Presented Paper at ISPRS Commission III Symposium 'Spatial Information from Digital Photogrammetry and Computer Vision', vol. 30, part 3/2, pp. 737 – 744, September 5 - 9, 1994, Munich, Germany,

OBJECT SPACE BASED SURFACE RECONSTRUCTION WITH DISCONTINUITIES – AN APPROACH

Martin Schlüter
Institute of Photogrammetry and Cartography, TH Darmstadt
Petersenstr. 13, D-64287 Darmstadt, Germany
ms@gauss.phgr.verm.th-darmstadt.de

KEY WORDS: surface reconstruction, object space based image matching, discontinuities

ABSTRACT

This paper discusses concepts for the reconstruction of surfaces with discontinuities. Possible extensions to object space based image matching approaches are pointed out on the basis of Facets Stereo Vision. Matching in object space avoids disadvantages of the widespread image space based reconstruction procedures, which are caused by making implicit assumptions on relations between image features and surface geometry.

First results with synthetic image data are presented. The main goal besides an improvement of pure surface reconstruction from stereo images is to reach a good starting point for scene understanding tasks in urban areas, especially for the extraction and the interpretation of buildings for three dimensional GIS applications.

1 THE BASICS OF FACETS STEREO VISION

In this section I give a brief introduction to object space based image matching, following the foundations of Facets Stereo Vision published by [Wrobel 87]. There are some more slightly different approaches to object space based image matching, please refer to [Weisensee 92] for an overview.

In Facets Stereo Vision the object space has to be modelled mathematically. Usually the surface is approximated by a regular grid, the so-called height facets. Within each single facet, we assume a bilinear interpolation of the heights. To describe the gray values of the surface in object space, each height facet is subdivided into several gray value facets. They are representated in the same way as the heights, using bilinear interpolation between the object gray value grid points.

The elements of the surface have to be set in relation to their pictures in the images with regard to radiometric and geometric characteristics. This leads to the least squares estimation procedure outlined in fig. 1. The unknown parameters of the object space were estimated by the gray value observations of some digital images. In this case the orientation parameters are assumed to be known. Refer to [Kempa 94] for the simultaneous estimation of orientation parameters.

As the observation equations are to a high degree nonlinear, good approximate values for the parameters are needed. By the use of an image pyramid the requests to the start values can be kept low, [Kaiser et al. 92]. Corresponding to the image pyramid a pyramidal structure is built up for the unknows of the object space: With higher resolution of the images the size of the facets can decrease.

$$\boldsymbol{l} + \boldsymbol{v} = \boldsymbol{A}\boldsymbol{x}$$

$\boldsymbol{x}$ → parameters:
- object heights
- object gray values
- radiometry

$\boldsymbol{A}$ → designmatrix:
- approximate parameter values
- inner and outer orientation

$\boldsymbol{v}$ → residuals

$\boldsymbol{l}$ → observations:
- gray values of n digital images, $n \geq 2$

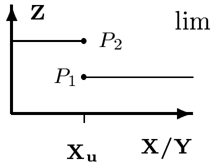
pseudo observations:
- adaptive curvature minimization

Fig. 1: Surface reconstruction with Facets Stereo Vision: observation equations

In order to overcome ill-posedness of the image inversion and to bridge areas, in which the gradients of the image gray value signal are too low, surface reconstruction requires regularization. Facets Stereo Vision uses adaptive curvature minimization of the object heights which is able to deal with smooth surfaces and gives good results in the presence of breaklines, [Wrobel et al. 92].

2 GOALS OF THE DETECTION OF DISCONTINUITIES IN OBJECT SPACE

Let us have a look at the mathematical formulation of a discontinuity, which gives us the name of our problem:



$$\lim_{\epsilon \to 0} f(X_u - \epsilon) \neq \lim_{\epsilon \to 0} f(X_u + \epsilon)$$

Fig. 2: Discontinuity

If we start with the simplified idea, that our ideally reconstructed surface in object space behaves somehow like the function in fig. 2, we can point out two problems that appear if we want to reconstruct this surface with a number of orientated digital images.

The first problem is that the functional surface representation $Z = f(X, Y)$ we usually apply in Facets Stereo Vision is not sufficient to describe the surface in object space, because the surface is not singlevalued in X_u. We are not able to describe the surface in X_u accurately because only one object height is allowed for each grid point.

