Summary

The estimation of relative spatial positions and orientations is one of the most important tasks of engineering geodesy. For example, we need these parameters in precision farming or controlling the driving direction of construction vehicles.

It is usual to use multi-sensor systems in these applications which are often a combination of GPS-sensors with Inertial Navigation Systems (INS). An optimal solution for the searched parameters could be achieved using filtering processes.

Besides the mentioned sensor types, camera systems are increasingly used for navigation applications. With photogrammetric methods it is possible to estimate the relative positions and orientations of mobile vehicles using the image sequences of cameras.

The advantage of this method is that we can estimate the spatial parameters of a relative motion without including other sensors. Furthermore, cameras are as applicable in outdoor as in indoor environments and comparatively cheap. However, the small field of view using conventional cameras could be considered as a disadvantage.

To meet this challenge omnidirectional or spherical camera systems were developed. With these systems it is possible to take up to 360 degrees views of the surrounding area at each camera position. In contrast to usual camera systems a much higher amount of information can be taken into account to estimate the relative motion.

Mosaic- based camera systems are a special realization of an omnidirectional camera system. They are made of an aggregation of several monocular cameras which are placed in different recording directions.

Due to the overlapping areas, a spherical image can be generated. High-resolution images allow these camera systems to be used, for instance, in street mapping applications.

In our work, we developed a procedure to estimate the relative positions and orientations of a mobile vehicle by using a mosaic-based camera system. We apply the camera system *Ladybug* 3 which was developed by the company Point Grey Research. Five cameras are positioned in a horizontal ring and one camera on the top pointing vertically to the sky. The system will be rigidly connected to a mobile object.

Due to the absence of an unique projection center, explicit ambiguities in the generated omnidirectional images exist. This is why these spherical images are not usable for our applications. We decide to treat each camera individually and to estimate the unknown values with the help of the image sequences of each monocular camera.

We focus on the estimation of the relative motion between two successive positions of the *Ladybug* 3. We refer the unknown motion to a relative reference system which is defined in the center of mass of the *Ladybug* 3. We specify the unknown motion by a spatial rotation matrix and a spatial translation vector. To estimate these values the coplanarity constraint and the precise geometry of the *Ladybug* 3 is used.

We describe the theoretical approach to geometrically determine the unknown relative motion in the next three steps:

1. The construction of the projection rays of corresponding points.

   In the first step, we construct the projection rays of corresponding image points referring to each camera system. In this approach, the corresponding image points are our observations.

   With the help of the relationship between the camera coordinate system and the image plane (called the *interior orientation)* we form the projection rays geometrically. A projection ray connects the object point, the image point and the projection center of a chosen camera.
   For the mathematical representation of the projection rays we use *Plücker* coordinates. The advantage of this representation is that we can easily transform the rays in other coordinate systems. Furthermore, we can test geometrical relations in an easy manner.

2. The transformation of the projection rays in the reference system.

   In the next step, we transform the projection rays in the reference system of the *Ladybug* 3 since the unknown motion should refer to this defined system. To carry out this step, we need the transformation parameters between each camera system and the reference system which is an additional part of the *interior orientation.* The transformation is represented by a spatial rotation matrix and a spatial translation vector.

   In our case, we use the calibration parameters provided by the company Point Grey Research.

   Since we are using *Plücker* coordinates the transformation step is realized by an easy matrix-vector-product. With this step, we get all corresponding projection rays referred to the reference system.

3. The intersection of corresponding projection rays.

   In the last step, we require that the transformed corresponding projection rays intersect in the adequate object point which we refer to as the coplanarity constraint. For this purpose, we set up a mathematical condition and, in addition, consider the unknown relative motion of the reference system between two successive positions.

Based on the three above presented steps, we derive a nonlinear functional model between the observed image points and the unknown parameters.

We must consider that the coplanarity constraint is invariant with respect to the spatial translation vector of a monocular camera along the direction which is defined between the projective centres of two positions. Hence, we must consider the consequence of this issue to a rigid aggregation of monocular cameras. Here, the coplanarity constraint for all cameras may never be breached, too.

