Survey of Optical Indoor Positioning Systems

Rainer Mautz and Sebastian Tilch
* ETH Zürich, Switzerland, rmautz@ethz.ch

Abstract—**Recent advances in CCD technologies, processing speed and image understanding have been driving the development of the camera-based positioning systems. An improved performance of optical systems has triggered image based positioning methods to become an attractive alternative for applications in industrial metrology as well as for robot- and pedestrian navigation. This paper provides a survey of current optical indoor positioning approaches. Different systems are briefly described and categorized based on how the images are referenced to the environment.**

Keywords — **Optical Indoor Positioning; Camera-based positioning; Indoor Photogrammetry**

I. Introduction

Optical positioning is currently becoming a dominating technique that covers a wide field of applications at all levels of accuracy, with its main application area in the sub-mm domain. The success of optical methods originates from improvement and miniaturization of actuators (e.g. lasers) and particularly advancement in the technology of the detectors (e.g. CCD sensors). In parallel there has been an increase in the data transmission rates and computational capabilities as well as profound development of algorithms in image processing.

Optical indoor positioning systems can be categorized into ego-motion systems where a mobile sensor (i.e. the camera) is to be located and static sensors that locate moving objects in the images. An answer is to be found how position and rotations in a 3D world can be computed where the primary observations are 2D positions on a CCD sensor. All camera-based system architectures measure image coordinates that represent angular information and exclusively built on triangulation by the Angle of Arrival (AoA) principle. The transformation between object space (X, Y, Z), the projection center of a camera (X_0, Y_0, Z_0) and the image coordinates (x'', y'') is given by

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_0 \\ Y_0 \\ Z_0 \end{pmatrix} + \lambda \mathbf{R} \begin{pmatrix} x'' \\ y'' \\ -c \end{pmatrix}, \qquad (1)$$

where λ is a scale factor, c the camera constant and $\mathbf{R}$ a rotation matrix describing the camera orientation angles ω, φ and κ. Typically, several images are taken from multiple cameras or alternatively from multiple views of a single camera. Depending on the application, (1) is used to determine the object coordinates or is rearranged for the determination of the exterior camera orientation parameters (X_0, Y_0, Z_0, ω, φ, κ) from known 3D object coordinates using a resection procedure. In a bundle adjustment, the orientation parameters of multiple images, multiple 3D object coordinates and camera calibration parameters (if not already known a-priori) can be adjusted commonly. For further details of the photogrammetric principles see Luhmann et al. (2006).

One fact that should be noted when relying on AoA only is that the scale factor λ (also denoted as distance or depth) differs for every pixel and is unknown if only a single image is available. Therefore, the transformation from the image space into the object space requires additional depth information.

Depth information of monocular images can be obtained by exploiting the motion of a camera. In this approach – known as synthetic stereo vision – the scene is observed sequentially from different locations by the same camera and the image depth can be estimated in a manner similar to the stereo-vision approach. However, the baseline between sequential images needs to be determined by a complementary technique. Therefore, the system scale λ cannot be determined from the images alone and requires a separate solution.

If a stereo camera system is used with a known baseline, the scale can be determined from the stereoscopic images. As a drawback, the performance of a stereo camera system is directly driven by the length of the stereo baseline. A too short baseline causes the geometry to be unfavorable for forward intersection and therefore a miniaturization for handheld devices is not applicable.

Alternatively, the distances can be directly measured with additional sensors, such as with laser-scanners or range imaging cameras. The latter return a distance value for every pixel of a 320 × 240 image at a frame rate of 100 Hz. In order to determine the scale roughly, the position of the autofocus can be used.

A decisive characteristic in the system architecture is the manner how reference information is obtained. Therefore, this survey of recently developed optical navigation systems takes the mode of reference as primary criterion for categorization. It is an expansion on the study carried out by Mautz and Tilch [1]. An overview of the here mentioned systems and their key parameters are given in Table I.

A comprehensive survey of older works can be found in DeSouza and Kak [2]. A more recent overview of video tracking systems has been carried out by Trucco and Plakas [3], where the main algorithmic approaches, namely window tracking, feature tracking, rigid object tracking, deformable contour tracking and visual learning are explained and 28 works of video tracking are discussed.

