The Potential for Automated Mapping from Geocoded Digital Image Data

Richard Swann, Dave Hawkins, Andrew Westwell-Roper, and *William Johnstone*
MacDonald Dettwiler, 3751 Shell Road, Richmond V6X 2Z9, Canada

ABSTRACT: Recent improvements in the resolution of available commercial satellite imagery, combined with new techniques for high-precision geometric image correction, now make production of accurate topographic maps from satellite imagery possible. This paper discusses the technical considerations of using imagery from existing satellites such as Landsat and SPOT.

INTRODUCTION

DIGITAL PROCESSING of satellite imagery has a number of possible advantages over conventional photogrammetric methods for generating geographic information:

- Satellite imagery provides constant global coverage, and is readily available from image archives.
- Data acquisition and ground control costs are substantially less than those for equivalent aerial photography, due partly to a decreased need for ground control.
- Satellite image data are already in the digital format necessary for computer processing. Computer processing reduces the labor intensity and, consequently, the costs of map compilation.

Until recently, use of satellite imagery for topographic mapping has been limited by the low resolution of available commercial satellite imagery, and a number of technical processing problems. However, with last year's launch of the SPOT satellite, monochromatic imagery with 10-m ground resolution and 20-m multispectral imagery are now available. Until the launch, the highest resolution imagery available was from the Landsat Thematic Mapper (TM) sensor, with 30-m resolution. Furthermore, rapid progress is being made on the technical processing problems.

This papers surveys recent advances in satellite image processing that will enable accurate topographic and thematic mapping from the new higher-resolution imagery. The aim of the research surveyed is to derive maps or mapping data with a maximum of ± 10-m RMS error in planimetric and height measurements. The paper considers important differences between digital satellite imagery and aerial photography, stages in the production of maps from satellite images, and areas of current technological development.

DIGITAL SATELLITE IMAGERY

The differences between photographic images and digital images from remote sensing satellites are somewhat obscured by the fact that satellite images are often presented in the form of film products. In this form they appear the same as photographs, and can provide no more information than photographs.

However, there are fundamental differences in the way satellite images and aerial photographs are generated, and in the information they contain. This section will explore those differences.

LINE SCANNING

An aerial camera can take a picture of an area of terrain in a single exposure. Landsat sensors, on the other hand, scan the Earth's surface with a scanning motion perpendicular to the satellite's trajectory, and form images from the series of line scans. SPOT sensors effectively take a "snap shot" of a single line of an image every few milliseconds. With both Landsat and SPOT, the satellite continuously changes attitude and proceeds several metres along its orbit between the times at which two adjacent pixel lines are imaged.

Neither aircraft nor satellites maintain ideal flight or orbiting paths, and corrections for non-ideal motion have to be made to aerial photographs and satellite images before any mapping information can be extracted. Techniques for doing this are well developed in the case of aerial photographs. A single correction matrix, based on the aircraft's attitude and position, can be applied to an entire photograph.

Satellite images, on the other hand, have to be corrected pixel-by-pixel because of the continuous variations in satellite attitude and position. This can be done by third order warping functions, but better results are obtained by reference to a "spacecraft model," a numerical representation of the orbiting satellite's dynamic parameters. The required correction varies continuously with time, so that the correction for any given pixel differs from that for pixels preceding and following it.

The only correction applied by photogrammetric devices such as stereoplotters is constant across a given photograph, so that these devices, as currently used, are inappropriate for correcting satellite imagery. It has been shown that an analytical stereoplotter can be adapted to correct film plots of satellite images, although it requires dramatic reprogramming (Denis, 1987).

DIGITAL MULTISPECTRAL DATA

Satellite imagery also differs from aerial photography in that the former is digital and the latter analog. Considered abstractly, a digital image is a two-dimensional array of elements or "pixels." Each pixel consists of a number of numerical values for different parameters. The image equivalent of a monochromatic photograph contains a single parameter value for each pixel. The image equivalent of a color photograph contains three parameter values for each pixel, one for each of the primary colors. However, digital imagery is not limited to three parameter values per pixel. Consequently, a digital image of an Earth surface region can contain far more information than a photograph of the same region.

For example, each pixel of data from the Landsat TM sensor contains a value for each of seven bands of Earth surface radiation. These bands span the spectrum from thermal infrared to visible light. The spectral band values, and ratios between them, can be used to extract detailed information about surface vegetation, and even surface minerals, that is unavailable from equivalent

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 54, No. 2, February 1988, pp. 187–193.

