4D SURFACE MATCHING FOR HIGH-SPEED STEREO SEQUENCES

R. Godding [a, *], Th. Luhmann [b], A. Wendt [b]

[a] AICON 3D Systems GmbH, Germany – robert.godding@aicon.de
[b] Institute for Applied Photogrammetry and Geoinformatics
University of Applied Sciences Oldenburg, Germany - luhmann@fh-oow.de

Commission V, WG V/1

KEY WORDS: high-speed camera, stereo, image sequence, free-form surface, matching, correlation

ABSTRACT:

In automotive industry, an increasing demand on 3D surface reconstruction from high-speed camera images can be observed, e.g. in car safety testing environments. The paper deals with the automatic photogrammetric measurement of free-form surfaces from synchronous stereo and multi-image sequences as they are required for dynamic area-based deformation analysis. Firstly, instrumental solutions for synchronous acquisition of multi-image sequences are discussed. For surface reconstruction different image-based matching methods are evaluated and compared with an object-based matching method based on the FAST Vision principle. For each step of the sequence, a three-dimensional surface model is generated, that can be enhanced to a 4D deformation model with time as fourth dimension. For visual analysis the photogrammetrically oriented image sequences can be resampled to sub-pixel accurate anaglyph movies.

1. INTRODUCTION

In car safety testing the analysis of three-dimensional motions and deformations is demanded by European legislation. The analysis includes the interior and exterior of a car, and the involved humans as well. Hence, car industry is forced to design their cars such that permitted limits for dimensional changes of the car are not exceeded in order to provide optimal safety for people involved in an accident. In this context, European standards define particular test environments that, on the one hand, have to be validated by measurements, e.g. for pedestrian safety testing (Luhmann & Raguse 2005). On the other hand, the tests must deliver reproducible measurement results for the analysis of a crash test. For both tasks image sequences from high-speed cameras are used, besides other sensors. In many cases the images are processed in 2-D, however, photogrammetric 3-D evaluations are increasingly used (Raguse et al. 2004).

Different professional products are available for processing of high-speed video sequences, usually suited for the detection and tracing of target points. As a result 3-D trajectories are generated. As examples, commercial systems are offered by FALCON or SIGNUM. They combine modules for 2-D image sequence processing with photogrammetry modules that are originally developed by suppliers of industrial photogrammetry systems. For special vehicle investigations single or multi-camera systems are available (e.g. AICON WheelWatch) and used for recording and processing of highly dynamic scenes. Depending on the configuration (single or multiple images) either single points can be determined in 3-D by spatial intersection, or the relative movement of two objects can be measured in 6 degrees of freedom. While standard high-speed cameras allow the recording of only a few seconds, intelligent cameras with real-time processing provide measurement and recording of infinitely long sequences (Wiora et al. 2004). The latter systems are most advantageous for the measuring on test sites or while a car is moving.

There are also commercially available solutions for the processing of multi-image sequences for the reconstruction of free-form surfaces. Usually they are based on stereo image correlation, hence restricted to two cameras whereby the two images may not have too large perspective variations. Strict matching procedures under consideration of the complete photogrammetric imaging model are known from research (e.g. Gruen & Baltsavias 1988, Heipke 1992), but, as to the knowledge of the authors, are not offered in commercial products for image sequences.

The above mentioned methods are only seldom used in car safety testing. As a matter of fact, severe practical problems exist for the recording of multiple high-speed image sequences. The imaging conditions are often restricted by narrow object spaces. In addition, difficult illumination conditions occur for high-speed cameras due to the extremely short exposure times (< 1 ms). In various applications additional problems arise due to moving objects and occlusions. The environmental parameters yield to problematic image matching conditions.

In a joint project between the companies Porsche, AICON and the IAPG currently different approaches for the measurement of dynamic deformations with high-speed cameras are investigated. The experiments are used to evaluate suitable imaging techniques, targeting methods and processing strategies that have to meet user requirements with respect to accuracy, processing times and practical usability.

2. STEREO IMAGES OF HIGH-SPEED CAMERAS

Three-dimensional analyses of fast dynamic processes (e.g. particle flow, car safety tests) are usually recorded by at least two synchronized cameras. The synchronisation of two or more cameras is not trivial, and often requires high technical effort.

