EXPLOITING THE MULTI VIEW GEOMETRY FOR AUTOMATIC SURFACES RECONSTRUCTION USING FEATURE BASED MATCHING IN MULTI MEDIA PHOTOGRAMMETRY

Kirsten Wolff and Wolfgang Förstner
Institute for Photogrammetry, Bonn University
Nussallee 15, D-53115 Bonn, Germany
Kirsten.Wolff@ipb.uni-bonn.de wf@ipb.uni-bonn.de

KEY WORDS: Matching Algorithm, Multi View Geometry, Multi Media Photogrammetry, Surface Reconstruction

ABSTRACT

In this paper we present a new method of a feature based matching algorithm for a 3D surface reconstruction exploiting the multiview geometry. The matching algorithm conceptually allows parallel processing treating all images equally. Especially the geometry of the image triplet is used, namely the trilinear relations between image features using the trifocal tensor. The method is transferred to multi media photogrammetry. The determination of the 3D point uses a direct method minimizing the algebraic error.

KURZFASSUNG

Dieser Artikel stellt einen neuen Matchingalgorithmus für die 3D Oberflächenrekonstruktion mit mehreren Bildern vor. Der Matchingalgorithmus erlaubt eine parallele Bearbeitung aller Bilder. Für die Zuordnung homologer Punkte wird die geometrische Beziehung zwischen einem Bildtripel in Form des trifocalen Tensors verwendet. Der Matchingalgorithmus wird auf die Mehrmediengeometrie angewendet. Die anschliessende Bestimmung der 3D Objektpunkte wird durch eine Methode zur Minimierung des algebraischen Fehlers durchgeführt.

1 INTRODUCTION

We present a new matching algorithm for feature based 3D surface reconstruction. It exploits the multi view geometry and includes the option of treating images from multi media photogrammetry. Especially the geometry of the image triplet is used, namely the trilinear relations between image features using the trifocal tensor (Hartley, 1995).

Our work is motivated by investigations on the generation of fluvial sediments. The goal of the interdisciplinary project *Geometric Reconstruction, Modeling and Simulation of Fluvial Sedimental Transport* in the Special Research Centre (Sonderforschungsbereich) SFB 350 *Continental Mass Exchange and its Modeling* is to derive a physical model of the generation process of sediments under water.

Two different approaches have have been published recently:

- One class of approaches aims at modeling the behavior of turbulent flow. They are used to model the process of sedimental transport. The results of these approaches, however, do not provide precise models for the sedimentation process itself, because they cannot directly include observations of the sedimentation process or the sedimental surface.
- Other approaches aim to investigate of fossile fluvial sedimentary structures. They observe the sedimental surface but they cannot observe the dynamical process of the sediment transport (Valdivia-Manchego, 1996).

To our knowledge observations of changes of the fluvial sediment surface have not explicitely been used for modeling. However, without an observation of the dynamical process no precise reconstruction and analysis appear to be possible. Only observational data filling the space time domain of the sedimentation process seeem to allow its adequate modeling.

Digital photogrammetry offers sufficient spatial and temporal resolution of the evolution of the sediment surface which is very useful for continuously monitoring essential parts of the sedimantation process.

The main features of the surface are (cf. Fig. 1)

1. The crestlines of moving ripples.

   They form typical patterns which evolve over time. Their spatial density and their form may be used to characterize the pattern which may be found in fossile sedimental structures.

2. Smooth surface structures between the ripples.

   On the one hand this allows simple interprolation, possibly taking the sedimentation process into account. On the other hand the surface shows enough texture in the images allowing reconstruction without specific texture projection.