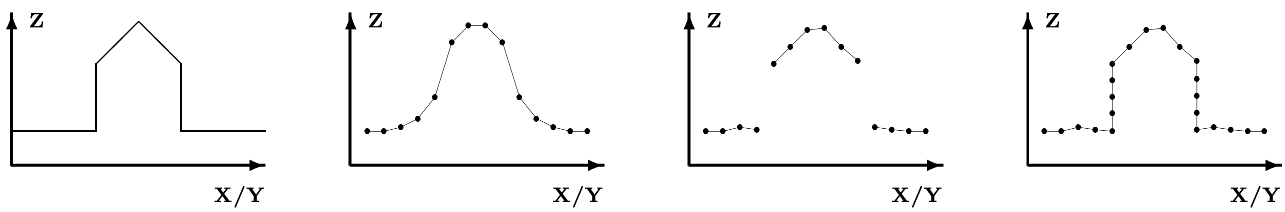
The second problem is, that the surface might not be reconstructed continuously. In fig. 2 for example, no statement could be done about the course of the surface between the points P_1 and P_2. This problem often appears at the presence of occlusion. The following case might occur that even large areas of the surface cannot be reconstructed, because they are not visible in at least two images taking part in the surface reconstruction. In other words, the mathematical description of height facets could have to be omitted due to a lack of information. This problem is in principle independent from the first problem, though both problems often appear together.

Both problems lead to the situation, that special areas of the surface can not be described accurately. The simplest way to deal with such areas is to detect (see section 3) and to omit (see section 4) them. Object space based matching offers the possibility to embed this detection and treatment of discontinuous areas of the surface in the matching procedure. The matching result of difficult terrain should be improved by this way.

Lets have a closer look at the course of action: If we try to reconstruct a surface containing a discontinuity with the classical Facets Stereo Vision approach, we are able to get a smooth approximation of this surface, see fig. 3a,b, if the height of the discontinuity step is small in relation to the resolution of height facets. The reason of this behaviour lies in the fact, that Facets Stereo Vision contains several smoothness constraints: Regularisation implies smoothness across the edges of the height facets. Interpolation within the height facets implies smoothness of the surface, too. In the same way smoothness of the gray values in one object gray value facet is assumed.

If the resolution of the height facets grows in the course of the multigrid process the contradiction between a discontinuity step of the real surface and the smoothness constraints of the modeled surface grows. Facets Stereo Vision offers some points of attachment for the detection of the discontinuity before the contradiction becomes so large, that the matching procedure fails. The discontinuity has to be localized as good as possible to keep the omitted area small. The expected result is characterized in fig. 3c. The surface is only described in those areas, where the assumed continuous model of the surface fits to reality.

For close-range applications it might be desirable to overcome the problem of the functional surface representation like it is shown in fig. 3d. In the following excursion, the idea of an extended surface representation is briefly introduced.



3a: Surface in object space

3b: Reconstruction by a smooth approximation

3c: Reconstruction of continuous areas only

3d: Reconstruction with extended surface representation

Fig. 3a-d: Quality steps of discontinuous surface reconstruction

If we are working with large scale aerial imagery, which is our main interest at the moment, the lack of information by occlusion might be too large in general. In this case, one should restrict oneself to reconstruct the continuous areas of the surface only. This kind of reconstruction does not only lead to a high quality DTM, but offers also a differentiated basis for further semantic approaches like the detection of houses and vegetation.

2.1 Excursion: Extended surface representation

The functional surface representation $Z = f(X, Y)$ does not allow any ambiguity. I we want to overcome this disadvantage, we should try to get a surface representation which is independent from any particular orientation of our coordinate system. This demand leads us to a differential orientation-based representation.

Further on we should keep in mind, that we are already dealing with an ill-posed problem. So the number of unknows should not be increased in a way that enlarges the solution space any more. This demand recommends an update at every iteration step instead of additional orientational surface-parameters of the least squares estimation.

