STEREO MATCHING FOR PUSHBROOM STEREO CAMERAS

Herbert JAHN
DLR
Institute of Space Sensor Technology and Planetary Exploration
Herbert.Jahn@dlr.de

Working Group III/2

KEY WORDS: Photogrammetry, Stereopsis, CCD Cameras, Image Matching, Parallel Processing.

ABSTRACT

A parallel stereo matching algorithm is presented which is mainly thought for the processing of images from pushbroom stereo cameras. The algorithm is designed for non-epipolar geometry, because of disturbances of flight attitude and velocity. Existing epipolar algorithms can give a first estimation of disparities in epipolar (x-) direction, but the recursive algorithm can also start with zero-disparity initial condition if the disparities are not too big. The algorithm minimizes locally a certain least squares distance of a stereo image pair using the method of steepest descent leading to a recursive disparity updating. To diminish ambiguities a pyramid with Gaussian image smoothing together with other measures (e.g. exploiting the ordering constraint and applying edge preserving disparity smoothing) is used. The presented matching algorithm is parallel in space and sequential in time. Therefore, when suitable parallel processing hardware (with one processing element assigned to each pixel) will be available then real-time stereo processing becomes possible. Some examples demonstrate the capabilities of the algorithm but also the remaining difficulties.

KURZFASSUNG

Ein paralleler Algorithmus zur Stereo-Bildzuordnung, der hauptsächlich für die Verarbeitung von Bilddaten von Pushbroom-Zeilenkameras gedacht ist, wird präsentiert. Der Algorithmus wurde wegen vorkommender Störungen der Fluglage und –geschwindigkeit für nicht-epipolare Geometrie konzipiert. Vorhandene epipolare Verfahren können für eine erste Schätzung der Parallaxen in epipolarer (x-) Richtung verwendet werden, aber der rekursive Algorithmus erbeitet auch ohne derartige Schätzwerte, wenn die Parallaxen nicht zu groß sind. Der Algorithmus minimiert lokal einen gewissen Abstand eines Stereo-Bildpaares durch Anwendung der Methode des steilsten Abstiegs. Dies führt zu einem rekursiven updating der Parallaxen. Zur Verminderung von Mehrdeutigkeiten wird eine Gaußsche Pyramide zusammen mit anderen Maßnahmen (Reihenfolgebeschränkung – ordering constraint, kantenerhaltende Glättung von Parallaxen) verwendet. Der Algorithmus ist räumlich parallel und zeitlich sequentiell und kann daher, wenn geeignete Parallelverarbeitungs-Hardware (mit einem Prozessorelement pro Pixel) verfügbar ist, Echtzeit-Stereoverarbeitung gewährleisten. Einige Beispiele zeigen die Fähigkeiten des Verfahrens und seine Mängel.

1 INTRODUCTION

The successful experiments with the digital stereo cameras HRSC (High Resolution Stereo Camera), WAOSS (Wide Angle Optoelectronic Stereo Scanner), WAAC (Wide Angle Airborne Camera) and with a prototype of the first commercial digital aerial camera ADC (Airborne Digital Camera) performed at the Institute of Space Sensor Systems and Planetary Exploration of the German Aerospace Center (DLR) have shown that high quality stereo reconstruction with pushbroom cameras is possible.

To generate cost efficient and high quality 3D data products the key problem continues to be the matching of two or more image stripes. One needs a very fast matching algorithm in order to process efficiently the huge data amounts generated.

Because of aircraft attitude and velocity variations the image geometry is not strictly epipolar. Therefore, available efficient epipolar algorithms, e. g. the algorithm of Gimel'farb (1999) using dynamic programming techniques, can be used only as a first approximation. For refinement a very fast and precise non-epipolar algorithm is needed.

