STEREO IMAGE MATCHING WITH SUB-PIXEL RESOLUTION

Borislav MARINOV
University of Architecture, Civil Engineering and Geodesy,
Faculty of Geodesy, Sofia, Bulgaria
MarinovB_fgs@uacg.acad.bg

Working Group III/2

KEY WORDS: Image matching, Orientation, High-resolution data/images, Edge extraction, Accuracy.

ABSTRACT

The automation of processes of the interior, relative and absolute orientation of stereo images requires solving the problem of image matching. A model for image matching is proposed that is based on the projective geometric transformation and the bi-linear grey level approximation. The enhancement of the correlation maximum is obtained by using of structural information about the contours in the matching windows. For that purpose a procedure for contour finding is applied. The contour information is used for generating the pattern image and weight matrix for the pixels in the matching window. The Least Squares Method is applied at the final stage of the processing to determine the parameters of matching at the sub-pixel level. The advantage of such approach is the possibility to determine not only the values of the parameters but also to estimate the accuracy of the results. It is provided the experimental comparison for combination of original images, contour image and weight matrix at the stages of correlation matching and Least Squares Matching. The procedure is adapted to the specific requirements for the steps of interior, relative and exterior orientation of the digital stereo images.

1 INTRODUCTION

The automation of processes of the interior, relative and absolute orientation in digital photogrammetry is impossible to be done without solving the problems of digital image matching. Different reasons prevent achieving accurate and reliable results. The influence of the image noise, the orientation parameters, the slope of the terrain and the sharp rise of the heights of some objects like buildings, produce a disturbance of the matching process. For these reasons a more sophisticated procedure for matching of corresponding areas in digital images is necessary to be used. The standard procedure for developing that process is application of digital correlation . There are known different methods for estimation of best correlation. By that reason there are a lot of investigations concerning the solving of that problem. The following main tasks have to be sold to achieve final results. The choice of fast and reliable method for initial position finding. The mathematical model for geometric transformation of two images. In that direction the orthogonal or affine transformations are most commonly applied (Kraus K., 1997). Different types of matching criteria are applied - absolute difference, correlation, least squares adjustment (Heipke Chr., 1996). The application of edge detection techniques is investigated from different authors (Alwan R.H., 1996). The maximum accuracy of relative position orientation of two images is goal of methods. In that direction the sub-pixel resolution of position determination is preferable.

2 MATHEMATICAL MODEL OF TRANSFORMATION

The degrees of freedom of mathematical model depend on stage of processing and used procedures. It is well known that the large number of parameters allow to ensure more accurate results. From the other side the increasing of the parameter number tends to worst stability of parameter determination. By that reason the control of used parameters is necessary to be applied.

2.1 The Model for geometric transformation

The geometric transformation is based on projective transformation: The procedure for transformation the target coordinates in coordinate system of pattern is applied. So the reverse transformation is used to obtain the coordinates of target pixels in coordinate system of pattern image. The equations describing that transformation are:

$$x_i^t = \frac{a_x^r x_i^e + b_x^r y_i^e + c_x}{d^r x_i^e + e^r y_i^e + 1}, \qquad y_i^t = \frac{a_y^r x_i^e + b_y^r y_i^e + c_y}{d^r x_i^e + e^r y_i^e + 1} \tag{1}$$

where x_i^e and y_i^e are coordinates of the pixel in the pattern image and

x_i^t and y_i^t are coordinates of the corresponding pixel position in the target coordinate space.

After determining the parameters of reverse transformation the parameters of forward transformation could be processed. This transformation will not be necessary to be done if the pixels of pattern image are transformed to the coordinate system of the target image. Such solution is not convenient because in that situation the position in pattern mage will change from iteration to iteration that will tend to instability of the process. Another consideration is that in case of determination the position of fiducial marks it is preferable to use the standard pattern of mark model at fixed position.