0099-1112/88/5402–187$02.25/0

aerial photography. Plates 2, 3, and 4 show examples of spectral information extraction.

Stereo Imagery

The SPOT satellite has two sensors suitable for stereo imaging, the Panchromatic Linear Array (PLA) with 10-m resolution, and the Multispectral Linear Array (MLA) with 20-m resolution. Both can be inclined up to 27° from vertical.

MacDonald Dettwiler's proof-of-concept work in extraction of height information from stereo imagery was based on Landsat TM imagery. The Landsat TM sensor is permanently pointed vertically at the Earth, and provides stereo imagery for only limited regions of the Earth's surface, *viz.*, those regions for which images from two adjacent orbits overlap. The elevation accuracy attainable from Landsat imagery is limited by the fact that the imagery has a small base-to-height ratio. SPOT, on the other hand, can produce stereo imagery with a base-to-height ratio of up to 1.0 if both images are acquired with maximum sensor inclination (Figure 1).

Height information is stored in stereo pairs of digital images in the same way as it is stored in stereo photographs, *viz.*, as stereo parallax between corresponding features in the two images. The main differences between aerial photography and satellite imagery as far as height information is concerned lie in the devices suitable for extracting the information. As mentioned, analytic stereoplotters need reprogramming to geometrically correct satellite imagery. Furthermore, satellite imagery is acquired in digital format, and so is degraded in conversion to an analog format, such as photographic film, for use in a stereoplotter. Extraction of height information from satellite imagery requires that parallax measurements be made with subpixel accuracy, which is very difficult to accomplish with imagery in film formats. Parallax measurements have been performed for well-defined targets in stereo SPOT imagery using an analytic stereoplotter (Denis, 1987). However, it does not appear that the human eye is capable of making the subpixel interpolation required for general height extraction applications. To maximize the accuracy of height information derived from digital satellite imagery, it is desirable to maintain the imagery in digital format, model the imaging process digitally, and derive corrections that can be applied to the digital image data.

MAP PRODUCTION FROM SATELLITE IMAGERY

Remote sensing satellites such as Landsat and SPOT transmit imagery that can be viewed and processed as it is received at an appropriate ground station. However, remote sensing satellite ground stations are very expensive to set up, and the normal route to acquiring imagery is to order specific scenes through agencies which have their own ground stations and archive all imagery received. Before this imagery can be used for map production, it must be geometrically and radiometrically corrected. The first step toward geometric correction is the identification of ground control in the imagery.

Ground Control

The most suitable ground control points (GCPs) for use with satellite imagery are road intersections and other features that can be seen in the imagery. These may be identified from GCP archives, or surveyed in for a specific mapping project. If no visible GCPs are available, ground control can be transferred to satellite images from archived aerial photographs. A high resolution aerial photograph of each GCP is digitized, and the pixels of the digitized photograph containing GCPs are identified. These higher resolution photographs can then be correlated with the satellite images, and GCP positions recorded in terms of pixel coordinates within the satellite imagery.

Landsat and SPOT transmit not only image data, but also ephemeris data from which their attitude and orbit position during acquisition of any given image can be inferred. In general, current satellite image correction techniques need only four to six GCPs per scene for correction to sub-pixel accuracy, and GCP placement is not particularly critical (Welch *et al.*, 1985). Using geometric correction methods based on spacecraft modeling, MacDonald Dettwiler has corrected SPOT scenes to subpixel accuracy using as few as three GCPs per scene (see Figure 2).

Compared to conventional aerial photogrammetry, this results in a substantial decrease in the number of GCPs required per unit area mapped. A standard-SPOT scene, requiring only three GCPs for correction to better than 10-m RMS, covers 360,000 hectares (60 km by 60 km), while an aerial photograph at 1:50,000 scale covers only 13,000 hectares.

Geometric Correction

There are three levels of geometric correction appropriate to satellite imagery. First, the imagery can be corrected for the effects of the satellite's changing attitude and irregularities in its orbit. At this level the perspective distortion remains, and the imagery is said to have been corrected to an "ideal satellite projection." Georeferencing, the next level, corrects for the satellite's perspective, and transforms the image to a map projection. However, georeferenced image pixels and scan lines are not aligned with the coordinate system grid, as shown in Figure 3.

Geocoding goes two steps beyond georeferencing. The imagery is rotated to align the scan lines with the map projection grid, and is resampled to a standard rectangular pixel size. This pixel size is selected from a set of exact multiples to allow for different sensor resolutions.