When discussing the performance of kinematic systems in terms of positional accuracy it is important to note that positions of a moving object are determined in 4D from spatio-temporal observations. Therefore, all involved

978-1-4577-1804-5/11/$26.00 ©2011 IEEE

sensors need to be synchronized – otherwise the modeled delay time of a sensor or the temporal difference between spatial observations of different sensors, e.g. between angular and distance measurements, will lead to additional deviations in space-time position. While the effect of insufficient synchronization is oftentimes neglected for low-accuracy and low object speeds, its compensation can become a crucial component within industrial measuring systems for tracking fast moving targets. In order to compensate for the delay times of kinematic optical systems, Depenthal [4] has developed a time-referenced 4D calibration system using by a tilting rotating arm. The system performances given in this survey however rely exclusively on the figures as stated by the authors. These figures are oftentimes assessed roughly without taking into account time delays. It is also important to note that the authors did not always state their positioning data rate. As an alternative, the camera's frame rate is given here – a parameter that does not necessarily correspond to the measurement rate due to the time required for the data processing.

II. DIFFERENT SYSTEM APPROACHES BY REFERENCE

Table 1 summarizes the key parameters of the optical systems discussed in this Chapter.

TABLE I.
OPTICAL INDOOR POSITIONING SYSTEMS

Name	Coordinate Reference	Reported Accuracy	Coverage	CCD Size [Pixel]	Frame Rate	Object / Camera Positioning	Camera Costs	Market Maturity
Kohoutek [5]	CityGML	dm	scalable	176 × 144	54 Hz	Cam., SR 4000	9000 $	suggestion
Hile [7]	floor plan	30 cm	scalable	640 × 480	0.1 Hz	Obj., Cell Phone	100 $	development
Kitanov [8]	vector model	dm	scalable	752 × 585	10 Hz	Cam., EVI-D31	£ 245	development
Schlaile [9]	edges segments	1 dm/min	scalable	752 × 582	50 Hz	Cam., VC-PCC48P	175 €	development
Ido [10]	images	30 cm	scalable	320 × 240 × 4	30 Hz	Cam., IEEE1394		development
Sjö [11]	images / scans	sub m	scalable	320 × 240	30 Hz	Cam., VC-C4R	700 €	development
Muffert [12]	images	0.15 gon/min	room	1616 × 1232 × 6	15 Hz	Cam., Ladybug3	>10000$	development
Maye [13]	images / landm.	1 %	scalable	16 × 16	2300 Hz	Cam., ADNS-2051	1.35 €	development
Sky-Trax [15]	coded markers	2 – 30 cm	scalable			Cam.		product
Mulloni [16]	coded markers	cm – dm	scalable	176 × 144	15 Hz	Cam., Cell Phone	low	product
AICON ProCam [17]	coded markers	0.1 mm	vehicle	1628 × 1236	7 Hz	Cam., ProCam	high	product
StarGazer [18]	coded markers	cm – dm	scalable		20 Hz	Cam.	980 $	product
Lee [19]	coded markers	dm	36 m^2	1280 × 1024	30 Hz	Cam.,VX-6000	40 $	development
naviSCAN3D	projection	50 µm	1.5–10 m	2448 × 2048 × 2	1 Hz	Obj., stereoSCAN	high	product
TrackSense [22]	projection	4 cm	25 m^2	640 × 480	15 Hz	Obj.,CamPro4000	200 $	development
CLIPS [23]	projection	0.5 mm	36 m^2	1032 × 778	30 Hz	Cam., GuppyF80	1000 €	development
Habbecke [24]	projection	mm	25 m^2	1280 × 960		Obj.	1000 €	development
Popescu [25]	projection	cm	25 m^2	720 × 480	15 Hz	Cam.	1500 $	development
NorthStar [27]	projection	cm - dm	36 m^2		10 Hz	Cam./Obj, IR	1400 $	product
DEADALUS [28]	none	0.04 mm	m - km	1024 × 768	30 Hz	Obj., GuppyF80	high	development
Boochs [29]	none	0.05 mm	4 m^3	2000 × 2000 × 4		Obj.,	high	development
Tappero [30]	none	dm – m	30 m^2	356 × 292	3 Hz	Obj., OV6620	20 $ US	suggestion
Soloviev [31]	GNSS	cm	scalable	1240 × 1024		Obj.		suggestion
Aufderheide [32]	image features		scalable		5-30 Hz	Cam.		suggestion
Liu T. [33]	scanner, image	1 %	scalable	1338 × 987 × 3	10 Hz	Obj.	low	development
Liu W. [34]	magnetic field	1 mm	1 m^3	768 × 576 × 4	25 Hz	Obj., Sony ICX	800 €	development

