A High-Precision Localization Device for Outdoor Augmented Reality

Marco Stranner[*] Graz University of Technology
Philipp Fleck[†] Graz University of Technology
Dieter Schmalstieg[‡] Graz University of Technology
Clemens Arth[§] AR4 GmbH

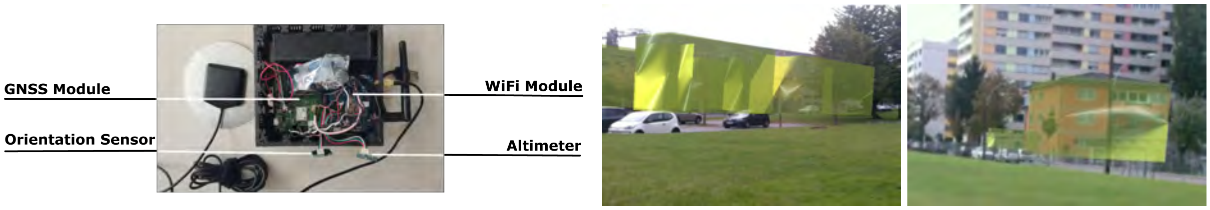


Figure 1: Left: First prototype of the sensor cube labeled with its components. Middle, right: GIS data Visualization of buildings at the university using Unity.

Abstract

In contrast to indoor tracking using computer vision, which has reached a good amount of maturity, outdoor tracking still suffers from comparably poor localization on a global scale. Smartphones and other commodity devices contain consumer-grade sensors for GPS, compass and inertial measurements, which are not accurate enough for augmented reality (AR) in most situations. This restricts what AR can offer to application areas such as surveying or building constructions. We present a self-contained localization device which connects wirelessly to any AR device, such as a smartphone or headset. The device gives centimeter-level accuracy and can be built out of commercial-of-the-shelf components for less than 500 EUR. We demonstrate the performance of the localization device using a variety of position and orientation sensing benchmarks.

Index Terms: Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

Augmented Reality (AR) can now address a mass audience, since model-free tracking solutions like ARKit and ARCore are available on commodity smartphones, and games like Pokemon Go are played by million of enthusiasts. AR can potentially benefit many industrial use cases, which profit from digital information being visualized at a task location. Such opportunities for increasing productivity and saving time and money are well received by industrial enterprises. Unfortunately, the adoption of AR in industry is slowed down by the fact that unmodified consumer devices often fail to deliver adequate performance in critical situations.

This is particularly problematic for applications in the area of logistics or construction. Even when visual-inertial tracking runs fine, it typically only delivers localization relative to an arbitrary reference point. Measurement of a reference point is required to establish a global coordinate system. Such a reference point is most easily established using GPS, but consumer-grade GPS often fails or is too inaccurate (error in the range tens of meters) to be useful.

The poor sensor quality is owed to the financial constraints of consumer products an not likely to change in the near future. Indeed, premium hardware, such as smartphones complying with the Google Tango specification, were unsuccessful in the marketplace. A similar constraint applies to headsets like the Microsoft Hololens, which delivers outstanding local tracking, but no global sensing at all.

In this paper, we report on the design and implementation of a sensor cube, a companion device intended to address the need for high-precision global localization. In a nutshell, the sensor cube combines a differential GPS receiver, an inertial measurement unit (IMU), an altimeter and a Wifi radio with a battery in a compact encasing. The IMU combines a 3D linear accelerometer, a 3D magnetometer, and a 3D gyroscope. The sensor cube streams its data over Wifi to a host computer and is otherwise fully self-contained. It can be easily attached to arbitrary hosts, such as on a headset or mobile phone, or worn on the body. Our device has a total weight of 40 grams and a net cost of less than 500 EUR for the parts. As we will show in this paper, it achieves very good localization and orientation estimation accuracy, compared to commercial mobile phone sensors.

2 Related Work

Outdoor AR has many possible applications [16]; Shin et al. [3] discuss several promising directions in construction. They identify building information modeling (BIM), underground infrastructure visualization, outdoor architectural designs, and geographic information systems as potential AR use cases what benefit from more precise localization and tracking.

Hardware-based approaches One example of an AR system for outdoor architectural preview was given by Hakkarainen *et al.* [14], who create photorealistic visualisations of architectural plans and environment feedback by use of BIM models within a complete mobile setup. Schall *et al.* [11] show how to visualize underground infrastructures in-situ directly from a geographic information systems (GIS) on a handheld device.

Due to continuing progress in hard- and software, outdoor AR systems got better and handier year after year. A lot of systems were developed in the last 20 years with several approaches. The first outdoor AR systems appeared over 20 years ago [**?**], and much progress in tracking has been made since then. According to Bostanci *et al.* [5], the most important improvements concern simultaneous localization and mapping (SLAM) and computer vision. However, outdoor systems also rely on alternative high-precision sensors, in particular, GPS and IMU. A combination of these sensors with vision is very common.