Geocoded imagery has several advantages over georeferenced imagery. Georeferenced imagery is difficult to combine with other data due to differences in image scan line orientation and pixel size. Geocoded imagery can be directly overlayed and used with other digital map data, or geocoded imagery from other sensors. Finally, geocoding allows graphics data, such as that stored in a digital mapping system or geographic information system (GIS), to be combined with image data. The ability to overlay digital satellite images is quite important and is a very different process from overlaying photographic images. Different kinds of image data can be combined to complement each other, as shown by Plates 1a and 1b.

Since the late 1970s, image analysts have been able to geocode images individually by specifying warping functions to carry out the transformation. However, in the early 1980s MacDonald Dettwiler developed a technique for automatically geocoding any satellite image (Friedman, 1981). This "operational" geocoding is necessary for the extensive image processing required to derive maps from images.

DEM Extraction

A digital elevation model (DEM) is a numerical representation of surface elevations over a region of terrain. DEMs provide the same sort of information as contour maps, but in a digital format suitable for processing by computer-based systems rather than in an analog format. Two broad kinds of DEM can be distinguished. "Gridded" DEMs record a surface elevation for every intersection in a two-dimensional coordinate grid covering the region under investigation. "Feature" DEMs record only a random distribution of elevations across the region, specifically those elevations along the boundaries of Earth surface "features," such as roads, rivers, and ravines. Any identifiable object on the Earth's surface can constitute a "feature" in this sense.

MacDonald Dettwiler has developed an automated process for digitally extracting gridded or feature DEMs from satellite

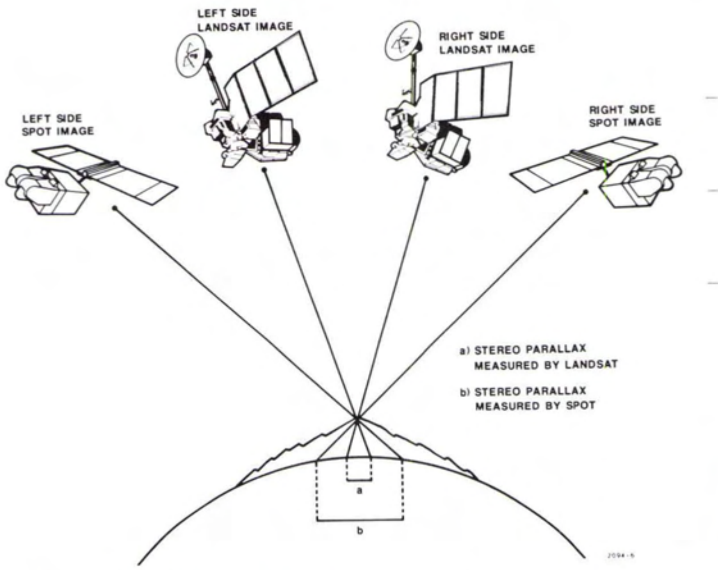


FIG. 1. SPOT's ability to incline its sensors up to 27° from vertical allows it to produce stereo imagery with a base-to-height ratio as high as 1.0. Landsat's sensors are permanently pointed downward, and so produce full stereo imagery for only limited areas of the Earth's surface.

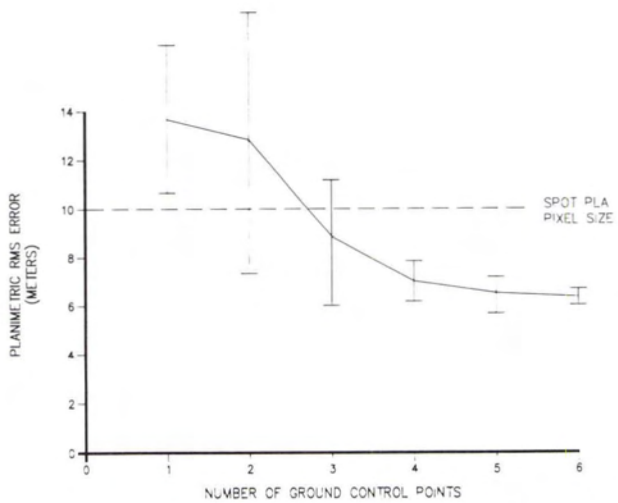


FIG. 2. RMS planimetric errors and standard deviations remaining after geometric correction of a SPOT scene by spacecraft mod techniques, as a function of the number of GCPs used. The results were obtained for a test area in the United Kingdom, and include an estimated 4- to 5-m RMS error in the reference points used.

