Synchronization of Image Sequences - A Photogrammetric Method

Karsten Raguse and Christian Heipke

Abstract

The three-dimensional photogrammetric analysis of image sequences is a growing field of application. For the analysis of dynamic processes one important precondition has to be guaranteed: All cameras have to be synchronized, otherwise the results are affected by asynchronism. In this article a new method is presented, which can determine the asynchronism of an arbitrary number of image sequences. In contrast to already existing methods, in the new approach the asynchronism is modeled in object space and then converted into an interpolation function containing a set of unknowns for each camera. In this form the asynchronism is introduced into an extended bundle adjustment, in which the unknowns are solved simultaneously with the image orientation parameters and the object coordinates of tie points. Therefore, the approach has no restrictions with regard to the number and the set-up of the cameras in the acquisition network. Furthermore, both the temporal and spatial analysis step are carried out simultaneously.

We have implemented the suggested method and have run a number of experiments in the context of vehicle impact testing. First, sequences with a frame rate of 1,000 Hz observing an object with a speed of up to 7 m/s and an asynchronism of 0.8 ms were analyzed. The accuracy of the object point determination could be improved by a factor of 10. Then, five sequences of a vehicle impact test with a speed of 15.6 m/s were investigated. Here, errors in the object coordinates of up to 30 mm could be eliminated using the new approach. Given the small tolerances in car development, this improvement in point accuracy is significant.

Introduction

Digital video and high-speed cameras offer a number of new areas for photogrammetric research and application, since they allow a rigorously investigation of observer and object motion in three dimensions. The three-dimensional analysis of a static object scene with a single moving camera and the analysis of a two-dimensional object motion with a single stationary camera have been reported in the literature (e.g., Pollefeys *et al*., 2004; Maas and Hampel, 2006). Also, several authors described work with a multi-camera set-up for the analysis of three-dimensional object movements. Examples include the analysis of three-dimensional wave surfaces (Santel *et al*., 2003), three-dimensional particle tracking velocimetry in fluids (Maas, 1992; Willneff, 2003) and gases (Putze, 2004), the analysis of high-dynamic object movements within vehicle impact tests in the car industry (Raguse *et al*., 2004; McClenathan *et al*., 2005), the analysis of human motion (D'Apuzzo, 2003), and the analysis of material testing (Schmidt *et al*., 2005; Maas and Hampel, 2006). All these applications use a multi-camera system for the acquisition of object motion, and they all have one common pre-condition: they need a synchronous acquisition of all image sequences. Otherwise, the results of the photogrammetric analysis suffer from the effects of asynchronism. These effects depend on the object speed, the object movement direction, the frame rate of the cameras, and the camera configuration.

Especially for applications with different types of cameras, synchronization is not always guaranteed. For high-dynamic applications, e.g., the analysis of vehicle impact tests, synchronism of the image sequences is indispensable. In a typical acquisition network of an impact test an asynchronism of 0.5 frames between two of the high-speed cameras can lead to a translation of an object point of up to 30 mm. The required accuracy of the object point coordinates for this kind of vehicle impact testing is about 5 mm. Thus, the effects of an asynchronism cannot be ignored.

In this contribution, we present a new approach for the photogrammetric analysis of asynchronously acquired image sequences. The remainder of this paper is organized as follows: In the next section different existing methods to synchronize image sequences are described, followed by the theory of our new approach. Results of practical tests are then described in the next two sections followed by the conclusions and an outlook to future research.

Related Work

Several methods for an accurate analysis of image sequences have been suggested in the past. Some methods have been developed, which are not affected by asynchronism due to the fact that they obtain three-dimensional object coordinates from only one image sequence, such as the 6-degrees of freedom method (Luhmann, 2005), a specific combination of active and passive sensors, e.g., the camera-projector systems (Maas *et al*., 2003) or the 3D-camera, which directly acquires three-dimensional object data (Oggier *et al*., 2004). The disadvantage of these methods is that they need special test conditions, which cannot always be guaranteed. In

Karsten Raguse is with Volkswagen AG, Research & Development, Department EKSE - Safety Test Facilities, Letter Box 1604, 38436 Wolfsburg, Germany (karsten.raguse@volkswagen.de).

Christian Heipke is with the Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hanover, Germany.

Photogrammetric Engineering & Remote Sensing
Vol. 75, No. 4, May 2009, pp. 535–546.

0099-1112/09/7504–0535/$3.00/0

addition, for many applications the accuracy of these methods is not high enough.

Alternatively, methods based on more than one image sequence exist. In order to obtain synchronous image sequence measurements, different solutions have been suggested. They can be divided into three main groups: methods using hardware components, methods using a combination of hardware and software, and methods using only software. These three groups are described in the following. The third group, the software methods, is described in more detail, because our new approach belongs to this group.

The methods of the first group use hardware components for the synchronization of the cameras. These hardware components are e.g., hardware trigger boxes or frame grabbers. They send synchronization trigger signals to the connected cameras (Santel *et al.*, 2003). The synchronization accuracy depends on the accuracy of the trigger device, which generates the synchronization signal. Other methods of this group use mirror systems, e.g., a beam splitter in front of a single camera (Putze, 2004; Hastedt *et al.*, 2005). With a stereo beam splitter, two virtual cameras are simulated, which acquire the scene exactly synchronously, however the available horizontal resolution per image is only 50 percent, and the camera acquisition set-up is not very flexible.

The methods of the second group use a combination of hardware and software. The image sequences are acquired asynchronously, but the times of image capture are registered and are used to correct the results in a postprocessing step. The registration of image data capture is achieved through different methods. In some approaches all cameras are connected to a high precision time system. The system time of image acquisition is inserted into each frame of the sequence, e.g., in form of a special time code (Narayanan *et al.*, 1995). In another method, the cameras all have to see the display of a high precision clock which shows the current system time. For both methods, restrictions with regard to the camera set-up exist: either all cameras have to be connected to a time system or they all have to see the display of the high precision clock.

The methods of the third group only use software to synchronize asynchronously acquired image sequences. Here, the cameras do not have to be physically connected to any kind of master system, and no special hardware device is necessary. In some applications the cameras used can have an arbitrary, also varying time offset, e.g., due to different frame rates. The values for the asynchronism are determined during the computation of the object point coordinates. In some cases, the temporal alignment is separated from the spatial alignment and is carried out in a preliminary step.

The third group, the software algorithms, can further be divided into three subgroups (Lei and Yang, 2005). This subdivision depends on the image information which is used for the determination of the asynchronism. The methods of the first subgroup are called intensity-based methods or direct methods. They use the intensities of all pixels of the image for the determination of the asynchronism (e.g., Caspi and Irani, 2000). The second subgroup contains feature-based methods, which use solely detected features for the calculation of the spatial and temporal alignment. These features can be an unordered collection of points (Stein, 1998; Lee *et al.*, 2000), lines (Lei and Yang, 2005), one or more point trajectories (Caspi *et al.*, 2002; Kuthirummal *et al.*, 2002; Pooley *et al.*, 2003; Zhou and Tao, 2003; Carceroni *et al.*, 2004; Tresadern and Reid, 2004; Tuytelaars and Van Gool, 2004; Wedge *et al.*, 2006), inflection points of a trajectory (Whitehead *et al.*, 2005), space-time interest point distributions (Yan and Pollefeys, 2004) or object silhouettes (Sinha and Pollefeys, 2004). A more detailed description of the intensity-based and feature-based methods is given by (Irani and Anandan, 2000) and (Torr and Zissermann, 2000). The third subgroup contains methods for cameras, which are joined together rigidly, thus the relative orientation between the cameras is fixed during acquisition. The methods of this group do not use pixels or features as parameters. Rather, within each image sequence the transformations between two successive frames are determined and these transformation parameters are used to compare and align the sequences (Caspi and Irani, 2001; Wolf and Zomet, 2002; Spencer and Shah, 2004) taking advantage of the known relative orientation parameters.

