Application of Image Pyramid for Surface Reconstruction with FAST Vision (=Facets Stereo Vision)

B.Kaiser, M.Schmolla, B.P.Wrobel
Institute of Photogrammetry and Cartography
Technical University Darmstadt
Petersenstr. 13
W-6100 Darmstadt
Federal Republic of Germany
Washington 1992, Comm. III

The performance of FAST Vision, which is an iterative method for the reconstruction of both object surface and orthophoto from two or more images, is determined by two measures: radius and rate of convergence. On one hand the radius of convergence should be as large as possible in order to make it possible to start FAST Vision with coarse approximate values for surface heights, which can be determined without much effort. On the other hand, a high rate of convergence is desired to reduce the time needed for reconstruction. Both aims can be achieved by the application of an image pyramid within the FAST Vision method. Experiments show, that by this way the rate of convergence can be increased by the factor of about 10. Furthermore, combining the image pyramid with the technique of object lifting delivers a method for obtaining very reliable starting values under very general circumstances.

Keywords: DTM, image matching, orthophoto, rectification, 3-D

1. Introduction

The non-linear basic equation of surface reconstruction by FAST Vision (Wrobel, 1987. Wrobel et al., 1992) has to be solved iteratively after carrying out a Taylor-linearization. The performance of iterative procedures is characterized by two measures:

a) radius of convergence, i.e. the maximum difference of the unknowns from their true values at the start of the iterative process.

b) rate of convergence, i.e. the number of iterations necessary to meet a specific break-off criterion.

The radius of convergence depends on the smallest wavelength λ_{min} of the grey value signal transmitted from object surface into image space (Korten et. al., 1988):

$$r \leq 0.5\ \lambda_{min}$$

The relation between pixel size Δp and wavelength λ_{min} can be obtained by Shannon's theorem: $\lambda_{min} \geq 2\Delta p$. If the image contains the lowest possible wavelength according to the sampling frequency, the radius of convergence will be only one pixel. Experience with real image material has shown, that this pessimistic assumption often proves to be true.

The rate of convergence decreases with the number of height facets in FAST Vision, provided the size of the facets remains constant. This fact makes the calculation of larger DTM's a time-consuming task. In order to reduce the number of necessary iterations, the start values for the unknowns have to be as close as possible to their real values.

On the other hand, the computation of start values for the heights should not be too complicated in order not to waste much time before the iterations can start. A simple method e.g. consists in the choice of equal values for all grid-points of the DTM to be reconstructed. This is a suitable possibility for small windows with a not too complicated surface geometry. In order to guarantee convergence for such a simple choice of start values, a larger radius of convergence may be necessary. This can be achieved by enlarging pixel size Δp - in other words: A coarser image resolution is chosen. In order to meet Shannon's theorem, a low-pass filtering has to take place before. Otherwise, aliasing could be the result.

The hierarchical ordering of digital images of successively lowered resolution is called image pyramid. In general, one starts at the resolution level of the original digitized image (level 0 of the image pyramid) and reduces the resoltution of the image in each of the two image coordinate system axes by the factor 2 (reduction of the image size to a quarter) after preceding low-pass filtering to get from level i to level (i+1) of the image pyramid. This procedure is continued, until the lowest required resolution (level n) is reached. A corresponding pyramidal structure is build up for the unknowns. Their number in each of the two directions of the object coordinate system is halved to get from level i to level (i+1). When applying the image pyramid to FAST Vision this means, that the number of height facets in each direction of the coordinate system has to be a potency of 2 on level 0 of the image pyramid (at least 2^{n-1}). After meeting the break-off criterion on level i of the image pyramid, the resulting height values will be prolonged by bilinear interpolation and be used as start values for level (i-1) of the image pyramid.

The basic theory of image pyramid in FAST Vision has been given in (Müller, 1990). In this paper we report an experimental application with that approach.

