ISE-TR-84-48

WINDOW SIZE OF MATCHING METHODS IN STEREO AERIAL IMAGES

by

Takashi Hoshi

Takayuki Matsushita

Yasuhiko Ikebe

OCTOBER 30, 1984

INSTITUTE
OF
INFORMATION SCIENCES AND ELECTRONICS

UNIVERSITY OF TSUKUBA

Window Size of Matching Methods in Stereo Aerial Images

Takashi Hoshi
Science Information Processing Center
University of Tsukuba
Ibaraki 305 JAPAN

Takayuki Matsushita
Fujisawa Development Laboratory
IBM Japan
Kanagawa 252 JAPAN

Yasuhiko Ikebe
Institute of Information Sciences and Electronics
University of Tsukuba
Ibaraki 305 JAPAN

Abstract

For automatic productions of digital terrain models, several methods in which computer software matches the corresponding pixels have already been known. This study developed four matching methods and analyzed their performance by applying them to actual photographic image data. Judging from the matching results, Cross-Correlation method and SSDA were said to be more applicable in this case than two Fourier methods, Fourier Cross-Correlation method and Fourier Phase-Delay method. And the matching window size proved to play an important role for a successful matching of correponding pixels. It also caused a significant influence upon the computing time to be processed as well as the matching accuracy.

Therefore, this study has been directed specifically toward developing an effective way to get the optimum size of matching window in Cross-Correlation method. The person concerned in its applications could hardly determine the optimum matching window size without many trials and errors so far, because it is considered to be changeable according to the image pattern. The optimum size must be set taking an account of texture information in connection with the image pattern. This paper deals with a way to determine the window size for matching successfully, based on the Fourier transform of image data. And another approach using the error limit in the resultant elevations is also described in detail.

Contents

1. Introduction

Photogrammetry has been applied to widely different fields, varying all the way from astronomic to microscopic measurement. The most frequent use in its applications has been in the production of digital terrain models. They are playing an increasingly important role in urban planning, in superhighway routing, and many map-related tasks. They can be produced in several ways from aerial photographs. Viewed in the light of economy and efficiency, there has recently been an especially significant advance in computer techniques to produce them.

It the productions of digital terrain models, it is important to identify the parts of the two images in stereo aerial photographs that show the same piece of the scene. It will lead to measure their stereo parallaxes and finally determine the elevations of objects in the scene based on these parallaxes. Accordingly, the basic problem to be solved in automatic digital stereo mapping is the determination of the precise geometric positions in stereo photographs. In the practical cases, it is necessary for computer software to match a pair of pixels corresponding to the identiacl ground point in stereo digital images, which are digitized to be amenable to computer analysis and are rectified by relative orientation [P.R.Wolf and L.L.Christenson(1978)] .

2. Matching Methods and Applied Results

The authors have already advanced this research on developing an efficient digital stereo mapping procedure for identifying and locating corresponding pixels of aerial photographs with an overlapping of about

60%.

The following four matching methods which can match the correponding pixels were developed and tested in performance [Y.Ikebe, T.Hoshi, T.Matsushita and K.Nakayama(1982)].

1) Cross-Correlation

2) SSDA (Sequential Similarity Detection Algorithm)

3) Fourier Cross-Corrlation

4) Fourier Phase-Delay

Since the authors already reported about these methods and applied results in detail in the 4'th Asian Conference on Remote Sensing, only their summaries are given in this report [T.Hoshi and T.Matsushita(1983)].

The basic philosophy behind the matching methods design is that the ideal situation for matching the stereo images is to match each pixel of one image individually with its corresponding pixel in the other image. However, matching one pixel with one pixel by simple comparisons between density values is impossible, because image noise prevents the presentation of the same density value even for identical patterns in the image. In addition to the image noise, the appearance of many pixels with the same density value probably troubles us so far as the correct selection of corresponding pixels [T.J.Blachut(1976)].