* Corresponding author.

State-of-the-art high-speed cameras provide 1000 frames per second at a resolution of about 1500 x 1000 pixels, e.g. offered by Weinberger, Redlake or NAC. Synchronisation of two or more of these cameras has to be ensured within about 50µs, which is often not given by real systems. Consequently, fast and synchronous stereo or multi-image recording is only possible with advanced technology and high financial investments.

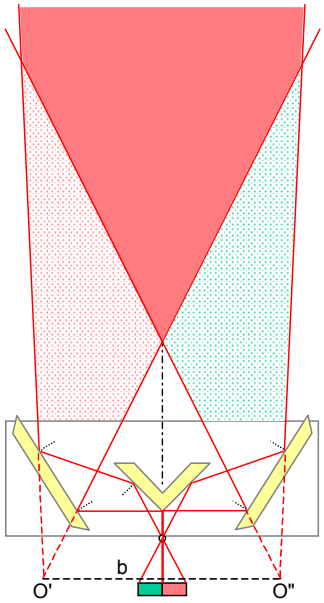

Fig. 1: Imaging principle of a mirror-based stereo camera

Recording of stereo image sequences without any synchronisation error is provided by single cameras that are equipped with optical beam splitters (Luhmann 2005). A prototype has been developed in a joint project with Volkswagen and Porsche where a high-speed camera was combined with an optical stereo splitter (Luhmann & Raguse 2005). Beam splitting is provided by a prism and two mirrors whose distance primarily defines the resulting stereo base length (Fig. 1). Here two virtual cameras are generated with an image resolution that is reduced by 50% in the horizontal direction. As an advantage, costs can be reduced by using only one (expensive) camera. As a drawback, the possible short base length and the reduced image resolution must be mentioned. It has already been shown that the optical set-up can be calibrated with high accuracy where, if necessary, non planar mirror surfaces can also be modelled (Luhmann & Raguse 2005).

The prototype shown in Fig. 2 consists of a virtual base length of about 325 mm. Absolute accuracies of 0.3 mm parallel to the images and 2 mm in imaging direction could be verified in a measuring volume of about 1.2 m x 1.2 m x 2 m. Hence, the accuracy specifications for pedestrian safety testing are fulfilled. Additional variations of the system are recently developed used for different applications and measuring volumes. A series production is in preparation.

Fig. 2: Prototype of a mirror-based stereo camera

In this investigation a test object of about 1 m x 1 m is subject to dynamic stress. The experiment is recorded by three camera systems simultaneously:

- two synchronized cameras NAC HiDCam (760 x 512 pixel, 1000 fps, focal length 12.5 mm),
- three synchronized cameras Weinberger Visario (1536 x 1024 pixel, 1000 fps, focal length 25 mm),
- stereo mirror device with Weinberger MotionFire camera (640 x 512 pixel per image half, 1000 fps, focal length 12.5 mm).

All imaging systems are calibrated and oriented using AICON's 3D Studio program system. Interior and relative orientations are calculated by the measurements of a three-dimensional carbon reference field with known 3-D coordinates. For verification test field points have been measured to an accuracy of 0.5mm (RMS).

As examples, Fig. 3 and Fig. 4 show subsets of the recorded image sequences. The visible surface pattern is explained in chapter 4. The five-point targets at the left and right image edges are located on stable concrete structures in order to provide a fixed coordinate system. The coded targets attached to the surface are used for automatic orientation and as control points for verification as well.

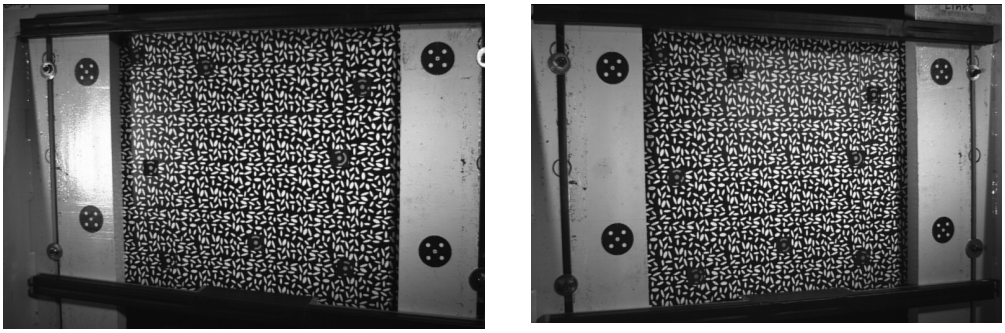

Fig. 3: Image pair of a stereo sequence (NAC HiDCam)

Fig. 4: Original image with stereo beam splitting

3. RESAMPLING OF STEREO IMAGE SEQUENCES

If interior and exterior orientations are given, the original images can be resampled as distortion-free images, normalised stereo images or anaglyph images.