2. Parameters, Image Material and Results

Two sets of images were used for the FAST Vision experiments with and without the application of an image pyramid:

a) three aerial pictures (1:12000, longitudinal overlap 60%, focal length 153 mm) with low contrasts, taken in the Dankelshausen area in the vicinity of Dransfeld (near Göttingen, Lower Saxony, Germany), showing a crossing of two paths in agricultural surroundings (fields). Surface heights are between 208m and 213m.

b) two computer-generated pictures, rich of contrast (only noise due to truncation to integer grey values, 1:12000, longitudinal overlap 60%, focal length 153 mm) of an object similar to a saddle roof (two inclined planes meeting at a 'ridge'). One plane ('exposed to the sun') contains grey values from 128 to 255, the other one ('in the shade') contains grey values from 0 to 127. Surface heights are between 1003m and 1011m.

In both cases, the area to be reconstructed had a size of 40mx40m. The pictures had a size of 240x240 pixels (pixel size: 20µmx20µm in image space). As the pictures were taken from an altitude of 1800m and with a base of 1350m, this pixel size corresponds to a minimum radius of convergence in Z-direction of 0.3 m, which is very small. It increases to 0.6 m on level 1 of the image pyramid, to 1.2 m on level 2 and to 2.4 m on level 3. Break-off criterion was the maximum difference of heights between two iterations. Convergence is assumed to be reached, when the differences of heights in all grid-points falls below the break-off criterion, which was chosen to be twice as high for the grid-points on the borders of the window to be evaluated (in comparison with interior grid-points) and four times as high for the four corner points (quoted in the result tables: break-off criterion interior / edge / corner). The number of iterations for each level of the image pyramid is given for each experiment (e.g. 4/3/8 iterations, i.e. 4 iterations on level 2, 3 on level 1 and and 8 on level 0). The number of Z-facets (i.e. height facets) given in the parameters is that on level 0 of the image pyramid; the number of grey value facets per Z-facets remains constant for all levels. Regularization was achieved by minimizing curvature with a regularization parameter (regularization parameter 0 = no regularization). If not stated otherwise, the regularization parameter was chosen to be 1000 and the break-off criterion was 5cm/10cm/20cm.

3. Choice of Mask for Low-Pass Filtering

There is no mask of finite dimension realizing ideal low-pass filtering. Therefore, the choice of such a mask has to be a compromise between approximation of ideal low-pass filtering and computational efficiency. A good approximation is characterized by almost complete elimination of higher signal frequencies while the lower ones have to be almost completely preserved. There are many suggestions for masks in literature reaching from a simple 2x2 mean value filter (Li, 1989 / Weisensee, 1991) to a (e.g.) 13x13 mask (Meer et. al., 1987). The latter one, which really is a 7x7 mask due to zero columns and rows, was used for low-pass filtering in the experiments in this paper. Fig.3.1 shows a comparison of this mask with two binomial masks (Jähne, 1989).

filter mask	no. of iterations	computation time for filtering (relative)	kernel no.
3×3	8/6/5/9*	1	1
5×5	10/5/5/10	2.78	2
13×13	15/7/5/9	5.44	3

*8/6/5/9= 8 iterations on level 3 / 6 on level 2 5 on level 1 /9 on level 0

$kernel=a \cdot a^T$	kernel no.
$a=0.25 \cdot (1,2,1)$	1
$a=0.0625 \cdot (1,4,6,4,1)$	2
$a=10^{-3} \cdot (0,51,0,-87,0,298,475,298,0,-87,0,51,0)$	3

Fig. 3.1: Comparison of number of iterations and computation time for different filter masks (image material: Dransfeld images, 4 pyramid levels, 8×8 Z-facets)

A low number of iterations on level 0 of the image pyramid is desired for the reconstruction of larger DTMs. Fig. 3.1 shows, that this number is the lowest for the 13x13 mask, although it can be as low for a simpler binomial mask (v. 3x3 mask). But as the computaional cost for filtering is low compared with that of an additional iteration on level 0, the 13x13 mask seems to be a good choice. Larger masks do not seem to be recommendable because of the large areas on the edges of the images, which cannot be filtered by them.

4. Comparison of the Experimental Results with and without Image Pyramid

It is necessary for the application of the image pyramid within FAST Vision, that the obtained results are of the same quality as those obtained without the application of an image pyramid. In order to be able to compare the results, we chose computer-generated images, because the exact heights of the object to be reconstructed are known. For the purpose of comparison a window of 8x8 Z-facets was reconstructed with and without image pyramid. The start value for all heights was 1007m, see section 2b). There was no regularization, no radiometric transfer function and the break-off criterion was 1mm/2 mm/4mm. In both cases the root mean square of differences between reconstruction results and real surface was 13 mm (i.e. 0.07°/₀₀ of flying altitude).