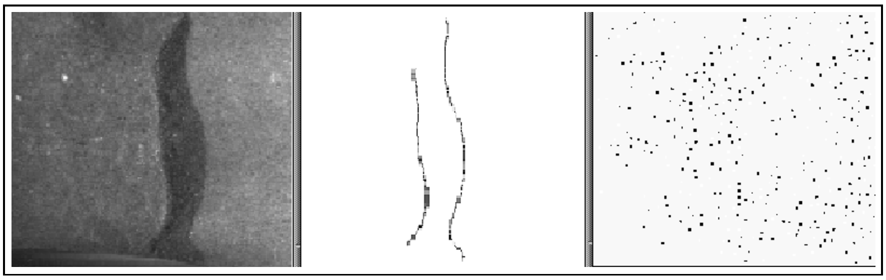

Figure 1: a) original image of a rippel b) extracted crestline (left line) and shadowline (right line) c) extracted points

In the following we discuss a matching algorithm for a reconstruction of the sedimental surface whithout taking the crestlines into account. Its essential part is a new matching algorithm which exploits the multi view geometry and is able to handle the multi media geometry in case of in situ measurements. We show the capability of the matching and reconstruction algorithm on a set of image points of the surface points.

2 THE ACQUISITION SYSTEM

Our image acquisition system is based on four CCD-cameras which are mounted above a hydro mechanic laboratory channel. The sediment in the channel can be observed without water in the channel or with water running over the sediment while changing its structure.

In the second case a perspex-sheet between the cameras and the sediment eliminates the waves of the flow in its area. This constellation leads to the standard case of multi media photogrammetry: the cameras are positioned in air, the observation process takes place in water and a plane-parallel perplex-sheet divides these two medias. For the experimental system see Figure 2.

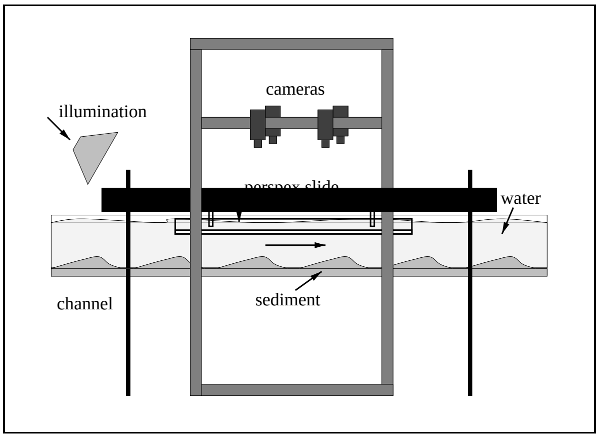

Figure 2: The aquisition system

3 POINT MATCHING FOR MULTIPLE VIEWS

We developped a new matching procedure for surface reconstruction. It had to fulfill the following requirements:

- the intrinsic and extrinsic parameters of the cameras are assumed to be known, e. g. by specifying the projection matrices of all images.

- for efficiency reasons features should be used as basic primitives of the matching algorithm.
- it should be able to handle an arbitrary number (≥ 2) of views.
- the result should not depend on the sequence or the numbering of the images. This conceptually may allow parallel computation.
- the geometry of the multiple view should be exploited for increasing performance, especially completeness and reliability.
- handling multi media geometry should be an option.

The first requirements have been already realized in several matching algorithms (e. g. (Schmid and Zisserman, 1997)). They mainly use interest lines and exploit the geometry of the image triplet using the trifocal tensor of three views. We want to follow a similar line, but include the option to handle the case of multi media geometry.

The structure of our algorithm is as follows:

1. extraction of interest points
2. determination of matching candidate tuples
3. enforcing consistency between the matching candidates
4. 3D point determination

3.1 Extraction of interest points

For the extraction of texture points in the images we use the feature extraction algorithm of Fuchs et. al. (Förstner, 1994, Fuchs, 1998).

Figure 3 shows the extracted image points in parts of four images. The four images on the left side are taken through the perspex-sheet and the running water. The four images on the right side show the surface without ray-refraction. The image quality is nearly the same. The point selection obviously is quite reliable and promises low confusion. As to be expected not all points appear in all images.

The greyvalues in the neigborhood of the image points are very similar for all points. Thus a greyvalue based correlation of these points would give nearly no information on the matching. This is the reason why we heavily rely on the crisp constraints of the multi view geometry.