2.2 Radiometric transformation of image mask

Another problem to be solved is the type of grey-level model. The usage of the nearest neighbour method is not appropriate due to the goal for reaching the sub-pixel resolution. The suggested model is based on bi-linear approximation between square of pixels in the target image. The model assumes different mean value, contrast and possible linear change of mean grey level in the square of pixels. Taking into account such types of dependency the final formula for grey level of target image corresponding to pattern pixel at position (x_i^e, y_i^e) has the form:

$$\begin{aligned} g(x_i^e, y_i^e) = {} & g_0 + h_x x_i^e + h_y y_i^e \\ & + h_g[g^t(j,k)(x_i^t - j)(y_i^t - k) + g^t(j+1,k)(j+1-x_i^t)(y_i^t - k) \\ & + g^t(j,k+1)(x_i^t - j)(k+1-y_i^t) + g^t(j+1,k+1)(j+1-x_i^t)(k+1-y_i^t)] \end{aligned} \tag{2}$$

where $g^t(j,k)$ is grey-level value of the target pixel at (j,k) position in the pattern space.

The term g_0 corresponds to the mean value. The h_x and h_y take into account the coordinate dependent mean value. The parameter h_g corresponds to the contrast ratio between pattern and target images. After expanding the expression for $g(x_i^e, y_i^e)$ and routing to the linear terms the linearized equations of measurement are formed that are used to form the normal system of least squares method. The formal presentation of residual equation is presented in the following form:

$$\begin{aligned} v_i = {} & c_{1i}\delta g_0 + c_{2i}\delta h_x + c_{3i}\delta h_y + c_{4i}\delta h_g + c_{5i}\delta a_x^r + c_{6i}\delta b_x^r + c_{7i}\delta C_x^r \\ & + c_{8i}\delta a_y^r + c_{9i}\delta b_y^r + c_{10i}\delta C_y^r + c_{11i}\delta d^r + c_{12i}\delta e^r + f_i \end{aligned} \tag{3}$$

The expressions for the first three coefficients have the form:

$$c_{1i} = 1 \qquad c_{2i} = x_i^e \qquad c_{3i} = y_i^e \tag{4}$$

If it has been applied preliminary contrast equalization the above parameters could not be taken into account. The expression for $\mathbf{c}_{4j}$ depends on the interpolated grey-level value in point (x_i^t, y_i^t).

$$\begin{aligned} c_{4i} = {} & g^t(j,k)(x_i^t - j)(y_i^t - k) + g^t(j+1,k)(j+1-x_i^t)(y_i^t - k) \\ & + g^t(j,k+1)(x_i^t - j)(k+1-y_i^t) + g^t(j+1,k+1)(j+1-x_i^t)(k+1-y_i^t) \end{aligned} \tag{5}$$

The expressions for coefficients from c_{5j} to c_{12j} are separated in three groups. The coefficients $\mathbf{c}_{5j}$, $\mathbf{c}_{6j}$ and $\mathbf{c}_{7j}$ depend only from derivatives of x_j^t:

$$c_{5i} = h_g\{[g^t(j,k) - g^t(j+1,k)](y_i^t - k) + [g^t(j,k+1) - g^t(j+1,k+1)](k+1-y_i^t)\frac{\partial x_i^t}{\partial a_x^r}\} \quad \text{where} \quad \frac{\partial x_i^t}{\partial a_x^r} = \frac{x_i^e}{d^r.x_i^e + e^r.y_i^e + 1} \tag{6}$$

The expressions for coefficients from c_{8j} to c_{10j} has the form:

$$c_{8i} = h_g\{[g^t(j,k) - g^t(j,k+1)](x_i^t - j)\frac{\partial y_i^t}{\partial a_y^r} + [g^t(j+1,k) - g^t(j+1,k+1)](j+1-x_i^t)\frac{\partial y_i^t}{\partial a_y^r}\} \quad \text{where} \quad \frac{\partial y_i^t}{\partial a_y^r} = \frac{x_i^e}{d^r.x_i^e + e^r.y_i^e + 1} \tag{7}$$