Here, a new method is presented which relies on a pixel-wise minimisation of a quadratic measure of the grey value deviation of two images (e. g. nadir and forward images) under some constraints to estimate the disparity in both directions. The 2D disparity vector is updated recursively in each image point of the nadir image until stability is reached, i. e. the disparity change remains below some threshold. As constraints at the moment the ordering constraint (Klette et al., 1998) is used. In image regions without occluded areas and with continuous surface the ordering constraint guarantees a coupling of disparity values over some distance which is essential for stereo matching within homogeneous areas where no stereo information is available. To transfer stereo information from edges, corners etc. into homogeneous regions Gaussian smoothing of the image pair is applied.

In order to prevent too big gray value deviations of the stereo image pair a coarse gray value fitting (equal mean value and standard deviation in both images) is applied as a pre-processing step. The initial condition of the non-linear algorithm can be given by the disparity values $s_x(i,j)$ (in epipolar direction, derived by an epipolar matching algorithm), $s_y(i,j)=0$ (perpendicular to epipolar direction). Using these values instead of $s_x(i,j)=0$ diminishes possible trapping in local wrong minima if big disparities occur in the image pair.
The resulting disparity images $s_x(i,j)$, $s_y(i,j)$ can be smoothed with a special edge preserving smoothing algorithm which assumes that the disparity inside image segments changes smoothly (i. e. without discontinuities). That assumption especially seems to be useful in order to assign disparities to partly occluded image regions. The used smoothing algorithm is an adaptation of an edge preserving smoothing algorithm for images of gray values (Jahn, 1998) and other features (Jahn, 1999).

In contrast to most other (sequential) methods the presented matching algorithm is parallel in space and sequential in time. As in the human visual system where layered neural processing structures solve the matching problem (Hubel, 1995) the new algorithm also can be implemented in special Multi Layer Neural Networks or in recurrent Neural Networks. Therefore, when suitable parallel processing hardware (with one neuron or processor element assigned to each pixel) will be available then real-time stereo processing becomes possible.

Because the disparity is computed in each pixel of the nadir image a dense disparity map is generated and there is no need for interpolation.

The method was tested with some airborne images with good success. Some results (disparity and matched gray value profiles, disparity images) are presented. Of course, the quality of the results must be enhanced further, especially in regions with occlusions.

The method will be described in chapter 2. Then in chapter 3 the results of some processed images are presented and discussed. Finally, conclusions for further investigations are drawn.

2 THE METHOD

Let $g_L(i,j)$, $g_R(i,j)$ (i=1,...N_x; j=1,...,N_y) be an image pair registered with a left and a right camera, respectively. In case of pushbroom line scanners the L-image corresponds to the nadir image and the R-image to one of the images taken by the forward or backward looking CCD lines. Between both images the following relation (approximately) holds:

$$g_L(i,j) = g_R(i + s_x, j + s_y) \,. \tag{1}$$

Here, $s_x = s_x(i,j)$ and $s_y = s_y(i,j)$ are shifts (disparities) in x- and y-direction, respectively. They are considered as functions of the coordinates (i,j) of the left image (the L-image is considered as a distinguished image having in mind the pushbroom scanner application where the L-image is the nadir image).

With the coordinate transform $i' = i + s_x/2,\ j' = j + s_y/2$ (1) can be written in the equivalent form

$$g_L\left(i'-\frac{s_x}{2}, j'-\frac{s_y}{2}\right) = g_R\left(i'+\frac{s_x}{2}, j'+\frac{s_y}{2}\right) \tag{2}$$

which can be used for obtaining better stability of the algorithm. Of course, the disparity can also be assigned to a centered (cyclopean) image as it is done sometimes (Belhumeur, 1996, Gimel'farb, 1999). This does not change the method.