All the above-mentioned methods for synchronizing image sequences with software algorithms have several restrictions. Similar to area-based image matching the methods of the first subgroup are sensitive to changes in brightness. Furthermore, the camera set-up is mostly restricted to two cameras which have to be positioned closely together and the scenes have to be acquired under nearly the same perspective conditions. The methods of the second subgroup are often restricted to two or three cameras, at least if linear spatial alignment functions are used. The main restriction of the methods of the third group is the requirement of rigidly joined cameras. In addition, the spatial and the temporal alignment of the sequences in the methods of all three subgroups are mostly separated into two independent steps. This can lead to problems in the analysis because potential correlations between temporal and spatial parameters cannot be properly considered.

New Approach for the Analysis of Image Sequences

Requirements and Basic Concepts

A reliable and robust algorithm for the solution of the asynchronism problem within image sequence analysis should be able to handle the following cases (Carceroni *et al.*, 2004):

- unknown frame rates of the cameras,
- arbitrary time shift between the sequences,
- arbitrary object motion and speed,
- unknown user-defined camera set-up, and
- absence of static points in the scene.

These requirements were the base for developing the new approach. The background of our work is the analysis of highly dynamic vehicle impact tests in the car manufacturing industry, keeping in mind, that the new approach should be used for tests, where the synchronization of the cameras with hardware components is not possible. We have to meet the stringent accuracy demands in this field and aim at handling an arbitrary number of different types of cameras, which can be positioned anywhere around the measuring volume without any restrictions. We can, however, allow for a number of stable points in the scene, which can be used as ground control points in the photogrammetric analysis.

To meet all these requirements we have combined the spatial and the temporal alignment and consider them simultaneously within the photogrammetric analysis. Image coordinates of these ground control points and of moving signalized points are used as input values for the algorithm. Therefore, the new method belongs to the subgroup of feature-based methods of the software algorithms. The problem of synchronizing image sequences is solved via a temporal correction function, which is converted to an interpolation factor in object space. This factor is introduced into the functional model of the bundle adjustment. The parameters of the interpolation factor are considered as unknowns.

Asynchronism as a Temporal Component of the Optical Data Channel

Like other components of the optical data channel the asynchronism between the cameras represents an important factor for the accuracy of the image sequence analysis (Raguse and Wiggenhagen, 2003). An asynchronism between the used cameras belongs to the group of temporal components of the optical data channel. These temporal components have different physical reasons and can be categorized as follows:

- accuracy and stability of the frame rates,
- accuracy and stability of the exposure times,
- different frames rates of the cameras,
- different exposure times of the cameras
- constant or dynamic time differences between the cameras, and
- object motion during exposure (motion blur).

To obtain correct and reliable results in time and space all these effects are considered simultaneously within the photogrammetric analysis independent of their actual physical reason and are denoted as asynchronism between the image sequences.

The reference for the temporal alignment is always a master time system. This master time system can be an external clock or one of the cameras of the acquisition network. All temporal calibration parameters are calculated with respect to this time reference.

Modeling the Asynchronism using Correction Functions

In this approach, the asynchronism of each camera is modeled by a temporal correction function. If the used camera has exactly the same frame rate as the reference frame rate and the frame rate is exactly constant over time, the correction function contains only a constant time offset. If the frame rate is different from the reference frame rate but constant, the asynchronism can be modeled by a linear correction function.

For each camera, a separate temporal correction function is introduced. The linear correction function reads:

$$\Delta t(t_i) = \Delta t_{Offset} + \Delta t_{Ratio}\,(t_i - t_0) \tag{1}$$

where $\Delta t(t_i)$ = asynchronism at time t_i, Δt_{Offset} = constant time offset, Δt_{Ratio} = ratio of the frame rate with respect to the reference frame rate, t_i = time step i of the image sequence, and t_0 = time step of the last synchronization pulse.

The ratio Δt_{Ratio} of the frame rate with respect to the reference frame rate is defined through:

$$\Delta t_{Ratio} = \left(\frac{f}{f_{ref}} - 1\right) \tag{2}$$

where f = frame rate of the camera to be synchronized, and f_{ref} = frame rate of the reference system.

The modeling of the asynchronism is not restricted to this linear correction function. In principle, an arbitrary functional approach can be used.

Introduction of the Asynchronism into the Bundle Adjustment

In this approach the temporal and the spatial alignment of the image sequences are considered simultaneously. Thus, the measurements of all time epochs and of all object points can be analyzed in one step with the advantage that possible correlations between space and time are automatically taken care of.

For the following explanation we only consider the measurements of one object point in two image sequences, keeping in mind that the method can be extended to an arbitrary number of object points and image sequences, just like conventional bundle adjustment. The left image subset in Figure 1 is regarded as the reference system in our example. If both image sequences shown in Figure 1 are exactly synchronous, the image points at epochs $i - 1$, $i, i + 1$, etc. in the two subsets are conjugate points. An asynchronism between the two cameras leads to a deformation of one trajectory with respect to the other. To overcome this problem the conjugate points in the right subset (Figure 1b) are interpolated with respect to the asynchronism between the two cameras. Assuming a relatively small distance between the two points of the trajectory, and thus a sufficiently high frame rate with respect to the object speed, a linear interpolation along the trajectory in object space is employed.

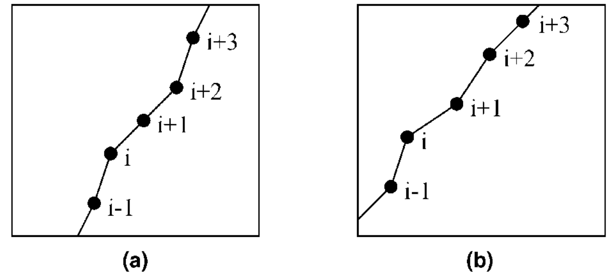

Figure 1. Subsets of two image sequences with a point imaged at different epochs $i - 1$, $i, i + 1$, etc. and the corresponding trajectories: (a) camera 1, and (b) camera 2.

The results of the extended analysis of the situation depicted in Figure 1 are shown in Figure 2. The black points in the right subset (Figure 2b) are the image points of the asynchronously acquired second image sequence. The image points $i - 1^*$, i^*, $i + 1^*$ (gray points in Figure 2b) are interpolated in order to eliminate the effects of asynchronism.

If the asynchronism is larger than the time interval between the acquisition of two consecutive images, the asynchronism (the frame time) has to be reduced by an integer multiple of the frame time (Equation 3):

$$\Delta t_{red}\,(t_i) = \Delta t(t_i) - n \cdot \frac{1}{f} \quad \text{with} \quad n = int\,[\Delta t(t_i) \cdot f] \tag{3}$$

where $\Delta t_{red}(t_i)$ = reduced asynchronism, and n = renumbering factor of the asynchronism.

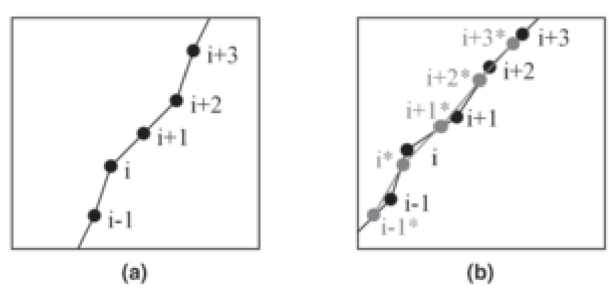

Figure 2. Subsets of two image sequences with the original point positions imaged at different epochs $i - 1$, $i, i + 1$, etc. and with the corrected positions: (a) camera 1, and (b) camera 2; original positions $i - 1$, $i, i + 1$, etc. and the corrected positions $i - 1^*$, i^*, $i + 1^*$, etc.