The expressions for coefficients $\mathbf{c}_{11j}$ and $\mathbf{c}_{12j}$ depend on derivatives of x_i^t and y_i^t.

$$\begin{aligned} c_{11i} = h_g\{ & [[g^t(j,k) - g^t(j+1,k)](y_i^t - k) + [g^t(j,k+1) - g^t(j+1,k+1)](k+1-y_i^t)]\frac{\partial x_i^t}{\partial d^r} \\ & + [[g^t(j,k) - g^t(j,k+1)](x_i^t - j) + [g^t(j+1,k) - g^t(j+1,k+1)](j+1-x_i^t)]\frac{\partial y_i^t}{\partial d^r}\} \end{aligned} , \tag{8}$$

$$\text{where} \quad \begin{aligned} \frac{\partial x_i^t}{\partial d^r} &= \frac{(a_x^r.x_i^e + b_x^r.y_i^e + c_x^r).x_i^e}{(d^r.x_i^e + e^r.y_i^e + 1)^2} \\ \frac{\partial y_i^t}{\partial d^r} &= \frac{(a_y^r.x_i^e + b_y^r.y_i^e + c_y^r).x_i^e}{(d^r.x_i^e + e^r.y_i^e + 1)^2} \end{aligned} \tag{9}$$

The usage of the expressions from (5) to (9) allow s to f orm a res idual equ ations and finally to f ind the parameters by Least-Squares method. This approach has advantages to determ ine m ore precis ely th e v alue of param eters. From practical point of view there are several problems that have to be solved. First of all increasing the number of unknowns tends to lower stability of solution. The large number of parameters decreases the speed of processing.

2.3 The Relief Disturbance

The estimation of relief influence could be done based on traditional projective relations. The exact matching takes part only for the centre of matching window. For pixels n ear th e w indow bou ndaries appears error du e to th e h eight differences. T he co llinearity eq uations co uld b e u sed to estim ate th e v alue o f th at d ifference. T he h orizontal parallax difference bet ween two poi nts $\mathbf{P}_i = (X_i, Y_i, Z_i)^t$ and $\mathbf{P}_j = (X_j, Y_j, Z_j)^t$ in matching window in case of the relative orientation model with basic left photo has the form:

$$\begin{aligned} \Delta p = p_i - p_j = x_{(1)i} - x_{(2)i} - (x_{(1)j} - x_{(2)j}) = c\frac{a_{11}^{(2)}(X_i - X_{B-}) + a_{21}^{(2)}(Y_i - Y_B) + a_{31}^{(2)}(Z_i - Z_B)}{a_{13}^{(2)}(X_i - X_B) + a_{23}^{(2)}(Y_i - Y_B) + a_{33}^{(2)}(Z_i - Z_B)} - c\frac{X_i}{Z_i} \\ - c\frac{a_{11}^{(2)}(X_j - X_B) + a_{21}^{(2)}(Y_j - Y_B) + a_{31}^{(2)}(Z_j - Z_B)}{a_{13}^{(2)}(X_j - X_B) + a_{23}^{(2)}(Y_j - Y_B) + a_{33}^{(2)}(Z_j - Z_B)} + c\frac{X_j}{Z_j} \end{aligned} , \tag{10}$$

where **c** is the camera constant, $a_{ij}^{(2)}$ are the coefficients of the rotation matrix for the right photo.

In cases of absolute orientation, after the transformation of right image to ensure working in epipolar lines, the expression for parallax difference takes the following simplified form

$$\Delta p = p_i - p_j = c\frac{X_i - X_B}{Z_i - Z_B} - c\frac{X_i}{Z_i} - c\frac{X_j - X_B}{Z_j - Z_B} + c\frac{X_j}{Z_j} =$$
$$= c\frac{(X_i + S_x).Z_i - X_i(Z_i + S_x\, tan\gamma)}{Z_i(Z_i + S_x\, tan\gamma)} + c\frac{(X_i - X_B)(Z_i + S_x\, tan\gamma - Z_B) - (X_i + S_x - X_B)(Z_i - Z_B)}{(Z_i - Z_B)(Z_i + S_x\, tan\gamma - Z_B)} = \tag{11}$$
$$= c\frac{S_x.Z_i - X_iS_x\, tan\gamma}{Z_i^2 + Z_i.S_x\, tan\gamma} + c\frac{(X_i - X_B)S_x\, tan\gamma - S_x(Z_i - Z_B)}{(Z_i - Z_B)^2 + (Z_i - Z_B)S_x\, tan\gamma}$$

