Spatial information of geography with Video Map

Mitsuaki MATSUYAMA, Toru INOUE
Technical Planning Department
AERO ASAHI CORPORATION
Minamidai 3-1-1, Kawagoe, Saitama, 350-1165
E-mail: mse@biglobe.ne.jp
JAPAN

Commission V, Working Group IC V/III

KEYWORDS: HDTV, Spatial information, DTM

ABSTRACT

In recent years, there is a growing tendency to use GIS (Geographic Information Systems) for facilities maintenance or in response to emergencies or in disaster control planning. Also, with the development of hardware, as practical and visual information, examples of using image data have become more prevalent.
Up to now, satellite photo images or aerial photo images were used as the source for these types of luster data, but have been criticized as not being very suitable for use when there occur sudden change in land form or other disastrous situations.
We have therefore developed "Video Map", a method of making luster data using relatively HDTV (High Definition Television) images. This is a method whereby, form the HDTV images received from a stabilized HDTV camera in the air pointed at the gravitational direction, a scanning spot of each field (1/60th of a second) is randomly used and geometric correction is made using the orientation parameters obtained at the same time as the filming.
In this thesis, while introducing the outline of this method, I would like to investigate the use of this method for expressing spatial information of geography.

1. AHVS(Asahi HDTV Videogrammetric System)

1.1 Outline

This system consists of an aerial measuring system, ground measuring system and ground analysis system.
The aerial measuring system records in a VCR the images taken by a HDTV (High Definition TV) installed in a helicopter, and at the same time records other data such as the ground altitude data, magnetometer data and the film direction data, which are the 3 axis gyro data, and the measurement data from the GPS receiver in a PC. The images and flight information are synchronized by inputting the VCR time code and the flight information into the PC at the same time. The images taken from the helicopter are not shaking because a stabilizer is used to balance the shaking of the helicopter. Also, by using a gyro and gimbal together, the camera lens is controlled to always point in a fixed direction in the inertial system (when producing the Video Map in the gravitation direction) to counteract the shaking around the roll, pitch and yaw axis angles. Using the magnetometer data, the camera direction angle from the airframe axis can be converted to the angle from magnetic north direction. Due to the GPS receiver, the camera position and height data can be obtained. The GPS position measurement method uses the DGPS method to take into account the high-speed movement over a wide area and interruptions in waves due to the shaking of the helicopter.
The ground measuring system uses a GPS antenna placed in a position already measured, and records position correction data received by the GPS receiver in a note PC. This operation is used for differential treatment so it must be recorded at the same time as the aerial measurement system.
The ground analysis system produces orientation data from the flight information, position correction data and the various calibration data. Also, the measurement of the image length and height, etc. and the production of the Video Map is done with the recorded image data and orientation data.

1.2 Orientation Data

1.2.1 Position and Inclination of the Camera: An ASHTECH GPS receiver is used for measuring the position of the camera. The antenna is attached to the right front of the helicopter. The correct position of the camera is determined, reflecting the relationship of the positions of the camera and the antenna. The inclination of the camera is determined by the output data of the 3 axis gyros. This is determined by the output of the angles (AZ, EL, RL) based on the airframe and angles (PITCH, ROLL) based on horizontal surface.

1.2.2 Principal Point: The video signals from the camera are first recorded on VCR tape in the helicopter. After returning to the ground the tape is played back on the ground analysis system VCR and put into a framegrabber. Because the actual measurements are done on the computer screen, it is necessary to have a principal point in the image data memory. The lens principal point is found by using the convergence of the image on the principal point when the zoom lens direction is changed from tele side to the wide side. A chart as shown in Figure 1 is filmed and recorded on a VCR tape. The image coordinates of the corners of the chart are measured with the ground analysis system, and with a suitable combination of linear regression equations the points of intersection are measured. The average value of the intersection is used as the principal point. The measurement results up to now have shown differences in about 2 picture elements.