A. Reference from 3D Building Models

This class of positioning methods relies on the detection of objects in the images and the matching of those objects with a building data base that contains position information of the building interior. Figure 1 visualizes such a data base for one room. The key advantage of these methods is that there is no requirement for the installation of local infrastructure such as the deployment of sensor beacons. In other words, the reference nodes are substituted by a digital reference point list. Accordingly, these systems have the potential for large scale coverage without significant increase of the costs.

Kohoutek et al. [5] use the digital spatio-semantic interior building model City Geography Markup Language, (CityGML, Gröger et al. [6]) at its highest level of detail (LoD 4) with the intention to determine the location and the pose of a Range Imaging Camera. In a first step the room where the camera is located is identified in the CityGML data base. From the 3D point cloud that is obtained by the range image sensor, fixed objects such as windows and doors are detected and their geometric properties are compared with the database. The second and last step consists of dm-level fine positioning of the camera based on a technique that combines trilateration and spatial resection.

Hile and Borriello [7] compare a floor plan with the current image of a camera phone. In a first step a rough location is determined by WLAN connectivity to limit the search area. In a second step extracted features from the images are used to find feature correspondences and compute the pose of the phone at decimeter level. Location based information can be displayed instantly.

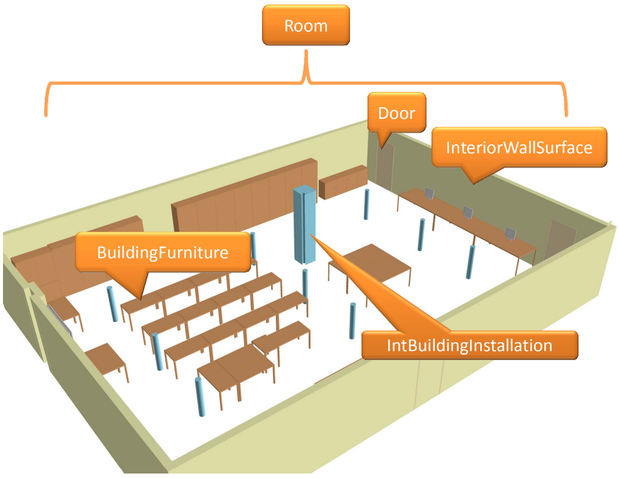

Figure 1. Model of a room in CityGML (Kohoutek et al. [5])

Kitanov et al. [8] compare image lines that have been detected in images of a robot mounted camera with a 3D vector model. The camera orientation is repeatedly computed from image sequences while the camera is in motion. An off-line optimal matching of rendered image lines and lines extracted from the camera images appears to achieve dm-level positioning accuracy. An odometer is used to stabilize the robustness.

The computer vision algorithm described by Schlaile et al. [9] also relies on feature detection in an image sequence. Here, the computer vision module is used as complementary aiding of an integrated navigation system that is mounted on a micro aerial vehicle.

B. *Reference from Images*

The so-called view-based approach relies on sequences of images taken beforehand by a camera along certain routes in the building, see Figure 2. In operation mode, the current view of a mobile camera (as shown in Figure 3) is compared with these previously captured view sequences. The main challenge of this approach is to achieve real-time capability. For the identification of image correspondences the computational load is particularly high since operability is assumed without deployed passive or active optical targets. Nevertheless, all systems require an independent reference source from time to time in order to control the accumulated error.

Figure 2. Example of a view sequence

Figure 3. Current view to be compared with the view sequence

In order to navigate a humanoid robot through office buildings, Ido et al. [10] carry out a template matching of images. In an initial phase, recording runs are captured by a camera mounted on the head of the robot. When the view sequences have been analyzed, compressed and stored, an autonomous navigation phase can commence. In this phase the correlation coefficients between the templates and the current view images are computed to determine the robot's pose. First trials indicate an accuracy of 30 cm.

