976), there is no reason to rectify the images at full resolution. Nor is there, in most cases, any reason to rectify images to extract features other than x-parallax. In fact, a variety of versions of the same digital picture can be readily presented to the photointerpreter in softcopy format for the several feature extraction tasks.

In an image domain system, feature extraction operations can be ordered to reinforce one another. Edge and line detectors can be used in conjunction with binary raster processing to aid the photointerpreter in extracting roads, creeks, drainage, field boundaries, etc. Simple two-component signatures can be used at the same time to sort out fields, forests, and roads, and to isolate built-up areas. These three operations: edge finding, two-component classification, and binary raster processing can be used in cooperation with one another and, if necessary, on both images of the stereo-pair (Crombie and Gambino, 1977). Note that field corners, road crossings, and high curvature points along lineal features produce photo-identifiable points which can be used to generate the needed match-points in image-areas where the automatic correlation breaks down (Lambird *et al.*, 1980). Note, too, that in an image-domain feature-extraction system, local mapping of image-to-image is a basic operation which is easily carried out once the x-parallax calculations have been computed. Image-to-image mapping will provide the photointerpreter a means for extracting detail from either one or both of the images for completeness and for edit purposes (Gambino and Crombie, 1979). For example, detail found on an image and stored in binary format can be mapped onto the image or onto its stereomate. The first mapping function would present visual verification of the feature extraction operation, whereas the second mapping function would enable the photointerpreter to fill in missing detail as well as to verify and evaluate the x-parallax function.

Rule-based feature extraction methods pertain to techniques that are being promoted by knowledge-base engineers and others from the artificial intelligence community (Stockman *et al.*, 1981). The installation and utilization of "smarts" into the computer system is an attractive idea which needs further development before it can be used effectively in the extraction and conversion of image primitives into useful cartographic primitives. Rules can be developed now to enforce consistency among derived cartographic primitives. For example, the lay-of-the-land, as determined by the digital terrain elevation matrix, can be used (and vice versa) to develop paths of creeks, drainage, and railroads. Note that, in this example, rules will be defined in object space. Rules can be developed to track along lineal features in order to find high curvature points and points where other lineal features cross for match-point data. Rules can be developed to determine corresponding buildings, and especially corners common to both images, for the same purpose. The possibilities are endless. An image-domain system will provide ample room and freedom of expression to exploit relationships between image and cartographic primitives when the relationships are expressible in logical, numerical, or relational forms.

Conventional Correlation

Conventional correlation methods using the linear correlation coefficient of statistics as the measure of similarity are well known and they generally produce acceptable results over much of the stereomodel area. No known method produces more reliable results in the mapping mode than a simple area match scheme where most of the y-parallax has been removed by a controlled reformatting exercise and where the imagery is shaped to reflect distortion due to local terrain undulations. The removal of y-parallax is carried out by a pixel resampling procedure which is regulated by known interior and exterior orientation data. Local x-parallax data, computed on the fly, is used both to predict match points and to shape one of the images.

The conventional process bogs down in very busy regions if the base-height ratio is such that detail on one image is noise with respect to the second image. The process also stalls in regions of little or no detail. In many cases, the problem can be resolved by blurring the imagery in the first case and by using very large windows in the second case. In either case, the process is slowed and generally the accuracy of the match is reduced.

Utilization of the linear correlation coefficient, R_{xy}, as the measure of similarity will model all additive and multiplicative differences between corresponding grey shades. Any nonlinear relative distortion due to atmosphere, perspective, film processing, scanner, pixel resampling, etc., will diminish the value of R_{xy} and consequently reduce confidence in the match. The grey shade values within a window are regarded as independent data in conventional correlation. Structural information is not accounted for in the process. In fact, if there is too much structure, the correlation function will peak only at, or very near, a perfect match. Since the "pull in" range for highly structured data is short, the process should, if possible, slow down and match nearly every point and utilize small windows. Even this procedure will fail if structure in one image, due to perspective and other distortions, is not evident on the second image. Certain structural features, especially lineal ones, are relatively immune to perspective when regarded as a stereo event. Portions of a road may not appear on both pictures, but high curvature points and intersections, if not obscured, will appear on both images and can be, in many cases, automatically identified and used as match points. At a higher level of computer sophistication, buildings can be isolated and corners, common to both, used as match points (Liebes, 1981).

The "pull in" range for low-frequency areas is larger than the pull in range for high-frequency areas, but the correlation function is flat, and precise matching tends to be uncertain. If the frequency content approaches zero, then there is no hope for conventional correlation. Manual intervention or other techniques must be employed. If the process is such that means to automatically identify the region are at hand, then that information can be used. For example, if the difficult region is a body of water (a small lake), then edge techniques can be used to determine the lakeshore line, and high curvature points can be used as match points. Similarly, if the region is identified as a field, then field corners and field boundary intersections can be located and used as match points.

The suggested process implies that regions of pixels must be identified wherever conventional correlation is not acceptable. All of this will take time, time that might not be available if the correlation-first imperative is imposed. A few simple calculations demonstrate that, if the feature extraction process takes place in image space, then the match process can be viewed in a much more sensible and relaxed manner. Consider Table 1.