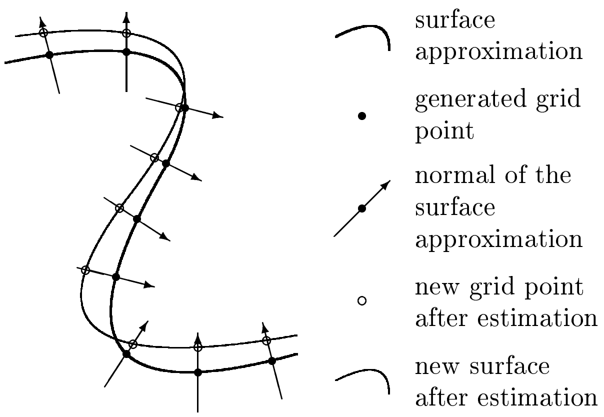


Fig. 4: Extended surface representation

The following surface representation would fulfill our requests: Bevor each iteration of the reconstruction process a grid in dependence on e.g. the curvature of the approximative surface has to be generated. For the details on the evaluation of surface curvature characteristics see [Besl 88], for the aspects of grid generation refer to [Schwarz 80]. In the iteration the direction of the estimated parameter increment for every grid point is the direction of the local normal vector – not as usual the direction of the global Z-axis, see fig. 4. By this way the surface representation is adapted to the estimated surface at every iteration step. Note that several interpolation methods between the grid points can be applied within the described surface representation – in fig. 4 a kind of smoothing interpolation is is used instead of the above mentioned bilinear interpolation!

With this method we lose the advantages of the regular quadratic grid structur, which gives us a very regular scheme of normal equations. Another problem is the management of the irregular grid, above all if we want to embed the concepts of omitting areas of the surface due to a lack of information as discussed in the following sections. [Liang and Taubes 94] give some hints in this direction.

3 DETECTION OF DISCONTINUITIES

3.1 Consistency check in respect of occlusion

Occlusion is present if several points or areas in 3-d object space are projected to identical positions in 2-d image space [Encarnação und Straßer 86]. Only the point in object space next to the projection centre will appear in the image. For areas in object space which are visible in only one image [Grimson 81] proposes an extrapolation of the estimated surface which imitates the human viewing system, which associates a depth value even for such occluded areas. In opposition to this suggestion the photogrammetric purposes demand a description of the surface only in those areas, where it is factually determinable.

Opposed to image space based matching approaches, where the detection of occlusion often just stands for the detection of regions of one image, which cannot be matched to the second image, see e.g. [Weng et al. 92], the object space based concept of Facets Stereo Vision implies to check wether a point in object space appears properly in an image without any occlusion by other regions of the surface.

This check has to be done after each iteration of the reconstruction process, after a first smooth approximation of the surface was evaluated at the first level of the multi-grid evaluation. By this way, consistency between the actually reconstructed surface and the requirements of visibility can be reached for the following iteration. We get the information whether an image gives information for the estimation of a surface facet or not and whether a surface facet is determinable (is visible in more than one image) or not. It can be seen that dealing with discontinuous surfaces the detection of occlusion is a necessary but not a sufficient tool.

The search for occlusions can be realized in the way, that the straight projection line of every image pixel is examined for another point of intersection with surface between the actual Z-value of the point of intersection with the surface facet and the actual maximum height of the DTM. Further on, e.g. a condition for the angle of intersection between the projection line and the surface normal in the point of intersection can be introduced to avoid a bad intersection.

An example shall illustrate the course of action: In fig. 5a-e the generation of some synthetic imagery is explained. For this example, only image 1 and 2 are used. It can be seen that a part of the surface on the right side of the discontinuity step is occluded in image 1, the surface can not be reconstructed in this area because there is only one image which shows the whole scene.

Starting with a global horizontal plane as approximation for the heights, the third level of the image pyramid was used to reach smooth approximation values for the heights similar as shown in fig. 7a. Fig. 6a-d show the results after reaching convergence on the next level of the image pyramid including the search for occlusion within every iteration. The above mentioned condition for the angle of intersection was approximated by the demand for a minimum number of image pixels per surface facet.

If this additional condition is not used, the surface won't be broken in all facets along the discontinuity step. This is not really a drawback, because the described detection of occlusion has only supplemental character to the detection of discontinuities by residual analysis as described in section 3.3.

3.2 Merging matching and surface analysis

If we want to reconstruct a discontinuous surface from stereo images it is obvious that we could try to use some hints by the course of our approximative surface or by the gray value contours in the approximative orthoimage to determine possible locations of discontinuities.

There have been lots of investigations in this direction. [Cochran and Medioni 92] and [Weng et al. 92] for example propose the connection of edge feature based - and area based matching in image space to get along with discontinuities in object space. The problem of this methods is that is it difficult to analyse whether a gray value edge feature belongs to a surface discontinuity or has some other reason.

