HIGH PRECISION RELATIVE ORIENTATION USING THE FEATURE BASED MATCHING TECHNIQUES

Mengxiang Li
Department of Photogrammetry
The Royal Institute of Technology
S-100 44 Stockholm, Sweden
Commission III

0. ABSTRACT

This paper presents a method for relative orientation using the feature based matching techniques. The interest operator and feature based matching are use to select points and measure parallax for relative orientation automatically. After the feature based matching, a fine least square matching is preformed to achieve high precision. The matched points are then used for the relative orientation. The precision of relative orientation depends on the final least square matching. An off-line experiment has been carried out. With the optimal window size, and inclusion of radiometric parameters in the least square matching, results of 2.08 µm for the standard error of unit weight(σ_0) of the y-parallax for a simulated image model and 3.83 µm for a real image model are obtained. The estimated precision of the feature based matching together with the least square matching for stereo parallax measurement is 0.98 µm, or 0.06 pixels. The experiment also shows the high potential for automating the relative orientation procedure.

1. INTRODUCTION

Digital photogrammetry provides the posibility for automating, speeding up photogrammetric processes as well as improving accuracy. Digital image matching is one of the basic image processing techniques used in digital photogrammetry. Genernally image matching can be classified into two types, namely, area based matching(ABM), and feature based matching(FBM). The area based matching, like the cross-correlation method, the least square matching method, has the advantage of high precision, while its drawback is the high requirement for the initial value (small pull-in range), and vast data processing. The feature based matching uses only some distinctive features instead of the whole image gray level, so the data processing can be fast, and also the initial value requirement is very low. But the feature based matching can not provide very high precision directly and no detail information either. By examining both the advantages and disadvantages of the ABM and FBM, one can see that a combined approach, which overcomes the drawbacks and promotes the advantages, will be a good solution. The method presented in this paper is such a combined approach.

Automation of conventional photogrammetric procedures, such as inner orientation, relative orienation, is the basic step in the automation of photogrammtery, because most of the photogrammetric processes are based on the oriented image model. An analytical plotter, supported with some image processing software and hardware, e.g. CCD cameras, is a typical system approach for such an automation. Semi-automatic inner, relative orientation has been implemented on the Kern DSR-11 analytical plotter using a classic correlation method. An algorithm for automatic inner orientation has also been developed on the ContextVision system GOP-300 (Stokes 1988).

The relative orientation is to reconstruct the image model and what is needed is some conjugate points from left and right images with their image coordinates. The problem to be defined in this investigation is to select points and measure parallax for the relative orientation automatically, replacing the human operator by the interest operator and FBM techniques. The method is based on the feature matching algorithm developed by Föstner(1986,1987). After the FBM procedures, a fine least square matching is performed in order to get high precision. As a part of the investigation, the least square mathcing has also been studied concerning the model fidelity and optimal window size. The final results are used for the relative orientation. An off-line experiment has been carried out on the DSR-11 analytical plotter. The results show that high precision can be achieved. Experiment also shows the potential for automatic relative orientation by using the feature based matching techniques.

2. THE FEATURE BASED MATCHING TECHNIQUES

A feature based matching algorithm usually consists of three steps, namely (1) feature extraction, (2) prelimilary matching between extracted features based on similarty, (3) consistency matching. The FBM is widely used in computer vision, pattern recognition(Barnard et.al 1980). And it has also been used in photogrammetry(Förstner 1986, Schewe 1986). The three steps are briefly described in the following.

2.1 Feature Extraction

Feature extraction is the extraction of some distinctive features out of the whole image. The extraction is done by a interest operator. There are different interest operators based on different principles, but the results are similar.(Luhmann 1986). The Moravec interest operator (Moravec 1977) and the Dershler interest operator are usually used by people in computer vision. In this study the Förstner interest operator is implemented. The Förstner interest operator selects distinctive points, like edge intersections, corners, circular centers. The result of the selection is a list of points with their descriptions, such as coordinates, weight or interest value. These descriptions will be used for the latter matching. The interest operators mentioned above are point type feature selectors. There are also some other types of operators for feature extraction. For example, the Laplacian-Gaussian operator(Marr1979, Grimson 1981,1985). The Laplacian-Gaussian operator marks out zero-crossings on the image, which is the location of significant changes of the orignal intensity function. The edge features are also quite often used for matching (Baker 1982). But in both these cases epipolar geometry has to be used because in the direction of the edge the matching is ambiguous.