To analyse the errors due to the relief, the four typical points, lying at the corners of matching window are examined. For points that differ in X direction the expression has the form:

$$\Delta p = c\frac{S_x.Z_i - X_iS_x\, tan\gamma}{Z_i^2 + Z_i.S_x\, tan\gamma} + c\frac{(X_i - X_B)S_x\, tan\gamma - S_x(Z_i - Z_B)}{(Z_i - Z_B)^2 + (Z_i - Z_B)S_x\, tan\gamma}\ , \tag{12}$$

where S_X is distance between points in X direction and γ is the angle of slope of terrain in X direction.

The simplification of expression could be made if the assumption that there is not difference between Z coordinates of the projection centres of photos. In that case the expression takes form:

$$\Delta p = p_i - p_j = c\frac{Z_BS_x - X_BS_x\, tan\gamma}{Z_i^2 + Z_i.S_x\, tan\gamma} \tag{13}$$

In case of points that has the same X coordinates and differ in Y coordinates the expression is:

$$\Delta p = \frac{-X_iS_Y\, tan\beta}{Z_i^2 + Z_i.S_Y\, tan\beta} + c\frac{(X_i - X_B)S_Y\, tan\beta}{(Z_i - Z_B)^2 + (Z_i - Z_B)S_Y\, tan\beta}\ , \tag{14}$$

where S_Y is the distance in Y direction and β is the angle of the terrain slope in Y direction.

The simplified expression for such points in case of $Z_B = 0$ takes the form:

$$\Delta p = p_i - p_j = c\frac{-X_BS_Y\, tan\beta}{Z_i^2 + Z_i.S_Y\, tan\beta} \tag{15}$$

The analyses of expressions for X and Y directions shows that the difference of p differ that can be modelled by usage of different scale for two directions. To analyse the influence more precisely the differentiating of final expressions could be made. For example for X direction the first derivative has the form:

$$\frac{\partial \Delta p}{\partial S_x} = c\frac{Z_B - X_B\, tan\gamma}{(Z_i + S_X\, tan\gamma)^2} \tag{16}$$

The expression (16) for first derivative is non-linear and could not be modelled by different scales but error is too small. For image with resolution 1200dpi, size of matching window 128x128 pixels, camera with view angle β=60° and photo size 23x23cm, and slope of the terrain γ=45° the ratio $Sx * tg\gamma / Z_1 = 0.0408$. That result allows to neglect the non-linearity of expression (13). The similar estimation could be made for expression in Y direction. As conclusion the terrain influence when the transformation by epipolar lines is done could be estimate by affine transformation. The situation when the terrain has different slope in matching window or there is a height jump in the window is more complicated and need more sophisticated modelling.

3 THE MATCHING PROCEDURE

The matching procedure consists of three stages. At the first stage the preprocessing the pattern image and weight matrix are prepared. The second stage of processing implements the procedure of finding the initial approximate position of target. The third step of the matching procedure performs the least squares matching.

3.1 The Preprocessing stage

At the preprocessing stage the edge enhancement procedure is used. Direct application of high pass filter does not give the satisfactory results. To avoid increasing the noise the regional contour detector of Hueckel type is applied. To ensure better results for contours with dense change of direction the modification of operator is used which uses small detector window (Marinov, 1994). The contour information is used to create the pattern image or weight matrix. The superposition in different ratio of original image and halftone vector image is generated. The line width of contour pattern depends on the parameters of low pass filter applied to the contour image. For pattern images of the fiducial mark and the artificial ground control points only the contour presentation of vector objects is used. For relative orientation procedure the enhancement of pattern image is applied.