[Terzopoulos 88] proposes to detect discontinuities by locally monitoring sharp deflections in the evolving surface to initialize his continuity control functions reconstructing a surface from scattered depth data. In the same sense [Hoff and Ahuja 89] argue that the steps of matching and surface interpolation should be merged. [Zheng 93] improves the matching result of a discontinuous surface by introducing the first and second derivation of the surface into the next iteration of the surface reconstruction process.

Concerning breaklines, the essentials of these ideas have already been realized within Facets Stereo Vision by the adaptive regularisation. With the following approach, the decision whether a discontinuity in object space exists or not shall not any more be dependent on the gradient or the curvature of the surface, but on the fact whether the reconstructed surface is consistent in relation to the input information of the images or not.

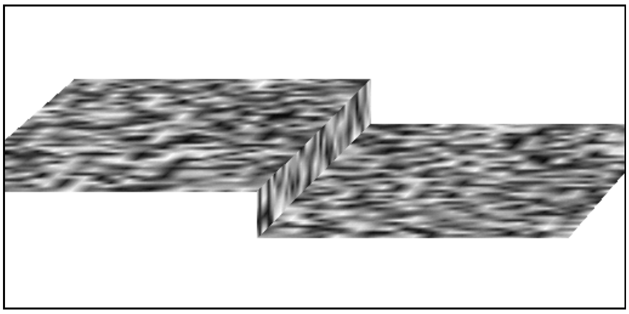
3.3 Detection of discontinuities by residual analysis

The residuals of the gray value observations of Facets Stereo Vision represent the difference between the gray values in the computed orthoimage and the gray values of the related image pixels. In the figures of this paper these residuals are represented in image space. There exists a second class of residuals related to the pseudo observations of the adaptive curvature minimization. We do not deal with them in this paper because their expressiveness towards gross error detection is very small in the case of adaptive regularisation.

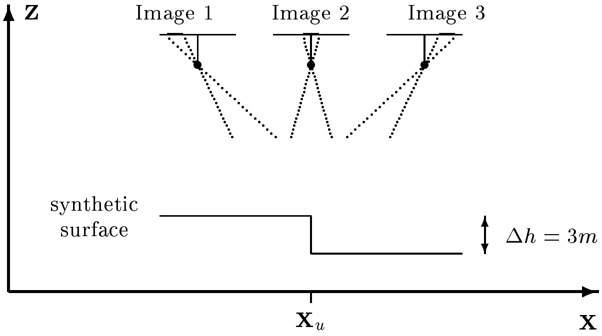
Early on [Helava 88] pointed out the possibility of gross error detection by residual analysis for object space based matching. [Heipke 89] shows that is possible to eliminate gross errors like reseau crosshairs or dust on single images in one of three images. [Schmolla 91] presents first considerations towards the detection of discontinuities by residual analysis.

The main idea is to distinguish whether only one image is concerned by a local disturbance in the residual images, which indicates a gross error like dust in the image, or whether all images are concerned by a local disturbance, which indicates a 'gross error in object space'. This 'gross error in object space' can be interpreted as a deficit of the mathematical model of the surface in object space. When this deficit is of a larger order than terrain noise, I argue to treat the corresponding area as a discontinuity.

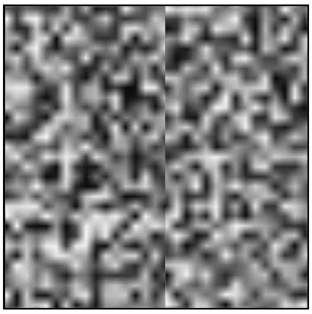
[Ebner et al. 93] calculate one orthoimage per input image and represent the difference between these orthoimages in object space, which is similar to the projection of the residual images in object space. They interpret the remaining areas of large differences by the presence of breaklines. Though they use a relative low resolution of the DTM in object space and smoothness constraints breaklines too have to be interpreted in this case as a deficit of the mathematical model of the surface. On the one hand, one can partially overcome this problem by adaptive regularisation and by a finer resolution of the height grid in relation to the image pixel size. On the other hand,



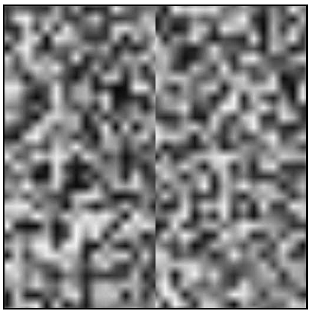
5a: Visualisation of the synthetic surface in object space and its random texture