Of course, (1) or (2) does not hold in regions which are occluded in one of the images. Therefore, in images with many occluded areas (e.g. in cities) each method which relies on a assumption like (1) in every pixel will not work very good in those regions. Future investigations are necessary to alleviate that problem. The smoothing procedure to be considered at the end of this chapter is one (preliminary) approach. Here, the problem is not considered in detail. But it must be stressed that the problem of occlusion generally cannot be solved exactly because of missing information. The same is true in homogeneous regions with constant gray value where no disparity information is available. A priori assumptions and interpolation processes are necessary preventing exact measurements in all image pixels. Precise measurements are possible only in image points which can be identified non-ambiguously in both images. Even the human visual system with its huge stereo processing capability is not able to measure precisely! Of course, using more than two images (as can be done with the above mentioned digital stereo cameras), the number of occluded image points can be reduced and the precision can be enhanced. Here, that possibility is not considered because here the development of the new parallel method is in the center of interest but in the future the method should be generalized to the multi-viewing case.

Furthermore, (1) does not hold exactly because of illumination changes, non-lambertian reflection properties of object surfaces etc.. That can be taken into account very roughly by adapting the mean gray values and standard deviations of both images. A better performance can be obtained if the image statistics is adapted locally using the local median and the local mean absolute difference instead of mean gray value and standard deviation. This must be studied more carefully in the future (see also Wei et al. 1998).

To determine the disparities $s_x = s_x(i,j)$ and $s_y = s_y(i,j)$ in each point (i,j) first the least squares measure

$$J_1(i, j; s_x, s_y) = \left[g_L(i, j) - g_R(i + s_x, j + s_y)\right]^2 \tag{3}$$

is considered. For every (i,j) J_1 is a function of two variables s_x, s_y. To look for the minimum of J_1 here the method of steepest descent is considered. Of course, there are many other minimization methods (Himmelblau, 1972) but here only the principle is essential and therefore comparisons of the existing methods are not performed.

According to the method of steepest descent starting from an initial point $\left(s_x^{(0)}, s_y^{(0)}\right)$ the disparities are changed in the direction of the negative gradient of J_1:

$$\mathbf{s}^{(t+1)}(i, j) = \mathbf{s}^{(t)}(i, j) - \boldsymbol{\kappa} \cdot \nabla J_1\left(i, j; \mathbf{s}^{(t)}\right) . \tag{4}$$

In (4) $\mathbf{s}^{(t)} = \begin{pmatrix} s_x^{(t)} \\ s_y^{(t)} \end{pmatrix}$ is the disparity vector at recursion level (or discrete time) t. The matrix $\boldsymbol{\kappa} = \begin{pmatrix} \kappa_x & 0 \\ 0 & \kappa_y \end{pmatrix}$ is determined later.

To compute the gradient of J_1 first the derivative in x-direction is considered:

$$\frac{\partial J_1}{\partial s_x} = -2\left[g_L(i, j) - g_R(i + s_x, j + s_y)\right] \cdot \left.\frac{\partial g_R(x, j + s_y)}{\partial x}\right|_{x=i+s_x} \tag{5}$$

This representation has the drawback that $\frac{\partial J_1}{\partial s_x}$ vanishes in points (i,j) where $\frac{\partial g_R}{\partial x} = 0$ even if there is a shift between g_L and g_R in that point. Therefore, according to (2) it is better to use

$$J_1 = \left[g_L\left(i' - \frac{s_x}{2}, j' - \frac{s_y}{2} \right) - g_R\left(i' + \frac{s_x}{2}, j' + \frac{s_y}{2} \right) \right]^2 \tag{6}$$

instead of (3) because then the derivative $\frac{\partial g_L}{\partial x}$ can be used too.

Now some words concerning the calculation of the derivatives of g_L and g_R are necessary. The (continuous) image in the focal plane of an diffraction limited optical system is band-limited and hence an analytical function which can be differentiated (Jahn, Reulke, 1995). If the sampling frequency is greater than the highest spatial frequency transmitted by the optical system then Shannon's sampling theorem is fulfilled and the intensity samples (gray values) represent the whole intensity function $g(x,y)$. Then computing the derivatives of $g(x,y)$ is possible using the sampling theorem. Of course, this ideal case is not fulfilled. Often the sampling condition is violated meaning that the sampling theorem is not fulfilled exactly. Furthermore, the gray values are digitized numbers and noise is present which prevents exact differentiation of $g(x,y)$. Therefore, only approximations of the derivatives of $g(x,y)$ are possible. Here, the approximation