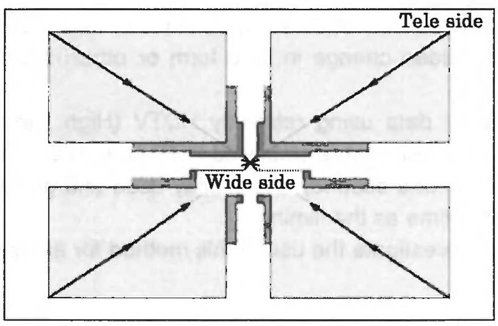

Figure 1. principal point measurement

1.2.3 Distortion: The lens distortion is calculated by interpolation using measurement data from the lens manufacturer. Figure 2 shows the radial distortion of the camera CCD surface according to the zoom lens output voltage . The smaller voltage is the wide-angle side.

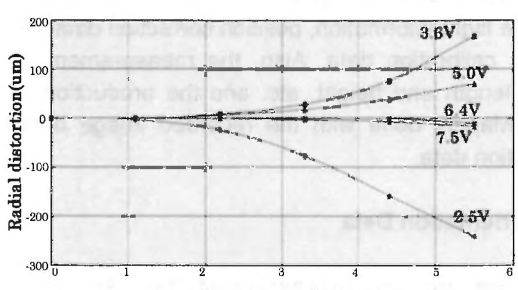

Figure 2. length and distortion

1.2.4 Principal Distance: To measure the length and position of the object filmed, it is necessary to know the accurate principal distance. Because a zoom lens is used in the aerial measurement system, the principal distance continuously changes. Therefore it is necessary to fix the zoom lens output voltage and the principal distance. This operation is done by filming a target, which the length is known of, while changing the magnification of the zoom lens, and at the same time recording the voltage data of the zoom lens. An image of the target is produced for measuring purposes, and by adjusting the distortion of the target's image coordinates, the length on the CCD surface of the target is found. As shown in Figure 3, using the relationship between the length of the target in the image and the distance to the target, the principal distance is obtained. Figure 4 shows the principal distance and output voltage relationship obtained by using the above method.

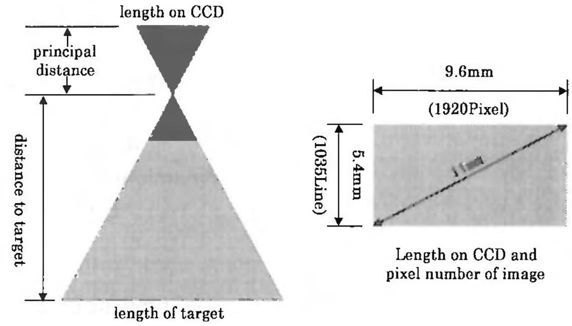

Figure 3. target length and principal distance

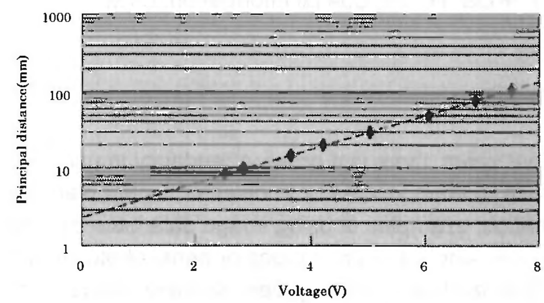

Figure 4. principal distance and voltage

2. Video Map

2.1 What is the Video Map?

The map image produced using the image filmed by the AHVS and the camera position and inclination information is called the Video Map(Figure 5), and it is produced by the steps shown in Figure 6. The Video Map has the following characteristics:

- The HDTV system is used from filming to producing the image, so the image is clearer compared to the presently used broadcasting standards(NTSC,PAL).
- Along the filming direction, the beginning and end points can be freely chosen, and because a continuous image along that sector can be achieved, it is suitable for the maintenance of railways, roads, rivers and power transmission line equipment, and land use surveys.
- With one session, a forward looking Video Map

image, nadir looking Video Map image and backward looking Video Map image can be obtained, so multi-base line stereo matching is possible.

Figure 5. Video Map

2.2 A/D Conversion

When loading images into a computer from VCR tapes for measurement purposes, generally the scenes are done in every frame or every field. Like ordinary photos, images obtained this way are central projection. As opposed to this, the original images of the Video Map are images using a single scan line continuously building on top of the output images, so they are images built up from central projection images for each line, the same as the line sensor images of LANDSAT and SPOT(Figure 7). Also, when producing a nadir looking Video Map, it uses the scan line of the principal point, for a forward looking Video Map a scan line higher than the principal point, and for a backward looking Video Map a scan line lower than the principal point.

