The definitive version is available at http://diglib.eg.org/EG/DL/WS/EGVE/JVRC12

Indoor Tracking for Large Area Industrial Mixed Reality

Fabian Scheer[1] and Stefan Müller[2]

[1]Daimler Protics GmbH,[2] University of Koblenz - Landau

Abstract

For mixed reality (MR) applications the tracking of a video camera in a rapidly changing large environment with several hundred square meters still represents a challenging task. In contrast to an installation in a laboratory, industrial scenarios like a running factory, require minimal setup, calibration or training times of a tracking system and merely minimal changes of the environment. This paper presents a tracking system to compute the pose of a video camera mounted on a mobile carriage like device in very large indoor environments, consisting of several hundred square meters. The carriage is equipped with a touch sensitive monitor to display a live augmentation. The tracking system is based on an infrared laser device, that detects at least three out of a few retroreflective targets in the environment and compares actual target measurements with a precalibrated 2D target map. The device passes a 2D position and orientation. To obtain a six degree of freedom (DOF) pose a coordinate system adjustment method is presented, that determines the transformation between the 2D laser tracker and the image sensor of a camera. To analyse the different error sources leading to the overall error the accuracy of the system is evaluated in a controlled laboratory setup. Beyond that, an evaluation of the system in a large factory building is shown, as well as the application of the system for industrial MR discrepancy checks of complete factory buildings. Finally, the utility of the 2D scanning capabilities of the laser in conjuction with a virtually generated 2D map of the 3D model of a factory is demonstrated for MR discrepancy checks.

Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Information Interfaces And Presentation]: —Artificial, augmented, and virtual realities I.4.8 [Image Processing And Computer Vision]: —Tracking

1. Introduction and motivation

In the last decades various tracking systems were presented to track the pose of a video camera in an indoor environment. Some are capable of covering more than just a limited volume, e.g. [SG05, NIH01, WBV*99], but the coverage of complete buildings with several floors is still a big challenge. A lot of different technologies have already been applied to solve the pose tracking problem, like ultrasonic, acoustic, mechanical or electromagnetic devices, optical and vision based systems or WiFi-based components [RBG00]. Each introduces its own opportunities and drawbacks and a user has to judge carefully to decide on a suitable solution for a given problem. In general, the user will be spoilt for choice to weigh between accuracy, installation and set-up times, costs, environmental requirements or superstructures and the usability of available systems.

In the context of industrial MR applications in really large areas, like factory buildings with a base area of several hundred square meters and multiple levels, further special requirements are posed on a camera tracking technology, that can be sumed up as follows:

- Coverage of several hundred square meters.
- Accurate tracking results, at least lying in acceptable tolerance levels.
- Minimal installation time and effort.
- Minimal set-up times of the system concerning transport, calibration or learning stages. Ideally, repeating efforts like recalibration or relearning stages should be avoided.
- At least no or minimal required changes to the environment, e.g. no placement of hundreds of markers.
- Flexibility, usability and reusability of the system.
- Costs of the system components.
- A cost efficient balance concerning the implementation of the former factors.

Against this background a lot of existing tracking technologies are difficult or even impossible to apply. Especially when rapidly changing environments, like a factory, are focused, where reconstructions can occur every day, enduring environmental modifications are hard to achieve or unwanted, a lot of occluding geometry exists and the daily workflow should not be impeded. Tracking in really large areas is a common problem in the robotics community and for several years now simultaneous localisation and mapping (SLAM) approaches have achieved an enormous progress to solve the problem for an autonomous navigation in an unknown scene (see [DWB06]). Even though SLAM has a big potential to solve the large area tracking (LAT) problem in the future, SLAM based systems have no deep understanding of the environment [KM07]. For an augmentation of location-dependent information an additional registration to a global coordinate system (CS) is mandatory. For SLAM a map of the environment is build simultaneously. Similar to natural feature tracking (NFT) algorithms for MR, e.g. [KS07, KM07], rapidly changing environments are a challenge and require a rebuild of the feature map in the worst case. Thus, additional set-up times may be introduced. Beyond that, the accuracy of recent SLAM systems for arbitrary environments is often not that precise as MR overlays would need [KSD*09].

Motivated by these issues we present a large area tracking system that fulfills industrial demands. A mobile system consisiting of a carriage equiped with an infrared laser based tracking system, a touch-monitor for collaborative sessions and a video camera mounted on a servo motor is used to cruise through large areas. A 2D laser tracking is extended to track the pose of a camera by a coordinate adjustment method. The laser relies on an inital measurement of sparsely set reflector targets. An evaluation of the system in a controlled lab setup and in a factory under real application conditions is presented. Finally, the application of the system for industrial MR discrepancy checks is shown.