In general, the renumbering factor n is obviously not known. We solve for n in an iterative way which can be considered a preprocessing step of the actual determination of the asynchronism: first, we run the analysis with $n = 0$. If the resulting asynchronism is larger than the frame time we compute n according to Equation 3. We repeat the procedure until the asynchronism is small enough.

To introduce the asynchronism to the functional model, the asynchronism, as a temporal term, has to be converted to a geometric term in image space to use it for the interpolation in the analysis (Equation 4):

$$\Delta sync(t_i) = f \cdot \Delta t_{red}(t_i) \tag{4}$$

where $\Delta sync(t_i)$ = interpolation factor of the asynchronism.

The use of the interpolation factor of the asynchronism leads to the following temporal correction terms for the object coordinates X, Y, and Z:

$$\begin{aligned} \Delta X_{sync}(t_i) &= |\Delta sync(t_i)| \cdot [X(t_{i+sign(\Delta sync(t_i))}) - X(t_i)] \\ \Delta Y_{sync}(t_i) &= |\Delta sync(t_i)| \cdot [Y(t_{i+sign(\Delta sync(t_i))}) - Y(t_i)]. \\ \Delta Z_{sync}(t_i) &= |\Delta sync(t_i)| \cdot [Z(t_{i+sign(\Delta sync(t_i))}) - Z(t_i)] \end{aligned} \tag{5}$$

These temporal correction terms are added to the collinearity equations:

$$\begin{aligned} x(t_i) &= x_o - c \cdot \frac{\begin{gathered} r_{11} \cdot (X(t_i) + \Delta X_{sync}(t_i) - X_0) \\ + r_{21} \cdot (Y(t_i) + \Delta Y_{sync}(t_i) - Y_0) \\ + r_{31} \cdot (Z(t_i) + \Delta Z_{sync}(t_i) - Z_0) \end{gathered}}{\begin{gathered} r_{13} \cdot (X(t_i) + \Delta X_{sync}(t_i) - X_0) \\ + r_{23} \cdot (Y(t_i) + \Delta Y_{sync}(t_i) - Y_0) \\ + r_{33} \cdot (Z(t_i) + \Delta Z_{sync}(t_i) - Z_0) \end{gathered}} + \Delta x_{dist.} \\ y(t_i) &= y_o - c \cdot \frac{\begin{gathered} r_{12} \cdot (X(t_i) + \Delta X_{sync}(t_i) - X_0) \\ + r_{22} \cdot (Y(t_i) + \Delta Y_{sync}(t_i) - Y_0) \\ + r_{32} \cdot (Z(t_i) + \Delta Z_{sync}(t_i) - Z_0) \end{gathered}}{\begin{gathered} r_{13} \cdot (X(t_i) + \Delta X_{sync}(t_i) - X_0) \\ + r_{23} \cdot (Y(t_i) + \Delta Y_{sync}(t_i) - Y_0) \\ + r_{33} \cdot (Z(t_i) + \Delta Z_{sync}(t_i) - Z_0) \end{gathered}} + \Delta y_{dist.} \end{aligned} \tag{6}$$

where c = calibrated focal length, x_o, y_o = image coordinates of the principal point, X_0, Y_0, Z_0 = object coordinates of the projection center, r_{ij} = elements of the rotation matrix between object space and image space, $X(t_i)$, $Y(t_i)$, $Z(t_i)$ = object coordinates at epoch t_i of the sequence, $\Delta x_{dist.}$, $\Delta y_{dist.}$ = correction terms for lens distortion, and $\Delta X_{sync}(t_i)$, $\Delta Y_{sync}(t_i)$, $\Delta Z_{sync}(t_i)$ = correction terms for asynchronism effects at t_i.

Discussion

The presented approach is an extension of the bundle adjustment. Thus, it has the same benefits as e.g., a simultaneous estimation of interior and exterior orientation parameters provided the necessary control information is available, the simultaneous analysis of an arbitrary number of image sequences and a high accuracy potential for the unknowns. The asynchronism is modeled through a linear function in object space. This implies the assumption that the object speed is constant within a short time interval, namely from one image to the next of the image sequence. If a sufficiently high frame rate is used for image acquisition this restriction does not have any practical consequences.

Furthermore, at least measurements of one image point in $n + 3$ consecutive time steps are needed within the analysis to calculate the asynchronism parameters, because we do not know *a priori* if the asynchronism is positive or negative (n is the renumbering factor). While the suggested determination of n must be considered as ad hoc and is not guaranteed to converge, we have found it to perform very well in all our experiments; convergence to the correct result was reached after only a few iterations. If, however, n is very large with respect to the frame time, the described procedure can fail. In this case a user-defined interval for the possible asynchronism is systematically analyzed in that way, that every possible asynchronism is used in an analysis and the best result is used for a further analysis to refine the results.

Finally, if the acquisition network consists only of two cameras, it is indispensable, that the object motion does not occur exclusively in the epipolar plane, because otherwise the asynchronism results in a systematic point shift in that plane since the two image rays intersect irrespective of the values for the asynchronism.

Test 1: Analysis of a Rotating Three-dimensional Test Field

The goal of the tests presented in this section is to demonstrate some characteristics of the new approach. In the first part, the theoretical accuracy for the object point coordinates is determined using error propagation. In the second part, the empirical accuracy of the new approach is shown. Due to the fact that reference positions for the object points are not known, we analyze the length of a rotating reference distance on a test field. In addition, the number of analyzed consecutive images is systematically reduced to a number, which is typical for the later application of this method, the analysis of vehicle impact tests.

Test Equipment and Test Conditions

In this test, a rotating three-dimensional stable test field was used. The object points on the test field have a maximum speed of 7 m/s. They were observed by two NAC Hi-DCam II digital high-speed cameras which acquire image sequences with a frame rate of 1,000 Hz. Within one complete rotation of the test field about 910 images are acquired. Each camera has a sensor size of 1,280 pixels × 512 pixels, a pixel size of 12 μm, and a lens with a focal length of about 16 mm. The base between the two cameras is about 28 cm, and the distances between the cameras and the rotating test field are about 1.9 m. Thus, the base-to-distance ratio of this camera constellation amounts to 1:7. The object coordinate system is aligned in the way that the rotation axis of the test field is parallel to the Z-Axis; see Figure 3. The stereo base is nearly parallel to the X-axis and the viewing directions of the cameras are tilted by about 30° with respect to the Z-axis of the coordinate system.

The parameters of interior and the exterior orientations had been determined automatically before the test and are assumed to be constant over the analyzed time interval. The image coordinates of each point were then determined using automatic target detection algorithms.

Calculation of a Reference Value for the Asynchronism between the Two Cameras

To evaluate the computed values for the asynchronism between the two image sequences, a reference value for the asynchronism is calculated using a high-precision clock. This clock is a LED panel of superior time accuracy, termed synchronometer, which is positioned in the view of the cameras. Counting the lit LEDs of the panel in each image of the sequence gives the current absolute master time of the image.