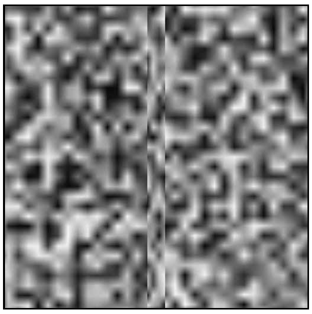
5b: Situation of the simulated exposure



5c: Image 1

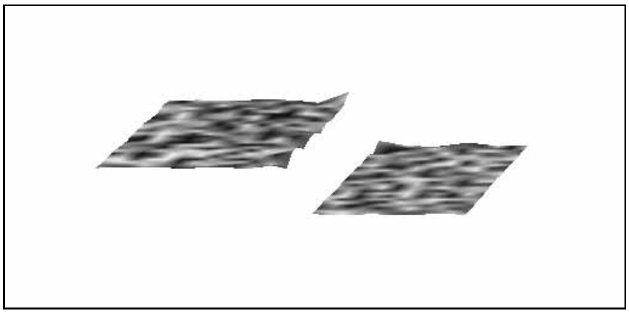


5d: Image 2

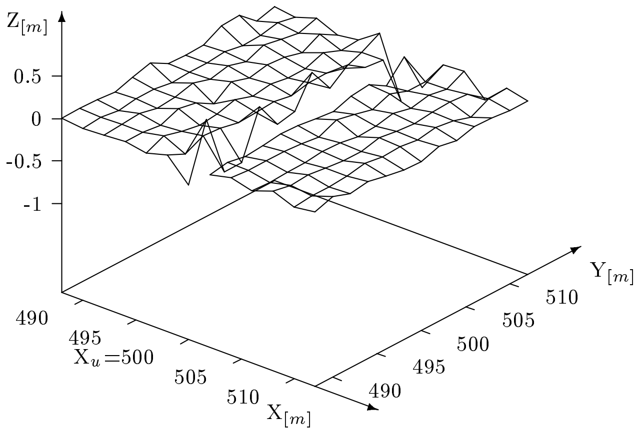


5e: Image 3

Fig. 5a-e: The synthetic input data for Facets Stereo Vision was generated under the assumption of aerial images with an image scale of $m_B = 1 : 4000$, a calibrated focal length of $c_k = 150mm$ and a picture overlap of $u = 60\%$. The images 1-3 are scanned windows from the aerial images with a pixel size of $30\mu m$, 256×256 pixels, 256 gray tones and a normal distributed noise with $\sigma_0 = 2$.

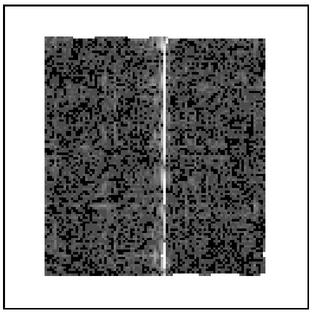


6a: Visualisation of the resulting surface

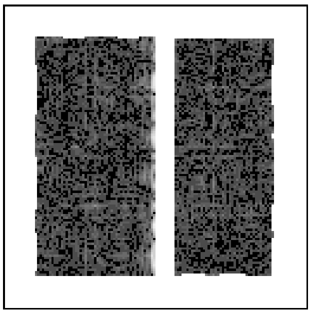


6b: Differences between the synthetic and the estimated heights

Remark on 6c,d: The gray value representations of the absolut values of the residuals are scaled in a logarithmic way. Small values, (black and dark gray tones), are visualized in a stretched way, whereas the larger values are visualized compressed with the light and white gray tones.



6c: Residuals of Image 1



6d: Residuals of Image 2

Fig. 6a-d: This example visualizes the possible effects of the selfdetection of occlusion. Two criterions for visibiliy are used: First, no part of a surface facet should be occluded by other parts of the surface. Second, a surface facet must be visible in a minimum number of image pixels.

it has to be pointed out here that the separation of breaklines and discontinuities is possible within the mathematical model of the surface, but it might be difficult to decide in reality.

The detection of a discontinuous area shall be illustrated by an example. The surface visualized in fig. 5a shall be reconstructed by the images 2 and 3, compare fig. 5b,d,e. Starting again with a global horizontal plane as approximation for the heights on the third level of the image pyramid, we get the smooth approximation of the surface shown in 7a after reaching convergence on the second pyramid level.