3.1 Distortion-free images

With given calibration parameters new images can be resampled that are free of principal point shifts and distortion. Images are rectified in such a way that the new principal point is located at (0/0) in the exact middle of the new image matrix. Distortion correction is calculated recursively (Luhmann et al. 2006). The principal distance of the distortion-free image is equal to the original image while all other camera parameters become zero. Hence, a pinhole camera image is generated.

Distortion-free images are required for program systems that do not enable a correction for distortion, or which use a different or unknown distortion model. In addition, distortion-free images can be used for faster image data processing.

3.2 Normalised stereo images

The normal case of stereo photogrammetry defines a theoretical imaging configuration with two parallel images where the imaging direction is perpendicular to the stereo base. Both images consist of equal principal distances, principal point coordinates and distortion are defined to be zero. In the normal case of stereo photogrammetry, epipolar lines are always parallel to the x' axis of the image coordinate system with identical y' coordinates. Hence, the search for corresponding image points is limited to single image lines, or a small band due to small errors of orientation data or image measurement, respectively. Height variations of the object occur only as parallaxes in x' direction. These normalised images are therefore a further step for an optimised access to image data.

The resampling process follows the principle shown in Fig. 5. The parameters of relative orientation are given as independent images, i.e. the x-axis of the model coordinate system xyz is defined by the two perspective centres. The resampled images possess the (virtual) image coordinate systems x'_n, y'_n, c_n and x''_n, y''_n, c_n, and the common epipolar line k. Each image point (x'_n,y'_n) and (x''_n,y''_n) is reprojected into the original image coordinates by means of the collinearity equations. In that position the acquired grey or colour value is used for the new image.

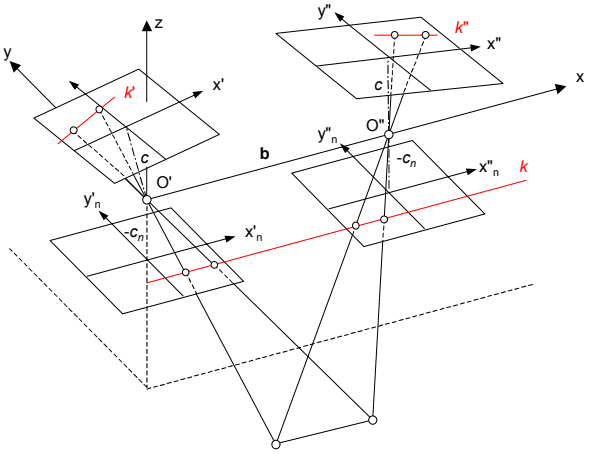

Fig. 5: Normalised images of a stereo image pair

The rectified normalised images do not have any rotations with respect to the model coordinate system. The rotation matrices of the normalised images $\mathbf{R}_n$ in the global coordinate system are calculated from the rotation matrix $\mathbf{A}$ of relative orientation and the rotation matrix $\mathbf{R}$ of the original image according to

$$\mathbf{R}_n = \mathbf{R} \cdot \mathbf{A}^T \quad (1)$$

The translations of exterior orientation remain unchanged.