2.2 Prelimilary Matching Based on Similarity

After the feature extraction, from each image patch a number of distinctive points are selected together with their descriptions. Each left point either corresponds to one right point or corresponds to none . But at the begining we do not known which one. The prelimilary matching is to find out all the possible corresponding candidate points for each left point based on the similarity measurement. This is actually done by two thresholdings. The first thresholding is based on the parallax. Although we do not know the exact parallax for each point, we may know the possible maximum parallax for the whole image patch. Setting up a certain threshold value for the parallax, the candidate points for a left point are limited certain right points. This parallax threshold value depends on the a priori knowledge of the relative position between the corresponding patches. It can be very large, e.g. 100 pixels. This can be considered as the pull-in range. The larger the pull-in range, the longer the searching will take. A proper threshold value will optimize the searching without loss of necessary information. The second thresholding is based on the correlation coefficient between two small windows around two points. If the correlation coefficient is less than a threshold value, say 0.5, then this candidate point is excluded from the candidate group. After these two thresholdings, the number of candidate points for a left image point is reduced to a very limited number. A further similarity measure could be the point type. Edge-intersection corresponds to edge-intersection, corner corresponds to corner, and so on. If the interest operator can provide the point type description, then the point can be matched further based on their type description.

2.3 Final Matching for Getting Consistency

After the similarity matching, the left and right image point are paired. The next step is to drive a consistent and unique matching between the left and right image patch. Förstner(1986) has developed a method for getting consistency based on affine transformation between the corresponding points together with the robust estimation techniques. This algorithm implies that the object has to be a plane or tilted plane, or a group of points located on a plane or tilted plane. The mathematical model is,

$$\left.\begin{aligned} px_i &= a_1 + a_2 * x_i + a_3 * y_i \\ py_i &= b_1 + b_2 * x_i + b_3 * y_i \end{aligned}\right] , \quad W_i = w_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (1)$$

where px, py are parallaxes between the left and right corresponding points, x, y are left image coordinates, W is the weight matrix. After the left-right point pairs have been constructed, each such pair is brought into the computation according to (1). A larger number of these pairs is incorrect and blunders with respect to the model. In order to detect this larger number of blunders, the robust estimation has to be applied. The following weight functions are used to detect blunders,

$$\left.\begin{aligned} w_1(v) &= 4*[\sqrt{(1 + v^2/2)} - 1]/v^2 \\ w_2(v) &= 2*\ln(1 + v^2/2)/v^2 \\ w_3(v) &= \exp(-v^2/2)/v^2 \end{aligned}\right] \qquad (2)$$

The computation is iterated until the required precision is obtained. The first few iterations uses w_1. After the solution has converged, w_2 or w_3 is used. Between each iteration the weight is updated and thresholded according to certain criteria. If the weight is thresholded out, the point is then excluded from the computation.

Barnard and Thompson (1980) have developed another algorithm for feature based matching based on the relaxation scene labelling technique. This algorithm works with piecewise plane object type. The basic idea of the algorithm is that if two points are near by on the left image, then their corresponding points on the right image should be near by too, or the parallax difference between these two points is close to zero. After the prelimilary matching, each left point is associated with a candidate points group. Within its candidate point group, a number, p is associated with every candidate point, which is interpreted as the estimated probability that this left point corresponds to this candidate point. The consistency is achieved by iterative updating the probabilities. In this investigation these two algorithms have been implemented. But experiment shows that the affine transformation gets more globally consistent results than the relaxation, if the object is plane. Results presented here are mostly based on the affine transformation.

3. PROCEDURE FOR THE EXPERIMENT

This method is suitable for analytical plotters equipped with CCD cameras for on-line relative orientation. But due to capacity of the computer on DSR-11, the experiment is done off-line on a Apollo system. The model is inner oriented first on the DSR-11. A patch of size 300x300 pixels is digitized by the CCD camera at each position of the six standard relative orientation points from both images. These image patches are transfered to the Apollo system together with the inner orientation and the CCD camera calibration data. The interest operator is applied patch by patch, and the FBM procedure is performed between each patch pair. Finally the least square matching is done for each matched point in order to get high precision. This is also a final check for the correctness of the feature based matching. If the LSM does not reach convergency, or the computed shift is greater than a threshold value, e.g. 5 pixels, then this point is regarded as a incorrectly matched point and deleted from the point list. The relative orientation is done as usual except with a larger number of points. This is the procedure for the off-line experiment. For an automatic, or semi-automatic on-line relative orientation using the feature based matching, the interest operator and the FBM procedure can be patch-wise applied. The results of the previous patch can be used to predicte the next patch's position. If the object is flat this prediction can be good enough for the FBM, provided the area contains some feature points.