Pioneering work on the Tinmith system was presented by Thomas *et al.* [17] and later expanded by Piekarski *et al.* [19], who developed a wearable application supporting the modeling of buildings and other large physical structures and the positioning and visualization of those objects in AR. The system setup consisted of a laptop, an Intersense IS-300 tracker for orientation sensing, a Garmin

[*]marco.stranner@icg.tugraz.at
[†]philipp.fleck@icg.tugraz.at
[‡]schmalstieg@tugraz.at
[§]arth@icg.tugraz.at

12XL DGPS receiver, a monocular display, a video camera, custom designed pinch gloves for controlling the system, a small display for debugging on the back and a battery for powering the system, mounted on a backpack. The accuracy of positioning is between 1 and 5 meters depending on conditions. King *et al*. [6] achieved even better accuracy with a newer version of the Garmin receiver in their system ARVino.

With improved hardware, Schall *et al*. [10, 12] developed a new handheld AR system relying on sensors for localization. Their system, SmartVidente, could register virtual objects in the real world with high geospatial accuracy. SmartVidente consisted of a tablet PC, which is equipped with a camera, a 3DOF orientation sensor and a Novatel OEMV-2 L1/L2 real-time kinematic receiver for achieving an accuracy at centimeter level accuracy. The system also had a laser range finder embedded into to setup to provide a 3D cursor for the system, allows the user to select and move objects. Differential correction data is supplied by the EPOSA reference system, which enables a position accuracy better than 10 cm, while the orientation sensor provides an accuracy up to one degree.

Hybrid Approaches Reitmayr *et al*. [7] presented a vision-centric approach, noting that GPS quality was not reliable enough due to shadowing from buildings in urban areas. Instead, their system used an edge-based tracker for accurate localization. For dealing with motion, they relied on an IMU . The setup achieved a really good accuracy in an optimal environment, with a deviation of 0.0979m in easting, 0.1463m in northing and 0.1577m in altitude. However, the problem with this setup is that magnetic field disturbances, for instance, from moving cars, can affect the accuracy. Moreover, their effort was considered very computationally intensive on the mobile hardware of 2006. For tackling this problem, they later extended their approach by using GPS data for (re-)initialization to recover from any failures of the vision component [8].

Ventura *et al*. [18] realized accurate localization for outdoor AR systems using SLAM, relying on computation instead of special sensors. This system did not need any specific hardware, but only an ordinary mobile device (Apple iPad Air in their experiments). SLAM alone does not provide a registration of the tracked pose with a global coordinate system. Therefore, initialization of SLAM and global registration was provided by using a city-scale detection method using large-scale matching of SIFT features on a cloud server. Arth *et al*. [2] improved this technique by initializing SLAM and global registration using facade outline matching based only only OpenStreetMap data, instead of feature databases that must be built after extensive 3D scanning.

Many approaches rely on hybrid tracking, in particular vision and IMU, for example, the work of Jiang *et al*. [1] which uses vision tracking and gyroscopes, or the system of Fong *et al*. [20], which combines GPS , orientation sensors and vision. While feature matching based on statistical classification is the main component of their system, the additional sensor serve initialization and re-initialization. The hybrid tracking approach of Karlekar *et al*. [13] use 3D models for improving tracking accuracy by edge and corner detection. The sensor pose is corrected by matching silhouettes of the 3D models using shape context descriptors. When tracking seems accurate enough, the system switches to vision tracking exclusively using an extended Kalman filter. Similarly, Artemciukas *et al*. [4] combine orientation sensors and Kalman filtering for robust orientation estimation. Today's commercial solutions, such as Apple's ARKit and Google's ARCore, can be classified as visual-interial hybrid tracking solutions using a SLAM approach.

Since pretty good localization can be achieved with low-cost DGPS modules, and IMU kits are getting better and better, a setup for useful outdoor AR applications can be built cheaply. For global registration, pure vision approaches are not scalable enough, since they generally lack coverage of visual reference data. Vision has its role after a coarse, but good enough prior is delivered by non-visual sensing. Making this prior better than possible with built-in sensors of commodity devices was the main motivation for our work.

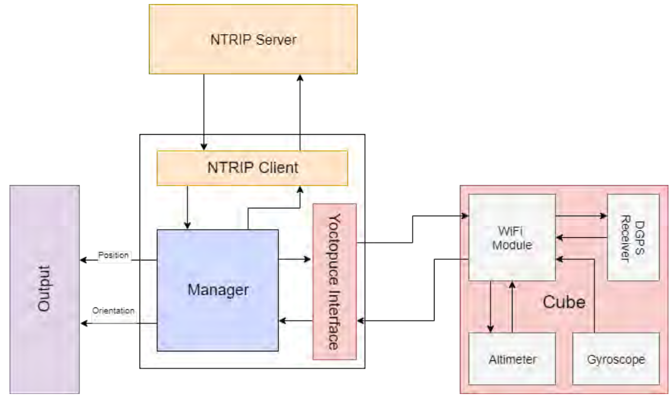


Figure 2: The general architecture of the system, showing its components and the communication of all parts.