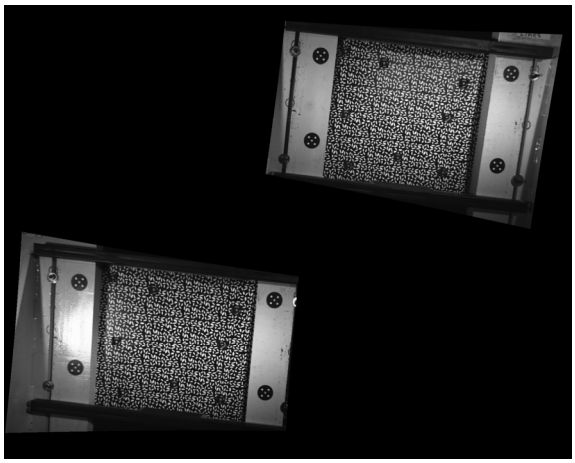

Fig. 6: Normalised images

For larger convergence angles the rectification of normalised images can lead to larger image areas without information (see Fig. 6). The images can be reduced to the actual image information if the corresponding shift of the principal point is stored with the image.

3.3 Anaglyph images

Anaglyphs are a simple way for visual inspection of stereo images. It is limited to b/w images. Optimal stereo viewing is provided if the existing parallaxes occur parallel to the base line of the eyes, i.e. if normalised images are given according to chapter 3.2. Fig. 7 shows the image pair of Fig. 3 displayed as an anaglyph picture. If all images of a stereo sequence are resampled in this way, the recorded scene can be displayed as a stereo movie.

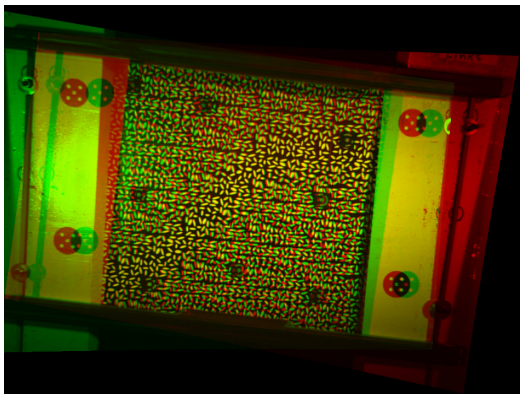

Fig. 7: Anaglyph image

4. SURFACE RECONSTRUCTION

4.1 Surface texture

In automotive industry most surfaces are free of texture, i.e. they have to be marked by an artificial texture. For the measurement of static surfaces usually fringe projection methods are used. In special cases, also raster projection methods, electrolytic marking techniques or methods based on random patterns are used that are also suitable for dynamic scene analysis.

Since high-speed cameras with high frame rates are often operated under poor lighting conditions, only textures with high contrasts can be used. In addition, they must provide gradients in all directions and must be resolved according to the desired object resolution. Furthermore a unique matching in different scales is required if image pyramids are used.

The pattern used in Fig. 3 consists of high contrast and gradients in all directions. The visible ellipses are not used for a single point measurement and can also be replaced by a different pattern with similar properties.

4.2 Stereo image matching

Stereo image matching based on normalised cross correlation is successfully used for a long time (e.g. Piechel 1991). Benefits of the algorithm are given by high robustness, independence of contrast and simple implementation. The similarity of two image patches is measured by the correlation coefficient of both pixel value data sets. The disadvantages of the methods are mostly given by the limited modelling of geometric deviations, e.g. by rotations or perspective images. Normalised cross correlation is always limited to two images.

The measurement of 3-D surface points by correlated image points can be performed by different strategies. Based on given approximate values of surface coordinates in object space the desired 3-D coordinates can be calculated by alternative methods:

a) Spatial intersection
Object coordinates are calculated by spatial intersection using the corresponding image coordinates. The method can not directly be used if a surface grid with constant (pre-defined) XY coordinates shall be measured, hence only a new Z value per point is required. Spatial intersection leads to new coordinates in all directions.

b) Vertical Line Locus
Starting with a given XY position in object space a vertical line in Z direction is created. With a given approximation for Z, new Z values are generated along the line within a given interval. For each Z value the corresponding image coordinates are determined by the imaging equations. For these positions the related image patches are correlated. The best Z value along the vertical line is given by the position of highest correlation. The chosen Z interval can be reduced iteratively so that a pyramidal approach is resulting. Since the XY values are not changed a fixed point grid can be measured directly.

For both strategies a refinement according least-squares matching can be applied (Gruen 1985). In this case an affine geometric and radiometric transformation is estimated. The approach can further be extended by epipolar constraints and multiple images (Gruen & Baltsavias 1988).