3.2 Determination of matching candidate tuples

3.2.1 Prediction of points with the trifocal tensor The trifocal tensor T_{ijk} (Hartley, 1995) describes the geometry of the relative orientation of three views. It may be used for either checking consistency between matching candidates, points or lines, or for predicting points or lines into a third image, given two homologeous points in two images

The prediction of a point with the trifocal tensor is equivalent to the intersection of the two epipolar lines in the third image. But, analogically to the epipolar line, it allows a direct prediction without reconstructing the 3D point first. In case of collinear projection centres the prediction using the intersection of epipolar lines does not work, as the three pairwise epipolar planes are identical. However, the prediction with the trifocal tensor is possible also in this case.

The elements of the tensor can be determined from the elements of three projection matrices P_1,P_2 and P_3 for the projection of a 3D point $P(\mathbf{X})$ onto the three image plane (Förstner, 2000). This projection determines the image point $P'(\mathbf{x}')$ using direct the linear transformation.

$$\mathbf{x}' = \mathsf{P}\mathbf{X} \qquad \text{with} \qquad \mathsf{P} = \mathsf{K}\boldsymbol{R}(\boldsymbol{I}| - \boldsymbol{X}_o)$$

where $(\cdot|\cdot)$ denotes concatenation. The 3×4 projection matrix P can be explicitly related to the 6 parameters of the exterior orientation and 5 parameters of the interior orientation namely the Euclidean coordinates $\boldsymbol{X}_o$ of the projection center $O(\boldsymbol{X}_o)$, the rotation matrix $\boldsymbol{R}$, the principle distance c, the coordinates (x'_H, y'_H) of the principle point, the shear

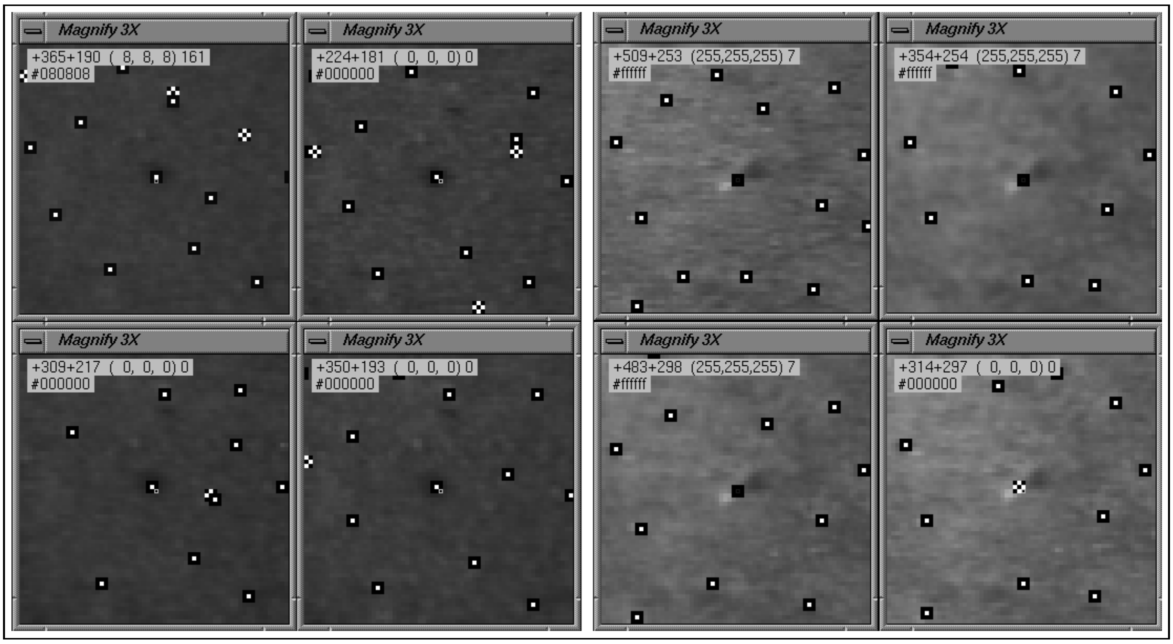

Figure 3: extracted points of four different views : the four left images show multi media images, the four right images show one media images

s and the scale difference of the x'- and the y'-coordinates. The parameters of the interior orientation are collected in the 3×3 calibration matrix

$$\boldsymbol{K} \doteq \begin{pmatrix} c & cs & x'_H \\ 0 & c(1+m) & y'_H \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