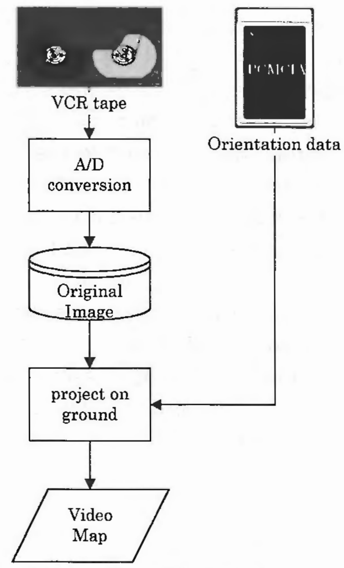

Figure 6. making process

With the ground analysis system for aerial HDTV, a VCR tape image is A/D converted through a framegrabber and the results are fed into the EWS hard disc. As well as being able to read a whole scene by each field, this framegrabber can also read scan lines selected freely. By using this ability, the same scan line is fed in by each field and built up line by line so that the output image line expresses the field of VCR time code to produce the original Video Map images .

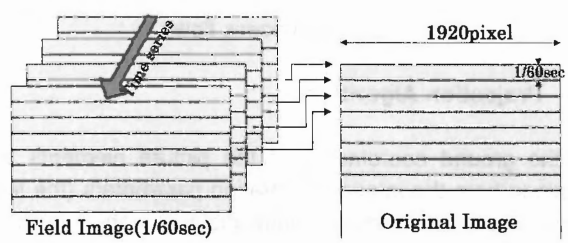

Figure 7. original image

2.3 Inclination of the Camera (κ)

Originally, it was planned to use a magnetometer sensor to obtain the inclination (κ) around the Z axis of the camera. However, its location environment (magnetic field, vibrations) affects this magnetometer sensor, and it was realized that it would be extremely difficult to decide the azimuth using this data. Therefore, since the camera is stable in the inertia space, and because the GPS data is being collected at the same time, it was decided to obtain κ by the method shown in Figure 8.

First, the field image to be used as the base and the field image 1 second later are fed in from the continually recorded VCR tape. Next, as shown in Figure 9, the direction in which the base field image near the principal point has moved in the image 1 second later is determined, matching by the value of correlation. From this result, as shown in Figure 10, the angle at the image coordinate (the angle, going clockwise, with 0 degrees being vertically upwards from the principal point) is θ_m, it is turned a further −180 degrees and changed to the direction of the helicopter (θ_h). Also, the direction of the helicopter is traveling in on the map is obtained from the GPS location data, and that is θ_g. The last angle needed to be found is, as shown in Figure 10, the inclination of the camera, κ, and this can be found by the difference between θ_g and θ_h. From hereon until the end, by moving it every second, the linear regression equation between the various angles and times was determined, and from this formula the κ for each field was calculated.

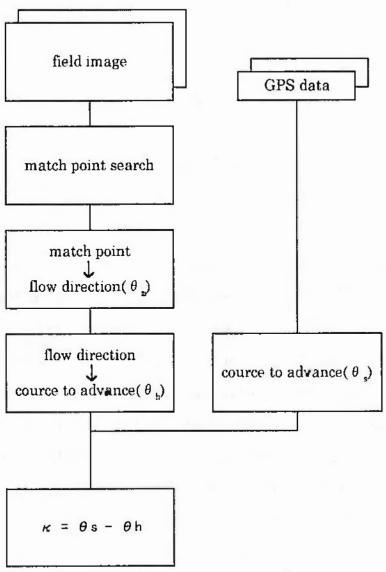

Figure 8. process flow

2.4 Projection Algorithm

All the ground coordinates of the picture elements are obtained from the interior orientation parameters (the lens principal point, distortion, image distance) obtained from the AHVS and the exterior orientation parameters (position, inclination). Because the original image of the Video Map is a single photograph of each line, it has orientation parameters for each of the lines, but apart from that it is the same as general photogrammetry. If the relationship between the converted photographic coordinates (x, y) and the ground coordinates (X, Y, Z) is used, absolute orientation is possible. For that purpose, the following well-known relationship (collinearity condition) should be used.