$$\frac{\partial g(x,j)}{\partial x}\Big|_{x=i} \approx \frac{1}{2}\left[g(i+1,j) - g(i-1,j)\right] \tag{7}$$

is used.
Of course, if one applies (5) with (7) to the original images then the recursion (4) often will be trapped in false (local) minima, especially when the initial disparities $\left(s_x^{(0)}, s_y^{(0)}\right)$ are too far from the real ones. Therefore, the recursion (4) is applied in a Gaussian pyramid (Jolion, Rosenfeld, 1994) starting with the coarse image

$$g_\sigma = G_\sigma \otimes g\,. \tag{8}$$

Here $\otimes$ means convolution, and the kernel G_σ is the normal distribution function $N(0,\sigma)$. The operation (8) has the advantage that it transfers disparity information (which is available at edges, corners etc.) into homogeneous regions. Therefore, if σ is big enough also inside large homogeneous regions a disparity can be measured and determined with the algorithm (4). Of course, the convolution (8) also blurs both images and fine detail is destroyed. In those regions wrong disparities are generated. To overcome this, the values of σ used in the pyramid must become smaller with increasing pyramid layer. Here, a pyramid with four layers l = 0, 1, 2, 3 was used and the σ - values were chosen according to

$$\sigma_l = 2^{3-l}\,. \tag{9}$$

Then the algorithm works as follows:

1. Initial disparities $\left(s_x^{(0)}, s_y^{(0)}\right)$ are chosen. Here the initial condition $\left(s_x^{(0)} = 0, s_y^{(0)} = 0\right)$ can be used. In case of pushbroom images the epipolar condition is fulfilled approximately. Then, using an epipolar matching algorithm (e.g. that of Gimel'farb, 1999), one can determine $s_x^{(0)}$. This reduces the number of pyramid layers (l = 0 may be sufficient), the number of iterations, and hence the computing time.
2. Layer l = 0: The smoothed L- and R-images are computed according to (8) with σ_0 = 8. Then the iteration (4) is started with the initial disparities $\left(s_x^{(0)}, s_y^{(0)}\right)$. The result are new disparities $\left(s_x^{(l=0)}, s_y^{(l=0)}\right)$ which are used as initial values in the next layer l = 1, and so on.
3. At the end of the iteration in layer l = 3 the result of the algorithm are the final disparity estimates $\left(\hat{s}_x = s_x^{(l=3)}, \hat{s}_y = s_y^{(l=3)}\right)$ in each point (i,j) of the left (nadir) image.

That algorithm has the drawback that it is slow on a *sequential* computer. To speed up the algorithm the smoothed images g_{σ_l} can be sub-sampled (for l = 3 one has $N_x \cdot N_y$ gray values, whereas in layer l there are only $N_x/2^{3-l} \cdot N_y/2^{3-l}$ sampling values). Then, the output of layer l are $N_x/2^{3-l} \cdot N_y/2^{3-l}$ disparity values, but in layer l+1 $N_x/2^{2-l} \cdot N_y/2^{2-l}$ initial values are needed. Therefore, interpolation is necessary. This works but it is not considered here in detail.

In order to transfer disparity information into the interior of large homogeneous regions the maximal value of σ and hence the highest pyramid layer must be chosen big enough. Adaptivity seems to be necessary, but this was not studied up to now.

Another method to enhance the performance of the algorithm is the use of

$$\tilde{g}_\sigma = \alpha_1 \cdot g + \alpha_2 \cdot g_\sigma \tag{10}$$

instead of the blurred images g_σ (8). The original image g used here besides g_σ preserves detail, and hence reduces the number of wrong disparities in image regions with fine structure. Some experiments showed that $\alpha_1 = \alpha_2 = 1$ seem to be good values, and that, using $\tilde{g}_\sigma$ instead of g_σ, the number of necessary pyramid layers can be reduced.