4. RESULT AND ANALYSIS

The method has been tested using two sets of data, one simulated image model and one real image model. The material used for the testing is a pair of aerial photographs with flying height of 600 m and photo scale of 1:4000. The photographs are diapositive copies of orignal photos. Results of relative orientation depend on the final least square matching. More detail discussions about LSM see next section. The results presented here are based on the optimal window size and model including radiometric parameters.

4.1 The Simulated Image Model

A simulated image model is constructed in the following way. Using one of the photos as both left and right image and introduing a base bx into the image coordinates, a fictitious image model can be reconstructed. This fictitious model is a normal case model and all its relative orientation parameters should be zero. The image is placed first on the left stage, the inner orienation is done based on the fiducial marks measurements by the operator. Then six patchs are digitized. Moving the image to the right stage, the same procedure is preformed. After that the interest operator and the FBM precedures are applied. The results are then used for the relative orientation. Table 1 shows the results of relative orientation of the simulated image model using the FBM and LSM.

As can be seen from the table all the parameters are close to zero but not zero. This is the effects of errors of measurement in the inner orienation and errors of the instrument.

Tabel 1. Result of Relative Orientation of Simulated Image Model

FBM measure	para.	by(mm)	bz(mm)	φ(gon)	ω(gon)	k(gon)
with LSM point=383 σ_0= 2.08μm	value	-0.0102	-0.0013	0.0017	0.0018	-0.0004
	S.D.	0.0001	0.0001	0.0002	0.0005	0.0003
without LSM point=389 σ_0= 9.01μm	value	-0.0125	0.0024	0.0005	0.0019	-0.0023
	S.D.	0.0004	0.0002	0.0009	0.0020	0.0015

The standard error of unit weight σ_0 is reduced after the LSM dramatically. For the simulated image model the left and the right image are exactly the same, the effect of the image errors is eliminated from the relative orientation. The final result of σ_0=2.08 μm is a combined effects of the method and the instrument used for the relative orientation. According to the error theory of image coordinates, and considering the result in y-parallax, which is the combined effect of left and right images, we have approximately the following variance components equation,

$$\sigma^2_{simu.} = \sigma^2_{FBM} + 2 * \sigma^2_{instr.} \qquad (3)$$

From the instrument calibration we have $\sigma^2_{instr.}$ = 1.30^2(μm), so the estimated variance component for the FBM method is,

$$\sigma^2_{FBM} = \sigma^2_{simu.} - 2 * \sigma^2_{instr.} = 0.98^2 (\mu m) = 0.06 \text{ pixels}$$

This is interpreted as the intenal precision of the FBM together with the LSM for maesuring stereo parallax as an operator.

4.2 The Real Image Model

The real image model used the material described at the begining of this section. The model is set up on the DSR-11. The inner orientation is done based on operator measurement. The operator also made the relative orientation on the instrument in order to compare the operator measurement and the FBM approach. The orientation is repeated five times and the average result is used for the comparsion as shown in Table 2 together with the results of FBM.

Tabel 2. Results of Relative Orientation of Real Image Model

	para.	by(mm)	bz(mm)	φ(gon)	ω(gon)	k(gon)
FBM measure with LSM point=137 σ_0= 3.83μm	value	-3.6585	0.6162	0.0019	-0.8662	-1.6539
	S.D.	0.0003	0.0002	0.0006	0.0013	0.0011
without LSM point=145 σ_0=12.62μm	value	-3.6824	0.6067	0.0011	-0.8567	-1.6476
	S.D.	0.0011	0.0005	0.0019	0.0040	0.0036
operator point=12 σ_0= 4.64μm	value	-3.660	0.611	-0.010	-0.866	-1.651
	S.D.	0.000	0.000	0.005	0.004	0.002