$$\left.\begin{aligned} x &= -f\frac{a_{11}(X-X_0)+a_{12}(Y-Y_0)+a_{13}(Z-Z_0)}{a_{31}(X-X_0)+a_{32}(Y-Y_0)+a_{33}(Z-Z_0)} \\ y &= -f\frac{a_{21}(X-X_0)+a_{22}(Y-Y_0)+a_{23}(Z-Z_0)}{a_{31}(X-X_0)+a_{32}(Y-Y_0)+a_{33}(Z-Z_0)} \end{aligned}\right\} \quad (1)$$

Here, f is the image distance, (X_0, Y_0, Z_0) are the ground

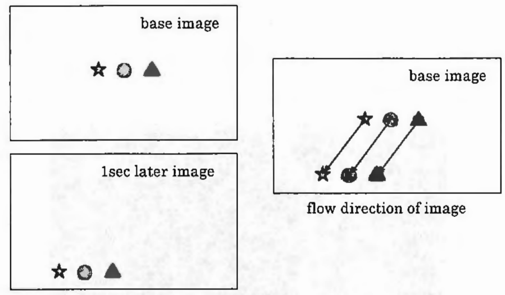

Figure 9. image flow direction

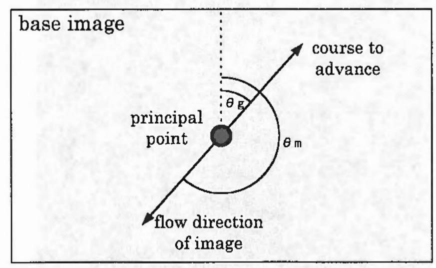

Figure 10. inclination of camera(κ)

coordinates of the photo projection center. Also, the 9 coefficients a_{11} ~ a_{33} are determined from the camera photographic axis inclination (ω, ϕ, κ) . Here, ω, ϕ and κ are positive when it is turning counterclockwise to the positive direction of the X, Y and Z axis. the inverse transformation is as follows.

$$\left.\begin{aligned} X &= (Z-Z_0)\frac{a_{11}x+a_{21}y-a_{31}f}{a_{13}x+a_{23}-a_{33}f}+X_0 \\ Y &= (Z-Z_0)\frac{a_{12}x+a_{22}y-a_{32}f}{a_{31}x+a_{32}y-a_{33}f}+Y_0 \end{aligned}\right\} \quad (2)$$

3. Producing a DTM using the Video Map

3.1 Outline

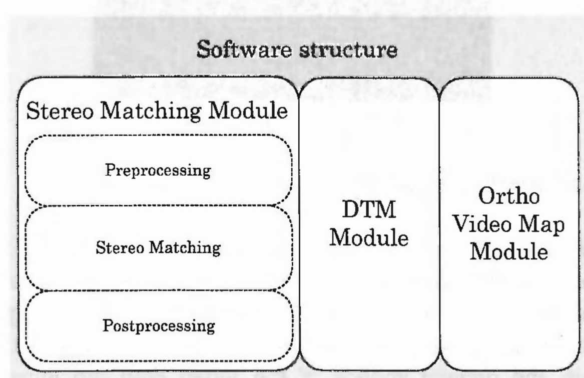

Figure 11. software structure

A DTM is produced by stereo matching the forward looking Video Map and the backward looking Video Map. The software is made up as shown in Figure 11, with the functions of being able to stereo match the Video Map stereo pair, being able to calculate the DTM from the parallax difference, and being able to make an Ortho Video Map from the DTM. Also, because of the high speed processing of the stereo matching and the prevention of any straying, multi-stage stereo matching is used, and the resultant parallax difference is fed into the matching grid points in the following stage.