3.2 Finding the Initial Position

The stage of initial position finding is very important to ensure high speed of processing and to avoid false matching. To solve that task the variable resolution of images is used. The resolution of pyramid images varies in power of 2. In practice only the resolution 2 and 4 from the original image resolution are applied. Second specific feature is the used model for correlation criteria. Here the well known normalised correlation is applied. It is described by the expression:

$$r(m,n) = \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} P(j,k).T(j-m,k-n)}{\sqrt{\sum_{j=1}^{J} \sum_{k=1}^{K} P^2(j,k)}.\sqrt{\sum_{j=1}^{J} \sum_{k=1}^{K} T^2(j-m,k-n)}}, \qquad (17)$$

where P(j,k) are grey-level values of pattern image and T(j,k) - for target image.

At the stage of initial position finding only the displacement parameters and rotation between images are determined. For variable steps of the matching window resolution only the displacement is determined. The rotation is determined only for maximal resolution. The nearest neighbour method is used for grey level approximation at this stage. Such approach suppose that the scale differences and the relative angles are to small and their influence could be neglected at that stage of matching procedure.

3.3 The Least Squares Matching

The Least Squares (LSQ) method is applied only at the final stage of parameters estimation. The specific feature is usage of mask that covers the pixels near the contour model. Mask usage allows to decrease the number of pixels taking part in equation forming. The mask is based on contour presentation of pattern image. Another specific feature is an adaptation of the procedure in terms of number of parameters. The number of parameters depends on the type processing - interior, relative, exterior and the stage of processing. In realisation of procedure the maximal number of parameters is used only at final iterations of relative orientation and control point identification. As final results of LSQ matching the parameters values and co-variance matrix of parameters are calculated. The coordinates of centres of pattern and target windows are calculated to obtain horizontal and vertical parallaxes in points of interest. The co-variance matrix values are used for estimation the accuracy of matching.

4 THE IMPLEMENTATION OF MATCHING PROCEDURE

The matching procedure could be used at different stages of processing. There are specific characteristics that appear at different stages. The matching procedure is applied for identification the position of fiducial marks, the relative position of corresponding points in relative orientation, the corresponding points in absolute orientation and the identification of position of artificial control points.

4.1 The Interior Orientation

The implementation of matching procedure requires the creation of fiducial mark pattern. Many authors like Lue (1977) discuss such patterns. The semi-automated procedure for preparing the patterns is suggested here. It is based on the generation of contour presentation based on utilisation of the edge detector and contour forming procedure. The pattern image is prepared from high-resolution raster image. When the vector model is used it is rasterized according with the resolution of processed image. The raster presentation is smoothed by low pass filter so the halftone image is produced. An example of pattern forming for Wild RC20 camera is shown on fig.1, fig.2 and fig.3. The halftone model generation from vector pattern is presented on fig.4.

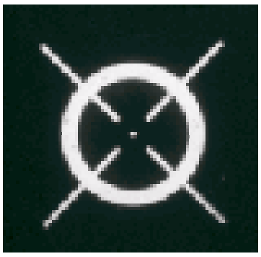

Figure 1. Fiducial mark grey level image

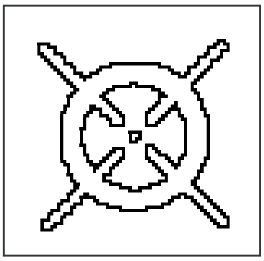

Figure 2. Contour image of mark formed from image

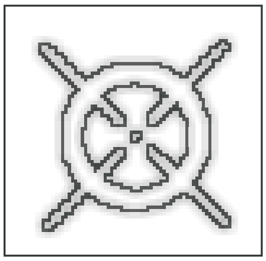

Figure 3. Grey level mark model from contour image

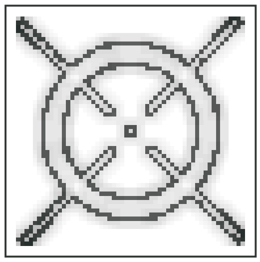

Figure 4. Grey level model from vector pattern

If the fiducial mark is rotationally asymmetric the pattern image have to be rotated accordingly with the actual position of fiducial mark. The initial position of target mark is used to determine the parameters of transformation in target image. On that stage of processing only the six parameters of affine transformation are used.