4.3 Multi image matching in object space

The Facet Stereo Vision approach (FAST Vision) was developed by Wrobel (1987) and Weisensee (1992). It is founded on the physical imaging process where the relationship of image points or image rays to their corresponding object surface elements is modelled as an inverse imaging process. In principle, the model allows for arbitrary imaging geometries or reflection models. Usually the unknown radiometric value and height value of a surface element are estimated by a Gauss-Markov least-squares adjustment. Iteratively the intensity and geometry model are determined, usually by a given grid. The area between the grid points is interpolated, e.g. by bilinear interpolation. The adjusted radiometric model is equal to an orthophoto generated by all images available. The approach can be used for multiple images and spectral channels.

During the last years the method was further developed at the IAPG (Wendt et al. 2004), especially under consideration of additional object data such as 3-D point clouds (Wendt 2002). The main advantage of the method is given by the simultaneous reconstruction of a complete surface area. Hence, regions without texture can be bridged, or neighbourhood conditions can be integrated by curvature constraints. As a drawback, large normal equation systems have to be solved that depend on size and resolution of the selected surface grid. Computing time can be saved by optimised matrix inversion procedures.

4.4 Surface deformation of the image sequence

Alternative strategies and methods are available for the measurement of dynamic surface deformations.

a) In each epoch independent surface models are measured. The model of epoch k-1 can be used as approximation of the surface

model in epoch k. If a constant XY grid is defined, only deformations in the Z direction are determined.

b) Equal to a), but the resulting points of epoch 0 are transferred to the subsequent epoch such that the physical object point is tracked through the sequence. For this purpose the image pattern of a point in the left image is matched to the same point in the left image of the next epoch. The corresponding points in the right image are measured by epipolar matching. The resulting deformation vectors have an arbitrary trace as given by the point motion in space.

4.5 Results

4.5.1 Surface modelling

The following figures show a selection of calculated surface models out of a variety of existing results. As an example, an image sequence of 100 image pairs is processed that was recorded by two NAC HiDCam (see Ch.2). In object space the pixel resolution is equal to 1.2 mm. The reconstructed object area is about 450 mm x 450 mm with a grid spacing of 10 mm.

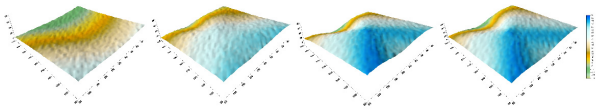

Fig. 8: Surface models of epochs 0, 15, 30 and 45 (fixed grid).

Fig. 8 shows examples of surface models with fixed grid. The object is tilted by about 10 degrees with respect to the XY plane. During the image acquisition it is deformed in the middle of the object. The maximum height variation amounts to ca. 44 mm to be observed between the epochs 30 and 45.

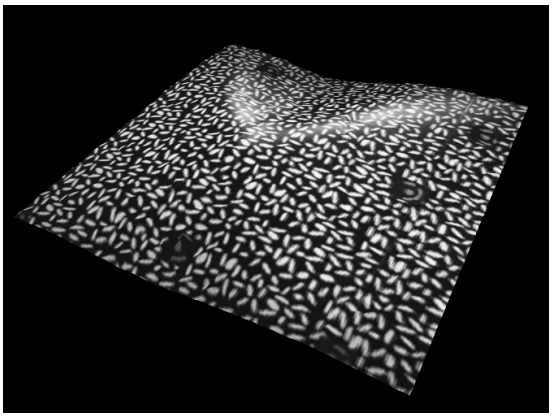

Fig. 9: Surface model superimposed by image texture

For visual inspection it is useful to visualize deformations as a spatially animated image sequence, e.g. as a dynamic VRML model. For this purpose the original image texture can be mapped onto the visualised surface model. Fig. 9 displays an example of texture mapping.

Since no reference surface data of superior accuracy could be provided, the achieved accuracy is verified by visual stereoscopic inspection on a photogrammetric stereo workstation, by comparison of surface coordinates of the point targets, and by comparison of the results between different measuring procedures.

The visual stereoscopic inspection shows slight deviations in the areas of large deformations. The effect must not be resulted by inaccurate height measures, but can also be created by possible errors of the interior and exterior orientation parameters.