The method can be enhanced further if instead of the gray values alone additional local image features are taken into account. Here, the image gradients in x- and y- direction have been used. Then, the algorithm (4) is applied to

$$\begin{aligned} J_2(i,j;s_x,s_y) = \gamma_1 \cdot \left[g_L(i,j) - g_R(i+s_x, j+s_y)\right]^2 + \gamma_2 \cdot \left[\nabla_x g_L(i,j) - \nabla_x g_R(i+s_x, j+s_y)\right]^2 \\ + \gamma_3 \cdot \left[\nabla_y g_L(i,j) - \nabla_y g_R(i+s_x, j+s_y)\right]^2 \end{aligned} \tag{11}$$

This gives considerably better results because some ambiguity is removed. The problem of choosing adequate features is not considered here. It seems that wavelet transform coefficients give good results which also have been applied in a coarse-to-fine pyramid (Kim et al., 1997). (11) shows how such and other features can be used in the algorithm described here. It was not investigated which the best choice of the parameters γ_1, γ_2, γ_3 is. The results presented in section 3 were obtained with $\gamma_1 = \gamma_2 = \gamma_3 = 1$.

The steps (8), (10), (11) for reducing ambiguity are not sufficient. Other measures are necessary. In the literature many constraints have been formulated (Klette et al., 1998) which can be exploited to reduce ambiguity further. Here, the so-called ordering constraint (Klette et al., 1998, Wei et al., 1998) is used. This means that (in x-direction) if $g_R(i+s_x(i,j),j)$ is matched to $g_L(i,j)$ and $g_R(i+1+s_x(i+1,j))$ to $g_L(i+1,j)$ then

$$i + s_x(i,j) \le i + 1 + s_x(i+1,j) \tag{12}$$

must be fulfilled. (The same holds in y-direction). This is equivalent to the condition that the distance

$$d_x(i,j) = 1 + s_x(i+1,j) - s_x(i,j) \tag{13}$$

of successive matched image points i', $i'+1$ ($i' = i+s_x(i)$) of the g_R – image must be non-negative.

The condition (12) can be violated if there exist objects which are located in front of other objects (or the background) such that there is no continuous surface (see e.g. Klette et al., 1998). In aerial images this case is met when single clouds are present. In cloud free conditions it is met seldom. Up to now (12) is used not very often to enhance matching. One of the exceptions is the algorithm of Gimel'farb, 1999 which is based on dynamic programming and which gives good results for epipolar geometry.

The condition (12) has to be fulfilled in each iteration of the algorithm (4). To guarantee this the step width $\Delta\mathbf{s}^{(t)}(i,j)$ must be limited. If e.g. $\Delta s_x^{(t)}(i,j) > 0$ then the point $\left(i + s_x^{(t)}(i,j), j\right)$ will be shifted in positive x–direction. To fulfill (12) even if the next point $\left(i+1+s_x^{(t)}(i+1,j), j\right)$ is shifted in negative x–direction the increment $\Delta s_x^{(t)}(i,j)$ can be limited as follows:

$$\Delta s_x^{(t)}(i,j) = \begin{cases} d_x^{(t)}(i,j)/2 & \text{if} \quad \kappa_x \cdot \nabla_x J\left(i,j;\mathbf{s}^{(t)}\right) > d_x^{(t)}(i,j)/2 \\ -d_x^{(t)}(i-1,j)/2 & \text{if} \quad \kappa_x \cdot \nabla_x J\left(i,j;\mathbf{s}^{(t)}\right) < -d_x^{(t)}(i-1,j)/2 \\ \kappa_x \cdot \nabla_x J\left(i,j;\mathbf{s}^{(t)}\right) & \text{elsewhere} \end{cases} \tag{14}$$

The condition (14) is essential for the reduction of ambiguity because it prevents outliers (which do not obey the order constraint and which can be very frequent without application of (14)) of the disparity generated by noise and other deviations of both images. Of course, because (14) reduces the possible step size in each iteration many iterations may be necessary if the disparities are big. This is a drawback which might be reduced if the directions of ∇J in the right and left neighbors of point i are taken into account but this was not studied up to now.