Sjö et al. [11] navigate their robot based on a low-resolution camera with zoom capabilities. To approximate the distances to objects they use the zoom position in a first step and then carry out SLAM (Simultaneous Localization And Mapping) by computing RFC Histograms (Receptive Field Coocurrence Histograms) from the current image and comparing them with histograms of previously captured images. In order to stabilize their SLAM method geometrically, the robot is also provided with laser scanning data.

Muffert et al. [12] determine the trajectory of an omnidirectional video camera based on relative camera orientation of consecutive images. If there is no additional control via other references positions or directions, the recorded path drifts away from the true trajectory, similarly to inertial sensors relying merely on dead reckoning. For an acquisition time of 40 seconds, standard deviations for the yaw angles of 0.1 gon are reported.

Based on the principle of optical odometry, Maye et al. [13] develop a low-cost optical navigation device using an optical mouse sensor. The only modification to a computer mouse is a different lens tailored to higher speeds (2 m/s) and ground clearance (5 cm). In order to correct the accumulated path errors, fixed landmarks are deployed where the system is updated. In addition, a magneto-inductive compass is used. The reported drift up to a velocity of 2 m/s is 1% of the travelled path length.

C. *Reference from Deployed Coded Targets*

Optical positioning systems that rely entirely on natural features in the images lack of robustness, in particular under conditions with varying illumination. In order to increase robustness and improve accuracy of reference points, dedicated coded markers are used for systems with demanding requirements for positioning. The markers serve three purposes for algorithmic development: a) simplification of the automatic detection of corresponding points, b) introduction of the system scale, c) distinction and identification of the targets by using a unique code for each marker. Common types of targets include concentric rings, barcodes or patterns consisting of colored dots, see Figure 4. There are retro-reflective and non-reflective versions.

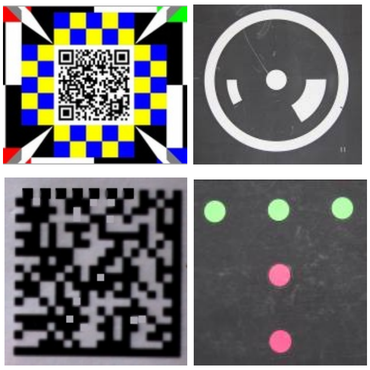

Figure 4. Examples of coded targets used for point identification and camera calibration

Mark [14] has designed a rectangular target (see Figure 4 upper right) such that the relative location and orientation of the camera can be determined from a single image. An array of these targets is used to navigate camera equipped vehicles or drones.

Sky-Trax Inc. [15] developed an optical navigation system for forklift trucks in warehouses. Coded reference markers are deployed on the ceilings along the routes. On the roof of each forklift an optical sensor takes images that are forwarded to a server where they are processed centrally. The position accuracy is reported as “between one inch to one foot”.

Mulloni et al. [16] developed a low-cost indoor positioning system for “off the shelf camera phones" by using bar-coded fiduciary markers. These markers are placed on walls, posters or certain objects. If an image of these markers is captured, the pose of the device can be determined with an accuracy of “a few centimeters”. Additional location based information (e.g. about the next conference room or the next session) can also be displayed.

AICON 3D Systems [17] developed the ProCam System for industrial applications in the sub-mm accuracy range. The mobile video camera is pointed to a pre-calibrated reference point field. In order to increase the robustness of the point detection, the camera emits active infrared light that illuminates the reference points. Tactile measurements are carried out manually with an integrated probe tip.

The StarGazer system of Hagisonic [18] is tailored for robot positioning and relies on retro reflective targets mounted on the ceiling. An infrared sensitive camera observes different point patterns that are actively illuminated by an infrared light source. From the points uniquely arranged on a 3 × 3 or 4 × 4 grid on the targets, a room can be identified and the pose of the roving camera can be determined within sub-dm accuracy.