The temporal resolution of the synchronometer in the used mode is 0.01 ms. The results of the reference measurements are shown in Figure 4. The determined time

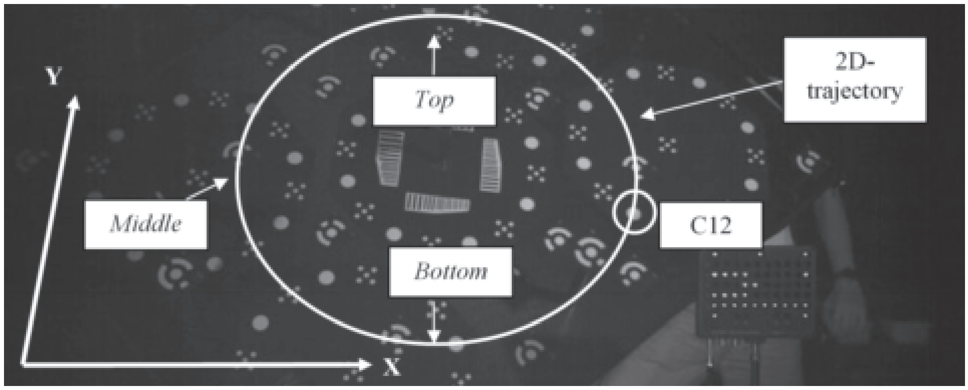

Figure 3. Test field with 2D-trajectory of object point C12.

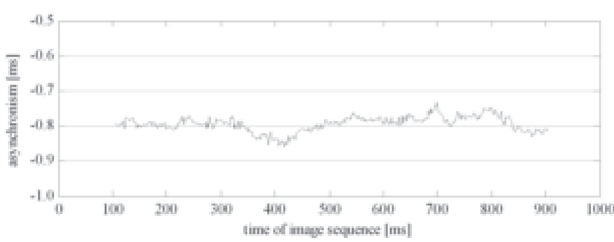

Figure 4. Asynchronism during the analyzed time interval as determined with the synchronometer.

differences between the two image sequences show a temporal offset with the mean value of −0.79 ms and a standard deviation of 0.02 ms. The results do not show any effects of different frame rates of the two sequences or of a frame rate drift. The measurements of the temporal offset are a bit noisy, but the results are accurate enough to use them for evaluating the results of our algorithm.

Determination of the Theoretical Accuracy of the New Approach

For the following explanation the focus is set to three special positions on the trajectory of object point C12 which is representative for the whole set-up (see Figure 3). The three selected positions are marked in the figure and are denoted as top, middle, and bottom. Due to the alignment of the test field with respect to the coordinate system, the Z-component of object point C12 does not change during the analyzed time interval, the object point only moves in the X-Y plane.

Analysis of the Test without Modeling the Asynchronism

First, the object space coordinates of point C12 resulting from a bundle adjustment, which neglects the asynchronism, are analyzed. The analysis is done for each time step separately and subsequently, three-dimensional trajectories of the object points are computed. The effect of the asynchronism on the three-dimensional point determination depends on the direction of movement. If the object point moves in the epipolar plane the asynchronism results in a translation of the object coordinates in this plane. If the object point moves in another direction, the asynchronism results in higher standard deviations for the object coordinates. The translation effects at the three positions are shown in Figure 5.

At the trajectory positions top and bottom the object point moves in the epipolar plane. At these positions the asynchronism leads to a translation of the object point in viewing direction. Depending on the direction of movement the calculated positions of the object point lie in front or behind the real position (see positions "top" and "bottom" in Figure 5). In this test the asynchronism is about −0.79 ms and the object point C12 moves with a speed of 2.9 m/s. Due to this constellation the object point C12 has moved about 2.3 mm between the acquisitions of two conjugate images. At the position bottom, this results in a translation of about 1.2 mm perpendicular to the viewing direction and a translation of about 15.7 mm in viewing direction. Due to the tilt of the viewing direction with respect to the Z-axis of about 30°, the translation affects the Y- and Z-component of the object coordinates. Thus, the correct coordinates at position bottom are translated about 14 mm in Z-direction and 8 mm in Y-direction (see Figure 5).

The dotted line in the left part in Figure 5 shows the plane in which the object point C12 actually moves. At the trajectory position middle the object point moves in a direction perpendicular to the epipolar plane. At this position the asynchronism results only in an increased standard deviation for the coordinates. The calculated position of the object point, however, is correct.

Analysis using the Extended Model

The new approach including the extended functional model is carried out using a constant temporal correction function. The results are shown in Figure 6. The calculated position for top, middle and bottom lie all on the plane (see Figure 6a). The remaining deviation of the Z-component (Figure 6b) is smaller than 5 mm and is probably caused by a small movement of the rotation axis of the test field. The calculated value for the asynchronism amounts to −0.78 ms, the standard deviation is 0.002 ms.

Comparison of the Results and Evaluation of the New Approach

For the position middle the results of the analysis are presented in Table 1. It can be seen that modeling the asynchronism resulted in an improvement of the theoretical standard deviation of the object coordinates of point C12 by approximately a factor of 10.

The differences between the calculated coordinates are shown in Figure 7. They correspond to the theoretical values for the translation with an asynchronism of −0.79 ms and an object speed of about 2.9 m/s. The strong correlation

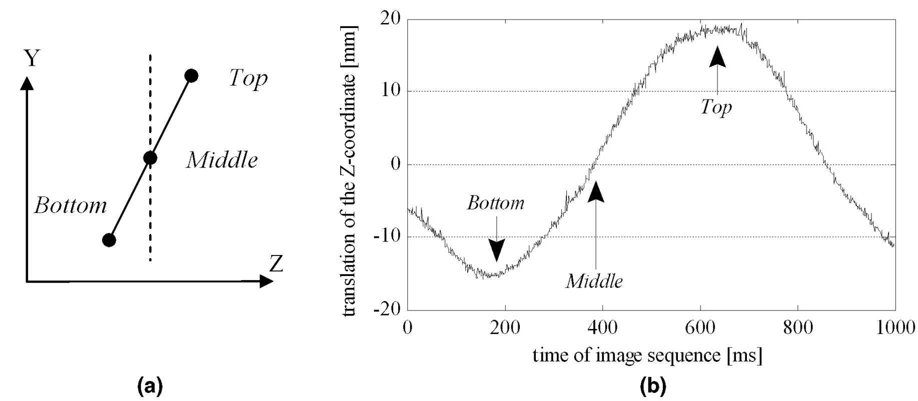

Figure 5. Translation of the Z-coordinate of object point C12 due to the effect of the neglected asynchronism in the analysis: (a) theoretical situation, and (b) analysis results.

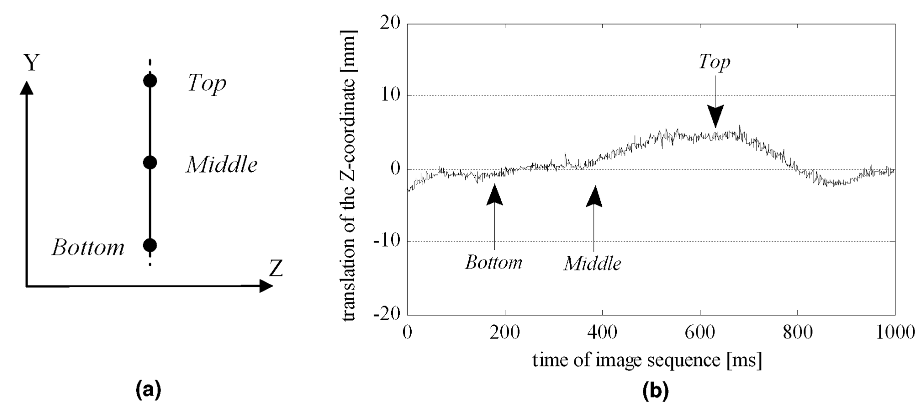

Figure 6. Accuracy of the analysis of the Z-coordinate of object point C12 when the asynchronism is taken into consideration: (a) theoretical situation, and (b) analysis results.