In this special experiment the position of the Z-facets was chosen in a way, in which there was a coincidence between the ridge and the border of the Z-facets (resulting in 8x4 facets situated on each half of the 'roof'), so that the reconstructed heights in each grid-point could be exactly compared with the heights on the real oject. The mean differences of the reconstructed heights from the real values on the object are equal with and without the application of the image pyramid. Thus, the quality of reonstruction is the same in both cases.

5. Advantages of the Application of the Image Pyramid within FAST Vision with Respect to Radius and Rate of Convergence

The advantages of the application of the image pyramid within FAST Vision have to be investigated. For this pupose, a DTM and an orthophoto in the Dransfeld area was reconstructed. Several experiments were carried out with and without the application of an image pyramid. There were also different regularization parameters and radiometric transfer functions chosen. Fig. 5.1 shows the number of iterations and the relative time for computation, until the break-off criterion was met.

regularization parameter λ		radiometric transfer funct		
λ=0	with pyramid	constant	8/4/8**	1.8*
	no pyramid	constant	wrong	
	with pyramid	linear	7/5/7	2.3
	no pyramid	linear	wrong	
λ=1000	with pyramid	constant	8/4/4	1.6
	no pyramid	constant	67	15.1
	with pyramid	linear	7/3/4	1.0
	no pyramid	linear	55	12.4
λ=8000	with pyramid	constant	8/4/2	1.0***
	no pyramid	constant	44	9.9
	with pyramid	linear	7/3/4	1.5
	no pyramid	linear	45	10.2
	wrong=convergence towards wrong heights			
	**8/4/8 = 8 iterations on level 2 / 4 on level 1 / 8 on level 0			
	+1.8 = computation time relative to ***			

Fig.5.1: Comparison of FAST Vision with and without image pyramid: no. of iterations & CPU-times (image material: Dransfeld images, 3 pyramid levels, 4×4 Z-facets)

As the start value for the height of 205 m was lower then the real heights in the area (bad start value), the application of the image pyramid resulted in an advantage of a factor of up to 10 in terms of computational speed. Fig. 5.1 also shows, that convergence can be reached by applying the image pyramid in cases, in which it is not reached when not applied (v. first column of fig. 5.1.: FAST Vision converges to wrong height values without application of image pyramid).

6. Choice of Break-Off Criterion on Higher Levels of the Image Pyramid

Regarding the number of iterations on the higher levels of the image pyamid (v. fig. 3.1), one realizes that it is generally greater on the higher levels of the image pyramid, which indicates that one does not have to choose the same break-off criterion on the higher level as on level 0. This becomes especially clear when realizing, that the resulting heights on the higher levels are used as mere start values for the next lower ones. In order to investigate, if it is necessary to choose the same break-off criterion on each level of the image pyramid, the following experiment was carried out with two methods of choosing the break-off criterion on each level:

a) The break-off criterion on each level is the same,

b) the break-off criterion is doubled from one level to the next higher one.

Doubling of the break-off criterion in b) has been chosen because of the proportionality of height errors with x-parallax errors, which are proportional to pixel size.

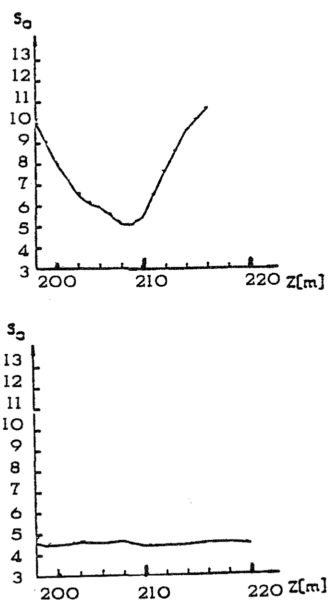
Fig. 6.1 shows the number of iterations on each level of the image pyramid. They are equal for level 0 for a) and b), but the values are lower for b) on the higher levels. Therefore, the start values obtained from the higher levels are exact enough even if the break-off criterion is doubled from one level to the next higher one. On the other hand, as the time needed for an iteration on a higher level of the image pyramid is much lower than that on level 0, the added advantage in terms of compuational time amounts to only 1% in case b)

case	iterations (level 3/2/1/0)
a)	8/5/4/4
b)	3/3/3/4

Fig. 6.1: No. of iterations for equal break-off criterion on each level of image pyramid (case a) and for break-off criterion doubled on each higher level (case b)