Lee and Song [19] exploit the same principle of retro reflective targets as the StarGazer system to locate and orientate a mobile robot. Here, the corners of the triangle shaped targets are used for approximate orientation estimation and six inner sectors for unique identification. In order to achieve algorithmic robustness the difference between one image with and one without active illumination are processed. According to stated results, the 2D accuracy is at sub-decimeter level.

Frank [21] describes the stereoScan-3D developed by Breuckmann GmbH for high precision mapping of industrial surfaces. Two mobile high resolution cameras with a fixed baseline are used to capture points with an accuracy of 50 µm. The precise positioning of the cameras is carried out in a calibration phase that consists in the capture of images with markers that are attached to the to-be-measured object.

D. *Reference from Projected Targets*

The projection of reference points or patterns spares the physical deployment of targets in the environment, making this method economical. For some applications the mounting of reference markers is undesirable or not feasible. Optionally, infrared light can be projected to attain unobtrusiveness to the user. In contrast to systems relying only on natural image features, the detection of projected patterns is facilitated due to their distinct color, shape and brightness. The principle of an inverse camera (or active triangulation) can be exploited where the central light projection replaces the optical path of a camera. The main disadvantage of active light based systems is that both, the camera and the light source require direct view on the same surface.

Köhler et al. [22] built an experimental model called TrackSense consisting of a projector and a simple webcam. A grid pattern is projected onto plain walls in the camera’s field of view as shown in Figure 5 (upper left). Using an edge detection algorithm, the lines and intersection points are determined. By the principle of triangulation – analogous to stereo vision – the distance and orientation to each point relative to the camera is computed. With a sufficient number of points TrackSense can determine the camera’s orientation relative to fixed large planes, such as walls and ceilings. The evaluation of TrackSense indicates that such a system can deliver up to 4 cm accuracy with 3 cm precision.

Tilch and Mautz [23] developed CLIPS (Camera and Laser based Indoor Positioning System) with the purpose to determine the pose of a mobile camera in respect to a laser rig. Since the rig emits laser-beams from a virtual central point, it can be regarded as an inverse camera. From the bright laser spots that are projected to any surface without any specific structure of the scene as shown in Figure 5 (upper right), the relative orientation between the camera and the laser rig can be computed. Point tracking is achieved at frame rates of 15 Hz and the accuracy of the camera position is sub-mm.

The video camera system of Habbecke and Kobbelt [24] is based on a mobile rig of laser pointers and a fixed camera. The laser rays on the rig have an arbitrary alignment without the need for a central point of intersection. In order to accomplish correct identification of the laser spots (shown in Figure 5 lower left) a greedy pairing algorithm is used. Via least squares minimization, the relative orientation between the camera and the rig is determined. Besides pose determination, the system is used for tracking and scene reconstruction. Reported accuracies indicate position deviations in the order of a few millimeters.

The laser rig of Popescu et al. ([25], [26]) is rigidly mounted to a hand-held video camera with the advantage that the relative orientation between laser source and camera remains constant. From a single laser source and a diffraction grating that acts as a beam splitter a grid of 7 × 7 laser spots is generated, see Figure 4.5 lower right. These 49 spots are located in each frame, and their 3D positions are computed by triangulation between the optical rays and the laser beams. When the camera with the laser rig is freely moved through a scene, the 3D positions of the laser spots can be used for scene modeling. The positional accuracy was reported to be better than 1 cm.

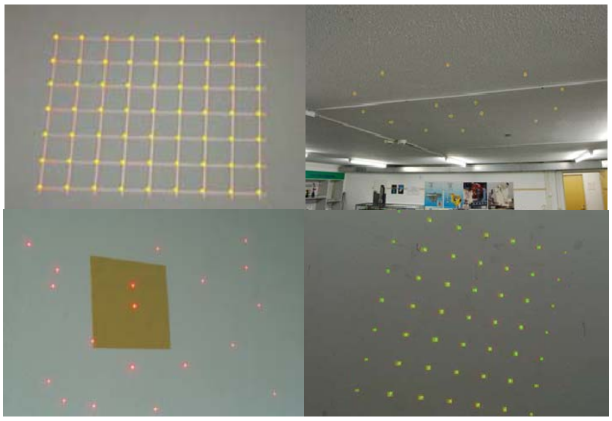

Figure 5. Projected reference patterns. Upper left: TrackSense Grid, upper right: CLIPS laserspots, lower left: laserspots of Habbecke, lower right: diffraction grid of Popescu.