As we can see the result of FBM with LSM is much better than without LSM and the operator. It should be noticed that the results shown here are computed in the relative orientation without any blunder detection. We also noticed that there are a few points with very larger residuals ($>3*s_0$) after the relative orientation. It is expected that result will be improved by using robust estimation for blunder detection. The difference between the simulated image model and the real image model in the sense of error source is that the simulated image model is free from image errors, while the real image model is not. We can use this to estimate the variance components of the image and the operator. we have approximately the following variance components equation:

$$\sigma^2_{total.FBM} = 2 * \sigma^2_{image} + \sigma^2_{FBM} + 2 * \sigma^2_{instr.} \qquad (4)$$

From the real image model we have $\sigma^2_{total.FBM} = 3.83^2 \mu m$, so the estimated variance component of the image is

$$\sigma^2_{image} = \frac{1}{2}\left[\sigma^2_{total.FBM} - \sigma^2_{FBM} - 2 * \sigma^2_{instr.}\right] = 2.27^2 \mu m$$

For operator measurement, the total effect is a sum of image, instrument and operator, so we have the following,

$$\sigma^2_{total.ope.} = 2 * \sigma^2_{image} + 2 * \sigma^2_{instr.} + \sigma^2_{ope.} \qquad (5)$$

So the estimated variance component, or the precision of operator for stereo parallax measurement is,

$$\sigma^2_{ope.} = \sigma^2_{total.ope} - 2 * \sigma^2_{image} - 2 * \sigma^2_{instr.} = 2.80^2 \mu m$$

This estimated precision is close to realistic situation. Comparing the precision of FBM and the operator for stereo parallax measurement, we can see the precision of FBM is much higher than the operator. We should say this estimation is approximate, because the precision of operator measurenemt depends on the image quality and the point type. So there is a high correlation between image quality and operator's precision therefor there is covariance between them, while the precision of the FBM is estimated without the effect of image errors. But we can still see that precision of FBM is higer than the operator.

5. MORE ABOUT LEAST SQUARE MATCHING

The least square matching has been very well developed and a lot of investigations have been done with it(Rosenholm,1986,1987). A simplified mathematical model, which uses only geometric shift parameters, is the following,

$$G_1(x,y) = G_2(x + x_0, y + y_0) + n(x,y) \qquad (6)$$

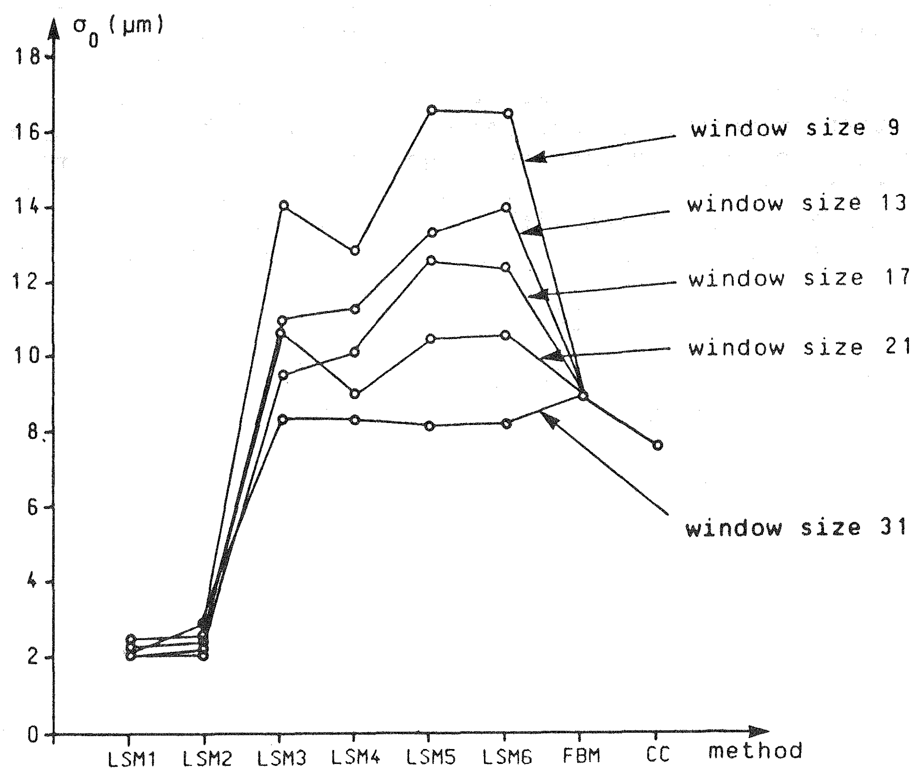
If the radiometric parameters are included, the model is,