2. Related work

Marker tracking methods as [WS07] can be used with multi-marker configurations to track a camera in a specific volume, but depend on having at least one marker in the field of view within a suitable small distance for a reliable detection. A huge modification of the environment by placing a sufficient amount of markers is required to achieve a result that is accurate enough for large tracking volumes.

[KS07] reduce the amount of necessary markers for larger areas by using a natural feature tracking in the space between two markers. Thus, they can locally avoid the line of sight problem, but suffer from drift problems. Still huge amounts of markers are necessary and a learning stage for natural features has to be applied. Since the approach relies on precalibrated features and feature correspondences, a rapidly changing environment is hard to handle and introduces repetitions of the feature learning stage.

[WBV*99] introduce an optical tracking system using active targets. The covered volume depends on the amount of panels with infrared light emitting diodes (LEDs), which are mounted at the ceiling. They report very precise tracking results. 3000 LEDs are used to cover an area of approximately 5.5 by 8.5 meters, implying huge installations and set-up times for large areas.

[SS08] uses the ARTTRACK system [Adv10] with 4 cameras to conduct an out-side-in tracking of a video camera mounted on a mobile cart in a limited quadratic volume of approx. 10 x 10 meters. To cover large areas the area is separated into the 10 x 10 m quadratic volumes (hot-spots) and the whole system has to be transported to the different hot spots, introducing major set-up times including transport and calibration procedures for each hot-spot.

Inspired by SLAM methods of the robotics community [KM07] present parallel tracking and mapping for small work places. Unlike SLAM methods, the mapping and tracking task is split in two separate threads allowing the application of computationally expensive batch optimisation techniques. In [CKM08] this technique is extended by using multiple maps to cover bigger environments, e.g. different rooms in a building. For a synthetic test scenario the accuracy of a calculated trajectory compared to ground truth had a standard deviation (SD) of 6 mm. Unfortunately, no evaluation of the accuracy in a real world scenario was conducted. The approach is promising for the future, but for a continuous indoor tracking in large areas issues like the management of massive feature data sets, the handling of rapidly changing environments and the achieved tracking precision have to be further investigated.

[JISD06] present a hybrid system. A handheld-PC is tracked via markers in the environment. If no marker is at sight an ultra-wide-band(UWB) tracking [SG05] is used for the position and an inertial sensor for the orientation. Even though large areas can be covered the precision of the UWB tracking is about 10-15 cm.

In [NIH01] a hybrid system consisting of an ultrasonic tracking combined with an inertial sensor is presented, beeing able to cover a large volume. Low update rates of the ultrasonic tracker are compensated by high update rates of the inertial sensor, whose long-term-drift is reversely corrected by the ultrasonic tracking updates. The system relies on a lot of ultrasonic receivers to cover a large area and suffers under the typical restrictions for acoustic tracking devices [RBG00].

In [FHP98] the hybrid tracking system Constellation is presented. It uses an inertial sensor and accelerometers for the tracking. To compensate the positional drift frequent updates of an ultrasonic tracking are used. Beeing conceived for a building-wide tracking the commercial version was typically used for room sized installations [WFK07]. Building-wide installations require severe set-up times and a lot of acoustic devices. Furthermore in an industrial factory setting interferences with other technologies , e.g. acoustic cleaning or measurement systems, may occur.

In [WFK07] an advanced system based on the inertial tracking of [FHP98] is introduced, aiming to allow a building-wide tracking. It is a hybrid system and fuses inertial and vision based data. A 2-5 mm RMS position and a 0.2° RMS orientation accuracy are obtained, but therefore a lot of passive markers have to be placed in the environment. [BHAP11] show a laser based tracking system for areas of 40 x 80 meters. Larger areas are partitioned into blocks of 40 x 80 meters. The set-up time and registration to a global CS takes about 20 minutes per sector. Since the laser tracking relies on a line of sight to a target attached to the camera, the system is primarily suited for widely opened areas. Elsewise, in densely developed areas, like running factories, the laser has to be repositioned when occlusions occur, thus introducing additional set-up and registration times.