Evolution Robotics [27] developed the indoor localization system NorthStar for navigation of shopping carts or robotic vacuum cleaners. The position and heading of the mobile unit is determined from infrared light spots, emitted from one or more infrared LEDs (Light Emitting Diodes). Each mobile unit can be equipped with an infrared detector and projector that allow determination of the relative orientation between mobile devices. The reported positioning accuracy is in the magnitude of cm to dm.

E. *Systems without Reference*

The purpose of systems in this class is to observe position changes of objects directly and therefore do not require external reference. The common approach is to track mobile objects with high frame rates in real-time by a single or multiple static cameras.

The DEADALUS system described by Bürki et al. [28] consists of a CCD camera that is clipped on a surveying totalstation. Due to the magnification of the telescope, high-precision horizontal and vertical angle measurements are possible for automated 2D monitoring of objects with high data rates. Generally, any object can be tracked, but in the normal use case illuminated targets are observed to enhance algorithmic robustness, see Figure 6. DEADALUS is a high-end system where reported accuracies reach 0.3 arc-seconds or 0.04 mm.

Boochs et al. [29] use multiple fixed calibrated and orientated cameras to track an illuminated target that is mounted on the head of an industrial robot. The target body consists of a sphere with 54 self-luminous infrared LEDs to allow robust tracking from all directions. First test results of this photogrammetric tracking approach have shown 3D coordinate quality of about 0.05 mm.

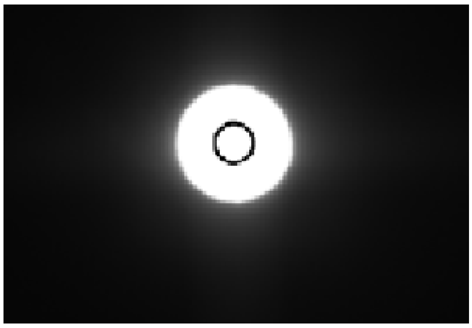

Figure 6. Tracked LED by DEADALUS with a circle indicating the center of mass.

Tappero [30] suggests a low-cost system for the tracking of people in an indoor environment. In order to accomplish real-time tracking using extremely cheap components the computational efficiency is optimized by detection of changes in the difference image of subsequent frames. A static camera is mounted at the ceiling is able to locate people and objects within an accuracy of some decimeters.

F. *Reference from other Sensors*

Optical sensor data can be fused with observations obtained by another measuring system. The purpose for using a hybrid sensor system can be manifold, e.g. to enhance accuracy, coverage or robustness. In some cases the additional sensor system has the capability to provide absolute coordinate reference.

Soloviev and Venable [33] combine vision data with GNSS (Global Navigation Satellite System) carrier phase measurements for GNSS signal-challenged environments. If less than the required number of satellites is visible, single carrier phase measurements in the accuracy level of sub-centimeters are used to support the feature extraction from video images. Thereby, sub-cm level range measurements from GNSS observed at the imaging positions provides a synthetic baseline between delta positions that is used to determine the system scale and facilitate the feature extraction by providing image depth initialization. There is no need to determine the integer carrier ambiguities because the unknown ambiguities are eliminated by differencing of the carrier phase measurements between successive positions.

Aufderheide and Krybus [32] propose a dual-track system combining inertial and optical measurements as a loosely coupled system. The camera pose is estimated from corresponding image features between successive frames from a monocular camera. Therefore, their approach could also be categorized as an optical system with reference from image sequences. The pose determination of the camera is used to bound the drift error of the inertial pose predictions for long-term sequences. Vice versa, pose predictions from IMU data can limit the search space for feature tracking. The development stage of the proposed system does not yet allow performance assessment.

The SLAM (Simultaneous Localization And Mapping) approach of Liu T. et al. [33] combines IMU, laser scanner and image based localization, all integrated in a human-operated backpack system that can be used to generate 3D models of complex indoor environments. The positions are determined from data capture based on two laser scanners and an IMU with 6 Degree of Freedom (DoF). An average positional error of 1 % of the travelled distance is reported. The localization performance can be improved by making use of the camera images in an offline phase and thereby refining the six values of the camera pose.