Large absolute values of the residuals mark the area around the discontinuity step, where the reconstructed surface does not fit well to the original surface in object space. They are represented in fig. 7c,d by light grey tones. The result represented in fig. 8a-d was calculated by simply omitting all the height facets from the estimation, which showed a mean absolute value of the residuals per facet higher than a certain limit, here e.g. $3\sigma_0$.

It shall be mentioned, that the simplest way to treat discontinuous areas and disturbances in the single images of course would be to omit the evaluation windows after negative result of a global analysis of the residuals of the whole window. But if we want to get as much information as possible about the course of the surface in object space from real image material, it will be necessary to replace the simple stategy of a treatment by height facet with a strategy based on a low level image processing of the residual images to get the discontinuous areas. The differentiation between local errors in one image and deficits of the surface model (= local errors in all images) has to be taken into account, too.

The results in fig. 8 show, that too much height facets were omitted by our simple criterion. One could try to overcome this problem by a simple test: The surface can be extrapolated from the facets at the border of the detected discontinuity. For this extrapolated surface facets the gray value residuals can be calculated without a new inversion of the whole normal equation system. On this foundation a decision can be made, whether the height facet should take part in the next iteration of the reconstruction process or not. By this way, the detected discontinuous area can be kept as small as possible.

It is obvious, that the discontinuity detection has to be done at every level of the image pyramid. On the highest level a small discontinuity step in object space won't disturb the smooth surface approximation - but it will disturb it on a lower level, when the resolution of the height facets grows. Starting from a global smooth approximation of the surface on the highest level, discontinuities can be found sucessively according to their order of magnitude. Of course, some texture within the surface gray values is in need, too. The standard deviations of the heights reflect the quality of the texture, but give no information in respect of model deficits, [Schlüter 93].

The detection procedure of occlusion as described in section 3.1 has no effects on this example because there is no occlusion. But if we let take part all three images represented in fig. 5c-e in the reconstruction, it can be seen that the detection of occlusion and by residuals work hand in hand as expected. The result is similar to the one shown in fig. 6.

4 HANDLING OF UNRECONSTRUCTABLE AREAS

In the last section we 'omitted' facets from the reconstruction process. Let us finally have a closer look at this omission.

We decided to omit those height facets from the estimation procedure, which cannot be determined because there is not enough information to determine them (e.g. they are only visible in one image), or for which our mathematical model of the surface is not flexible enough to give a sufficient represention.

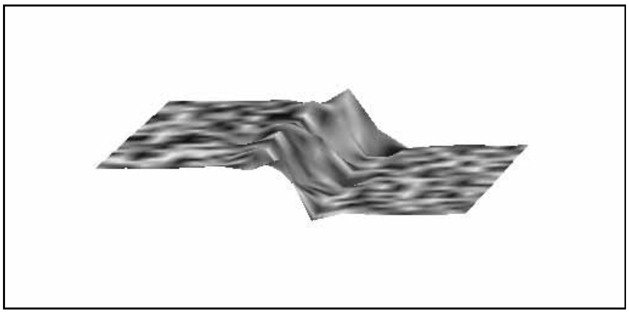
If we declare a height facet to be not reconstructable, we have to omit all the object gray value grid points within this height facet from the estimation procedure. They are not determinable, because we cannot assign a spatial position to them. In the same manner height grid points are not determinable, if all four height facets in the direct neighbourhood are declared to be not reconstructable. In the software implementation one can define these undeterminable unknowns by an additional fictitious observation equation to preserve the simple structure of the normal equation system.

Concerning the pseudo observations of the adaptive curvature minimization, it has to be guaranteed that there is no interpolation across undeterminable areas.

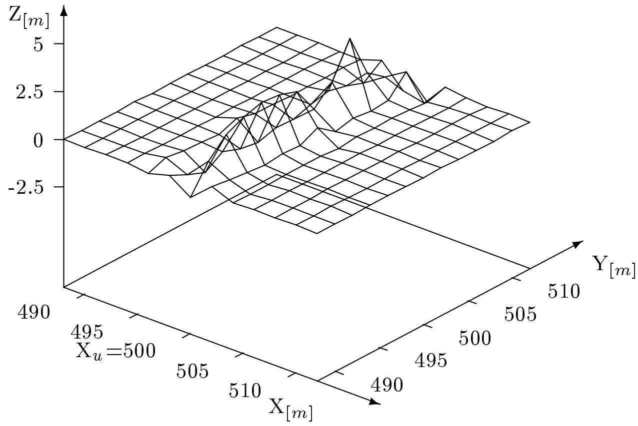
By this treatment the accuracy of the heights at the border of the discontinuous areas show the same behaviour as at the border of the evaluation windows: Accuracy in the corner point is less than accuracy in the edge point which in turn is less than accuracy within the evaluation window, see [Wrobel et al. 92]. This effect is visible in fig. 6b and 8b.