In case the projection matrices of three images are

$$\mathbf{P}_1 = (\boldsymbol{I}|\mathbf{0}) \qquad \mathbf{P}_2 = (\mathbf{R}|\mathbf{r}) \qquad \mathbf{P}_3 = (\mathbf{S}|\mathbf{s})$$

The trifocal tensor is defined as

$$T_{ijk} = R_{ij}s_k - r_jS_{ki}$$

For arbitrary $\mathbf{P}_1 = (\mathbf{A}|\mathbf{a})$ which can be achieved by transforming the 3D coordinates by

$$\overline{\mathbf{X}} = \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{X}$$

The prediction of the third image point from the image coordinates of the points in the first and second image are given by (Hartley, 1995)

$$x'''_l = \sum_{k=1}^{3} x'_k (x''_i T_{kjl} - x''_j T_{kil}), \qquad i, j = 1, 2, 3 \tag{2}$$

Prediction with the trifocal tensor is much more efficient than going via 3D space, especially if we normalize the images such that the rotation matrices are $\boldsymbol{R}_i = \boldsymbol{I}$ (cf. (Mikhail, 1963)).

3.2.2 Point prediction and multimedia photogrammetry In the case of multimedia photogrammetry the imaging process cannot be represented as a projective mapping, as it does not map straight lines to straight lines. This would prevent to exploit projective geometry. But we may partition the object space and in a corresponding way the image space such that for every part an approximated trifocal tensor is determinable. This significantly increases efficiency of prediction.

The necessary multimedia geometric models for the projection from the object space into the image space are described in (Maas, 1995). He used a strict multimedia geometric model based on Snell's Law for the effect of a ray being twice broken due to different refractive indices on the optical path through water, glass and air. Knowing the camera calibration parameters, the interface between the different optical media and the refractive indices the effect can be modeled strictly. The solution for the spatial intersection of two image points is described in (A. Okamoto, 1972).

3.2.3 Finding matching candidates The algorithm for finding matching candidates assumes $n > 2$ images to be given with their projection matrices and the extracted points. It is divided into the following steps.

1. For all points $P'_{i'j}$ in a starting image $j \in \mathcal{N}$ determine the epipolar lines $l''_{i'jk}$ in a second image $k \in \mathcal{N}$, where $\mathcal{N}$ denotes the set of all image numbers and i' the number of all points in the starting image. .
2. Find a set $\{P''_{i''k}\}$ of point hypotheses in the second image k which lie close to the epipolar line $l''_{i'jk}$
3. Predict these point pairs $(P'_{i'j}, P''_{i''k})$ point into all other images $m \notin (j,k)$ yielding the set $\{\tilde{P}'''_{i'''m}\}$.
4. Find the sets $\{P'''_{i'''m}\}$ of N homologeous points lying close to the predicted points $\tilde{P}'''_{i'''m}$.

Every detected combination of sets of homologeous points $\{P'''_{i'''m}\}$, $(i''' = 1..n,\ m = 1....N,\ m \notin (j,k))$ belonging to the point pair $P'_{i'j}$ and $\{P''_{i''k}\}$ is combined in one *point group* (see example below).

For the prediction of homologous points we use the above mentioned geometric relations between three images. The third image is actually used to verify the point hypotheses of the first and second image. It is possible that the homologous image point of an object point is only detected in two images. In this case the verification is not possible. However, this correspondence might be detected starting with the second or the third image, as the prediction is not symmetric with respect to the image numbering.

For the case of four images we have the situation sketched in figure 4). For the selection of the starting image there are four different possibilities, for the second image three possibilities and for the prediction in a third image exist two possibilities. Alltogether we get twelve different combinations for the *matching order* of four images.