Liu W. et al. [34] combine optical tracking and magnetic localization in order to overcome the occlusion problem that optical systems face. A permanent magnet is tracked by a dense array of sensors that measure the magnetic flux intensity in three dimensions within a cubic shaped magnetic field. The advantage of using a magnetic system component is that line of sight between magnetic sensor and the tracked object is not a requirement. The optical module is a redundant system component that consists of 4 video cameras and is used to enhance the robustness to disturbing ferromagnetic objects within the working volume. The reported positional accuracy for the combined system is 1 mm within a volume of 1 m^3.

III. CONCLUSIONS

Current optical indoor positioning approaches achieve accuracy levels between a couple of µm and dm, where the high precision systems offer solutions for surveying and industrial metrology. The covered area of the systems presented here (not including microscale or nanoscale technologies) differs between 4 m^2, room sizes and arbitrarily scalable systems. High update rates of typically more than 10 Hz allow for kinematic applications such as precision-navigation, real-time mapping and pose estimation. The performance of an optical positioning system can be improved by fusion of image data with data from other sensors, such as INS, GNSS or magnetic sensors. With the abundance of computing power and CCD sensor chips, low-cost positioning solutions are in view that have the potential to serve the mass market.

REFERENCES

[1] Mautz, R. and Tilch, S. (2010): “Innenraumpositionierung mit optischen Methoden”, AVN, no. 7, pp. 250-255.

[2] DeSouza, G. and Kak, A. (2002): “Vision for Mobile Robot Navigation: A Survey”, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 2, pp. 237-267.

[3] Trucco, E. and Plakas, K. (2006): “Video Tracking: A Concise Survey”, IEEE Journal of Oceanic Engineering, vol. 31, issue 2, pp. 520-529, April 2006.

[4] Depenthal, C. (2009): “Entwicklung eines zeitreferenzierten 4-D Kalibriersystems für optisch kinematische Messsysteme”, DGK vol C, no. 627, 136 p.

[5] Kohoutek, T.K., Mautz, R. and Donaubauer, A. (2010): “Real-time Indoor Positioning Using Range Imaging Sensors”, SPIE Photonics Europe, Proc. of SPIE Vol. 7724, 77240K.

[6] Gröger, G., Kolbe, T., Czerwinski, A. and Nagel, C. (2008): "OpenGIS City Geography Markup Language (CityGML) Encoding Standard" Version 1.0.0, International OGC Standard. Open GeospatialConsortium, Doc. No. 08-007r1.

[7] Hile, H. and Borriello, G. (2008): “Positioning and Orientierung in Indoor Environments Using Camera Phones”, IEEE Computer Graphics and Applications, July/August 2008, pp. 32-39.

[8] Kitanov, A., Biševac, S. and Petrović, I. (2007): “Mobile robot self-localization in complex indoor environments using monocular vision and 3D model”, IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Zürich, Switzerland.

[9] Schlaile, C., Meister, O., Frietsch, N., Kessler, C., Wendel, J. and Trommer, G. F. (2009): “Using natural features for vision based navigation of an indoor-VTOL MAV”, Aerospace Science and Technology, vol. 13, no. 7, pp. 349-357.

[10] Ido, J., Shimizu, Y., Matsumoto, Y. and Ogasawara, T. (2009): “Indoor Navigation for a Humanoid Robot Using a View Sequence”, The International Journal of Robotics Research, vol 28, no. 2, pp. 315-325.

[11] Sjö, K., López, D., Paul, C., Jensfelt, P. and Kragic D. (2009): “Object Search and Localization for an Indoor Mobile Robot”, Journal of Computing and Information Technology, vol. 17, no. 1.

[12] Muffert, M., Siegemund, J. and Förstner, W. (2010): “The estimation of spatial positions by using an omnidirectional camera system”, Proceedings of the 2nd International Conference on Machine Control & Guidance, pp. 95-104.

[13] Maye, O., Schaeffner, J. and Maaser, M. (2006): “An Optical Indoor Positioning System for the Mass Market”, WPNC'06, IEEE Xplore, pp. 111-115.