TABLE 1. COMPARISON OF THE STANDARD DEVIATION IN IMAGE SPACE s_0 AND THE COORDINATES OF OBJECT POINT C12 DETERMINED USING ERROR PROPAGATION, POSITION *MIDDLE*

Object point C12 on the test field		No. modeling of asynchronism	With modeling of asynchronism
s_0		7.2 μm (0.60 Pixel)	0.7 μm (0.06 Pixel)
Asynchronism		./.	−0.78 ms
Middle	σ_x	1.86 mm	0.17 mm
	σ_y	5.25 mm	0.49 mm
	σ_z	8.37 mm	0.77 mm

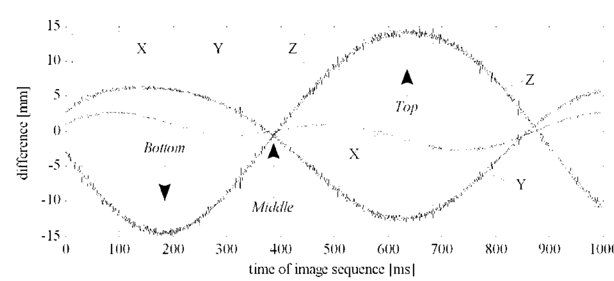

Figure 7. Differences of the object coordinates of point C12 during the analyzed image sequence caused by the different types of analysis.

between the asynchronism and the translation of the object coordinates in viewing direction only appears in a test set-up with two cameras acquiring an object which moves in the epipolar plane as in our experiments. If image sequences of more than two cameras are used the asynchronism results in an increase of the standard deviation of the calculated object point coordinates, as additional experiments have actually demonstrated.

In summary, in our test set-up the modeling of the asynchronism has led to correct object point coordinates and to an improved accuracy. In the analysis an asynchronism of −0.78 ms with a standard deviation of 0.002 ms was computed. The reference value for the asynchronism obtained through the master clock was −0.79 ms with a standard deviation of 0.02 ms. Thus, our results can be regarded as correct.

Analysis of a Reference Distance on the Test Field

In the second part of the test series the accuracy of the algorithm is tested empirically. Due to the fact that we do not precisely know the reference positions of the objects points, the focus is on the computed length of a reference distance. The analyzed reference distance is defined on the test field between the two object points B12 and C12 (see Figure 8). Note that the center of rotation does not lie on the line through B12 and C12. Obviously, the length of the reference distance is constant in 3D space. To obtain the empirical accuracy of the new approach, we use the differences between the computed length and the reference length. The length of this reference distance is calibrated to 520.16 mm with a standard deviation of 0.01 mm. The object coordinates of the two points were calculated for each time step of the image sequence analysis with and without the synchronization term of Equation 5. The results over the complete time interval of about 910 ms are shown in Figure 9.

Figure 8. Reference distance between the object points B12 and C12 on the test field.

The analysis which neglects the asynchronism leads to significant systematic changes in the calculated length (gray line in Figure 9). The maximum changes are up to 14 mm. The effects of the asynchronism on the length of the reference distance with respect to the orientation of the length can also be seen in Figure 9. Note that both object points do not have the same object speed. The object point B12 lies closer to the centre of rotation and therefore the object speed is only 2.1 m/s in comparison to the speed of 2.9 m/s of object point C12. Except when the points move perpendicular to the epipolar line, they are somewhat translated in the viewing direction. This translation affects the length in a systematic way (see sine-like gray curve in Figure 9). The exact form depends on the camera geometry, the object speed and the object movement direction.

The curve in Figure 9 shows two maxima and two minima which are different in their values. The minima belong to the images of the sequence where the reference distance is approximately perpendicular to the epipolar planes of the two object points B12 and C12. In these positions the two points are systematically shifted: the upper endpoint is translated in the direction towards the cameras and the lower endpoint is translated in the opposite direction (see Figure 10). In the used camera configuration the viewing directions of the cameras are tilted by about 30° to the *Z*-axis of the coordinate system of the test field. Due to this set-up and the translation effects, the determined length of the reference distance is too short. Due to the different speed of the two object points the minimum positions differ in their values.

The two maxima of the curve belong to the positions where the reference distance lies approximately in the two epipolar planes. At the lower maximum both points are translated in the direction towards the cameras, and thus the computed length is shortened. At the higher maximum both points are translated away from the cameras and the computed length is extended. Independently of the object speed of the two points, the two maxima are symmetrical to the real length of the reference distance.

Using the new approach the asynchronism of −0.78 ms with a standard deviation of less than 0.01 ms is found.

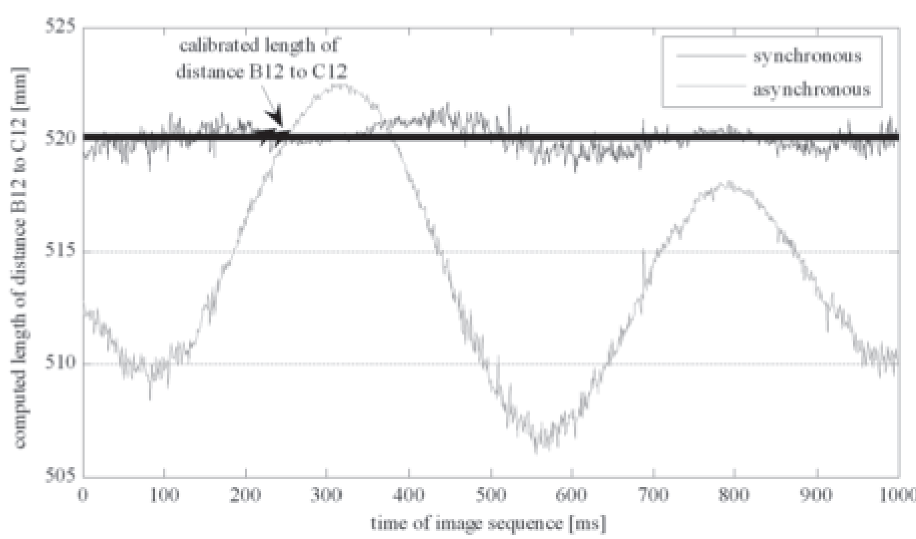

Figure 9. Length of the reference distance B12-C12 in the analyzed time interval of 910 frames with and without considering the asynchronism.

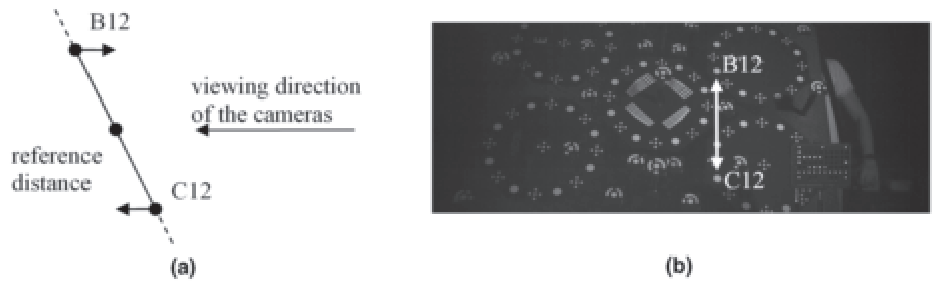

Figure 10. Translation effect due to the asynchronism at the position where the reference distance is nearly perpendicular to the epipolar plane: (a) theoretical situation, and (b) image of the sequence.

Thus, there is no significant difference to the reference value for the asynchronism of −0.79 ms. The calculated mean value for the length is 520.15 mm with a standard deviation of 0.51 mm (black line in Figure 9). The difference to the calibrated length is 0.01 mm, which can be neglected. The changes of the length over the analyzed time interval (see Figure 9) show a small systematic effect. This effect can possibly be traced back to the slightly varying temporal offset between the two sequences (see again Figure 3).

Due to the modeling of the asynchronism the length of the reference distance B12 to C12 could be determined correctly. Without modeling the asynchronism the length of the reference distances shows systematic errors of up to 14 mm.