Then, a measure of the 'agreement' between two pixels must be done by computing inside two small sub-areas including the two pixels. And it is at best hoped to maximize some measure of the agreement of match between patterns of the two images. These sub-areas are termed 'Matching Windows' Namely, the agreement of two dimensional image patterns within two matching windows is calculated as the agreement between two pixels. When the image contents in two matching windows are approximately the same, the

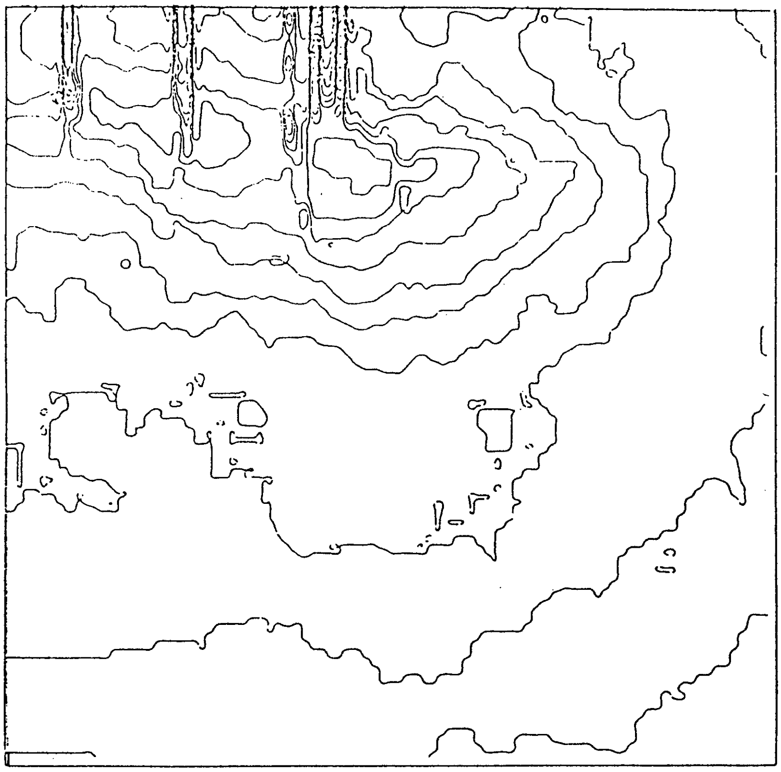
Fig-1 Countor Map produced by Cross-Correlation Method

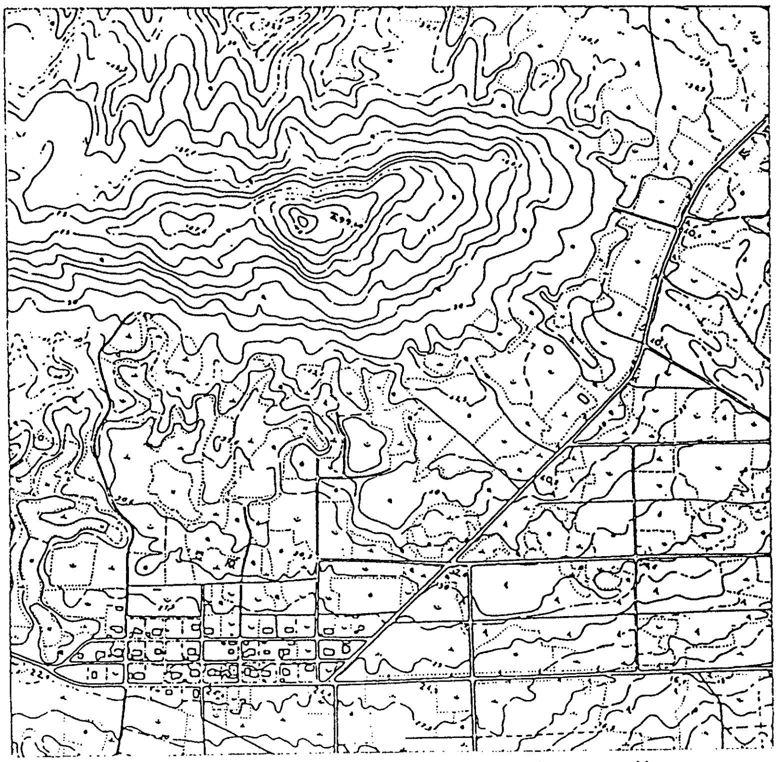
Fig-2 List of Original Topographic Countor Map

measure of image agreement is maximized. [Y.Ikebe,T.Hoshi,T.Matsushita and K.Nakayama(1981)] .