4.2 The Relative Orientation

The relative orientation requires determination of at least six points in the stereo overlapping zone. After initial determination of point position it is applied the procedure for finding the most convenient position. It is selected by calculation the width of auto-correlation function. The searching area is restricted to the size of matching window. The second step is finding the initial position at pixel level. The third second step is initialising the parameters of geometric transformation. The relative orientation procedure is processed in two steps. After initial finding the coordinates of corresponding points the pass of relative orientation is done. Before reaching the final accuracy the determination of relative position by LSQ method is applied again.

4.3 The Exterior Orientation

The finding of control points position has several specific features. The matching procedure is similar to those at the stage of relative orientation. The finding of control points depends on the type of control point. The point could be natural object. In that situation the automatic determination could not be made. Some types of the signalized control points are discussed by Hahn (1996). In some cases the signalized point consists not only of circular element but contain linear objects. The procession of such type of signalized point is shown on the following figures.

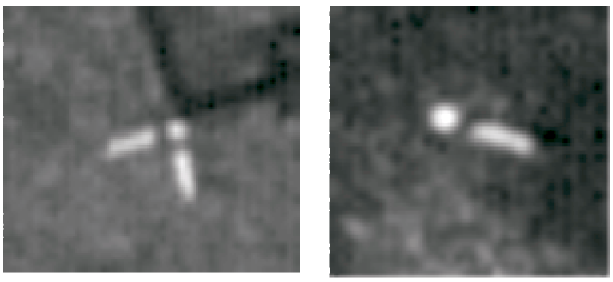

Figure 5. Control points with pointing strips

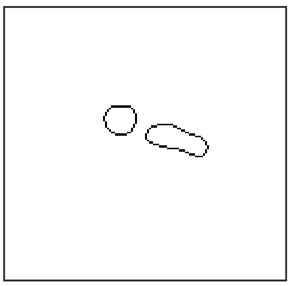

Figure 6. Control point after applying the contour detector

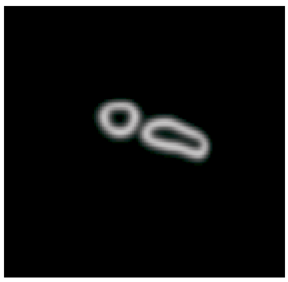

Figure 7. Grey level model of control point

The contour model of control point is used to prepare normalised model of control point. The halftone pattern of control point is generated that is used in procedure for automatic point finding. It could be used as pattern image or as weight matrix in LSQ method. If the point is marked on terrain there are a lot of distortion due to the projective transformation.

If the point is marked in one photo there are not distortion due to the terrain. In that situation arises the problem with the matching. To solve the problem a special type of weigh matrix is used that avoids the usage of pixels in marked area.

4.4 Experimentation of the method

The analyses of procedures for automatic determination the fiducial mark position and the identification show that the mainly used types of objects are circles and line strokes. By that reason the accuracy and reliability of suggested method for such type of objects is very important. To investigate the properties of method for such type of objects are examined circle objects with 5 pixel diameter and strip objects with the width of 5 pixels. For such types of objects are calculated normalised correlation function determined by (17), and the error function for least squares condition:

$$pvv = \sum_{i=1}^{n} p_i v_i^2 \ , \tag{18}$$

where $\mathbf{v}_i$ is the grey level residual for the corresponding pixels of pattern and target windows.

The correlation function is calculated for several combinations of target and pattern images. The analyses is done for the following variants: pattern and target images without any changes, target image and contour model of pattern, contour models of target and pattern images.