	NAV200		NAV350	
Avg. distance	**position**	**angle**	**position**	**angle**
3 m	± 4 mm	±0.1°	-	-
5 m	± 8 mm	±0.1°	-	-
10 m	± 12 mm	±0.1°	4 mm	±0.1°
20 m	± 15 mm	±0.1°	10 mm	> ±0.1°
28.5 m	± 25 mm	±0.1°	-	-
30 m	-	-	15 mm	±0.25°
50 m	-	-	25 mm	> ±0.25°

Table 1: *Specification of the laser tracker precision [SA12].*

3. Overview mobile tracking sytem

Figure 1: *AR-Planar with the calibration pattern. The video camera and the servo motor (mid). To attach the cylindrical targets surveying elements are used (right). The target height is telescopically adjustable to ease the transport. A correct alignement is achieved by a water level attached to the rod (see mid bottom). The targets have a diameter of 10 cm.*

The concept of the carriage like mobile device, called AR-Planar, is based on the work of [SS08]. The poster [SM10] presented ongoing work and introduced the laser tracker and the servor motor for the view adjustments to the initial AR-Planar platform. This paper presents an enhanced nearly ready for production version of the AR-Planar device, that is extended by the enhanced laser NAV350 and an overall enhanced component design, including an enhanced mount of the camera to the servo motor. The monitor depicts a view into the virtual or augmented world for collaborative sessions of multiple users. The device consists of a robust chassis and attachment for the laser and the video camera, a wireless keyboard, a large pen-based touch display for collaborative sessions, a video camera with a resolution of 1024x 768, an attachement with a servo motor, pluggable automotive batteries to maintain a round-the-clock mobile usage, electronic elements to control the servo motor and the laser tracking device (see fig. 1).Tracking in really large environments also requires a suitable data supply and a handling of massive data sets, because of the possible augmentation at any place in the large area. The approach of out-of-core visibility guided rendering (VGR) [KBHP07] is able to handle and display millions to billions of polygons and gigabytes of data in highly interactive frame rates by reloading necessary data from the harddrive. This method is used in this paper being able to augment 3D models of whole factories, consisting of several million polygons, e.g. >50 millions. To compare large virtual 3D models with large areas in the real world, a transparent adjustment is used by gradually blending between the two worlds. In order to avoid computationally expensive reorderings of the triangles for a correct transparent representation, the video image is rendered by an orthographic projection on a view aligned 2D plane lying at the near plane of the actual view frustum. By simply blending the video image of this plane expensive reorderings can be avoided and both worlds are easily compared by an adjustment of the α-value. Thus, large 3D models in large real world areas can be displayed at interactive frame rates.

The used infrared laser systems [SA12] NAV200 and NAV350 are commercial off-the-shelf products and typically used in the logistics sector for driverless transport systems. Therefore a 2D tracking is sufficient. The laser rotates 360° with $8 \pm 5\%$ hertz. The laser tracks retroreflective cylindrical or flat targets in the environment and compares them with a target map to compute its 2D position and orientation by bundle adjustment methods. The target positions for the map can either be successively scanned by the laser or measured by a seperate device, e.g a theodolite. The accuracy of the laser tracking reaches millimeter precision and depends on the distance of the laser to the targets. Table 1 shows the manufacturer informations concerning the tracking accuracy for the lasers NAV200 and NAV350. Since we present the results of a longer research period, both laser trackers are used for the evaluations in this paper. The newer device NAV350 is advanced in accuracy and is also able to deliver a 2D contour scan of the environment. The NAV200 is used for the evaluation in the lab. Currently the AR-Planar is equiped with the advanced following device NAV350 that is used for the evaluation in a large factory. The NAV200 has a maximal tracking range of 28.5 m to a target and the NAV350 a range of up to 70 m (see table 1). The NAV200 can internally

handle 1280 and the NAV350 up to 12000 targets. Thus, assuming an average target distance of 10 to 15 m to obtain a suitable 2D tracking precision, very large areas of several hundred square meters can be covered. A fixed placement of the targets at specific positions modifies the environment. For a setup in a lab this is acceptable and also used for the evaluation in the lab in this paper.
However in a productive rapidly changing environment, like a factory, targets may be overpainted or dismantled due to necessary reconstructions and the attachement to machines or pillars can afford the usage of more targets since occluding geometry like pipes, cables and auxiliary constructions exist around the attachment points. On this account, for the usage of the system in productive environments a flexible and mobile setup of the targets is chosen. At a surveying point a pillowed rod with a cylindrical target is placed and aligned by checking a water level attached to the rod (see fig. 1). The laser can only compute its position if three targets are visible. Due to occlusions and to have an overdetermined setting to guarantee an area-wide tracking six to eight targets are sufficient to use the system in a factory. Once a target is far away or permanently occluded from the actual position, it is picked up and placed at a new surveying point along the driveway. Hence, with this mobile target design a large area tracking can be conducted with only six to eight targets and the flexible placement of the targets at specific surveying points allow very short set-up times. The laser is mounted at a height of approximately 2,23 m at the AR-Planar. Since the laser scans in a 2D plane, occlusions by persons, moving vehicles and other geometry can be avoided.
By moving the AR-Planar around the view can be manipulated in a horizontal direction. To also allow a vertical view manipulation the video camera is mounted on a servo motor [Gra12] and can be adjusted out of the application.