[14] Mark, U. (2011): http://www.thingiverse.com/thing:5465 , Thingiverse, last accessed May 2011.

[15] Sky-Trax Inc. (2010): http://www.sky-trax.com/, last accessed December 2010.

[16] Mulloni, A., Wgner, D., Schmalstieg, D. and Barakonyi, I. (2009): “Indoor Positioning and Navigation with Camera Phones”, Pervasive Computing, IEEE, vol. 8, pp. 22-31.

[17] AICON 3D Systems (2010), http://www.aicon.de, last accessed December 2010.

[18] Hagisonic (2008): “User’s Guide Localization system StarGazerTM for Intelligent Robots”, http://www.hagisonic.com/, last accessed March 2010.

[19] Lee, S. and Song, J.B. (2007): “Mobile robot localization using infrared light reflecting landmarks”, Proceedings of ICCAS’07, pp. 674-677.

[20] Luhmann, T., Robson, S., Kyle, S. and Harley, I. (2006): “Close Range Photogrammetry”, Whittles Publishing, 500p.

[21] Frank, M. (2008): “Eine Idee nimmt Gestalt an 3D-Daten von physischen Modellen und Bauteilen erzeugen”, QZ. Qualität und Zuverlässigkeit, vol. 53, no. 12, pp. 53-55.

[22] Köhler, M., Patel, S., Summet, J., Stuntebeck E. and G. Abowed (2007): “TrackSense: Infrastructure Free Precise Indoor Positioning Using Projected Patterns”, Pervasive Computing, LNCS, vol. 4480, pp. 334-350.

[23] Tilch, S., Mautz, R. (2010): “Development of a new laser-based, optical indoor positioning system”, ISPRS Commission V Mid-Term Symposium, no. 98, pp. 575-580.

[24] Habbecke, M. and Kobbelt, L. (2008): “Laser brush: a flexible device for 3D reconstruction of indoor scenes”, Symposium on Solid and Physical Modeling vol. 2008, pp. 231-239.

[25] Popescu, V., Sacks, E. and Bhamutov, G. (2004): “Interactive modeling from dense color and sparse depth”, Proceedings of 3DPVT, pp. 430–437.

[26] Popescu, V., Bahmutov, G., Mudure, M. and Sacks, E. (2006): “The modelcamera”, Graphical Models, vol. 68, pp. 385–401.

[27] Evolution Robotics (2010): http://www.evolution.com, last accessed December 2010.

[28] Bürki, B., Guillaume, S., Sorber, P. and Oesch, H. (2010): “DAEDALUS: A Versatile Usable Digital Clip-on Measuring System for Total Stations”, Proceedings of IPIN 2010, ETH Zurich, Switzerland, pp. 32-41.

[29] Boochs, F., Schütze, R., Simon, C., Marzani, F., Wirth, H. and Meier, J. (2010): “Increasing the accuracy of untaught robot positions by means of a multi-camera system”, Proceedings IPIN 2010, ETH Zurich, Switzerland.

[30] Tappero, F. (2009): “Low-cost optical-based indoor tracking device for detection and mitigation of NLOS effects”, in Procedia Chemistry, vol. 1, no.1, pp. 497-500.

[31] Soloviev, A. and Venable, D. (2010): “When GNSS Goes Blind - Integrating Vision Measurements for Navigation in Signal-Challenged Environments”, GNSS Inside, Oct 2010, pp. 18-29.

[32] Aufderheide, D. and Krybus, W. (2010): “Towards Real-Time Camera Egomotion Estimation and Three-Dimensional Scene Acquisition from Monocular Image Streams”, Proceedings of IPIN 2010, ETH Zurich, Switzerland, pp. 13-22.

[33] Liu, T., Carlberg, M., Chen, G., Chen, J., Kua, J. and Zakhor, A. (2010): "Indoor Localization and Visualization Using a Human-Operated Backpack System", Proceedings of IPIN 2010, ETH Zurich, Switzerland, pp. 890-899.

[34] Liu, W., Hu, C., He, Q., Meng, M. and Liu, L. (2010): "An Hybrid Localization System based on Optics and Magnetics", Proceedings of the International Conference on Robotics and Biomimetics, Tianjin, China, pp. 1165-1169.