One can see from (14) that it is not useful to use too big values of κ_x such that $\kappa_x \cdot |\nabla_x J|$ is much bigger than $d_x/2$.Therefore, at each pyramid level κ_x can be chosen adaptively according to some average values of $|\nabla_x J|$ and $d_x/2$, e.g. from $\kappa_x \cdot \langle |\nabla_x J| \rangle = 1/2$ or similar relations.

Further measures to reduce ambiguity are edge preserving smoothing of both images as a pre-processing step and a smoothing procedure applied to the disparity increments $\Delta s_x^{(t)}(i,j)$ (before (14) is applied) in order to reduce noise in the original images and in the disparity increments. For edge preserving smoothing of the original images the algorithm described in (Jahn, 1998) was applied. A modified algorithm for disparity smoothing is described now:

$$\Delta s_x^{(n+1)}(i,j) = \Delta s_x^{(n)}(i,j) + c \cdot \sum_{k,l=-1}^{1} w^{(n)}(i,j;k,l) \cdot \left[\Delta s_x^{(n)}(i+k,j+l) - \Delta s_x^{(n)}(i,j)\right] \tag{15}$$

($c = \dfrac{1}{\sum_{k,l=-1}^{1} w^{(n)}(i,j;k,l)}$).

With the special weights $w = 1$ the algorithm (15) computes the ordinary mean value recursively which blurs the disparities. But, the hypothesis to be used is that the disparity is constant or smoothly changing inside homogeneous image regions. Therefore, as for edge preserving smoothing of the original images (Jahn, 1998) the weights are chosen according to

$$w^{(n)}(i,j;k,l) = \frac{\mu^{(n)}}{\mu^{(n)} + \left[g_L(i+k,j+l) - g_L(i,j)\right]^2} \quad . \tag{16}$$

These weights favor contributions to (15) where $g_L(i+k,j+l) \approx g_L(i,j)$. $\mu^{(n)}$ is a parameter which can be kept constant or which can be chosen to tend to zero with increasing n in order to accelerate convergence. (15) generates a coupling of image lines because there is smoothing in y-direction. This diminishes errors in single image lines which sometimes can be seen as stripes in disparity images generated with pure epipolar algorithms.

Some general words concerning the recursive algorithm (4). If one looks at anaglyph images with big parallaxes then one observes that it takes some time until the final 3D impression is obtained and that this time increases with increasing disparity. That behavior can be explained with recursive algorithms of type (4) where the disparity increment is limited by some constraint as e.g. (14). Furthermore, the iterative determination of the disparities also reduces trapping in wrong maxima (as compared with "one-step" algorithms). Another argument for algorithms of type (4) is that the disparities in every image point (i,j) of the left image can be calculated in parallel (this holds also for the order constraint (14) and the smoothing procedure (15)). Therefore, when appropriate parallel processing hardware will be available then real-time stereo matching becomes feasible. Of course, this does not mean that the algorithm which is used here is an ultimate one. It only shows to a certain direction of further research.

When big disparities in both directions are present in the image pair, then the x- and y- disparities must be estimated together as proposed by (4). If the geometry is strictly epipolar then an one-dimensional version of the algorithm can be used ($\kappa_y = 0$). But, if the geometry is near-epipolar (only small y- disparities are present) then the algorithm can be first applied in epipolar direction (or another epipolar algorithm can be used) to estimate s_x coarsely with subsequent s_y estimation with the same algorithm applied in y- direction (with recursive application if necessary). This seems to be a good processing scheme for pushbroom stereo image pairs.

3 RESULTS

A few results to be presented now show the capabilities of the method and the difficulties which should be overcome by future research. First, a stereo pair of a rural region with a small village generated with the airborne camera WAAC is considered. Figure 1 shows the image pair.