4. Coordinate system adjustments

Tracking in large environments requires the registration of a tracker to a global CS in order to obtain a precise location-dependent augmentation. The used laser merely passes 2D tracking values. A correct augmentation requires a six DOF pose of the video camera, that can be achieved by an extension of the 2D tracking with the correct geometric transformation between the laser and the image sensor.

4.1. Registration to a global coordinate system

According to the tracking environment the layout of the target positions should be designed carefully. It is obvious that geometrically dense populated areas require more targets than widely opened areas, due to occuring occlusions. Furthermore, it should be checked if permanently attached targets are possible or if mobile target setups can be used. In this paper we present both a permanent and a mobile target attachment. Fixed positions are used for the evaluation of the tracking system in a lab.
Once the target positions are defined they are measured by a theodolite. For the fixed targets at least three points have to be measured, e.g. on a pillar, to determine a surface normal. By this normal the mid point of an attached cylindrical target with a specific radius can be computed. The mid points are used to set up a 2D target map that is further used as input for the laser tracker. The laser device internally compares actual measurements of detected targets with this map and computes its position and orientation by bundle adjustment methods. Thus, the laser outputs tracking values in the CS of the target map. This can be exploited to register the laser tracking to a global CS by simply setting up the map in the global CS.
For industrial scenarios, factory premises are referenced in a global CS. This CS can be used to setup surveying points by theodolite measurements in buildings on the factory premises. Since this system is also used for the CAD 3D models, a map of the surveying points is sufficient to register the laser tracking to this global CS. To reuse the surveying points non disturbing anchor bolts can be set in the floor. They can be reused as long as the floor of the factory exists. Thus only a one-time effort is needed to setup the registration to a global reference system.

4.2. Registration of the laser to the image sensor

By setting up the target map in a global CS the tracking of the laser is globally registered, but merely delivers 2D position and an orientation value. To track the pose of the video camera attached to the AR-Planar, the relation of the image sensor to the laser scanner has to be known for all possible view adjustments of the camera by the servo motor. Since the view adjustment is done by a seperate control and not by pointing at a desired location common for handheld MR approaches, e.g. [JISD06], the adjustable angularities can be limited to some defined positions, e.g. discretization of the half-space in front of the AR-Planar in steps of $20°$. This reduces the complexity and calibration effort of the system and in consequence the set-up time.
To register the coordinate sytem of the laser and the image sensor for the defined view perspectives a calibration pattern is used. During the registration the AR-Planar is placed in such a distance to the pattern that at least four markers are visible for every adjustable perspective. The calibration pattern is depicted in fig. 1. It consists of two flat retroreflective targets on top of the pattern to determine the position and orientation in relation to the laser tracker. To obtain the relation between the pattern and the camera a marker tracking [WS07] is used. A multimarker setup is applied, assuring that for every camera view adjustment at least one marker is at sight and to minimize tracking errors by the additional information of tracking multiple markers. During the registration the laser measures the midpoint of both flat targets. Since the geometric relation between this mid point and the origin of the multimarker setup (fig. 1, top left marker) is known due to the construction of the pattern, the transforma-

tion between the laser and the image sensor can be computed for every defined adjustment of the camera perspective.

4.2.1. Algorithm overview

First the two flat targets are measured by the laser and the 2D position and the orientation of their midpoint are computed in the laser CS. These 2D informations can be extended to a 3D transformation matrix by

$$T_L = \begin{pmatrix} & & t_x \\ & R_y & 2230 \\ & & t_z \end{pmatrix} \tag{1}$$

where R_y is the rotation matrix in Euler angles of the laser in the xz-plane of the 3D world CS. To map the laser coordinates to the world CS the measured 2D position is extended by a height component according to the height of the laser on the AR-Planar at 2,23 meters. Then, the first camera perspective is adjusted by setting the according step width of the servo motor and all markers in the camera view are tracked. The marker tracking is repeated for a user given repetition value, e.g. 50 times, and the poses for a specific marker are identified by a unique marker ID. To reduce any jitter, common for marker tracking, the mean of the translational $(\vec{t})$ and rotational part (R) of the poses for each marker is computed. Thereby spherical linear interpolations (slerp) are used for the rotational part, yielding

$$T_{marker} = \begin{pmatrix} slerp(R_{1...n}) & \frac{1}{n}\sum_{i=1}^{n}\vec{t} \end{pmatrix}, \tag{2}$$