The entries in the table pertain to the number of hours of correlation time needed to produce 100 square miles of match data at a specified spacing for a given processing rate. Because the unit of time associated with other kinds of feature extraction is on the order of a week or perhaps a month, the image matching exercise with processing time measured in hours should not be allowed to dominate the total feature extraction process. If the entire exercise is performed in image space, where most parts of the operation can begin at the discretion of the manager, then it is quite conceivable that processing speeds of twenty points-per-second or less will be more than satisfactory. Regions that are poorly correlated can be completed later in an interactive and rule-based regime by using structured image primitives.

Summary

In the event digital or digitized images exist, and they can be reformatted to remove y-paral-

Table 1. Processing Time (Hours) For 100 Square Miles

Processing Speed (Pts/Sec)	Ground Spacing (Feet)		
	300	100	50
10	0.902	8.123	32.500
20	0.451	4.063	16.250
30	0.301	2.708	10.833
40	0.226	2.031	8.125
50	0.181	1.625	6.500
75	0.120	1.083	4.333
100	0.090	0.813	3.250
200	0.045	0.407	1.625

lax, then an improvement in the image-matching function of a feature extraction system can be achieved by dismissing the correlation-first imperative and by reorganizing the process in order to coordinate the several functions to their mutual advantage. The various feature extraction operations, including the determination of x-parallax, along with the required control mensuration, can be performed in almost any order, concurrently or sequentially, and from a variety of soft copy terminals as long as the exercise takes place in an image domain. A regular grid of points will be defined on one image, and those points that can be matched by using automatic methods will be processed in a background mode, while the much more difficult job of extracting structured primitives will be performed in an interactive mode. Those structural features that provide well defined corresponding points will be used to complete the image-to-image mapping in areas where the automatic mode breaks down.

References

Crombie, M. 1976. *Stereo Analysis of a Specific Digital Model Sampled from Aerial Imagery*, ETL-0072, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Crombie, M., and L. Gambino, 1977. *Digital Stereo Photogrammetry*, Computer Sciences Laboratory, US Army Engineer Topographic Laboratories, Fort Belvoir, Va., prepared for Congress of the International Federation of Surveyors (FIG), Commission V, Stockholm, Sweden.

Crombie, M., and R. Rand, 1977. *An Evaluation of the Method of Determining Parallax from Measured Phase Differences*, ETL-0145, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Gambino, L., and M. Crombie, 1979. *Manipulation and Display of Digital Topographic Data*, US Army Engineer Topographic Laboratories, Fort Belvoir, Va., prepared for the Second Symposium on Automation Technology in Engineering Drawing, Naval Postgraduate School, Monterey, Ca.

Helava, U., 1976. *Digital Correlation in Photogrammetric Instruments*, Bendix Research Laboratories, Southfield, Michigan, Prepared for XIII International Congress for Photogrammetry, Helsinki, Finland.

Lambird, B., D. Lavine, G. Stockman, K. Hayes, and L. Kanal, 1980. *Study of Digital Matching of Dissimilar Images*, ETL-0248, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Liebes, S., Jr., 1981. *Geometric Constraints for Interpreting Images of Common Structural Elements: Orthogonal Trihedral Vertices*, Artificial Intelligence Laboratory, Computer Sciences Department, Stanford University, Stanford, California, prepared for the proceedings of the DARPA Image Understanding Workshop.

Norvelle, F., 1981. *Interactive Digital Correlation Techniques for Automatic Compilation of Elevation Data*, ETL-0272, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Rosenberg, P., K. Erickson, and G. Rowe, 1974. *Digital Mapping System: Mathematical Processing*, ETL-CR-74-6, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Stockman, G., B. Lambird, D. Lavine, and L. Kanal, *Knowledge-Based Image Analysis*, ETL-0258, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

USAETL, 1980. *Feature Extraction Plan (FY80-85)*, Computer Sciences Laboratory, US Army Engineer Topographic Laboratories, Fort Belvoir, Va.

Microwave Signatures in Remote Sensing

Toulouse, France
16-20 January 1984

This meeting—a joint effort of Centre National d'Etude Spatiales, which will sponsor the symposium, and the International Scientific Radio Union Commission F, with the assistance of Centre d'Etude Spatiale des Rayonnements—is intended for specialists who are working directly in the field of radar backscatter and microwave emission from the earth and sea (theory and experiment). The focus is on the interaction of microwaves (passive and active) with the surface of the earth, including its natural cover (vegetation, snow, etc.) and of the ocean (both water and ice surfaces) as this interaction affects remote sensing. Attendees are expected to present papers and/or participate in the discussions.

For information on the technical program, contact

Europe

Dr. Erwin Schanda
Universität Berne
Institute of Applied Physics
Sidlerstrasse 5
3012 Berne, Switzerland

Other Continents

Dr. Richard K. Moore
Remote Sensing Laboratory
University of Kansas Center for Research, Inc.
2291 Irving Hill Drive—Campus West
Lawrence, KS 66045, USA