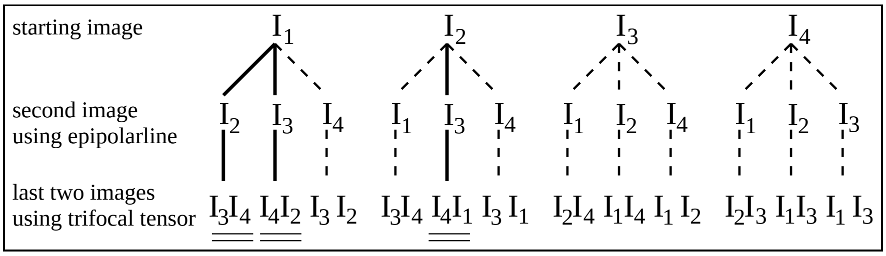

Figure 4: matching orders for point matching

To get all points which exist in three or four images not all combinations need to be calculated. For the selection of the necessary combinations of the four images we look at the possible extraction of homologous points: an image point of an object point can be extracted in

1. two images (not sufficient)
2. three images (possible combination: $(I_1, I_2, I_3)(I_1, I_2, I_4)(I_1, I_3, I_4)(I_2, I_3, I_4)$
3. four images (I_1, I_2, I_3, I_4)

To get all possible combinations of case 2 and case 3 it is sufficient to regard only the three solid drawn cases in figure 4. If the orientation parameters are not precise, it could happen that a group of homologous points cannot be detected using only these three matching orders but it is detected using another matching order. Or it is possible that an ambiguity which has to be examined is only detected by using another combination.

Example we want to find the homologeous points in four images for one matching order. For the point 2 P'_{43} of the starting image (image j=3) we find the point hypotheses $\{P''_{21}\}$ in the second image (image k=1). Predicting a point for this point pair P'_{43} $\{P''_{21}\}$ in the third (image o=2) and fourth image (image p=4) we get the set of homologeous points $\{P'''_{92}\}$, $\{P'''_{82}\}$ and $\{P'''_{74}\}$. Now we combine all combinations to different point groups, in the table below every row represents one point group : group1 = (P'_{43},$\{P''_{21}\}$,$\{P'''_{92}\}$,$\{P'''_{74}\}$) and group2 = (P'_{43},$\{P''_{21}\}$,$\{P'''_{82}\}$,$\{P'''_{74}\}$)

The result of the matching process is a list of all detected point groups (*matching list*) of homologous points. The matching list contains for every point group and for every image wether a point is detected or not and the detected point numbers (cf. table in 3.3).

image	group1	group2	group3	group4	group5	group6
1	2*	2*	1*	5	5	1*
2	7	7	7	-	5	7
3	4	4	4	1	-	4
4	9*	8*	8*	2	2	9*

3.3 Enforcing consistency between the matching candidates

After the initial matching process the matching list has to be checked for consistency, as a point may occur in more than one point group. All contradictions which belong together are detected as one *contradiction group*. The example in figure 3.3 contains one contradiction group which consists of the point group 1,2,3 and 6 (the contradictions are marked (*)). The point group 4 and 5 are add together as one group - they contain no contradiction and represent the same object point.

For every contradiction group which is detected in the matchinglist one solution - if possible - has to be chosen. A robust adjustment calculates exactly one solution for one group. Another possible approach is not to find one solution for all point of a contradiction group but to determine a accuracy of every detected combination and to use this accuracy for the selection of one combination out of all.

3.4 3D point determination

For every point group with no contradiction of the matching list a spatial forward intersection must be calculated. There exist different methods for the intersection of three or four warped space rays. We use a linear least square method to get the numerical least squares solution for the problem.

We represent the space lines using the homogeneous coordinates of two line points $\mathbf{X}_1 = (X_1, Y_1, Z_1, 1)^{\mathsf{T}}$ and $\mathbf{X}_2 = (X_2, Y_2, Z_2, 1)^{\mathsf{T}}$. The two points lead to the line parameters (Plücker coordinates)

$$L = \begin{pmatrix} X_2 - X_1 \\ Y_2 - Y_1 \\ Z_2 - Z_1 \\ Y_1 Z_2 - Y_2 Z_1 \\ Z_1 X_2 - Z_2 X_1 \\ X_1 Y_2 - X_2 Y_1 \end{pmatrix}$$

and a matrix $\mathsf{B}(\mathbf{X})$ depending on the parameters of the line (cf. (Förstner, 2000)).

$$\mathsf{B}(\mathbf{L}) = \begin{pmatrix} 0 & L_3 & -L_2 & -L_4 \\ -L_3 & 0 & L_1 & -L_5 \\ L_2 & -L_1 & 0 & -L_6 \\ L_4 & L_5 & L_6 & 0 \end{pmatrix}$$

A point $\mathbf{X}$ lies on a line $\mathbf{L}$ when $\mathsf{B}(\mathbf{L})\mathbf{X} = \mathbf{0}$.