Only the result of Cross-Correlation method which showed the most suitable one from the view point of matching accuracy is listed in Fig-1. Here, comparing the contour map produced by Cross-Correlation method in Fig-1 with its original topographic map in Fig-2, it can be seen they almost correspond with each other in rather flat areas though there are some parts in which elevation errors were produced. In 'Mountains and Forests' area, however, there are considerable elevation errors owing to mismatchings. At a glance, three, inaccurate, overcrowded contour line can be recognized. The reasons why it is difficult to match corresponding pixels in this area may be estimated. The most important one is the fact that matching is suscepiable to the influences of random textures in high frequency domain. Second, a delicate difference of photographed objects between both images exists due to the shadows of trees and the changes of the elevations. The third point is the fact that considerable distortions are introduced into the flight direction at steep slope points.

3. Discussions on matching window size

As a result of applying the method to the stereo image data, it has become evident that the matching window size plays a significant role for the successful matching. Since the CPU time needed for stereo aerial image processing is consumed most to match the corresponding pixels, the system performance is considerably dependent on the optimum matching window size selection. Namely, if the window size is too small, matching two pixels is difficult because of the lack of pattern information within the matching window. If too large a window is adopted, though the matching accuracy is increasing in terms of the abundance of pattern information, the computing time to match may increasingly become longer. Accordingly, it is necessary to find the minimum matching window size which can match for corresponding pixels while maintaining a requisite matching accuracy. In this study, the range of matching window size is drawn an inference from the error limit in resultant elevations and the texture of image pattern [T.Matsushita(1984)] .

3-1. Maximum window size based on error allowance

It has already been stated that a pair of corresponding pixels can be determined based on the agreement within two matching windows. Here, it is assumed that less than one pixel error exists in stereo photographic images passing through two matching windows, that is, between an image pattern in one of them and another corresponding image pattern distorted with a factor $cos\theta$. θ is a slope angle against the horizontal line. Since the photo scale is very large, two bundles of ray passing through the window edge are approximately pallalel, and two bundles through the right window is perpendicular to the ground surface as a matter of convenience. While ground image patterns corresponding to the ray passing through the left window are equal to $Wcos\theta$, it may be W in the case of right one, as shown in Fig-3. From the condition that their distortions of image patterns between two matching windows must be at most less than one pixel, we get