as an averaged pose for a marker, where n denotes the number of repetitions. The averaged poses refer to the midpoint of each visible marker. Since the geometric relations between the markers are known due to the multimarker setup, the averaged poses are further improved using a least-squares-optimisation with the observed average poses and the multimarker model as input. The resulting residuals indicate the quality of the marker tracking. If large residuals occur, the process can be repeated until satisfying values are obtained. The resulting pose T_{mm} describes the origin of the multimarker setup (fig. 1, first top left marker) in the camera CS and the inverse $T_{mm}{}^{-1}$ the camera in the multimarker CS. In order to have all transformations in the multimarker CS the translational offset (known by the pattern geometry) between the computed midpoint of the retroreflective targets and the multimarker CS is added by

$$T_{L_{mm}} = T_L T_{offset}{}^{-1}. \tag{3}$$

Based on this information the final transformation describing the relation between the laser scanner and the image sensor for a camera perspective can be computed:

$$T_{L \Rightarrow IS} = T_{L_{mm}} T_{mm}{}^{-1} \tag{4}$$

During tracking $T_{L \Rightarrow IS}$ is used as an offset to extend the 2D laser tracking values to obtain a 3D transformation describing the six DOF pose of the video camera. Repeating this coordinate adjustment for all camera perspectives yields a list of transformations. When a user changes the camera view during runtime the servo step width is determined and the according matrix from the list is loaded as the actual offset. Before the respective offset is added during runtime the 2D laser tracking values are extended by a height component, similar to the coordinate adjustment stage. For ground floors this is the height of the laser on the AR-Planar at 2,23 m. For other levels in a building the 2,23 m must be added to the height of the respective floor. By this way the six DOF pose of the video camera can be computed during runtime.

5. Evaluations and results

To evaluate the system two setups were chosen. First the system was evaluated with the NAV200 in a controlled lab setup, but still simulating industrial application conditions. Due to the controlled setup, the several error sources of the system can be analysed in detail. The second evaluation was conducted with the new and more precise laser NAV350 in the large area of a factory under real application conditions.

5.1. Evaluation in the laboratory

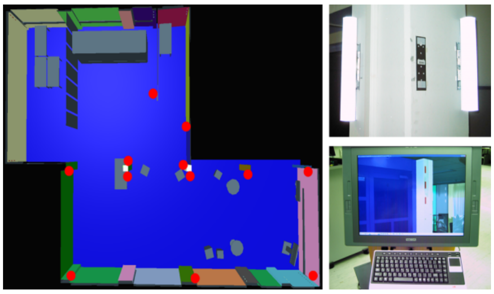

Figure 3: *3D model of the laboratory with targets marked as red circles (left). Flexible target attachment at pillars (right top) and a live image from the evaluation (right bottom).*

The base area of the lab is about 20 x 18 meters (fig. 3). Fixed reflector attachments at pillars are used with glued metal strips. 12 targets are used to avoid occlusions and to have more than three targets in sight for every possible position. Even though the laser manufacturer recommends a diameter of $> 80mm$ and a height of $> 500mm$ for cylindrical targets in combination with the NAV200, we used targets with a diameter of $50mm$ and a height of $350mm$ to stress test the system (fig. 3). To obtain merely slight modifications of the environment the targets are plugged to the fixed metal strips by hook-and-pile fasteners. This allows an easy assembly and disassembly and was primarily favored to place targets in industrial environments. The corner points of the metal strips are measured with a theodolite, to generate the required 2D target map for the laser. Furthermore, they serve as ground truth data for the evaluation of the augmentation, since the dimensions of the rest of the lab were manually

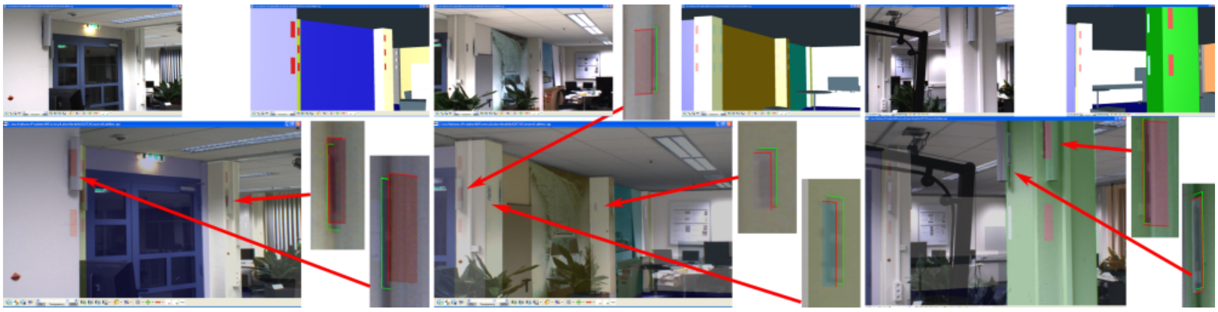

Figure 2: *Three perspectives of the lab evaluation. For every view the video image, the virtual image and a bigger overlay image are shown. The overlay of the metal strips is magnified and red arrows point to the associated image regions. For a better observation of the pixel difference the metal strips are outlined in green and the virtual overlay in red in the magnification.*