Reduction of Consecutive Measurements

In the prior test series all measurements of about 910 consecutive images were used for the determination of the asynchronism parameters. The aim of this part of the test is to find the smallest number of images to still obtain correct results with respect to the restriction that the object movement is not allowed to occur in the epipolar plane. In order to achieve this goal, we reduce the number of consecutive image coordinate measurements used simultaneously. The analysis with the reduced number of measurements is then carried out repetitively at different equally spaced positions within the time interval. The results are listed in Table 2.

It can be seen, that the results are only slightly worse when reducing the number of measurements. The changes in the calculated asynchronism are in the range of 0.1 ms. The changes of the reference lengths are approximately 1.5 mm.

In this set-up, the reduction of the consecutive measurements to only ten images does not reduce the accuracy of the results significantly, because at least one of the two image points always moves outside the epipolar plane.

Test 2: Analysis of a Vehicle Impact Test

In the second experiment, a vehicle impact test is analyzed. In this test a vehicle is propelled into a deformable barrier with a speed of 15.6 m/s. Images are acquired with a circular set-up of eight digital high-speed cameras (see Figure 11). In the photogrammetric image sequence analysis reported here, the five cameras L1, L3, R1, R3, and O2 are introduced (see Figure 12). The other cameras (L2, R2, and O1) are only used for special detail analysis and only acquire the movement of parts of the car and the dummies. As in the first test, the parameters of interior and exterior orientations had been determined automatically prior to the test and are assumed to be constant during the analyzed time interval.

In the first part of the analysis, the computed asynchronisms are tested for plausibility. In the second part the five sequences are analyzed with and without the consideration of asynchronism, and the results are compared with respect to a distance between a target on the a-pillar of the car and a target on the dummy head.

Plausibility Check of the Calculated Asynchronism of the Cameras

To check the calculated asynchronism for each camera, different constellations of two and three sequences are analyzed. The calculated asynchronism of the different constellations are compared to each other and tested for plausibility.

Table 2. Results of the Reduction of the Number of Consecutive Images on the Asynchronism Δt and the Length of the Reference Distance B12 to C12

No. of images	No. of repetitions	Δt [ms] min	Δt [ms] max	length [mm] min	length [mm] max
910	1	−0.78		520.15	
200	40	−0.74	−0.83	519.59	520.33
100	80	−0.73	−0.85	519.17	520.98
50	150	−0.72	−0.86	519.07	521.10
20	450	−0.71	−0.87	518.87	521.32
10	900	−0.70	−0.88	518.68	521.51

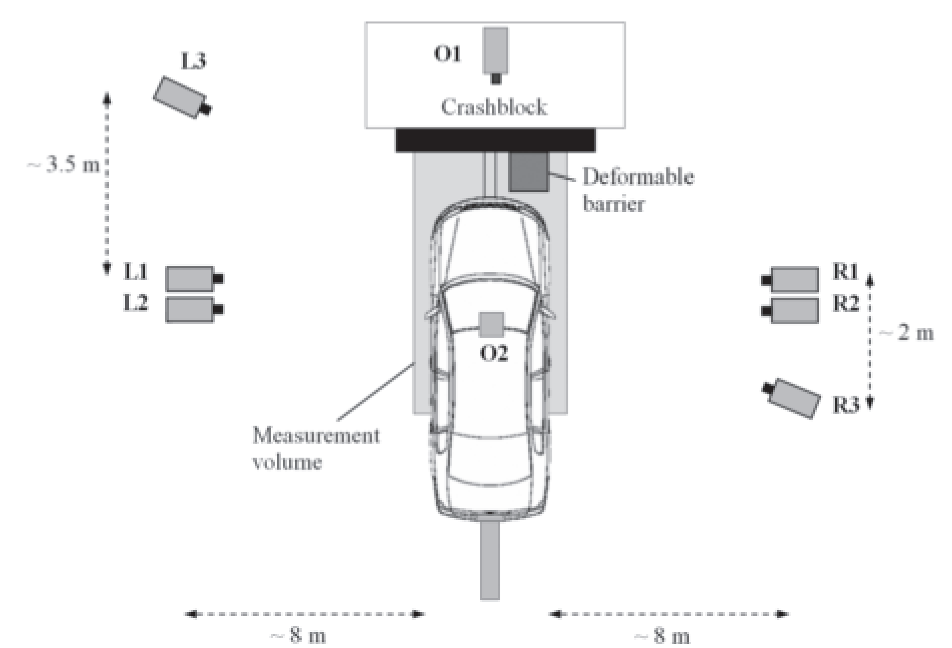

Figure 11. Schematic view of the camera constellation of a frontal vehicle impact (view from above, not scaled).

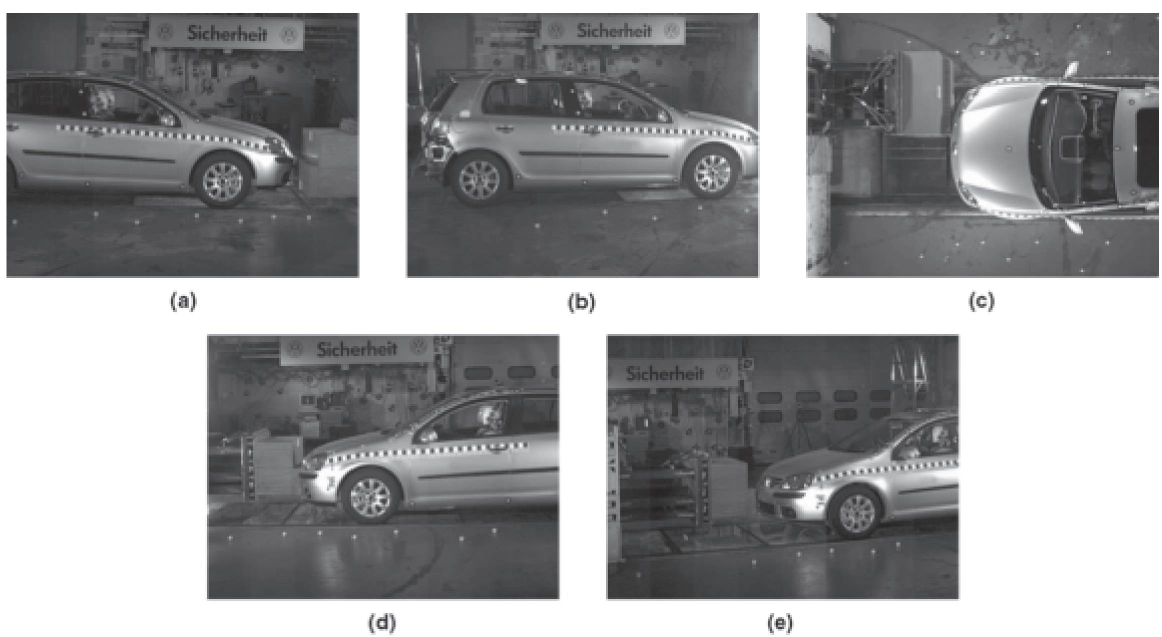

Figure 12. Snapshot of the five sequences used in the test: (a) Image of sequence L1 at *T0*, (b) Image of sequence L3 at *T0*, (c) Image of sequence O2 at *T0*, (d) Image of sequence R1 at *T0*, and (e) Image of sequence R3 at *T0*.

First, the cameras L1, L3, and O2 are considered, and then the computations are repeated using cameras R1, R3, and O2. The two-camera-constellations L1-L3 and R1-R3 cannot be analyzed due to the fact that in these constellations the object points move exclusively within the epipolar plane. The derived values for the camera L1, L3, and O2 are listed in Table 3; the network error is the sum of the mean values for the three-camera constellations and is 0.000 ms. The camera R1, R3, and O2 yielded very similar results. The results show differences of up to 0.014 ms for the two-camera-constellations and 0.009 ms for the three-camera-constellations; these differences can be neglected.