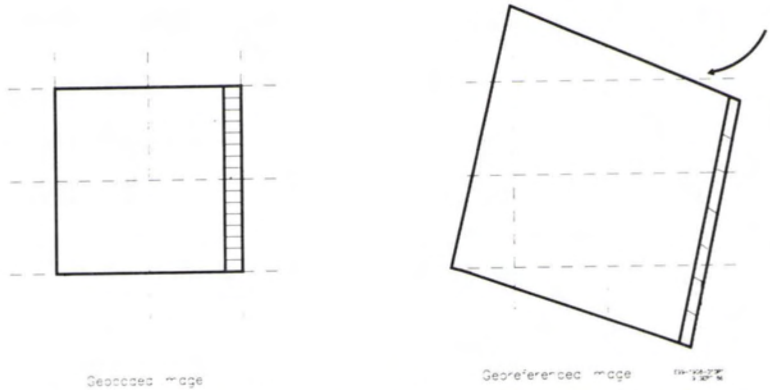


FIG. 3. Geocoded imagery is resampled to a uniform pixel size, and rotated to align with the coordinate grid of a chosen map projection. Georeferenced image pixels are registered to a map projection, but may represent regions of different area, and will not generally be aligned with the coordinate grid. Consequently, it is very difficult to combine different digital images, or digital imagery and graphics data, if the imagery has been georeferenced but not geocoded.

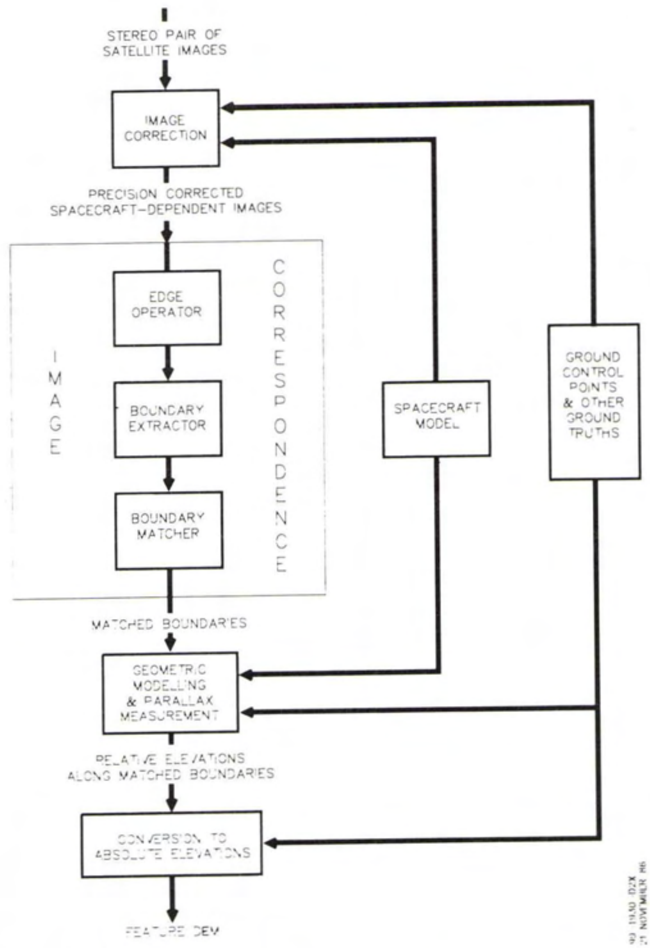
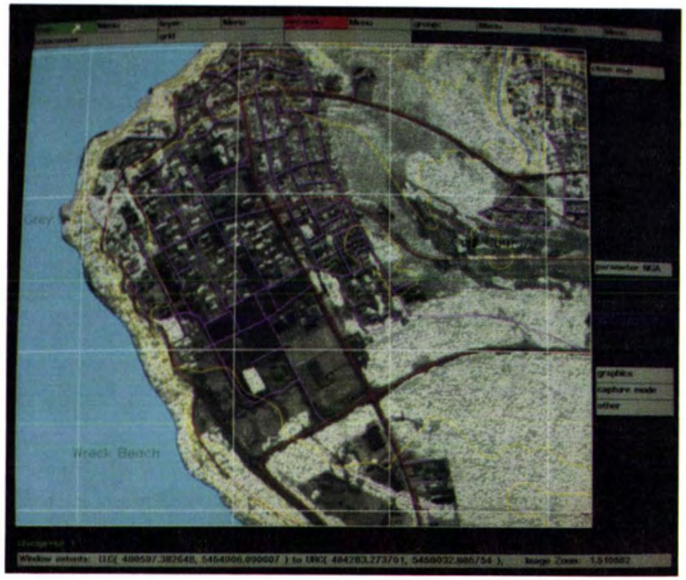


FIG. 4. MacDonald Dettwiler's automated DEM extraction process, based on machine vision techniques.