$$G_1(x,y) = G_0 + S * G_2(x + x_0, y + y_0) + n(x,y) \qquad (7)$$

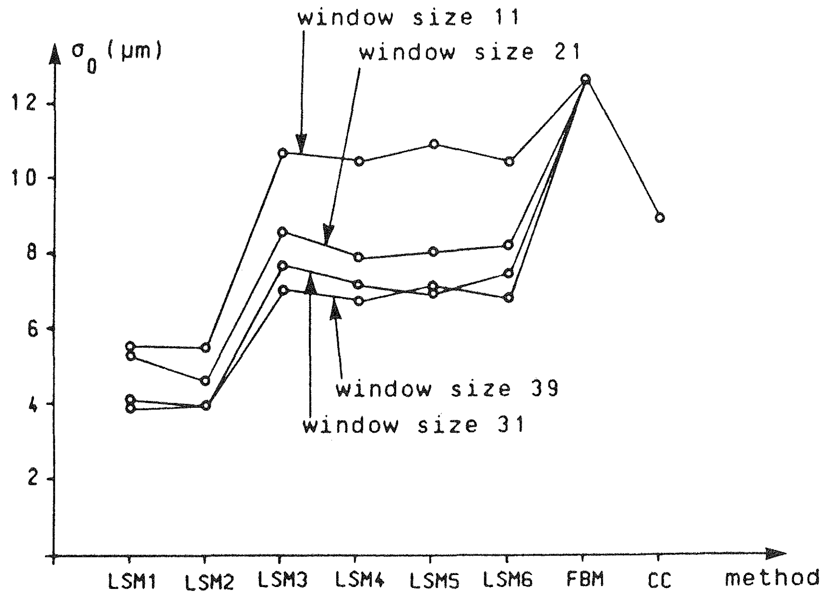
where G_1, G_2 are left and right image gray level values, G_0 and S are gray level shift and scale, x, y are left image coordinates, and x_0, y_0 are geometric shift paratmeter. G_0, S, x_0 and y_0 are the parameters to be estimated. In a practical implementation, the right side of (5) or (6) has to be linearized. The solution is iterated until convergency is reached. After each iteration the right image has to be resampled. There are many ways for resampling. Usually two are used, the bilinear interpolation using four neighbours, or simply taking the nearest neighbour. The derivatives or gradient images, G_{2x} and G_{2y} for the linearization can also be replaced by G_{1x} and G_{1y}, which do not need to be updated each iteration. Rosenholm(1987) has studied the optimal window size, the effect of radiometric parameters as well as the effect of derivative image using one or both images. In this investigation, the following aspects have been considered,

1. How significant are the radiometric parameters in (6) if no a priori gray level normalization.
2. What is the optimal window size in this specific example.
3. Does the bilinear interpolation and taking the nearest neighbour for resampling make any difference.
4. Does replacing G_{2x}, G_{2y} by G_{1x} and G_{1y} make any difference.

To the first aspect, the results show that the radiometric parameters are very important as long as the radiance difference exists and there is no a priori gray level normalization. Experiment by Rosenholm shows that if a priori gray level normalization is done, then the radiometric parameters does not effect the precision of matching significantly while a slight improvement on reliability. If there is no a priori gray level normalization, how does the radiometric parameters effect the matching? Fig. 1 shows the results of the relative orientation using different models without a priori gray level normalization. As can be seen from the figures, the results are improved very much by using the radiometric parameters in both the simulated model and real model. As the window size increases, the results of LSM without radiometric paramters(LSM3, LSM4,LSM5, LSM6) get better, but still far from the results of LSM1 and LSM2, which are results with radiometric parameters. So we can say that the radiance difference between images is very important. One has to take care of them by either a priori gray level normalization or inclusion of radiometric parameters in the least square matching.