Table 3. Asynchronism between the Cameras L1, L3, and O2 Calculated from the Two- and Three-Camera Constellations ; the Mean Value is Calculated from the Three Camera Constellations

Camera constellation	Δt_{Offset} [ms] between L1 und L3	Δt_{Offset} [ms] between L3 und O2	Δt_{Offset} [m between O2 und L1
L1-O2	–	–	−0.685
O2-L1	–	–	−0.696
L3-O2	–	1.505	–
O2-L3	–	1.519	–
L1-L3-O2	−0.810	1.507	−0.696
L3-L1-O2	−0.806	1.501	−0.695
O2-L1-L3	−0.813	1.510	−0.697
Mean value	−0.810	1.506	−0.696

Effects of Taking the Asynchronism into Account

In impact tests several distances between components of a car and a dummy are of special relevance. In the last test, the effects of neglecting the asynchronism in the analysis are investigated. For this purpose object point coordinates of two targets, one on the a-pillar of the car (APR2) and the other one on the dummy head (DHR2), are computed with and without considering the asynchronism (see Figure 13).

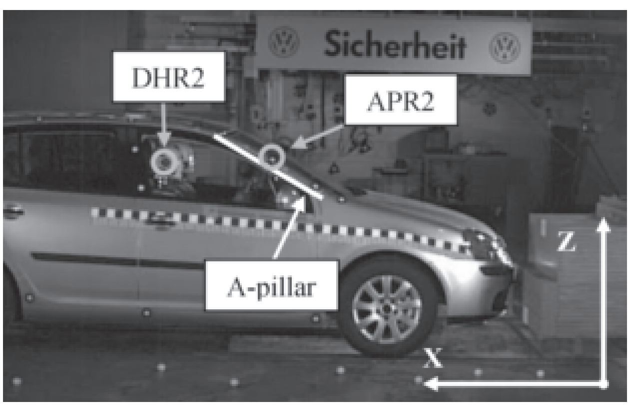

Figure 13. Image of the sequence from camera R1 with the two targets on the a-pillar of the car (APR2) and on the head of the right dummy (DHR2).

The comparison shows a significant difference in the Y-coordinate of DHR2 (see Table 4), whereas the X- and Z-coordinates of DHR2, and the coordinates of APR2 show no significant differences. The reason is that while APR2 lies in the sequences R1, R3, und O2, DHR2 is only depicted in R1 und R3. Since the movement occurs mainly in the epipolar plane of R1 and R3, DHR2 is systematically displaced by approximately 30 mm. As a result the distance between the two points also changes by 9.5 mm. Also the standard deviation of the DHR2 Y-coordinate is significantly higher than that of the other coordinates which is caused by the small base-to-height ratio.

The required object point accuracy for a three-dimensional image sequence analysis of an impact tests is about 5 mm. Thus, the difference in the coordinates of the object point at the dummy head of approximately 30 mm show the necessity for taking the asynchronism into account.

Conclusion and Outlook

In this article, a new method is presented, which permits the photogrammetric analysis of asynchronously acquired image sequences. The asynchronism is modeled by a temporal correction function in object space. It is then converted to an interpolation factor and is introduced into the functional model of the bundle adjustment. This extension of the bundle adjustment leads to a significant improvement of the results of the image sequence analysis.

In various tests we could show the advantage of the new approach: the object point accuracy was improved by factor of 10 and systematic errors due to displaced points could be detected and eliminated. Using the new approach the required object point accuracy for a three-dimensional analysis of a vehicle impact test of about 5 mm was reached, whereas if the asynchronism is not considered differences in the object coordinates of up to 30 mm were found. In future research the applicability of the new approach has to be further investigated. Tests will be carried out with different types of cameras, where the asynchronism cannot be assumed to be constant. Further tests will also address the simultaneous determination of the interior and exterior orientation and the asynchronism of a larger number of cameras. Tests are also planed with a varying exterior camera orientation, opening up the way to a multitude of additional possibilities. Nevertheless, we can state already at this point that the new approach is an important means to increase the accuracy of point determination in photogrammetric image sequence analysis to a level necessary in applications such as car impact testing.

Table 4. Comparison of the Calculated Coordinates of the Targets APR2 und DHR2 and the Distance between the Two Targets

Camera set-up		No. consideration of async.		With consideration of async.		Difference
O2-L1-L3-R1-R3		Coord.	Std.dev.	Coord.	Std.dev.	Coord.diff
	X	2087.5 mm	3.4 mm	2087.0 mm	1.0 mm	0.5 mm
APR2	Y	−3844.5 mm	5.7 mm	−3845.3 mm	1.7 mm	0.8 mm
	Z	−350.1 mm	4.1 mm	−350.0 mm	1.1 mm	−0.1 mm
	X	2785.7 mm	4.4 mm	2784.2 mm	1.2 mm	1.5 mm
DHR2	Y	−4117.3 mm	35.4 mm	−4145.9 mm	11.3 mm	28.6 mm
	Z	−324.2 mm	4.4 mm	−325.2 mm	1.2 mm	1.0 mm
s_0		12.2 µm (0.8 Pixel)		3.6 µm (0.2 Pixel)		–
Dist. APR2-DHR2		750.1 mm		759.6 mm		−9.5 mm

References

D'Apuzzo, N., 2003. *Surface Measurement and Tracking of Human Body Parts from Multi station Video Sequences*, Ph.D. dissertation, Technische Wissenschaften ETH Zürich, No. 15271, Mitteilungen/Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich, Nr. 81, Switzerland, 147 p.

Carceroni, R.L., F.L.C. Pádua, G.A.M.R. Santos, and K.N. Kutulakos, 2004. Linear sequence-to-sequence alignment, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27 June - 02 July, Washington, D.C., (1):746–753.

Caspi, Y, and M. Irani, 2000. A step towards sequence-to-sequence alignment, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 13–15 June, Hilton Head Island, South Carolina, 2:682–689.

Caspi, Y., and M. Irani, 2001. Alignment of non-overlapping sequences, *Proceedings of 8th International Conference on Computer Vision*, 09–12 July, Vancouver, B.C., Canada, 2:76–83.

Caspi, Y., D. Simakov, and M. Irani, 2002. Feature-based sequence-to-sequence matching, *Proceedings of Workshop on Vision and Modelling of Dynamic Scenes*, 27 May - 02 June, Copenhagen, Denmark, unpaginated CD-ROM.

Hastedt, H., T. Luhmann, and K. Raguse, 2005. Three-dimensional acquisition of high-dynamic processes with a single-camera system and stereo-beam splitting, *Proceedings of Optical 3-D Measurement Techniques VII* (A. Grün and H. Kahmen, editors), Institute of Geodesy and Geophysics, Vienna University of Technology, 2:175–184.

Irani, M., and P. Anandan, 2000. About direct methods, Vision algorithms: Theory and practice, *Proceedings of the International Workshop on Vision Algorithms 1999* (B. Triggs, A. Zisserman, and R. Szeliski, editors) 21–22 September 1999, Corfu, Greece, Vol. 1883 of LNCS, Springer-Verlag Berlin Heidelberg, pp. 267–277.

Kuthirummal, S., C.V. Jawahar, and P.J. Narayanan, 2002. Video frame alignment in multiple views, *Proceedings of IEEE International Conference on Image Processing*, 22–25 September, Rochester, New York, 3:357–360.

Lee, L., R. Romano, and G. Stein, 2000. Monitoring activities from multiple video streams: Establishing a common coordinate frame, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):758–767.

Lei, C., and Y.-H. Yang, 2005. *Line Feature based Multiple Video Synchronization*, Technical Report TR05-20, Department of Computing Science, University of Alberta, July, Edmonton, Canada, 8 p.