The point closest to four non-intersecting lines minimizes the algebraic error

$$\mathsf{G}\mathbf{X} = \begin{pmatrix} \mathsf{B}(\mathbf{L}_1) \\ \mathsf{B}(\mathbf{L}_2) \\ \mathsf{B}(\mathbf{L}_3) \\ \mathsf{B}(\mathbf{L}_4) \end{pmatrix} \mathbf{X} = \mathbf{w}$$

Minimizing $\mathbf{w}^{\mathsf{T}}\mathbf{w}$ under the constraint $\mathbf{X}^{\mathsf{T}}\mathbf{X} = 1$ leads to minimizing the ratio (Hartley, 1998)

$$\lambda = \frac{\mathbf{X}^{\mathsf{T}}\mathsf{G}^{\mathsf{T}}\mathsf{G}\mathbf{X}}{\mathbf{X}^{\mathsf{T}}\mathbf{X}}$$

Therefore the best point $\mathbf{X}$ ist the eigenvector of $\mathsf{G}^{\mathsf{T}}\mathsf{G}$ to the smallest eigenvalue. This simple solution is not the optimal in a statistical sense in presence of noise, but experiments show that it is a very good approximation. The accuracy is often sufficient, otherwise a maximum likelihood estimate could be performed based on this approximation.

4 RESULTS

4.1 Extraction of interest points

Figure 5 shows the results of the point extraction for one multi media image. The number of extracted points and their spatial point distribution is sufficient for the multi media case to represent the sediment surface. In figure 3 we have seen, that homologous points exist in all four images.

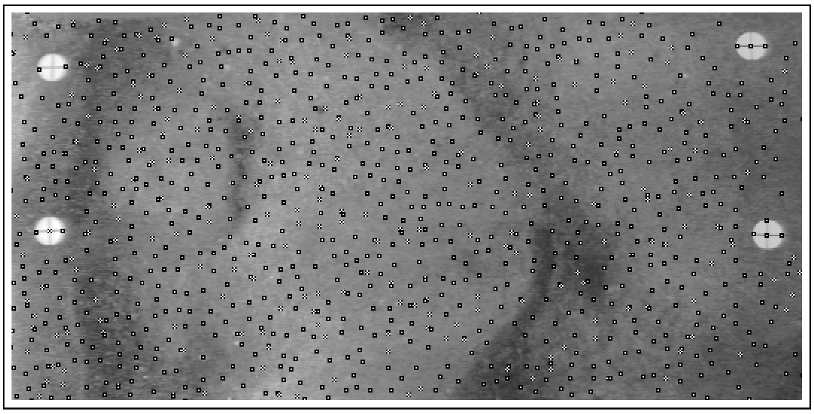

Figure 5: Result of point extraction with multimedia geometry

4.1.1 Determination of matching candidate tuples For a test of the matching algorithm we focussed on the image points of six objectpoints. The objectpoints are well distributed in the objectspace and one of them is a control point (cf. Fig 6). An imagepoint of point 1 was not extracted in image 1 and another point was extracted in image 2 near by the image point of point 4 (this might be a contradiction).

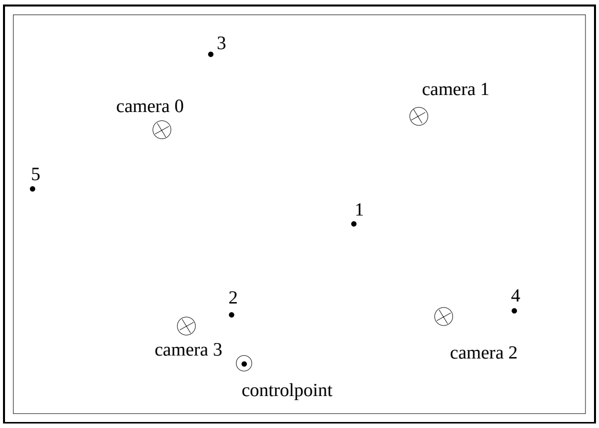

Figure 6: Objectpoints and Camerapositions projected onto the X/Y plane of the object system