3.2 Stereo Matching Preprocessing

The raw material of the Video Map images is data in RGB space, and each of these RGB have 256 gradation. However, it is not practical to use matching by the value of correlation for all the RGB bands for the stereo matching considering the time involved. In the template image, 1 RGB band is selected, or various RGB bands are selected in turn, and the stereo matching can be undertaken, but in this system, the following method was selected because of its simplicity, converting the RGB colorimetric system to XYZ colorimetric system, and using Y (brightness).

$$\left.\begin{aligned} X &= 0.49000R + 0.31000G + 0.20000B \\ Y &= 0.17697R + 0.81240G + 0.01063B \\ Z &= 0.01000G + 0.99000B \end{aligned}\right\} \quad (3)$$

By converting to XYZ colorimetric system, tristimulus vales of X, Y and Z became the real values. Therefore, there became more bands than the simple RGB space bands, leading to a higher precision in the stereo matching.

3.3 Stereo Matching

There is an element of error in the DTM, matching error, the cause of which is thought to be a lack of dynamic range in the template image. If the template image can be judged to the unsuitable for stereo matching beforehand, this will result in less matching errors. For this purpose the histograms of template images and standard deviation was obtained, and when there was a good standard deviation this was considered not to have enough dynamic range, and was not used for stereo matching.
Also, with stereo matching, the smallest units for determining the corresponding points are the picture elements, so the parallax difference cannot be of finer precision than the picture elements. So for cases with stereo matching where there was a need for high precision, the stereo matching of generally used sub-pixel units was used. Here, the maximum value of the correlation coefficient and the correlation coefficients before and after the maximum value were used as the vertical axis, and for horizontal axis as the picture element, the points where the regression curve of secondary degree obtained by method of least squares has maximum values were used.

3.4 Stereo Matching Postprocessing

Parallax difference postprocessing is undertaken in order to remove any matching error. In order to do this, the parallax difference is positioned in 2 dimensions, the spike noise is taken out by using a median filter, and smooth processing is undertaken. Interpolation was conducted where the stereo matching was not done. The median filter can effectively take out spike noises, without damaging important information such as the edge, and unexpected matching errors can be eliminated.

4. Ortho Video Map Verification

In order to verify the position precision of the Ortho Video Map, an air-to-ground mark was set up on land near this company and GPS measurement and Video Map filming was carried out. Stereo matching was done with the forward looking Video Map images and backward looking Video Map images as shown in table 1, and an Ortho Video Map as shown in Figure 12 was made using the obtained DTM in order to measure the position coordinates of the air-to-ground mark. The position of the air-to-ground mark from the GPS measurements is shown in table 2, and the coordinates from the image measurements are shown in table 3.

Data	Date	Scan Line
forward Video Map	1997/03/07	216
backward Video Map	1997/03/07	716

Table 1. use data

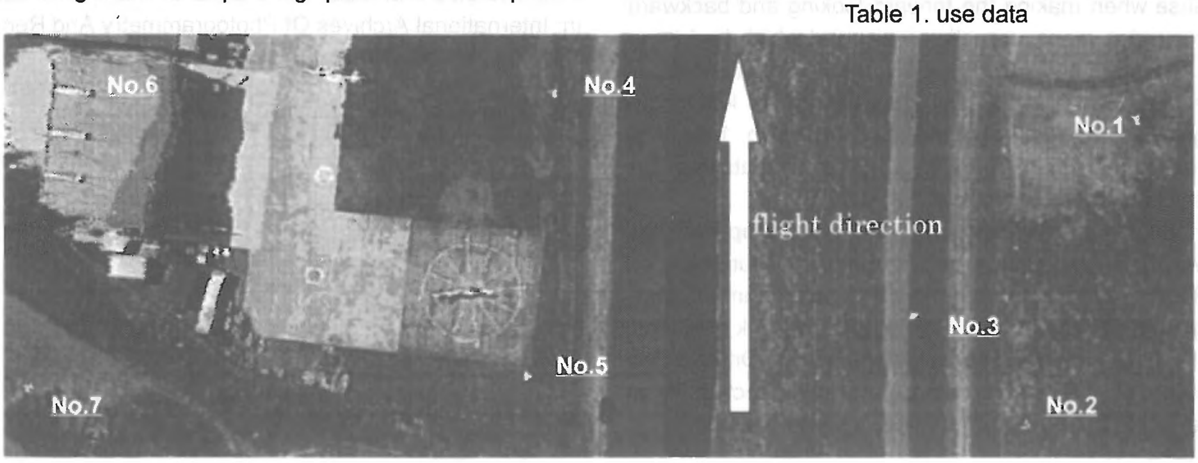

Figure 12. target (ortho image)