We observe that we can only estimate the direction of the translation vector by a straight forward drive of the mobile object. As soon as the mobile object rotates, we can determine the complete translation vector.

Since we do not exclude straight forward drivings as well as very small rotations we embrace an additional constraint. We require that the distance derived from the estimated spatial translation vector corresponds to an additional observed distance. We assume this additional observed distance to be given by a relative distance measuring system. Including this condition and a linearization of the functional model we set up a Gauss-Helmert Model with constraints of the unknowns.

We need approximate values for the unknowns which can be estimated by geometrical considerations directly or by taking into account the theoretical moving behavior of the mobile object.

Since we expect a high rate of bad correspondences we first perform a robust estimation using a RANSAC produce to detect a priori outliers. Because this produce is not effective enough we use an additional detection which is based on a Maximum-Likelihood-Type Estimation.

Our first experiments are based on simulated observations. We generate synthetic normally distributed image points for two admission situations for all six cameras of the *Ladybug* 3. We determine the relative postions- and orientation parameters according to the number of the involved cameras.

In addition, we simulate multiple measurements to derive empirical accuracies. Because the ground truth is given we can control if the estimate values of the adjustment are correct.

The theoretical accuracies of the orientation parameters grow with increasing number of involved cameras as was expected. The theoretical accuracies of the estimated distance is influenced by the accuracy of the additional observed distance.

Our required computing time (ca. 50 seconds for 1700 intersection conditions) makes clear that an almost real time capability cannot be achieved by the used software. The computing time can be reduced for instance by implementing the used software to a C++ algorithm.

The achieved results suggest a correct implementation of the developed concept. In particular, the significance tests support this statement. However, since outliers were not modeled the robustness of the implementation could not be checked.

In the next step, we test the extended implementation on the basis of real image sequences which were recorded with a multi-sensor system. It consists of the *Ladybug* 3, an Inertial Navigation System (INS), a GPS-sensor and a relative distance measuring system. A time referencing of all measuring data is realised.

Due to image processing and detection and matching of prominent image points an uniform workflow from the data transmission to the adjustment of the unknowns could not be achieved.

To get corresponding points we use the developed software *Aurelo* which applies the *SIFT-operator* to the detection. The outlier proportion is up to 63%.

We achieve an average accuracy of ca. 3 millidegrees for the yaw, pitch and roll angles which we did not anticipate. We observe that the accuracy of the orientation is mainly dependent on the quantity of the intersection conditions which was already shown in the simulation estimations. Since in some cases only a few intersection conditions (less than 20) are provided, the adjustment becomes very unstable and the estimate values have a bad precision.

The average accuracy of the relative distance is 0.8 cm which coincides with the given accuracy of the relative distance measuring system.

With the help of the INS, we achieve independent reference measurements which we compare to our estimated relative orientations.

The estimated Euler angles show similar characteristics as the reference Euler angles which is at first satisfying. Nevertheless, there is a different qualitative behaviour of the several Euler angles.

Taking advantage of the absolute deviations between the estimated yaw angles and the reference values, we achieve an average deviation of 0.05 degreees which we consider a good result. In contrast, we achieve an average deviation of 0.11 degrees and 0.12 degrees for the pitch and roll angles.

Herefrom, we assume systematic differences between the pitch and roll angles referred to the inertial measured values. We cannot ultimately identify the reason for these differences.

Possibly, a different definition for chaining the element rotation matrices is available which would define the whole rotation matrix in a different way.

Therefore, a confirmation of the theoretical precision is not available.

In the final step, we estimate the relative position and orientation parameters by different image resolutions (using different image pyramids) of the same sequence. We achieve a nominal-actual value comparison.

The results of the significance tests yield that the differences between the parameters are significant in many cases. This leads us to the conclusion that an optimal robust adjustment could not be realized.

Therefore, the reasons could be a false outlier detection and possible significant errors in the given transformations between the camera systems and the reference system.