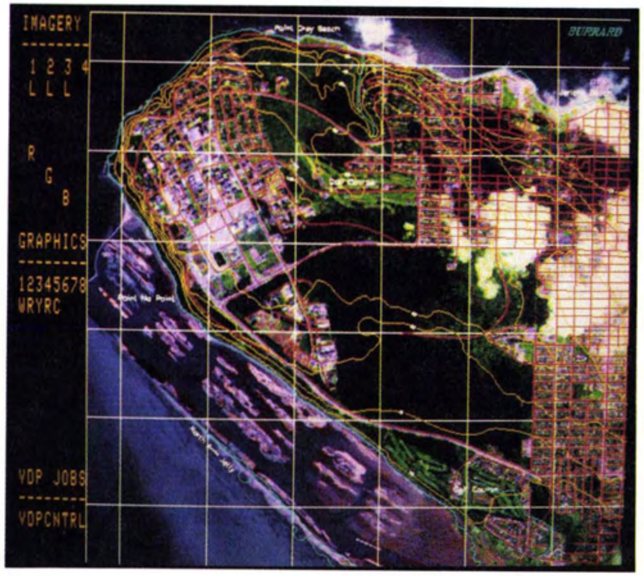
imagery. The major steps for deriving feature DEMs are described in the following paragraphs, and illustrated in Figure 4.

The stereo imagery is geometrically corrected to only an ideal satellite projection, to leave height information intact. Corresponding stereo images are matched using techniques originally developed in the field of computational vision. Image correspondence includes three steps: application of an edge operator, boundary extraction, and boundary matching.

The edge operator locates those lines in an image where there is a sharp change in image intensity. These lines are assumed to represent the edges of features on the Earth's surface. Next, a boundary extractor builds feature descriptions from the detected edges. It locates and links together edges to form boundaries, and gauges the shapes of the boundaries to determine the features represented in the images. The linked boundaries are then screened to separate those due to image noise from those due to actual imaged features. As a final step, a boundary matcher determines for each boundary in one image the corresponding



(a)



(b)

PLATE 1. Plate 1a is a 10-m monochromatic SPOT PLA image of Vancouver, Canada. Plate 1b is a 30-m multispectral Landsat TM image of the same scene, modulated with Plate 1a. This has been achieved by geocoding both images to the same map projection, and overlaying their corresponding pixels. Thus the composite image, Plate 1b, combines all the advantages of the multi-spectral information of a Landsat TM image with the higher resolution of SPOT PLA imagery. Both images have also been overlaid with a digitized map of the area, another process for which geocoding is essential.