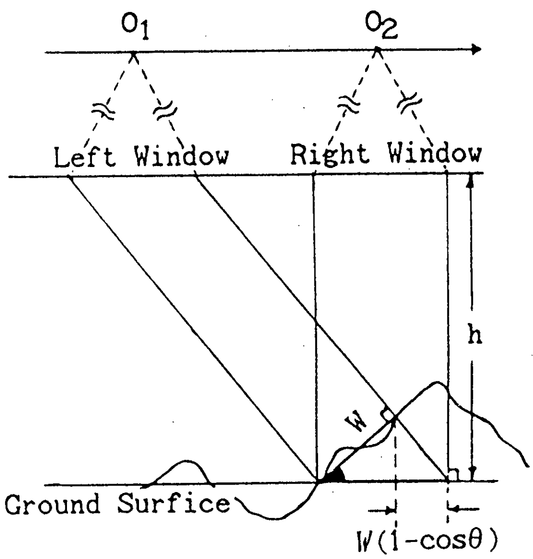

Fig-3 Difference between Distorted Left Window and Right Window

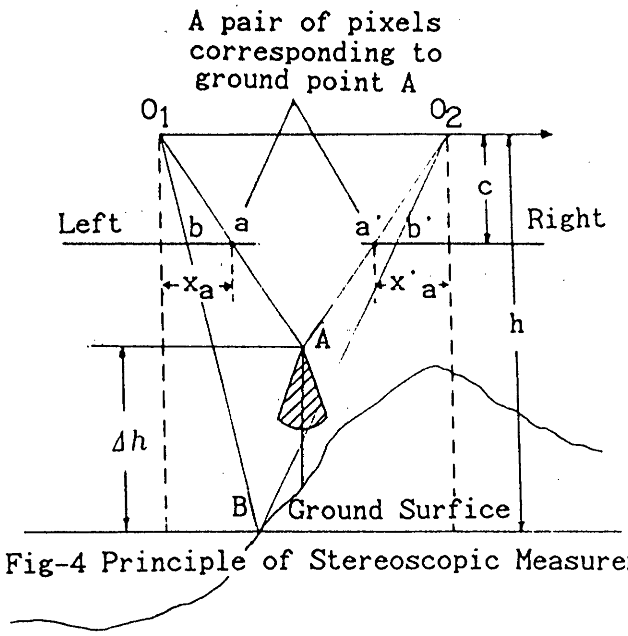

Fig-4 Principle of Stereoscopic Measurement

$$W - W\cos\theta \leqq 1. \tag{1}$$

Actually, since θ is small and $\cos\theta = 1 - \theta^2 / 2$, then Ineq-1 becomes $W \leqq 2 / \theta^2$. Substituting approximate equation $\theta = b / h$ which is introduced from $\tan\theta = b / h$, Ineq-1 is finally given by

$$W \leqq 2 \cdot \left(\frac{h}{b} \right)^2, \tag{2}$$

where b implies the length of base and h is flying height.

On the other hand, the elevation of the ground point can be determined by the locations of its corresponding pixels in the image. Fig-4 illustrates the nature of parallax on overlapping vertical photogrphs taken over a tree in varied terrain. The horizontal parallax of point A is given by $p_a = x_a - x'_a$, and the parallax difference $\Delta p = p_a - p_b$ between the two points A and B are directly dependent upon the elevation difference Δh. The numerical relation between Δp and Δh can be given by

$$\Delta h = \frac{\Delta p h^2}{bc} \cdot \left(1 - \frac{\Delta h}{h} \right). \tag{3}$$

Eq-3 is most useful for determining the elvations based on the parallaxes, and moreover, it is of very great significance in connection with the theory of errors. Any error in the parallax difference, no matter whether it originates from deformations in the image material, distortions, errors in orientation rotation elements, or something else, is transformed into errors in the elevations according to the same formula, i.e., approximately with the factor h^2 / bc. Hence in addition to Ineq-2, it is assumed that the parallax difference correspondeding to the error limit in the measured elevations contain must be beyond one pixel size. The error limit is h*, and sampling size in image digitization, i.e., one pixel size is d, then this assumption is given by

$$\frac{h^2}{b} \leqq \frac{c\Delta h^*}{2d}. \tag{4}$$

Thereupon, substituting Ineq-4 into Ineq-2, the maximum size of matching window is given by Ineq-5 based upon requisite accuracy in the elevations.

$$W \leqq \frac{c\Delta h^*}{bd} \tag{5}$$

However, since two assumptions derived from Ineq-5 is lacking in various information about image pattern characteristics, it should be utilized as one possible standard. Substituting the photographic data, i.e. b, c, d into Ineq-5, it is equivalent to

$$W \leqq 1.82 \cdot \Delta h^*.$$

In other words, assuming the maximum of elevation error is 10m, it will imply that the matching window size must be smaller than 18 pixels.

3-2. Minimum size based on image texture

Judging from many applied results, the optimum window size is apt to change according to the textures of the image pattern, and the amount of texture information included within the matching window is closely related to success of matching. Though it is too difficult to determine its size correctly, this study tries to infer the texture information the matching windows contain.

The authors take advantage of the Fourier transform of image data on the assumption that the texture of image pattern is considered to be the variance and that covariance of two sub-images is interpreted as the convolusion of two Fourier transforms of the images. The minimum size of the matching window can be approximately estimated based on these assumptions. The following descriptions are its theoretical explanations.

Taking the numerical series as shown in Fig-5 from the digital image containing M pixels in the x-direction and N pixels in y-direction, it may be regarded as the cross-sectional view of the image pattern. Namely, the