7. Object lifting and Image Pyramid

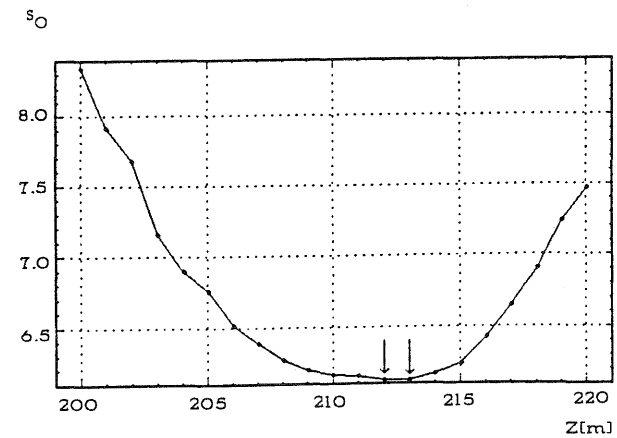
Object lifting (Kaldun 1988, Müller 1990) is a procedure, which delivers start values for the heights of one Z-facet. The start values are varied and one iteration is carried out for each start value. Very often, there will be a minimum standard deviation of unit weight s_o with one of the chosen start values and a change in the sign of the mean height differences after this single iteration. This value is a suitable start value for FAST Vision. Object lifting only works, if there is such a minimum standard deviation of unit weight and such a change of sign, which is not the case in areas of low grey value contrasts (v. fig.7.1)



equal start values Z for the one Z-Facet on original resolution

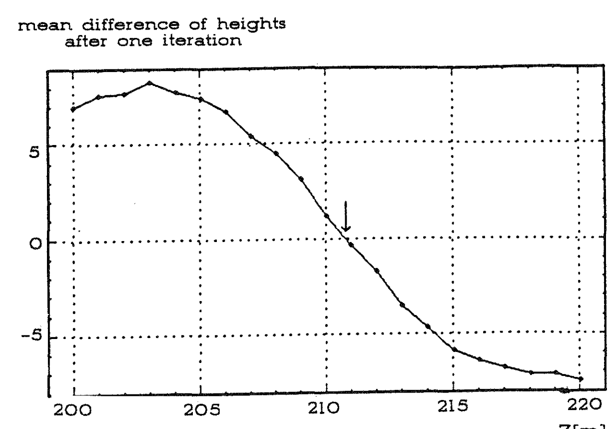
Fig. 7.1 s_o just after one iteration for different start values of heights (above: good contrasts in the images, below: bad contrasts)

This difficulty should not occur, if object lifting is applied in combination with the image pyramid. Here, object lifting is used only to determine the start values for the heights on the highest level of the pyramid. If there is no minimum of the standard deviation of unit weight and no change of sign of the mean height differences even on this level, it means that the images do not contain coarse textures. In that case, simple matching of the images would be difficult and of course the application of FAST Vision, too. Fig. 7.2 shows the results of object lifting for one Z-facet on the highest level (level 2) of the image pyramid of the three Dransfeld pictures. There is a minimum for the standard deviation of unit weight, although its situation cannot be located very clearly (v. fig. 7.2a). There is also a distinct change in the sign of the mean height differences after that one iteration. A good choice for the start value of that facet seems to lie between 211 m and 213 m. As it can be seen in fig. 7.3, there is only a difference in the number of iterations on level 2 of the image pyramid for the choice of 211 m, 212 m and 213 m. Thus, the conclusion can be drawn, that a choice of a start value in the vicinity of the minimum of the standard deviation of unit weight is sufficient for convergence of FAST Vision.