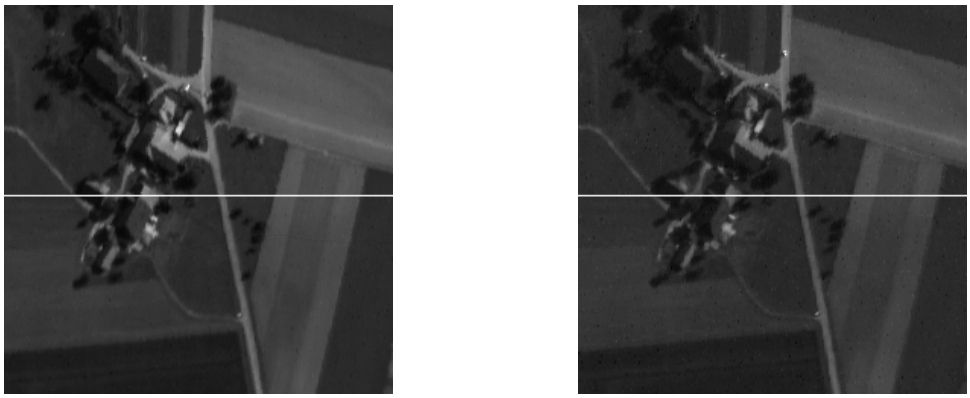

Figure 1. WAAC stereo pair “Village”

The bright horizontal lines show the places where a cross section of both images were taken. Figure 2 shows the gray value profiles along those lines. One can see the varying disparity. Figure 3 shows the profiles after matching. A good coincidence of both profiles now is observed.

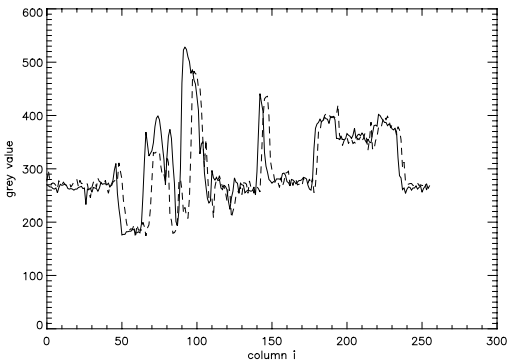

Figure 2. Gray value profiles

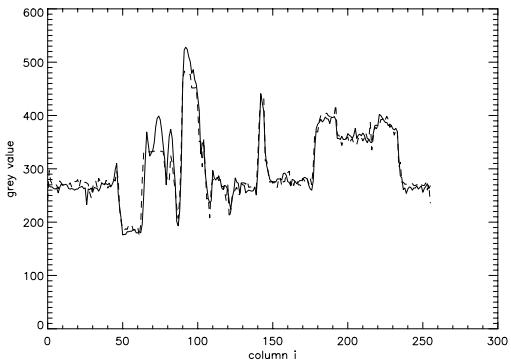

Figure 3. Gray value profiles after matching

The disparity s_x along the horizontal line is displayed in figure 4. This seems to be a useful result although ground truth is not available. Figure 5 shows the s_x- image.