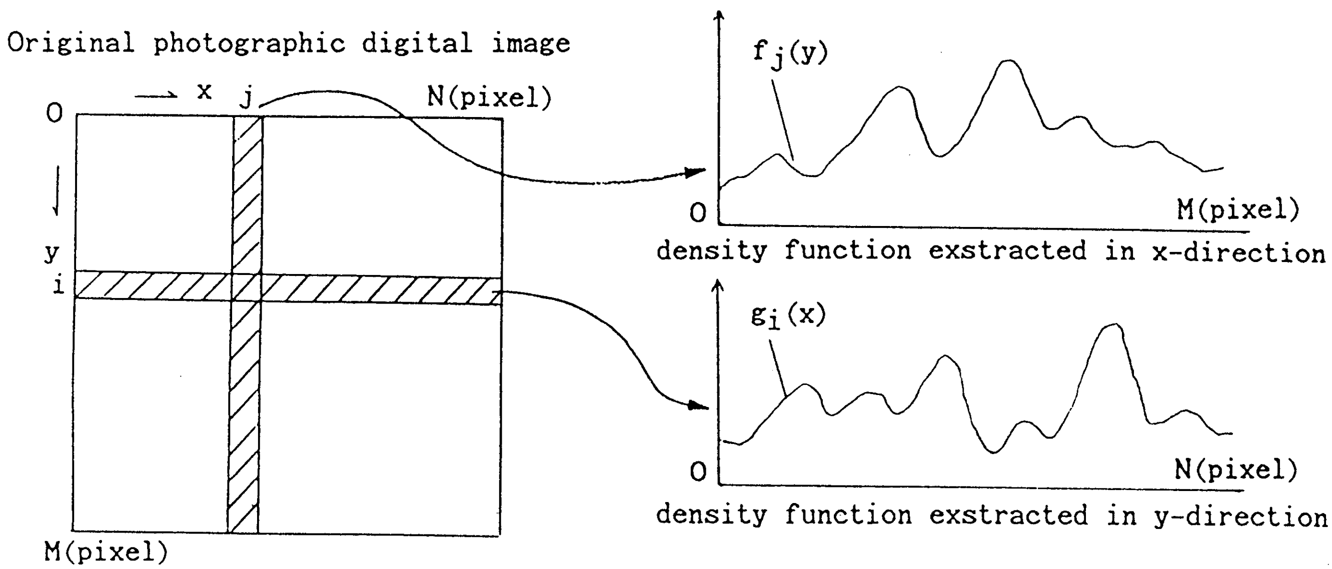

Fig-5 Series Exstraction from Two-dimensional Image Data cross-sectional views in the x-direction are $g_i(x)$ in the finite region $0 < x \leqq N$, and others in the y-direction are $f_j(y)$ in the finite region $0 < y \leqq M$. Matching of corresponding pixels is roughly considered to be a two dimensional problem searching for the image most similar image to any sub-images within a larger original image, and it will be reduced to the case of one dimension. The problem determining the optimum matching window size can be mathematically interpreted as follows.

The function $\varphi'(x)$ is apt to be similar to another function $\varphi(x)$ defined in finite interval $(0, l)$. When the sub-function $\varphi\triangle(x)$ with the matching window interval $\triangle$ is defined in the function $\varphi(x)$, we determine enough minimum interval $\triangle$ to uniquelly find the most similar area with $\varphi\triangle(x)$ within the function $\varphi'(x)$. As a matter of convenience, a function $\varphi(x)$ has non zero density values within $(0, l)$, while the definition interval is extended to $(-\infty, \infty)$. Then the Fourier transform of $\varphi(x)$ is given by

$$\Phi(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\varphi(k)\exp(-ikx)dx$$

and the Fourier transform of the function $\varphi\triangle(x)$ at finite interval $(-\triangle/2, \triangle/2)$ is given by

$$\Phi\triangle(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi(k')D(k-k')dk', \tag{6}$$

$$D(k-k') = \frac{\sin(k-k')\triangle/2}{(k-k')/2}.$$

The function $D(k-k')$ decreases in both directions, having a finite expanse. It has the maximum value at $k=k'$, and zero at the point $(k-k')\cdot\triangle/2 = \pm n\pi$ $(n=1,2,3,\ldots.)$. The half interval between the first zero points ($n=\pm 1$), denoted by $\varDelta k$, is given by

$$\varDelta k = 2\cdot\frac{\pi}{\triangle}$$

In proportion as $\triangle$ spreads, the interval of its expanse will be closed. It will finally tend to zero as $\triangle \to \infty$ and $(2\pi)^{-1/2} \cdot D(k-k')$ will become the Dirac $\delta(k-k')$ function. Eq-6 indicates that the finite interval $\Delta k = 2\pi/\triangle$ fades out each line-spectrum making up the spectrum of $\varphi(x)$ while the spectrum of $\varphi(x)$ is completely constructed from innumerable line-spectrums as $\triangle \to \infty$. If the interval $\triangle$ is infinitely extended, i.e., taking all functional intervals into consideration, a continuous-spectrum can be reproduced in the integral form multiplying delta function with each line-spectrum. Note that this Dirac delta function is generally called a point source input with property