(a) Simulated Image Model



(b) Real Image Model

Fig.1 Results of Relative Orientation Based on Different Models.
LSM1 uses (6) and G_{2x} and G_{2y}. LSM2 uses (6) but G_{1x} and G_{1y}.
LSM3 uses (5) with G_{2x}, G_{2y} and linear interpolation for resampling.
LSM4 is the same as LSM3 but takes nearest neighbour for resampling.
LSM5 uses (5) with G_{1x} G_{1y} and linear interpolation for resampling.
LSM6 is the same as LSM5 but takes nearest neighbour for resampling.
FBM is result without LSM fine matching.
CC estimate optimal position by correlation coefficient around a window of 9x9 pixels, taking the maximum cc without interpolation.

To the second aspect, the experiment shows the optimal window size in this example is around 34X34 pixels in the real image model case, which is similar to the one obtained by Rosenholm. While in the simulated model the optimal window size is rather smaller, around 20X20 pixels. The LSM is done without affine transformation parameters. As can be seen from Figure 2(results based on LSM1), the window size has to be larger than a certain number in order to get a reliable solution, 20 pixels in the real image model This has also been observed by Rosenholm. If the window size is too small, the information covered by the window is too little to make comparsion between two windows. In the simulated image model, the rotation between the two images is significat(result from FBM affine transformation). This is because the image ia placed on the stage individually at different ways. In the LSM the affine parameters are not included. Rosenholm shows that when the window size is larger than 30x30 pixels, the affine parameters are very effective. If the affine parameters are included the results are expected to be better, and the optimal window size for the simulated model is also expected to be larger. This should be test further.

To third and fourth aspects, experiment shows that the bilinear interpolation or taking the nearest neighbour does not effect the result of matching. And so is the replacing G_{2x} and G_{2y} by G_{1x} and G_{1y}. This has also been shown by Rosenholm.