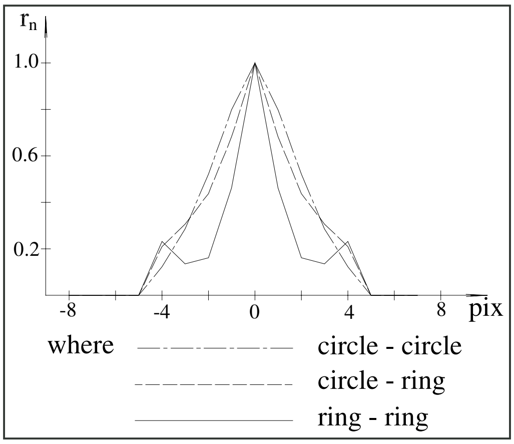

Figure 8. Correlation function for circle

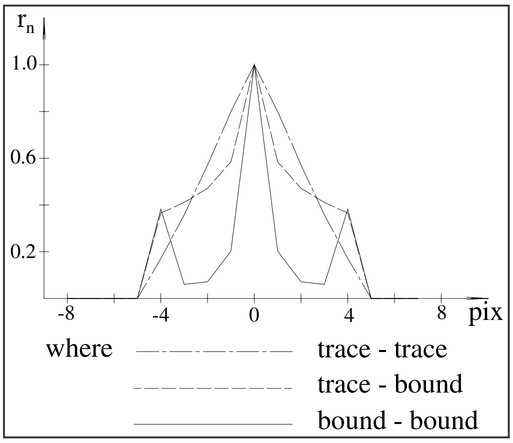

Figure 9. Correlation function for strip

For cases of the correlation function usage the width of maximum is wider if there is not applied contour image. The usage of contour image for target and pattern tends to the false maximums outside of the main one. The most appropriate combination is usage of the original image for target and the contour image for pattern. The error function for least squares method is calculated with weight function based on the pattern object and without the weight function. The combinations of target and pattern images are the same as for the correlation function. They are presented by the line types in the same way as in fig. 8 and fig. 9. The graphical presentations of results are shown in figures from 10th to 13th.