$$\int_{-\infty}^{\infty} \delta(k)dk = 1,$$

while any output image is given by

$$f(x) = \int_{-\infty}^{\infty} f(\alpha)\delta(\alpha-x)d\alpha.$$

Applying this theory to the matching methods, the agreement calculations between two functions f_W, g_W are interpreted as that of the Fourier elements F(k)G(k). In practice, this is performed with faded window spectrum $W\triangle(k)$ on account of a finite matching window. And the fact that the Fourier spectrum of $\varphi(x)$ is faded out with a factor of finite interval Δk is approximately equivalent to existing unreliability Δk within frequency k. Then, Δk is represented with the unreliability of wave-length by

$$|\Delta k| = |\Delta(\frac{2\pi}{\lambda})| = 2\pi \cdot \frac{\Delta\lambda}{\lambda^2}$$

If we must take a minimum wave-lengths until λ^* into consideration, counting from the longest wave-length in order to maintain the pattern $\varphi(x)$ contains, accordingly the Fourier elements of smaller wave-length than λ^* can be ignored, and minimum window size derived from λ^* is enough to match successfully.

The method of determining λ^* based on a function $\varphi(x)$ is as follows. The Fourier series of $\varphi(x)$ at an interval $(0, l)$ is given by

$$\varphi(x) = \sum_{i=1}^{N} (a_i \cos k_i x + b_i \sin k_i x) + a_0$$

$$a_0 = \frac{1}{l}\int_0^l \varphi(x)dx = \varphi \qquad k_i = \frac{i\pi x}{l} = \frac{2\pi x}{(2l/i)}$$

$$a_i = \frac{1}{l}\int_0^l \varphi(x)\cos k_i x dx \qquad b_i = \frac{1}{l}\int_0^l \varphi(x)\sin k_i x dx$$

In this study, the authors adopt the variance, a measure of the spread of the φ as the information which indicates the power of the pattern, i.e., the intensity of pattern changes.

$$E = \int_0^l | \varphi(x) - \varphi |^2 dx = \frac{1}{2}\sum_{i=1}^{N} (a_i^2 + b_i^2) = \frac{1}{2}\sum_{i=1}^{N} C_i^2 \tag{7}$$

In Eq-7, the sum of informations until m-th elements is defined by

$$E^{(m)} = \frac{1}{2}\sum_{i=1}^{m} C_i^2$$

Eq-8 implies the rate of $E^{(m)}$ to E, which is equivalent to the sum of all elements.

$$\alpha = \frac{E^{(m)}}{E} \tag{8}$$

Accordingly, minimum mathing window size is equivalent to value m, which satisfies the condition Eq-9 for given α.

$$\lambda_m = \frac{2l}{m} \tag{9}$$

Then, we can get what percentage of total texture information are included within any window size based on this approach. Taking on account of this approach, this study tries to determine the amount of texture information the optimum window has. Unfortunately, there is no telling how much the optimum size may have so far as its texture information. It can not be derived from any theory. Thus, the authors experimentally determine the minimum window size based on the resultant elevations in three test sub-areas through many implementations varying the matching window size. In the implementation, three sub-areas which are regarded as 1)Farm, 2)Residential Land, and 3)Mountains and Forests, respectively, are extracted from from the photographic image. Next, the authors refer what percentage of the texture information within each experimental window bears to that of the total image. As the result of these implementations, the experimental optimum window size and the percentage of texture information it has in each test sub-area can be referred to in Table-1.

The optimum window size are diffrent according to three sub-areas, and it is chiefly caused by the fact that there are great differences in texture information among them. In the case of 'Farm' and 'Residential Land' sub-areas, the percentage each window has is about 81%, while 84% in 'Mountains and Forests' sub-area.

Table-1 Optimum Window Size and its Texture Information

Test-area / Item	Farm	Residential Land	Mountains and Forests
Optimum Window Size	15 x 15	15 x 15	17 x 17
Percentage of Texture Information $E^{(m)}/E$	81.3%	80.6%	83.8%

Judging from only these results, more than 80% of image texture information within the matching window are necessary to match successfully. Though the fact derived from this approach is not supported by any theoretical considerations, accumulating these implementations will lead to find an experimental guide-line connecting the optimum window size with the percentage of texture information.