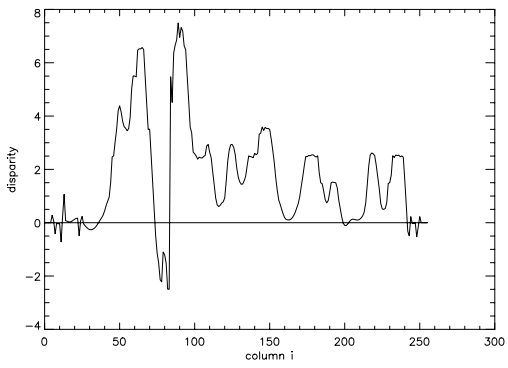

Figure 4. Disparity profile

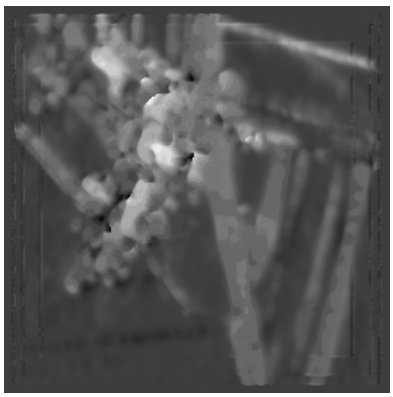

Figure 5. Disparity image

The stereo pair of figure 1 is not strictly epipolar. After matching in x- direction as was shown here there remain small shifts in y- direction in some image regions caused by aircraft attitude disturbances. These shifts can be nearly removed by applying the algorithm in y- direction. A red-green coded overlay of both images reveals this but that is not shown here because of limitations of printing space. Some areas with matching errors remain especially near buildings where occlusions occur. Furthermore, the disparities seem to be too smooth.

The well-known image pair Pentagon (figure 5) shows the limits of the present status of the algorithm better because in that image pair are many occlusions and especially small non-overlapping structures with big disparities which the algorithm cannot handle satisfactorily up to now. The x- disparity image shown in figure 6 reveals this.

Figure 5. Stereo pair "Pentagon"

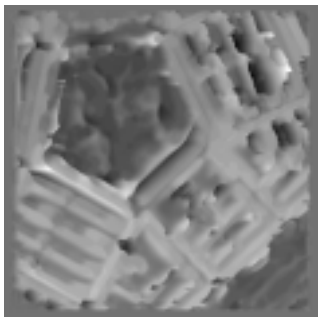

Figure 6. Disparity image

4 CONCLUSIONS

The few results presented here show that the algorithm works in some image pairs generated with aerial stereo cameras. In section 2 some remarks concerning necessary future research were already made which shall be supplemented now. Because of its local gradient computation (5), (7) the algorithm gives only good results when the disparities are small, i.e. if there is an overlapping of structures in both images. To cope with big disparities the gradient computation must be extended to a non-local operation. First experiments have given promising results but more investigations are necessary. Furthermore, the *J*-function (11) to be minimized has to be generalized. More geometric rather than radiometric information should be included because some disturbances which occur in only one image of the stereo pair (e.g. reflections of sun light) can lead to wrong disparities.

REFERENCES

Belhumeur, P. N., 1996. A Bayesian Approach to Binocular Stereopsis, Int. J. of Computer Vision, Vol. 19, pp. 237-260

Gimel'farb, G., 1999. Stereo Terrain Reconstruction by Dynamic Programming, in: *Handbook of Computer Vision and Applications* (eds. B. Jähne, H. Haussecker, and P. Geisser), Vol. 2, Academic Press, San Diego, pp. 505-530

Himmelblau, D. M., 1972. Applied nonlinear programming, Mc Graw-Hill, New York

Hubel, D. H., 1995. Eye, Brain, and Vision, Scientific American Library, New York

Jahn, H., Reulke, R., 1995. Systemtheoretische Grundlagen optoelektronischer Sensoren, Akademie Verlag, Berlin

Jahn, H., 1998. A neural network for image smoothing and segmentation, Proc. of Joint IAPR Int. Workshops SSPR'98 and SPR'98, pp. 329-338; Lecture Notes in Computer Science 1451, Springer

Jahn, H., 1999. Feature grouping based on graphs and neural networks, Proc. of CAIP'99, Lecture Notes in Computer Science 1689, pp. 568-577, Springer

Jolion, J.-M., Rosenfeld, A., 1994. A Pyramid Framework for Early Vision, Kluwer Academic Publishers, Dordrecht

Kim, Y.-S., Lee, J.-J., Ha, Y.-H., 1997. Stereo matching algorithm based on modified wavelet decomposition process; Pattern Recognition, Vol. 30, pp. 929-952

Klette, R., Schlüns, K., Koschan, A., 1998. Computer Vision. Three-Dimensional Data from Images, Springer

Wei, G.-Q., Brauer, W., Hirzinger, G., 1998. Intensity- and Gradient-Based Stereo Matching Using Hierarchical Gaussian Basis Functions, IEEE PAMI, Vol. 20, pp. 1143-1160