Servo	$\mu(x,y,z)$	$\mu(\alpha,\beta,\gamma)$	$\sigma(x,y,z)$	$\sigma(\alpha,\beta,\gamma)$
$-30°$	(159.7, -936.1, 640.4)	(-30.3,2.4,0.2)	(3.4, 2.6, 2.2)	(0.2, 0.2, 0.1)
$-19°$	(156., -928.6, 638.1)	(-19.2,2.4, 0.1)	(4.2, 3.0, 1.2)	(0.2,0.4,0.1)
13°	(156.6, -899.2, 643.2)	(13.1,3.3,0.0)	(4.2, 5.8, 2.0)	(0.5,0.4,0.1)
26°	(154.9, -893.4, 651.0)	(26.9,3.5,0.0)	(2.4, 3.9, 2.7)	(0.3,0.1,0.1)
40°	(155.0, -886.9, 659.7)	(40.7,3.8,0.3)	(2.7, 5.7, 3.6)	(0.3,0.1,0.0)

Table 2: *Means and SDs of the marker tracking for used camera perspectives. Positions (x,y,z) are given in millimeters and the angles (α,β,γ) in Euler representation.*

measured by a handheld laser distance measurement and a folding rule. Thus for an evaluation of the tracking accuracy the pixel difference between the augmented 3D model and the real metal strips in the video is the most reliable source. The 3D model was built with 3D Studio Max in millimeter precision. Unfortunately, the file conversion to the VGR format leads to observed errors of up to 2 mm. Since camera and laser are not synchronised the augmentation can be delayed during fast movements. The adjustable views of the camera by the servo motor are limited to predefined positions (see table 2). Table 2 depicts the mean values of the tracked marker poses and their SD. The distance of the two targets on the calibration pattern is 64 cm. 30 measurements with the NAV200 at a distance of approx. 1 meter of the AR-Planar to the pattern yielded $\mu(636.4mm)$ with $\sigma(0.9mm)$. Together with table 2 these values give an overview of the dispersion about the mean of the laser and the marker tracking during the coordinate adjustment process. Overall the accuracy of the system is influenced by several already described error sources that are listed as follows:

- Theodolite measurement, error of up to 2-3 mm.
- Errors occuring during the file conversions of up to 2 mm.
- The manual attachment of the targets with the hook-and-pile fasteners introduces an error of approx. 1 to 5 mm.
- Errors of the coordinate adjustment between the laser and the camera, including marker and laser tracking errors.
- Possible errors of the camera intrinsic parameters.
- Errors introduced by the laser tracker (see table 1). It should be noted that the chosen small diameter of the targets can also cause an additional error.

Due to the several error sources, whereof some can occur more or less distinctive and some are location dependent, an exact appointment of the accuracy is difficult. Nevertheless, the pixel difference of the virtual overlay on the video image can be used to determine the accuracy of the system. Since the metal strips are measured by a theodolite with an error of only 2-3 mm and are also modeled in 3D, the evaluation concentrates on the assessment of their overlay precision.
Figure 2 depicts three screenshots of the evaluation. Each metal strip has a width of 40 mm and a height of 160 mm. The image regions with metal strips are magnified and show that the overlay difference lies in the range of nearly 0 to 30 mm. Overlay precisions in this range were observed for the whole evaluation in the lab. Thereby, the AR-Planar was moved in the whole lab and the pixel differences of the overlay were evaluated. We also tested the precision with intentionally bad placed targets and could directly see a degradation of the overlay precision. For industrial factory planning scenarios the accuracy lies in the tolerance level of up to 5 cm. For the evaluation the NAV200 was used. The following advanced laser device NAV350 is evaluated in the next subsection and achieves a better tracking precision and therefore a better overlay quality.