equal start values Z for the one Z-Facet on level 2 of the image pyramid

Fig. 7.2a: s_o dependent on start values Z (one iteration carried out for the one facet on level 2 of the image pyramid)



equal start values Z for the one Z-Facet on level 2 of the image pyramid

Fig. 7.2b: mean difference of height after one iteration dependent on start value Z (one iteration carried out for the one facet on level 2 of the image pyramid)

st.height	iterations	st.height	iterations
200	8/6/3*	207	6/6/3
201	8/6/3	208	6/5/3
202	8/6/3	209	6/5/3
203	8/6/3	210	6/5/3
204	8/6/3	211	5/5/3
205	7/6/3	212	7/5/3
206	7/6/3	213	7/5/3
st.height	**iterations**		
214	8/5/3		
215	9/5/3		
216	7/5/3		
217	8/5/3		
218	6/5/3		
219	6/5/3		
220	13/5/3		

*8/6/3 = 8 iterations on level 2 / 6 on level 1 / 3 on level 0

Fig.7.3: No. of iterations for different start values for the four heights (st.height) of the one facet on level 2

8. Conclusion

The application of an image pyramid and a corresponding hierarchical structure for the unknowns within FAST Vision results in improved radius and rate of convergence. Most of the approximation of the real height values takes place on the higher levels of the image pyramid, where the computational time needed for one iteration is lower due to the significantly smaller number of variables. Start values for the highest level of the image pyramid can be obtained by applying object lifting on that level.
FAST Vision, image pyramid and object lifting can be combined in the following manner in order to get an object reconstruction, when an interval of possible start values for heights is given.

Construction of an image pyramid by succesive reduction of image resolution after low-pass filtering
Choice of an interval of possible start values for heights and determination of the optimal one by the application of object lifting on the highest level of pyramid (if there is no significant change of sign of the mean differences of heights after one iteration object lifting has to take place on the next lower level of the image pyramid)
For each level of the image pyramid from level n to level 0
Application of FAST Vision on that level until break-off criterion is met
Prolongation of heights to the next lower level of the image pyramid

Fig.8.1 A method for getting start values for heights for FAST Vision under global circumstances

Literature

Jähne, B.: Digitale Bildverarbeitung. Springer Verlag Berlin Heidelberg New York London Paris Tokyo, 1989

Kaldun, M.: Entwurf und Erprobung eines automatischen Startverfahrens mit kleinen Fenstern zum Start von FAST Vision. Unveröffentlichte Diplomarbeit am Institut für Photogrammetrie, TH Darmstadt, 1988

Korten, Th. /Wrobel, B. / Franek, M. / Weisensee, M.: Experiments with Facets Stereo Vision for Object Surface Reconstruction. ISPRS Congress, Comm. III, Kyoto, 1988

Li, M.X.: Hierarchical Multi-Point Matching with Simultameous Detection and Location of Breaklines. The Royal Institute of Technology, Department of Photogrammetry, Stockholm, 1989

Meer, P. / Baugher, E.S. / Rosenfeld, A.: Frequency Domain Analysis and Synthesis of Image Generating Kernels. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No. 4, July 1987

Müller, J.: Möglichkeiten der Bestimmung von Näherungswerten für die Objektflächenrekonstruktion mit FAST Vision. Close-Range Photogrammetry Meets Machine Vision, A. Gruen, E. Baltsavias (eds.), Proc. SPIE 1395, p. 1082-1091, 1990

Weisensee, M.: Modelle und Algorithmen für das Facetten-Stereosehen. Deutsche Geodätische Komission, Reihe C, Nr. 374, München, 1992

Wrobel, B.: Digital Image Matching by Facets Using Object Space Models. 4th International Symposium on Optical and Optoelectr. Appl. Science and Engineering, March 30th - April 3rd 1987, The Hague, Netherlands, SPIE 804, p. 325-333

Wrobel, B. / Kaiser, B. / Hausladen, J.: Adaptive Regularization - A New Method for Stabilization of Surface Reconstruction From Images. Presented Paper, XVIIth Congress of ISPRS, Comm. III, Washington, 1992

These investigations are supported by Deutsche Forschungsgemeinschaft.