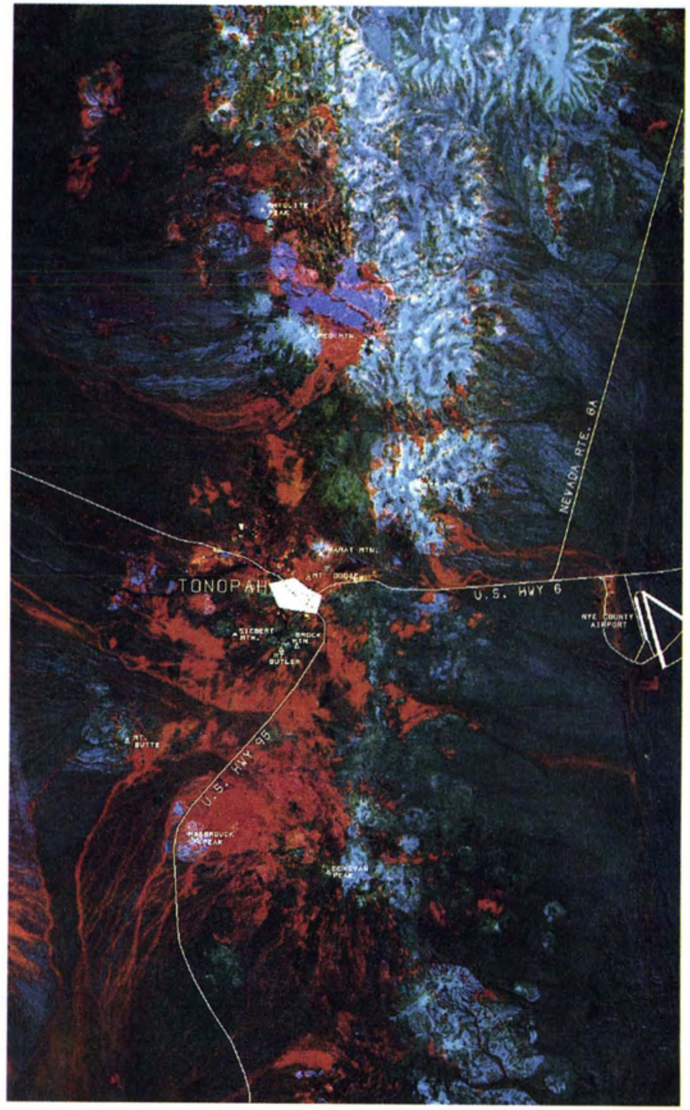


PLATE 2. This image of the Tonopah, Nevada region of the United States shows how ratios between TM band values can help in geographical analysis. The image has been processed to display band ratios which highlight the presence of hydroxyl-bearing minerals. Regions of surface material containing these minerals are indicated in red.

PLATE 3. In this Landsat TM orthoimage of the Adam River area of Vancouver Island, British Columbia, the white areas are covered in snow, the dark green areas are heavily forested, and the light green areas are open patches. Brown areas are bare of vegetation. This image has been processed to provide something like the equivalent of a color photograph taken from the satellite.

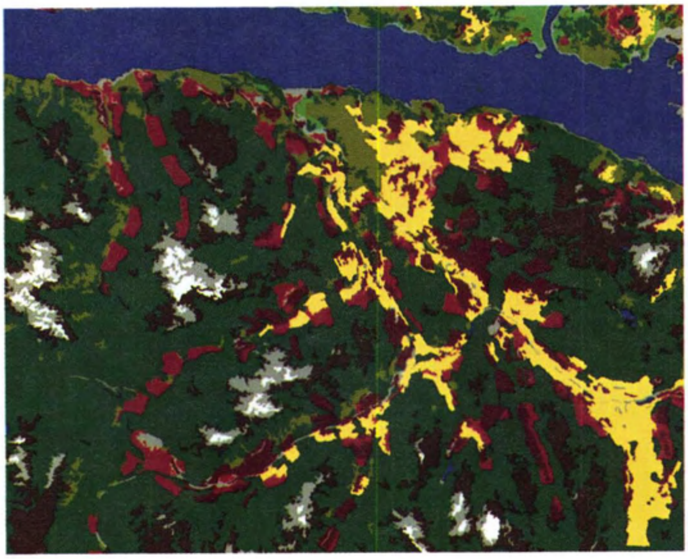


PLATE 4. This image has been derived from the same raw image data as Plate 3, but processed to exploit the multispectral information it contains. The image shows how fine distinctions between different kinds of vegetation can be discerned by automated radiometric classification. Different kinds of surface vegetation have been identified by their spectral signatures, and represented by colored polygons. Yellow represents clear-cut areas covered in young hemlock, orange represents clear-cut areas covered in huckleberry, and the light green areas are mature hemlocks.

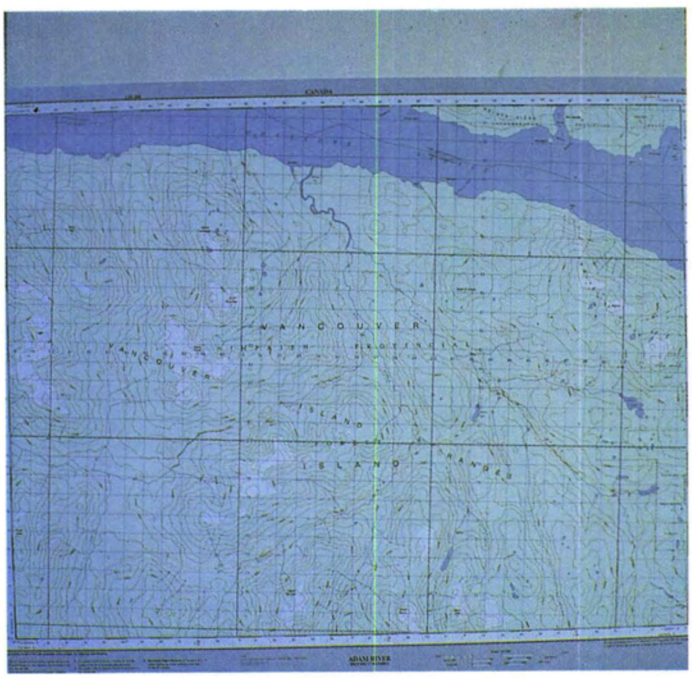


PLATE 5. This topographic map of the Adam River area (Plates 3 and 4) has been derived entirely from Landsat TM imagery.

boundary in the other image (if any). Once it has determined the correspondence, it calculates the disparity or parallax between the corresponding boundaries.