No	Easting(m)	Northing(m)	Hight(m)
1	362145.73	3979787.54	12.44
2	362202.53	3979766.62	12.40
3	362182.16	3979745.67	16.63
4	362141.02	3979678.51	13.11
5	362193.81	3979673.67	13.13
6	362134.56	3979586.46	31.45
7	362196.35	3979580.25	14.06

Table 2. target position(GPS positioning)

No	Easting(m)	Northing(m)	Hight(m)
1	362146.75	3979792.75	20.8
2	362204.50	3979771.75	17.6
3	362183.50	3979750.75	21.0
4	362141.75	3979682.75	16.6
5	362193.00	3979677.75	17.0
6	362135.25	3979589.25	31.7
7	362197.25	3979583.75	18.1

Table 3. target position (image measurement)

No	Easting(m)	Northing(m)	Hight(m)
1	-1.02	-5.21	-8.36
2	-1.97	-5.13	-5.20
3	-1.34	-5.08	-4.37
4	-0.73	-4.24	-3.49
5	-1.19	-4.08	-3.87
6	0.31	-2.79	-0.25
7	-0.90	-3.50	-4.04

Table 4. difference

5. Consideration

After making a DTM from the stereo matching of the forward looking video map and the backward looking video map, each of their air-to-ground mark positions was compared with the height above sea level obtained by the GPS measurement. The comparison shows all the heights of the air-to-ground mark positions in the DTM were higher than those by the GPS measurement. This is thought to be because when making the forward looking and backward looking video maps, a section was used which had large lens distortion, and inaccurate distortion correction values were used causing the lens distortion to be left in the image. Also, the reason for the residuals of No. 1 and No. 2 to be bigger than the others was because matching was done on the trees in the area.

Furthermore, after making an Ortho Video Map from the derived DTM, the results of the image measurement on the Ortho Video Map and the GPS measurement were compared with regard to the air-to-ground mark positions. All the air-to-ground mark positions of the Ortho Video Map were out of line to the right in the flight direction. What is thought to be the reason for that is shown below.

Data	Precision
GPS	0.1-1m (PDOP<4)
Gyro	±0.2 degrees
flight height	400m

Table 5 Margin of Error Factors

Thinking from the GPS recording time and the PDOP, it is thought that the precision of the GPS position was less than 0.5m. On the other hand, with the precision of the gyro at ±0.2 degrees, calculating the ground position from a height of 400m, there should be a margin of error of ±3m, and therefore having a deteriorating effect on the Ortho Video Map. Also, the standard deviation of the residuals was E:0.64m, N:0.85m.

This time, matching of buildings was also tried, but the results were not good enough to be practically used. For this reason, in the present form, not buildings but natural landscape was used as the object, and because digital images can be made from video recordings of these areas in a short amount of time, it is more than effective for use in times of natural disaster for measuring the extent of sediment discharge, etc.

In the future, in order to increase the precision of the Ortho Video Maps, it is planned to consider the following points.

- Improving the precision of the gyro
- Using kinematic measurement (GPS position measurement method)
- Improving the flight method: filming at lower altitudes
- Improving the precision of the DTM by multi base line processing

6. References

Chikatsu, H., 1997, Image Sensing, pp.57-73.

Star, J., Estes, J., 1990, GEOGRAPHIC INFORMATION SYSTEMS An Introduction, pp.1-22.

Takagi, M., Simoda, H., 1991, Handbook Of Image Analysis, pp.91-100.

Murai, S,. 1989, Analytical Photogrammetry, pp.34-56.

Heipke, C. Kornus, W. Gill, R. Lehner, M., 1990 MAPPING TECHNOLOGY BASED ON 3-LINE CAMERA IMAGERY In: International Archives Of Photogrammetry And Remote Sensing, Tukuba, Vol28, Part4, pp.314-323.