The comparison of measured surface points with the coordinates of target points, known from a photogrammetric evaluation of high resolution imagery, results in a mean deviation of 0.7 mm. This result corresponds to the theoretical height accuracy for the normal stereo case

$$dh = \frac{h^2}{b \cdot c} dpx' \qquad (2)$$

where, with b=885mm, h=1250mm and dpx'=6µm (0.5 pixel) a theoretical height precision of 0.8 mm is obtained.

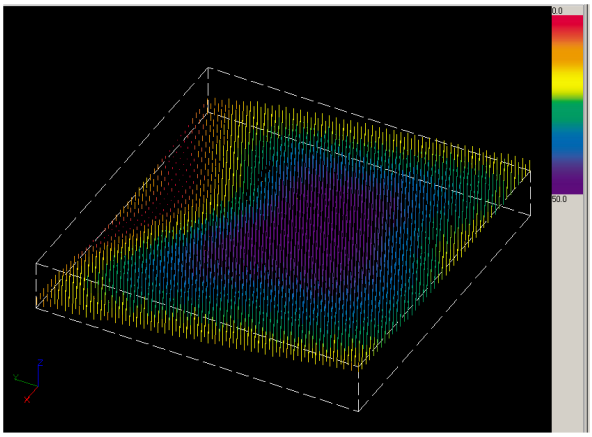

Fig. 10: Deformation vectors between epoch 0 and epoch 25 (fixed grid)

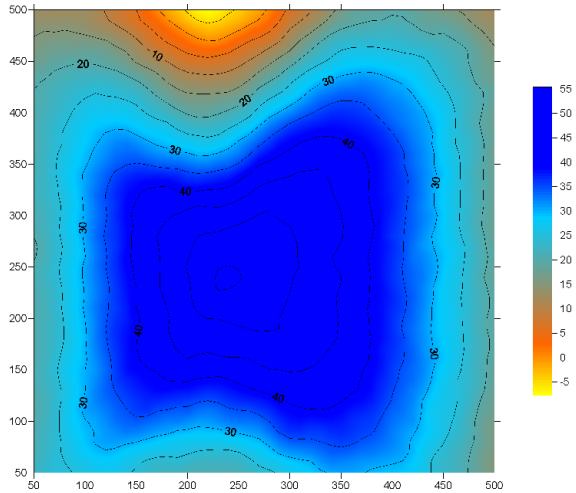

Fig. 11: Deformations between epoch 0 and epoch 30 (fixed grid)

4.5.2 Surface deformation

The following figures show examples of calculated deformations between two epochs. Fig. 10 shows trajectory vectors with fixed point grid, i.e. the vectors are vertical. The vectors are colour-coded with respect to their length. Alternatively, the deformations can also be displayed as differential models according to Fig. 11.

For the analysis of material characteristics it is desired to display the trajectories of distinct object points of the surface.

For this purpose, in each epoch a surface model with physically tracked object points is required. Point tracking is performed by the following scheme: Starting with a fixed point grid in epoch 0, a first initial surface model is computed. For each point at position (x',y') = f(X,Y,Z) an image template is extracted that is transferred into the left image of the next epoch by least squares matching (LSM). The corresponding point in the right image is measured by correlation and LSM at position (x",y"). The new surface point is calculated by spatial intersection.

The computed trajectories correspond to the real object deformation, thus have arbitrary directions. Fig. 12 shows an example of unfiltered deformation vectors between epoch 0 and epoch 25. The visible outliers are caused by object areas that are covered by the coded targets, which obviously are less suitable for matching.

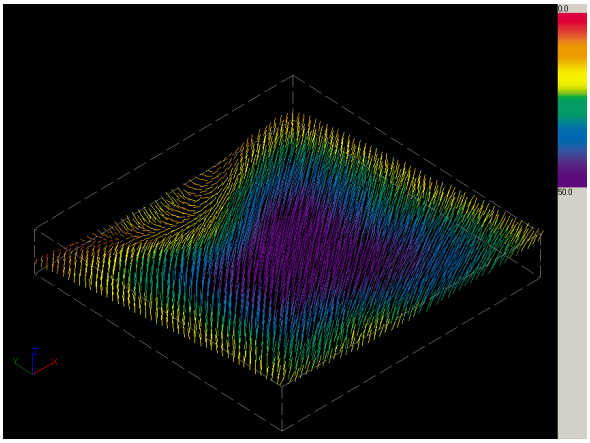

Fig. 12: Deformations between epoch 0 and epoch 25 (variable grid)

5. CONCLUSIONS AND FUTURE WORK

The paper discusses possibilities for recording and measurement of dynamic surface deformations applied to car safety tests. Multi-image sequences are recorded by high-speed cameras. After calibration and orientation the image sequences are processed by image matching methods in order to generate surface models. Surface without appropriate textures must be prepared by artificial patterns that provide high contrasts and unique matching properties in different scales as well.