Boundary elevations are calculated from the boundary correspondence using geometric modeling algorithms. Geometric modeling relates the stereo imaging geometry and the parallax measured by the boundary matcher to the desired elevations. Spacecraft models derived in the image correction phase and the GCP information are used in the geometric model.

Parallax is measured to subpixel precision to obtain sufficient accuracy in the DEM. To obtain high precision, several models must be used to compensate for Earth, image, and satellite effects. The geometric modeling allows calculation of elevations along the matched boundaries. Because the boundaries represent Earth surface features, these elevation data constitute a feature DEM. At this stage elevations are relative to a floating datum plane defined by zero parallax. Ground control is then used to convert relative elevations to absolute. If desired, gridded DEM can be derived from the feature DEM using an interpolation algorithm.

MacDonald Dettwiler has used this process to prove the concept of DEM extraction using Landsat TM imagery (Cooper *et al.*, 1987). However, to date SPOT is the only commercial satellite to provide imagery with sufficient resolution and stereo coverage for practical DEM applications. MacDonald Dettwiler's recent rigorous tests with SPOT imagery over a wide range of terrain types in both Southern Europe and the United Kingdom have demonstrated RMS accuracies better than 6 m, which will allow contour mapping with a 20-m contour interval at full accuracy (see also Kauffman *et al.* (1987)). Similar results have been reported by Denis (1987).

Orthoimage Production

"Orthoimages" are the image equivalents of orthophotos. They can be produced by geocoding satellite imagery using an imaging model that incorporates DEMs for the terrain. MacDonald Dettwiler's Geocoded Image Correction System (GICS) performs this function, and Plates 3 and 4 are examples of orthoimages produced by it.

Planimetric Feature Extraction

Orthoimages are analyzed to extract planimetric feature information. The analysis techniques used include visual enhancement and classification, and can only be applied to images in digital format. Visual enhancements, such as pseudo-coloring, contrast stretching, and edge enhancement highlight the most interesting features in the imagery. Image classification is used to extract themes or classes representing land use/cover. These automated extraction processes have to be supplemented with height information from the DEM.

A wide variety of map products can be derived from the planimetric information. Graphics data from vector sources such as digital mapping systems and GISs can be overlayed on orthoimages and plotted as hardcopy "image maps." Thematic information about areal, lineal, and point features can thus be represented by colored polygons, lines, and conventional point symbols.

Finally, paper topographic or thematic maps can be printed, incorporating contours derived from image-based DEMs. The topographic map of the Adam River area shown as Plate 5 was derived entirely from Landsat TM imagery (Rose *et al.*, 1986).

CURRENT DEVELOPMENTS

This section discusses current research into improvements in the map production techniques presented in the previous section. The improvements are concerned with "pass processing," a technique for processing a series of image scenes as if they were a single scene, automated extraction of DEMs from satellite imagery, and the use of artificial intelligence (AI) techniques for planimetric feature extraction.

Pass Processing

Geometric correction of satellite imagery requires knowledge of the position and orientation of the satellite during the imaging period. MacDonald Dettwiler's GICS stores this information in the form of a spacecraft model. A preliminary version of the spacecraft model is constructed from auxiliary data transmitted with the image data. This model is then refined by comparing the known location of GCPs with their locations in the imagery.

Pass processing is used both to reduce the number of GCPs required and to give more flexibility in their location. The technique can process a series of about ten scenes as if it were a single scene, so that the normal GCP requirement of about three to six GCPs per scene is averaged out to less than one GCP per scene. Also, the GCPs can be located anywhere within the pass. For example, in Australia GCPs can be located along the populated northern and southern coasts, and used to correct a pass of imagery including scenes of the center, for which control would be expensive to establish.

A MacDonald Dettwiler research team has investigated the processing technique for a pass consisting of 15 scenes of Landsat TM imagery. The spacecraft model was successfully extended (from its usual range of one scene) to cover the whole pass with accuracy comparable to that of single scene processing (less than $\pm$ 20-m RMS). The spacecraft models were refined with GCPs from throughout the pass but with no more than would have been required to correct a single scene model. This demonstrated that an order of magnitude reduction in the number of GCPs required is possible using pass processing techniques. Not only was the total number of GCPs used no higher than for single scene processing, but 11 scenes containing no GCPs at all were corrected.