To observe the behavior of the matching algorithm all twelve matching orders mentioned in section 3.2.3 are realised. Thus for one objectpoint maximally twelve point groups could be found. The number of point groups belonging to one objectpoint is different for the different points. Here are some results:

- It was to be expected, that in twelve point groups for the controlpoint the right homologous points were detected for all images.
- The homologous points of point 3 for were found only five times in four images and 4 times in three images.
- It is conspicuous that the points 1,5 and 3 were not detected in the image 2. We suggest that the reason for this is a non precise orientation of the camera.
- The point 1 which was not detected in image 1 was only 4 times detected in the other three images.

Alltogether the results show, that the result of matching points using multi views depend on the matching order of the images and that it is necessary to calculate more than one matching order not to loose any point group.

4.1.2 Enforcing consistency between the matching candidates The matching algorithm has detected the contradiction of the two points in image 2.

4.2 3D point determination

The algoritm for the 3D determination was tested by reconstruction the control point which was detected for all twelve matching orders. The error of the reconstructed control point was for all three object coordinates smaller that 0.1 mm with a observing distance of 300 mm.

5 CONCLUSION

We have seen that the matching and reconstruction algorithm for a 3D point of the fluvial sediment surface using multi views and multi media moduls is usable. The necessity for the use of more than one matching order is an important point of the matching algorithm. The next step is to test the algorithm for the whole surface point set of the fluvial sediment. This will not be a problem if the orientation parameters of the cameras and the refraction indices of air, perspex and water are precise enough.

With the reconstructed object points and the reconstruction of the crestlines of the sediment ripples The surface can be represented by this features together with an interpolation between these features.

Having reconstructed the surface for different points in time the dynamical changes of the fluvial sediment surface could be detected and analyzed. Therefore the movement of the 3D crestlines have to be tracked through a surface sequenze. The following determination of the parameter of the model for the sedimentation transport will be implemented by collaboration within the Special Research Centre SFB 350 'Continental Mass Exchange and its Modeling'.

REFERENCES

A. Okamoto, J. H., 1972. Allgemeines analytisches orientierungsverfahren in der zwei- und mehrmedien-photogrammetrie und seine erprobung. Bildmessung und Luftbildwesen.

Förstner, W., 1994. A Framework for Low Level Feature Extraction. In: J. O. Eklundh (ed.), Computer Vision - ECCV 94, Vol. II, LNCS, Vol. 802, Springer, pp. 383–394.

Förstner, W., 2000. New orientation procedures. In: ISPRS 2000 : Geoinformation for all.

Fuchs, C., 1998. Extraktion polymorpher Bildstrukturen und ihre topologische und geometrische Gruppierung. DGK, Bayer. Akademie der Wissenschaften, Reihe C, Heft 502.

Hartley, R., 1995. A Linear Method for Reconstruction from Lines and Points. In: Proc. ICCV, pp. 882–887.

Hartley, R. I., 1998. Minimizing algebraic error. In: Philosophical Transactions: Mathematical, Physical and Engeneering Sciences, Vol. 356number 1740, pp. 1175–1192.

Maas, H.-G., 1995. New developments in Multimedia Photogrammetry. A. Gruen, H. Kahmen, chapter Optical 3-D Measurement Techniques III.

Mikhail, E. M., 1963. Use of two-directional triplets in a sub-block approach for analytical aerotriangulation. Photogr. Eng 29, pp. 1014–1024.

Schmid, C. and Zisserman, A., 1997. Automatic line matching across views. Proceedings Conference on Computer Vision and Pattern Recognition, Puerto Rico pp. 666–671.

Valdivia-Manchego, M., 1996. Stereophotogrammetrisch - sedimentologische Aufnahmen und 3D-Modellierung von Schraägschichtungsgefügen am Beispiel einer pliozänen fluvialen Abfolge in der Niederrheinischen Bucht. PhD thesis, Geologisches Institut der Universitä Bonn.