Luhmann, T., 2005. On the determination of object rotation and translation in 3d space (6 dof) by a single camera, *Proceedings of Optical 3-D Measurement Techniques VII* (A. Grün and H. Kahmen, editors), Institute of Geodesy and Geophysics, Vienna University of Technology, Vol. 2, pp. 157–166.

Maas, H.-G., 1992. *Digitale Photogrammetrie in der Dreidimensionalen Strömungsmesstechnik*, Ph.D. Dissertation, Technische Wissenschaften ETH Zürich, No. 9665, Mitteilungen / Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich, Nr. 50, Switzerland, 142 p.

Maas, H.-G., B. Hentschel, and F. Schreiber, 2003. An optical triangulation method for height measurements on water surfaces, *Videometrics VII - Electronic Imaging 2003* (S.F. El Hakim, A. Grün, and J.S. Walton, editors), 21–22 January, Santa Clara, California, SPIE Proceedings Series, Vol. 5013, pp. 103–109.

Maas, H.-G., and U. Hampel, 2006. Photogrammetric techniques in civil engineering material testing and structure monitoring, *Photogrammetric Engineering & Remote Sensing*, 72(1):39–45.

McClenathan, R.V., S.S. Nakla, R.W. McCoy, and C.C. Chou, 2005. Use of photogrammetry in extracting 3d structural deformation/dummy occupant movement time history during vehicle crashes, *Proceedings of SAE World Congress and Exhibition, Session: Safety Test Methodology and Structural Crashworthiness*, 11–14 April, Detroit, Michigan, SP-1937, pp. 81–88.

Narayanan, P.J., P. Rander, and T. Kanade, 1995. *Synchronous Capture of Image Sequences from Multiple Cameras*, Technical Report CMU-RI-TR-95-25, The Robotics Institute, Carnegie Mellon, University, December, Pittsburg, Pennsylvania, 11 p.

Oggier, T., M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, 2004. An all-solid-state optical range camera for 3d real-time imaging with sub-centimeter depth-resolution (SwissRanger™), *Optical Design and Engineering* (L. Mazuray, P.J. Rogers, and R. Wartmann, editors), 30 September - 03 October 2003, St. Etienne, France, SPIE Proceedings Series, Vol. 5249, pp. 534–545.

Pollefeys, M., L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, 2004. Visual modelling with a hand-held camera, *International Journal of Computer Vision*, 59(3):207–232.

Pooley, D.W., M.J. Brooks, A.J. Van Den Hengel, and W. Chojnacki, 2003. A voting scheme for estimating the synchrony of moving-camera videos, *Proceedings of IEEE International Conference on Image Processing*, 14–18 September, Barcelona, Spain, Vol. 1, pp. 413–416.

Putze, T., 2004. Einsatz einer Highspeedkamera zur Bestimmung von Geschwindigkeitsfeldern in Gasströmungen, *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation* (E. Seyfert, editor), Vol. 12, pp. 325–332.

Raguse, K., and M. Wiggenhagen, 2003. Quality parameters of the optical data channel used in impact tests, *Proceedings of Optical 3-D Measurement Techniques VI* (A. Grün and H. Kahmen, editors), Institute of Geodesy and Photogrammetry, ETH Zurich, Vol. 2, pp. 252–258.

Raguse, K., P. Derpmann-Hagenström, and P. Köller, 2004. Verifizierung von Simulationsmodellen für Fahrzeugsicherheitsversuche, *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation* (E. Seyfert, editor), Vol. 12, pp. 367–374.

Santel, F., W. Linder, and C. Heipke, 2003. Image sequence analysis of surf zones: Methodology and first results, *Proceedings of Optical 3-D Measurement Techniques VI* (A. Grün and H. Kahmen, editors), Institute of Geodesy and Photogrammetry, ETH Zurich, Vol. 2, pp. 184–190.

Schmidt, T., J. Tyson, K. Galanulis, D. Revilock, and M. Melis, 2005. Full-field dynamic deformation and strain measurements using high-speed digital cameras, *Proceedings of the 26th International Congress on High-Speed Photography and Photonics* (D.L. Paisley, S. Kleinfelder, D.R. Snyder, and B.J. Thompson, editors), 20–24 September 2004, Alexandria, Virginia, SPIE Proceedings Series, Vol. 5580. pp. 174–185.

Sinha, S.N., and M. Pollefeys, 2004. Synchronization and calibration of camera networks from silhouettes, *Proceedings of 17th International Conference on Pattern Recognition*, 23–26 August, Cambridge, United Kingdom, Vol. 1, pp. 116–119.

Spencer, L., and M. Shah, 2004. Temporal synchronization from camera motion, *Proceedings of the 6th Asian Conference on Computer Vision*, 27–30 January, Jeju Island, Korea, Vol. 1, pp. 515–520.

Stein, G.P., 1998. Tracking from multiple view points: Self-calibration of space and time, *Proceedings of DARPA Image Understanding Workshop* (G.E. Lukes, editor), 20–23 November, Monterey, California, Morgan Kaufmann, Vol. 2, pp. 1037–1042.

Torr, P.H.S., and A. Zisserman, 2000. Feature based methods for structure and motion estimation, *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms 1999 Proceedings* (B. Triggs, A. Zisserman, and R. Szeliski, editors), 21–22 September 1999, Corfu, Greece, Vol. 1883 of LNCS, Springer-Verlag Berlin Heidelberg, pp. 278–294.

Tresadern, P., and I. Reid, 2004. Uncalibrated and unsynchronized human motion capture: A stereo factorization approach, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27 June - 02 July, Washington, D.C., Vol. 1, pp. 128–134.

Tuytelaars, T., and L. Van Gool, 2004. Synchronizing video sequences, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 27 June - 02 July, Washington, D.C., Vol. 1, pp. 762–768.

Wedge, D., D. Huynh, and P. Kovesi, 2006. Motion guided video sequence synchronization, *Proceedings of Computer Vision - ACCV 2006: 7th Asian Conference on Computer Vision, Part II*

(P.J. Narayanan, S.K. Nayar, and H.-Y. Shum, editors), 13–16 January, Hyderabad, India, Vol. 3852 of LNCS, Springer-Verlag Berlin Heidelberg, pp. 832–841.

Whitehead, A., R. Laganiere, and P. Bose, 2005. Temporal synchronization of video sequences in theory and in practice, *Proceedings of 7th IEEE Workshop on Applications of Computer Vision and IEEE Workshop on Motion and Video Computing*, 05–07 January, Breckenridge, Colorado, Vol. 2, pp. 132–137.

Willneff, J., 2003. *A Spatio-temporal Matching Algorithm for 3D Particle Tracking Velocimetry*, Ph.D. Dissertation, Technische Wissenschaften ETH Zürich, No. 15276, Mitteilungen / Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich, Nr. 82, Switzerland, 123 p.

Wolf, L., and A. Zomet, 2002. Sequence-to-sequence self calibration, *Proceedings of Computer Vision - ECCV 2002: 7th European Conference on Computer Vision*, Part II (A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors), 28–31 May, Copenhagen, Denmark, Vol. 2351 of LNCS, Springer-Verlag Berlin Heidelberg, pp. 370–382.

Yan, J., and M. Pollefeys, 2004. Video synchronization via space-time interest point distribution, *Proceedings of the 6th Conference on Advanced Concepts for Intelligent Vision Systems*, 31 August - 03 September, Brussels, Belgium, 5 p., unpaginated CD-ROM.

Zhou, C., and H. Tao, 2003. Dynamic depth recovery from unsynchronized video streams, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 16–22 June, Madison, Wisconsin, Vol. 2, pp. 351–358.