5 OUTLOOK

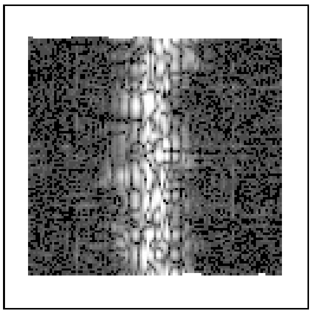
We presented concepts to extend the object space based image matching approach to a treatment of discontinuous surfaces and illustrated the main ideas by examples using synthetic images.

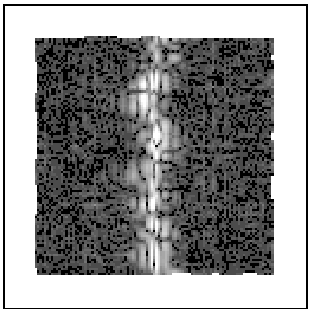


7a: Visualisation of the resulting surface



7b: Differences between the synthetic and the estimated heights

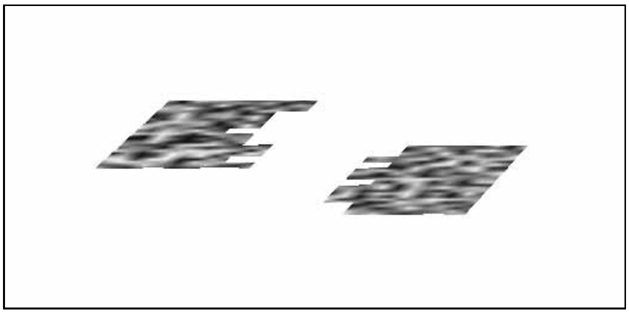




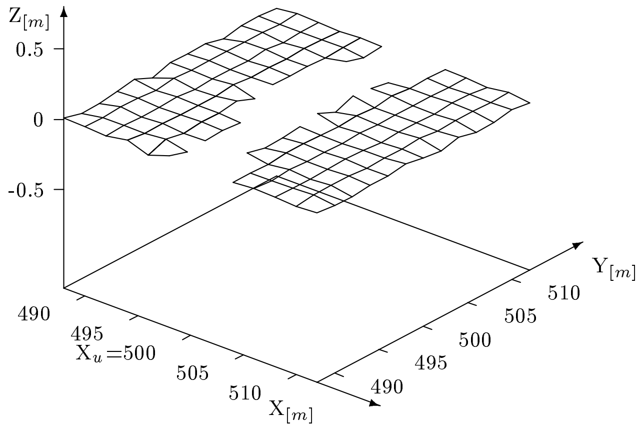
7c: Residuals of Image 2 **7d:** Residuals of Image 3

Remark on 7c,d: The gray value representations of the absolut values of the residuals are scaled in a logarithmic way. Small values, (black and dark gray tones), are visualized in a stretched way, whereas the larger values are visualized compressed with the light and white gray tones.

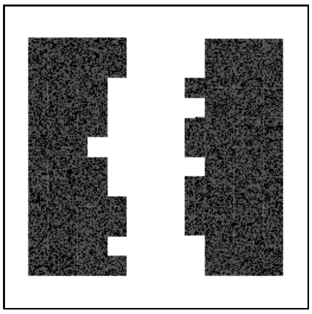
Fig. 7a-d: This example visualizes the reconstruction of a smooth approximation of a discontinous surface with Facets Stereo Vision and adaptive regularisation.

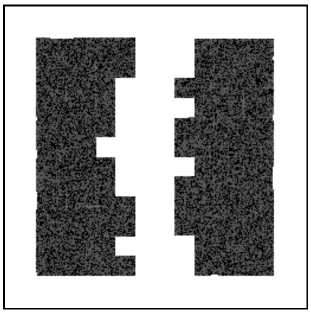


8a: Visualisation of the resulting surface



8b: Differences between the synthetic and the estimated heights





8c: Residuals of Image 2 **8d:** Residuals of Image 3

Remark on 8c,d: The gray value representations of the absolut values of the residuals are scaled in a logarithmic way. Small values, (black and dark gray tones), are visualized in a stretched way, whereas the larger values are visualized compressed with the light and white gray tones.