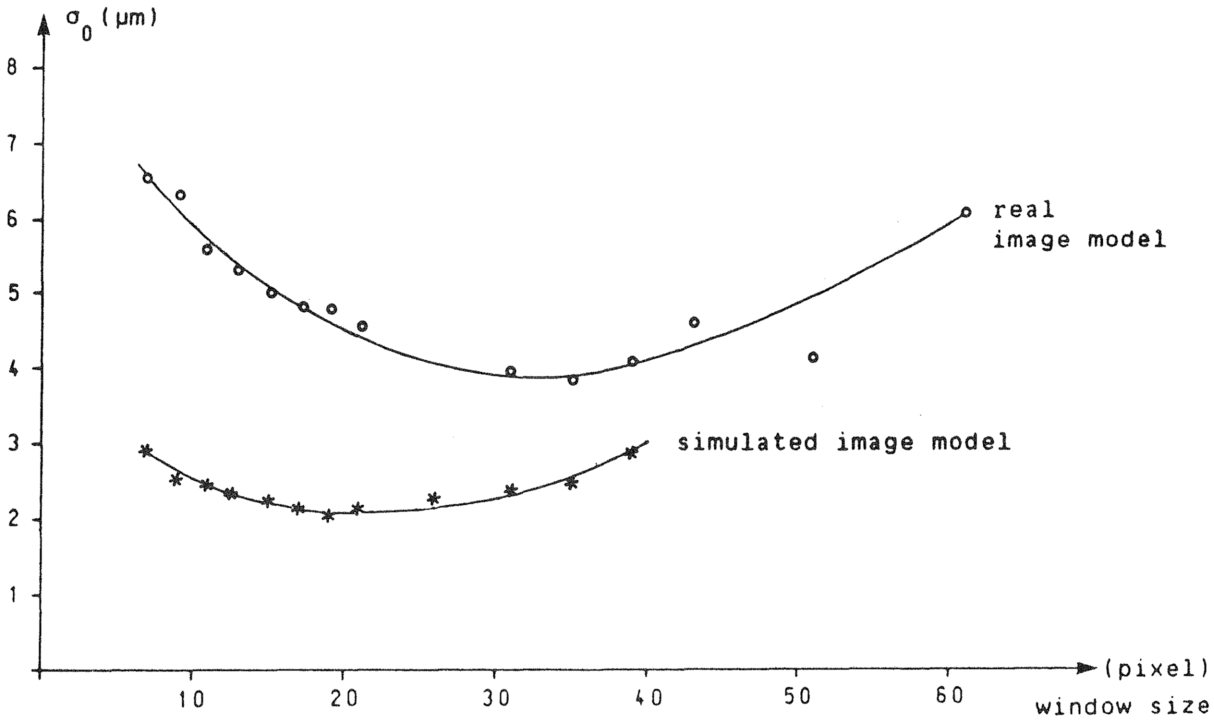


Fig. 2 Results of Relative Orientation Using Different Window Size

6. DISCUSSION

The interest operator together with the FBM procedure, usually can provide larger number of points(more than 100 usually) for the relative orientation, so the redundancy is extremely high. The relative redundancy r/n is close to 1. The relative redundancy is a measure of reliability(Torlegård 1981), so the method can provide high reliabilty. In this experiment, the relative orienation is done without blunder detection in order to see the real performance of the approach. But it is quite possible that results from FBM contain blunders, especially when one uses the relaxation method. So in general, the relative orientation should done with blunder detection (e.g. robust estimation).

Although this experiment is an off-line experiment, it shows the potential for on-line automatic realtive orientation. The computing time depends on the computer and the patch size. In this investigation, the algorithm is implemented on an Apollo system. The time consumption and program optimization is not very much considered. The time taken by the interest operator to select feature points from 300x300 pixels image is about 20 seconds, and the FBM takes about 15 seconds for such an image pair. No doubt that time can be reduced to a pratical level by program optimization and dedicated sytem software and/or hardware, for example, for computing the correlation coefficients, which is rather time comsumpting in this case. Results by other researchers also shows that higher speed can be achieved.

7. CONCLUSIONS

1. Using the FBM together with the LSM techniques for relative orienation, the precision and reliability are both improved.
2. The the relative orientation procedure can be automatically or semi-automatically realized by using the FBM and ABM techniques.
3. The radiance differences between images are very important for the LSM. It should be considered in the LSM.

8. ACKNOWLEDGEMENT

The author wants to thank Prof. K. Torlegård for his valuable suggestions and comments for this experiment. The author is also grateful to Mr. E. Dauphin for his help with the langauge.

9. REFERENCES

BAKER, H.H., BINFORD,T.O., 1981:
"Depth from Edge and Intensity Based Stereo". The 7'th International Joint Conference of Artificial Intelligence (IJCAI), 1981, pp631-636.

BARNARD,S.T., THOMPSON,W.B. 1980:
"Disparity Analysis of Images". IEEE PAMI-2, No.4, 1980, pp333-340.

FÖRSTNER, W. 1986:
"A Feature Based Correspodence Algorithm for Image Matching". International Archives for Photogrammetry and Remote Sensing, Vol.26, Com.III, Rovaneimi 1986.

FÖRSTNER, W., GUELCH, E. 1987:
"A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centers of Circular Features". ISPRS Intercommission Workshop, interlaken, 1987.

GRIMSON, W.E.L., 1985:
"Computational Experiments with a Feature Based Stereo Algorithm". IEEE PAMI-7, No.1, 1985, pp17-33.

LUHMANN, T. 1986:
"Interest-operator for Image Matching". International Archives for Photogrammetry and Remote Sensing, Vol.26, Com.III, Rovaneimi 1986.

MARR, D., POGGIO,T. 1979:
"Computaional Theory of Human Stereo Vision". Proceedings of the Royal Society of London, B.204, pp301-327.

MORAVEC, H.P. 1977:
"Towards Automatic Visual Obstacle Avoidance". The 5'th International Joint Conference of Artificial Intelligence (IJCAI), 1977, p584.

ROSENHOLM, D. 1986,
"Accuracy Improvement in Digital Matching for Evaluation of Digital Terrain Models". International Archives for Photogrammetry and Remote Sensing, Vol.26, Com.III, Rovaneimi 1986.

ROSENHOLM, D.1987.
"Empirical Investigation of Optimal Window Size using the Least Square Image Matching Method". Submitted for publication in Photogrammetria.

ROSENHOLM, D. 1987,
"Some Results from Experiments with the Least Square Matching Method". Accepted for publication in Photogrammetric Record.

STOKES, J. 1988:
"A Photogrammetric Monoplotter for Digital Map Revision Using the Image Processing System GOP-300". Inter. Archives for Photogrammetry and Remote Sensing, Vol.27, Com:II, 1988.

TORLEGÅRD,K. 1981:
"Accuracy Improvement in Close Range Photogrammetry". Schriften -reihe, Wissenschaftlicher Studiwngang Vermessungswesen Hochschule der Bundeswehr, Heft5. 1981, 68p.