Different approaches for surface reconstruction are evaluated that are founded on image-based stereo correlation or object-based multi-image matching. The practical usefulness is a function of achievable accuracy, completeness of the surface model, robustness against artefacts, computing times, ease-of-use and availability in professional environments.

The user has specified a maximum computing time of 2 s per image pair which can only be achieved by stereo correlation. The different approaches do not differ significantly with respect to the achieved accuracy. The object-based matching approach can easily be parameterized such that too smooth surfaces result. However, it is advantageous for the use of more than two cameras, and for the potential consideration of disturbing objects in object space.

REFERENCES

Gruen, A., 1985: Adaptive least squares correlation – a powerful image matching technique. - South African Journal of Photogrammetry, Remote Sensing and Cartography, 14 (3): 175-187.

Gruen, A., & Baltsavias, E.P., 1988: Geometrically Constrained Multiphoto Matching. - Photogrammetric Engineering & Remote Sensing (42) 3: 633-641.

Hastedt, H., Luhmann, T. & Raguse, K., 2005*:* Three-dimensional acquisition of high-dynamic processes with a single-camera system and stereo-beam splitting. - Optical 3-D Measurement Techniques, Vol. 2: 175-184.

Heipke, C., 1992: A global approach for least squares image matching and surface reconstruction in object space. - Photogrammetric Engineering & Remote Sensing (58) 3: 317-323.

Luhmann, T., Robson, S., Kyle, S., Harley, I., 2006: Close-range photogrammetry. Whittles Publishing, UK, 500p.

Luhmann, T., 2005*:* Zum photogrammetrischen Einsatz von Einzelkameras mit optischer Stereostrahlteilung. - Photogrammetrie-Fernerkundung-Geoinformation, Heft 2/2005: 101-110.

Luhmann, T. & Raguse, K., 2005: Synchronous 3-D High-Speed Camera with Stereo-Beam Splitting. Sensor+Test, Nürnberg: 443 – 448.

Piechel, J., 1991: Stereobild-Korrelation. - Bähr/Vögtle (ed.): Digitale Bildverarbeitung – Anwendung in Photogrammetrie, Kartographie und Fernerkundung, Wichmann Verlag, Heidelberg: 96-132.

Raguse, K., Derpmann-Hagenström, P. & Köller, P., 2004: Überlagerung der Bildinformationen von Berechnungsanimationen und Highspeed-Filmsequenzen mit Methoden der 3D-Bildmesstechnik. - Bonfig (ed.): Messen, Prüfen, Automatisieren, Band 5, Sensoren Signale Systeme, b-Quadrat Verlag, Kreuztal: 199-208.

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

Wendt, A., 2002: Gemeinsame Ausgleichung von Laserscannerdaten und digitalen photogrammetrischen Bildern. - Photogrammetrie-Fernerkundung-Geoinformation , Heft 2/2002: 103-110.

Wendt, A., Luhmann, T., Riede, R. & Weisensee, M., 2004: Multiple Bildkorrelation im dreidimensionalen Objektraum mit konvergenten Aufnahmen. - Publikationen der DGPF 16: 59-66.

Wiora, G., Babrou, P., Willbold, M., Kofler, R. & Mergenthaler, E., 2004: WheelWatch – Online-Photogrammetrie mit FPGA-Bildverarbeitung. - Luhmann (ed.): Photogrammetrie – Laserscanning - Optische 3D-Messtechnik, Wichmann Verlag, Heidelberg: 272-278.

Wrobel, B., 1987: Digitale Bildzuordnung durch Facetten mit Hilfe von Objektraummodellen. - Bildmessung und Luftbildwesen, Heft 3/1987: 93-101.