Fig. 8a-d: This example visualizes the results of a discontinuity detection on the basis of residual analysis.

The concepts have to be proved with real imagery. First tests with large scale aerial images (image scale 1:4000) are in preparation. These images should for example allow a separation of house-roofs and ground terrain and possibly vegetation.

On the basis of the experiences from real imagery the concepts should be refined especially concerning the quality of a regional residual analysis. Further on, it has to be investigated if the results from the residual analysis concerning the detection of discontinuities can be supported by the inclusion of information from surface parameters like surface curvature and gray value information like edges.

REFERENCES

Besl, P.J. 1988.
Surfaces in Range Image Understanding.
Springer, New York

Cochran, S.D., G. Medioni 1992.
3–D Surface Description from Binocular Stereo.
IEEE PAMI, Vol.14, No. 8, 981 – 994

Ebner, H., C. Heipke, M. Holm 1993.
Global image matching and surface reconstruction in object space using aerial images.
Integrating photogrammetric techniques with scene analysis and machine vision, Orlando, Florida, SPIE 1944, 44-57.

Encarnação, J., W. Straßer, 1986.
Computer Graphics.
Oldenbourg Verlag, München

Grimson, W.E.L., 1981.
From images to surfaces.
MIT Press

Heipke, C., 1989.
An Integral Approach To Digital Image Matching and Object Surface Reconstruction.
in: **Gruen, A., H. Kahmen (eds), 1989.**
Optical 3 (three) d measurement techniques.
Wichmann–Verlag, Karlsruhe, 1. Auflage

Helava, U.V., 1988.
Object Space Least Squares Correlation.
PhEng, Vol. 54, no. 6, 711 – 714

Hoff, W., N. Ahuja 1989.
Surfaces from Stereo: Integrating Feature Matching, Disparity Matching, and Contour Detection.
IEEE PAMI, Vol.11, No. 2, 121 – 136

Kaiser B., M. Schmolla, B. Wrobel 1992.
Application of Image Pyramid for Surface Reconstruction with FAST Vision (=Facets Stereo Vision).
IAPRS B 29, Comm. III, Washington, 341-345.

Kempa M. 1994
Hochaufgelöste Oberflächenbestimmung von Natursteinen mit dem Facetten-Stereosehen einschließlich Orientierung der Bilder.
Diss., TH Darmstadt (in preparation).

Liang, P., C.H. Taubes 1994.
Orientation-Based Differential Geometric Representations for Computer Vision Applications.
IEEE PAMI, vol. 16, no. 3, 249-258.

Schlüter, M. 1993.
Erweiterung des mathematischen Modells der Facettenmethode für unstetige Objektoberflächen.
unpublished report to DFG

Schmolla, M. 1991.
Bruchkanten und Unstetigkeiten bei der Objektrekonstruktion mit der Facettenmethode.
unpublished report to DFG

Schwarz, H.R. 1980.
Methode der finiten Elemente.
Teubner, Stuttgart

Terzopoulos, D., 1988.
The Computation of Visible-Surface Representations.
IEEE PAMI, Vol. 10, No. 4, 417 – 438

Weisensee M., 1992.
Algorithmen und Modelle zum Facetten-Stereosehen.
DGK C 374, München.

Weng, J., N. Ahuja, T.S. Huang 1992.
Matching Two Perspective Views.
IEEE PAMI, Vol.14, No. 8, 806 – 825

Wrobel B. 1987.
Digital Image Matching by Facets Using Object Space Models.
4th Int. Symp. on Optical and Optoelectr. Appl. Science and Engineering, The Hague, Netherlands, SPIE 804, 325-333.

Wrobel B., B. Kaiser, J. Hausladen, 1992.
Adaptive Regularization Of Surface Reconstruction By Image Inversion.
in Förstner W., Ruwiedel St.: Robust Computer Vision. Wichmann Verlag, Karlsruhe, 351-371.

Zheng, Y.J. 1993.
Inverse und schlecht gestellte Probleme bei der digitalen photogrammetrischen Objekt-Rekonstruktion.
DGK C 390, München.