4.Conclusion

Fortunately the approach to determine the optimum window size proposed in this report appears promising in matching corresponding pixels. The authors state the conclusions of this paper in the following.

Noise in the image trends to corrupt the image agreement measure. This situation can be improved somewhat by increasing the size of the matching window. Doing so reduces the resolution of the resulting image, however, since large windows tend to smear over any abrupt change in the photographed objects. Thus the window size should be as small as possible, while maintaining a low probability to mismatch.

This is the reason why this study has been directed specifically toward developing an effective way to get the optimum size of matching window. It can even be said that the solution of the window size may prove to be the solution of matching. The approach based on texture analysis is not at all perfect, indeed, and both the approach and its practice do not necessarily go together, but it will have great advantages for applications. Even in a pair of photographs, there are many image patterns. So it is not efficient to match with the same window size in all areas. This approach of analyzing the image texture can offer a guide-line to an appropriate selection of the window size according to its image pattern. Namely, a analyzer can tell by experience that the optimum size based on the image texture analysis before processing, and he may avoid many mistakes by this simple test of this approach.

Acknowledgement

The authors are indebted to a number of individuals who, either through suggestions or discussions or by providing the facilities for our study, have contributed to the presentation of our report. The authors are also especially grateful to Professor KAZUHIKO NAKAYAMA, Director of the Information Processing Center, who kindly guided and supervised our study.

Reference

Y.Ikebe,T.Hoshi,T.Matsushita and K.Nakayama(1982); "Assessmenta and Problems of Searching Algorithm on Aerial Stereo Photographs", Proc. 25th annual conference of Information Processing Sociaty of Japan

Y.Ikebe,T.Hoshi,T.Matsushita and K.Nakayama(1981); "Searching of Correlation Method on Aerial Stereo Photographs", Proc. 24th annual conference of Information Processing Sociaty of Japan

T.Hoshi and T.Matsushita(1983); "Reseach of Matching and/or Searching Methods of Identical Point on Stereo Aerial Images and its Application", Proc. 4th Asian Conference on Remote Sensing

T.Matsushita(1984); "A Study of Methods for Matching Corresponding Pixels in Stereo Aerial Images", A thesis for master program of the University of Tsukuba

P.L.Owen and W.F.Blank(1980); "SURFICE - Terrain Modeling and Contor Plotting", Technical Papers of the American Congress on Surveying and Mapping; 40th Annual Meeting

G.Konecny(1979); "Methods and Possibilities for Digital Differential Rectification", Photogrametric Engineering and Remote Sensing, Vol.45,No.6

P.R.Wolf and L.L.Christenson(1978); "Application of Photogrammetry in Land Surveying", Proc. of the American Congress on Surveying and Mapping; Fall Technical Meeting

INSTITUTE OF INFORMATION SCIENCES AND ELECTRONICS
UNIVERSITY OF TSUKUBA
SAKURA-MURA, NIIHARI-GUN, IBARAKI 305 JAPAN

REPORT DOCUMENTATION PAGE	REPORT NUMBER ISE-TR-84-48

TITLE

Window Size of Matching Methods in Stereo Aerial Images

AUTHOR(S)

Takashi Hoshi (University of Tsukuba)

Takayuki Matsushita (IBM Japan)

Yasuhiko Ikebe (University of Tsukuba)

REPORT DATE	NUMBER OF PAGES
October, 30th, 1984	14
MAIN CATEGORY	**CR CATEGORIES**
Image Processing, Photogrammetry	4.0, 3.63, 3.23

KEY WORDS

Matching Window, Stereo Aerial Image

ABSTRACT

This study has been directed specifically toward developing an effective way to get the optimum size of matching window in Cross-Correlation method. The person concerned in its applications could hardly determine the optimum matching window size without many trials and errors so far, because it is considered to be changeable according to the image pattern. The optimum size must be set taking an account of texture information in connection with the image pattern. This paper deals with a way to determine the window size for matching successfully, based on the Fourier transform of image data. And another approach using the error limit in the resultant elevations is also described in detail.

SUPPLEMENTARY NOTES