Automatic Feature Extraction

Current MacDonald Dettwiler research into feature extraction aims to use "rule based" AI techniques to encode rules used by human interpreters in a form which can be used by a computer. Image processing techniques for spectral classification, segmentation, shape analysis, structural and contextual recognition, and knowledge based image interpretation have been investigated and developed. A development system has been implemented in PROLOG and interfaced to MacDonald Dettwiler's MERIDIAN image analysis system. The complete system is installed on a MicroVAX II.

The most recent work has been concerned with automating extraction of planimetric features such as roads, rivers, lakes, and shorelines from satellite imagery. To recognize a road, a photointerpreter uses structural, spectral, and contextual knowledge. A computer system that is to extract features with any significant degree of competence must be able to use these knowledge sources also.

One method tested has incorporated structural knowledge of roads. Edges are extracted from the imagery and approximated by straight lines, and potential road segments are inferred from parallel lines separated by road widths. Collinearity and connectivity rules are then applied to join road segments. Road sections can be extracted by optimally joining the road segments obtained by this approach using graph searching techniques. This method can accurately extract the centerlines of roads, as required in topographic mapping. Bridges can be recognized by applying a knowledge based rule which describes a bridge as a road segment bounded by water (Yee, 1987).

CONCLUSION

Once the technical problems have been resolved, cartographic standard maps are expected to be produced from satellite imagery for a fraction of the cost of conventional methods. Automated image processing is not the only reason that map production from satellite imagery will be less expensive than conventional photogrammetry. Satellite images cover far larger areas, and can be obtained for a fraction of the unit area cost of aerial photography. Also, far fewer GCPs are needed for satellite mapping. Depending on the terrain and location, the current cost of establishing a GCP in Canada, for example, ranges from $400 to $8,000.

The potential cost advantages of deriving mapping information from satellite imagery are currently balanced by the low resolution of satellite sensors compared to aerial cameras. This makes identification of the complete range of features normally found on a 1:50,000-scale map sheet difficult.

Nevertheless, digital processing techniques which make use of spectral analysis allow identification of features that are not visually identifiable in the imagery, in that they cannot be identified by their shapes alone. Further research into AI techniques for scene analysis are particularly important here.

Despite the problems, satellite imagery and digital image processing hold great promise for topographic and thematic map production. Applications such as DEM extraction with less than 10-m RMS error were impossible with Landsat imagery, but are now possible with SPOT. Satellite image mapping should support low cost mapping operations in countries where accurate mapping has not yet taken place. It should also allow low cost updating of existing maps and creation of digital topographic data bases. In this respect it is particularly suited to thematic applications, such as forest resource mapping.

REFERENCES

Cooper, P.R., D.E. Friedmann, and S.A. Wood, 1987. The Automatic Generation of Digital Terrain Models from Satellite Images by Stereo. *Acta Astronautica*, Vol.15(3), pp. 171–180.

Denis, P., 1987. User Activities. *Nouvelles de SPOT (SPOT Newsletter)*, No.9, pp. 7–11.

Friedman, D.E., 1981. Operational Resampling for Correcting Images to a Geocoded Format. *Proceedings of the Fifteenth International Symposium on Remote Sensing of Environment*, pp. 195–212.

Kauffman, David S., and Scott A. Wood, 1987. Digital Elevation Model Extraction from Stereo Satellite Images. *Proceedings of the 1987 International Geophysical and Remote Sensing Conference*, Vol.1, pp. 349–352.

Rose, D.R., I. Laverty, and M. Sondheim. 1986. Base Map Production from Geocoded Imagery. *Proceedings of the 7th Symposium of the International Society for Photogrammetry and Remote Sensing* (Commission VII: Interpretation of Photographic and Remote Sensing Data), Vol.26(7/1), pp. 67–71.

Welch, R., T.R. Jordan, and M. Ehlers, 1985. Comparative Evaluations of the Geodetic Accuracy and Cartographic Potential of Landsat-4 and Landsat-5 Thematic Mapper Image Data. *Photogrammetric Engineering and Remote Sensing*, Vol.51(9) pp. 1249–1262.

Yee, Ben, 1987. An Expert System for Planimetric Feature Extraction. *Proceedings of the 1987 International Geophysical and Remote Sensing Conference*, Vol.1, pp. 321–326.