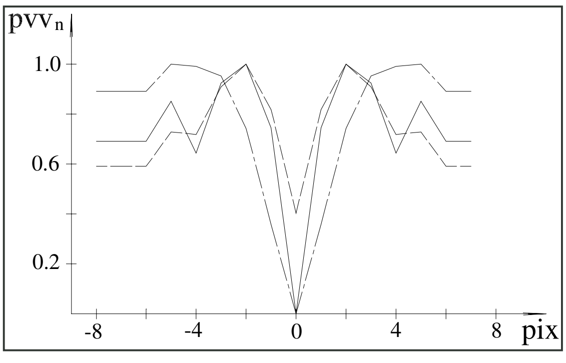

Figure 10. Error function for circle
without weight matrix

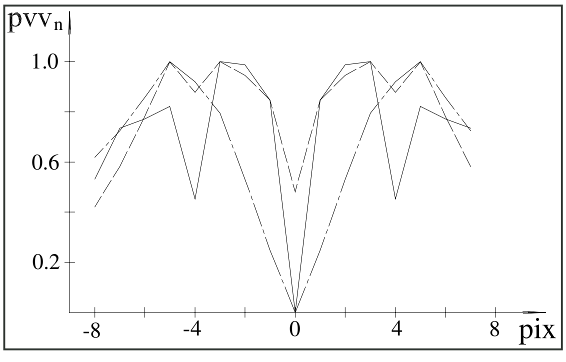

Figure 11. Error function for strip
without weight matrix

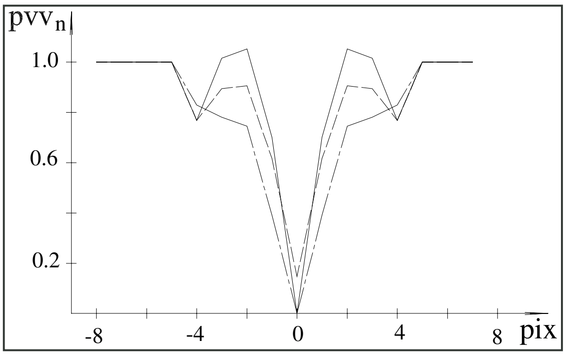

Figure12. Error function for circle with weight matrix

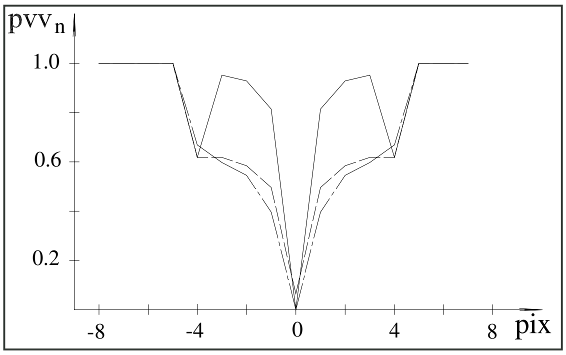

Figure 13. Error function for strip with weight matrix

The examination shows that the usage of LSQ method without weight matrix gives very poor results even in cases of usage the images without any changes. The combination of contour pattern with original target has false minimum in the case of circle determination.

5 DISCUSSION AND CONCLUSIONS

The application of the model is investigated at the stages of interior, relative and absolute orientation of stereo images. The control of the used parameters is proposed based on the availability of preliminary information and the corresponding stage of processing. The projective model is used only at the initial steps of relative orientation when the parameters of orientation are not well known yet. At the stage of interior orientation it is not necessary to apply projective model but only an affine geometric transformation. Such model is enough to be used at the stage of absolute orientation if preliminary image transformation is applied to ensure working on epipolar lines. The influence of the terrain slope does not allow applying a model with lower order. To obtain the initial position of the matching windows with pixel resolution more simple methods for image matching is suitable usage of more simple methods based on absolute difference and normalised correlation, Direct application of LSQ method does not gives good results due to the false extremes in optimisation function. The most appropriate is usage of original target and contour pattern at initial stage of correlation matching. At the stage of least squares matching the usage of original images with weight matrix is preferable.

The proposed procedure allows achieving more reliable and accurate results at different stages of stereo image processing. This procedure is adaptive depending on the stage of processing and characteristics of the images in the matching window. It ensures the sub pixel resolution and allows the error estimation to be used for creation of weight matrix of measurements at the following steps of analytical processing.

REFERENCES

Alwan R.H., 1996, Automatic Stereo Image Matching Using Edge Detection Technique, International Archives of Photogrammetry and Remote Sensing, Vienna, Austria, Vol.XXXI, Part B3, pp.29-35.

Hahn M., 1996, Semi-Automatic Measurement of Signalized Ground Control Points at Digital Photogrammetric Workstations, International Archives of Photogrammetry and Remote Sensing, Vienna, Austria, Vol.XXXI, Part B3, pp.291-296.

Heipke Chr., 1996, Automation of Interior, Relative and Absolute Orientation, International Archives of Photogrammetry and Remote Sensing, Vienna, Austria, Vol.XXXI, Part B3, pp.297-311.

Kraus K., 1997, with contributions by J. Jansa, H. Kager, Photogrammetry, vol.2: Advanced Methods and Applications, Bonn, “Ferd. Dümler Verlag”, pp.172-193.

Lue Y., 1977, One Step to a Higher Level of Automation for Softcopy Photogrammetry Automatic Interior Orientation, ISPRS, Journal of Photogrammetry & Remote Sensing, vol.53, No 3, pp.103-109.

Marinov B.D., 1994, High Accuracy Estimation of Area Boundaries from Digital Urban Images, Proc. of the International Conference “Remote Sensing and GIS in Urban Waters”, 12-16 Sep 1994, pp.137-144.