5.2. Evaluation in a factory

The factory building in which the evaluation under applications conditions was conducted has a base area of 250 x 160 meters and seven floors. The according 3D model of the building consists of 50.8 million polygons and takes 4.5 gigabytes on harddrive. Six telescopically adjustable mobile targets (fig. 1) are used and placed at surveying points on the floor that are measured by a theodolite. The surveying points are used to build a 2D target map for the NAV350 laser. We experienced that six mobile targets are almost sufficient for an area-wide tracking, but recommend to use eight targets in geometrically dense populated areas. The arrangement and

Figure 4: *Evaluation of the system in a factory. The virtual, overlayed and video image are shown.*

distance of the targets to each other are geared to the pillar grid on a shop floor, that typically lies in the order of approx. 10 to 20 meters. In contrast to the lab evaluation the file conversions between CAD tools and the VGR format introduced no additional error. The correct placement of the mobile targets is checked by aligning a water level. Thus, an uncareful placement can introduce an error.
The accuracy of the NAV350 is advanced in contrast to the NAV200. The range finding of the laser is afflicted with a systematic, e.g. depending on temperatures, and a statistic error. The compensation of systematic errors can be complex and hard to achieve, but statistic errors can simply be reduced by smoothing repeated measurements. Since users tend to stand at a fixed position during the examination of certain issues with the AR-Planar, a smoothing of repeated measurements can easily be accomplished. Therefore, a filter working on a tracking queue was implemented. If the five last values do not differ for a certain position or angular treshold the AR-Planar stands still and a queue of tracking values is filled. When 30 values are captured the average position and angular result of these values is set as the actual tracking value. Thus, the statistical noise is reduced, resulting in an overlay with nearly no jitter. Three camera perspectives are registered to the laser and very small residuals of $(0.2, 0.1, 0.2), (0.1, 0.2, 0.4), (0.1, 0.4, 0.1)$ are observed. The SDs are observed in the submillimeter range. Similar to the metal strips in the lab, for the factory, pillars and the targets itself can be used as ground truth data to assess the tracking accuracy by the overlay. For surveying points spheres at a height of 2,23 m are placed in the 3D model and should overlay targets in the video image. In general the position of pillars in a factory and the according 3D model are very reliable, so that they can be used to assess the overlay precision. Figure 4 depicts screenshots of the evaluation. The advanced precision of the NAV350 is noticeable at the pillars and the overlay of the spheres with the targets. In general the overlay accuracy in the near range nearly fits perfect, but in areas far away rotational errors of the laser to image sensor adjustment can lead to deviations. This is obvious, but it should be noted that errors in the orientation of the computed pose can cause overlay errors which linear grow with the distance of an observed object to the AR-Planar (also see [SS08]). Thus, serious judgements should only be made for objects in the near or mid range to the AR-Planar. We experienced that the precision in the near and mid range mainly depends on a correct alignment of the targets at the surveying points. Furthermore, the registration between the laser and the image sensor has to be done accurately to obtain a precise overlay. Considering these issues, we experienced the overlay accuracy in the tolerance levels of factory planning applications of 5 cm. With an accurate setup and registration the accuracy lies in the range of 0 to 20 mm for near and mid expectation ranges. A carefully planned target layout is essential to guarantee an area-wide tracking. Nevertheless, for our setup we still encountered some regions where further surveying points are required due to occlusions.

6. Industrial applications of the system

The system can be used for MR discrepancy checks of complete factory buildings, consisting of several hundred square meters. Building acceptances, like [SS08], can be conducted as well as a quality management of the CAD models of the factories, e.g. when a model was built based on 3D scans. Found discrepancies can directly be annotated via touch input and are documented as textures in the 3D model (fig. 5). By storing the viewpoint, the annotations and the video image of an examination, the checks can later on be repeated in a bureau for the specific location without the need to visit the building again. Since, the laser also captures a 2D contour scan of the environment, this data can be location-

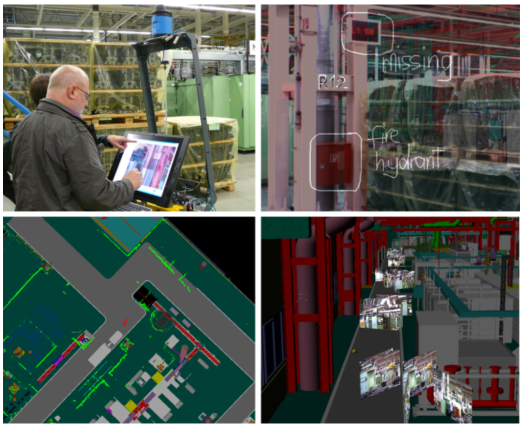

Figure 5: *Application of the system for discrepancy checks (top right). Scan data (green dots) overlayed over a map of a factory (bottom left). Conducted discrepancy checks are documented in the 3D model (bottom right).*

dependently overlayed over a rendered 2D map, as depicted in fig. 5. Thus, a user can easily and fast detect wrongly placed objects by comparing the scan with the map.

7. Conclusions

This paper presents a method to track the pose of a video camera in very large indoor environments. A coordinate adjustment method is shown to compute the transformation between an image sensor and a 2D laser tracker. An enhanced version of the mobile monitor-based AR-Planar is presented to collaboratively interact with virtual data in a large tracking volume. The system is evaluated in a controlled lab setup and in a large factory. For the latter the achieved tracking precision allows an accurate augmentation in the near and mid range. The obtained accuracy lies in tolerance levels of factory planning applications and is well suited for large area industrial MR applications. Beyond that, the system opens the door for new research and a variety of MR applications in really large environments.

Acknowledgements

The authors would like to thank Jörg Schad, Sven Gaukler and Martin Schlitter for their support in the factory evaluation and Ralph Schönfelder for including the described improvements and the tracking system to the AR-Planar platform.

References

[Adv10] Advanced Realtime Tracking GmbH (A.R.T.): ARTtrack1, http://www.ar-tracking.de, May 2010.

[BHAP11] Bade C., Hoffmeyer A., Alberdi A., Paul G.: Entwicklung und Einsatz eines AR-Systems mit laserbasiertem Large-Area-Tracking. In *10. Paderborner Workshop Augmented & Virtual Reality in der Produktentstehung* (2011).

[CKM08] Castle R. O., Klein G., Murray D. W.: Video-rate localization in multiple maps for wearable augmented reality. In *Proc 12th IEEE Int Symp on Wearable Computers, Pittsburgh PA, Sept 28 - Oct 1, 2008* (2008), pp. 15–22.

[DWB06] Durrant-Whyte H., Bailey T.: Simultaneous localization and mapping: part I The essential algorithms. *IEEE Robotics & Automation Magazine 13*, 2 (June 2006), 99–110.

[FHP98] Foxlin E., Harrington M., Pfeifer G.: Constellation: a wide-range wireless motion-tracking system for augmented reality and virtual set applications. In *Proc. of the 25th ACM SIGGRAPH* (1998).

[Gra12] Graupner: DSR8801, http://www.graupner.de, 2012.

[JISD06] Joseph N., István B., Stürzinger A., Dieter S.: Wide area tracking tools for augmented reality. In *Adjunct Proc. of the 4th International Conference on Pervasive Computing 2006* (2006), pp. 143–146.

[KBHP07] Kasik D., Brüderlin B., Heyer M., Pfützner S.: Visibility-guided rendering to accelerate 3d graphics hardware performance. In *ACM SIGGRAPH '07 courses* (2007).

[KM07] Klein G., Murray D.: Parallel tracking and mapping for small AR workspaces. In *Proc. 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'07)* (2007).

[KS07] Klopschitz M., Schmalstieg D.: Automatic reconstruction of wide-area fiducial marker models. In *ISMAR '07: Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality* (2007).

[KSD*09] Kümmerle R., Steder B., Dornhege C., Ruhnke M., Grisetti G., Stachniss C., Kleiner A.: On measuring the accuracy of SLAM algorithms. *Journal of Autonomous Robots* (2009).

[NIH01] Newman J., Ingram D., Hopper A.: Augmented reality in a wide area sentient environment. In *Proc. of the IEEE and ACM International Symposium on Augmented Reality (ISAR'01)* (2001).

[RBG00] Rolland J. P., Baillot Y., Goon A. A.: *A Survey of tracking technology for virtual environments*. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 2000, ch. 3.

[SA12] Sick AG: http://www.sick.com, 2012.

[SG05] Steggles P., Gschwind S.: The ubisense smart space platform. In *Advances in Pervasive Computing, Adjunct Proceedings of the Third International Conference on Pervasive Computing 2005* (2005).

[SM10] Scheer F., Müller S.: Large area indoor tracking for industrial augmented reality (poster). In *Proc. of the 9th IEEE International Symposium on Mixed and Augmented Reality(ISMAR)* (2010).

[SS08] Schoenfelder R., Schmalstieg D.: Augmented reality for industrial building acceptance. In *VR (Proc. of IEEE Virtual Reality Conference 2008)* (2008), IEEE.

[WBV*99] Welch G., Bishop G., Vicci L., Brumback S., Keller K., Colucci D.: The hiball tracker: high-performance wide-area tracking for virtual and augmented environments. In *VRST '99: Proc. of the ACM symposium on Virtual reality software and technology* (1999), ACM.

[WFK07] Wormell D., Foxlin E., Katzman P.: Advanced inertial-optical tracking system for wide area mixed and augmented reality systems. In *IPT-EGVE* (2007).

[WS07] Wagner D., Schmalstieg D.: Artoolkitplus for pose tracking on mobile devices. In *In Proc. CVWW* (2007).