Component	Name	Price
GPS Module	uBlox M8P-C94[1]	€150
IMU	Yocto 3d[2]	€50
Altimeter	Yocto Altimeter[3]	€30
WiFi Module	YoctoHub-Wireless-g[4]	€100

Table 1: Hardware components of the first prototype and their price.

3 Outdoor Localization Prototype

For developing a low budget outdoor tracking system, the first step was to find inexpensive, but good enough sensors. A series of prototypes was built, with increasingly better performance, After selecting individual hardware components, a first prototype of the sensor cube was built to be attached to mobile devices and used in outdoor AR applications.

3.1 Hardware

The first iteration of the sensor cube consisted of a differential GPS receiver and an IMU with three-axis for magnetometer, accelerometer and gyroscope. The IMU can use sensor complementarity to compensate most of drift problem. Keeping the IMU away from the host computer already reduces magnetic interference which, otherwise, substantially affects the magnetometer's yaw axis. Additionally, an Altimeter was integrated to compensate the GPS 's poor height positioning ability. These sensors are connected to a WiFi module which provides platform-independent streaming of positioning data to any kind of mobile device such as head-worn or handheld devices. The sensor cube can also be attached to objects, e.g., cars, to give them self-tracking ability.

On the top of Fig. 1, the sensors of the prototype are listed. Tab. 3.1 reveals the technical details of the setup. Our sensor cube including all sensors and a 3D-printed casing was built for less than €500. For comparison, a highly precise DGPS receiver, like the Novatel OEMv2 used in SmartVidente[5], costs more than €10,000.

3.2 Software

Our goal was to provide a C++ library with easy drag-and-drop integration into Unity3D, to ensure usability and cross-platform support. A flow chart of our software solution is depicted in Fig. 2. The individual components are described in the following.

[5]Novatel OEMv2 Receiver: `https://www.novatel.com/products/gnss-receivers/oem-receiver-boards/oemv-receivers/oemv-2`

Figure 3: GPS track of a walk through the park, comparing mobile built-in GPS to the uBlox receiver. Green: uBlox Red: LG G5

NTRIP Client. The NTRIP client is responsible for communication with an NTRIP server which provides GPS corrections. Requests with the current GPS position are sent to the server and responses with correction data are obtained.

Yoctopuce interface. Yoctopuce is a software interface for receiving data from the sensor cube over WiFi. Callbacks are triggered by the interface when sensor data arrives. Furthermore, correction data and GPS initialization are sent to the cube.

Manager. The manager is responsible for passing data between the NTRIP Client and the Yoctopuce interface, converting the NMEA position to the desired output format and providing the current device orientation. It can also be useful for preprocessing correction data, postprocessing the NMEA Data, or user-defined computations.

WiFi Module. The WiFi module is a hub that wirelessly communicates the sensor readings from the cube. Its special-purpose design intended for Internet-of-Things infrastructure makes it very power efficient; in addition to out-of-the-box Wifi streaming ability, it has several convenient features. For instance, communication with the uBlox receiver is implemented via a Yoctopuce serial connection, and the WiFi module is able to parse NMEA from the receiver.

The communication with the NTRIP server relies on HTTP over TCP. After sending an HTTP request for correction data, the server starts streaming RTCM Messages. The correction data is passed in RTCM form to the Wifi module and forwarded to the GPS receiver. The receiver decodes RTCM and reports its measurements as NMEA strings. The WiFi module decodes these NMEA strings, combines them with orientation changes from the IMU and altitude changes reported by the altimeter, and forwards data only if changes occur.

3.3 Vision tracking extension

Since sensor tracking alone is still not robust and accurate enough for AR, it is desirable to combine vision tracking with the sensor cube. For the first implementation we wanted to stuck to the SLAM approach of Arth *et al*. [2], but driven by proprietary GIS data instead of OpenStreetMap, which was originally used.

Liu *et al*. [15] implemented such a feature-based monocular SLAM system. Their approach proposes an instant initialization algorithm providing an accurate metric camera registration and an optimization method aligning large scale SLAM reconstructions with the associated camera trajectory using 2.5D maps and 6DoF tracking data. Their system was combined with our sensor cube and fed with GIS data dealing as 2.5D map input. Although, GIS data only provides 2D ground view of the buildings and their for the height was just guessed to provide 2.5D data, the system still improves accuracy and especially the robustness. However, the approach is not real-time ready on mobile devices yet.

4 Experiments

To show how accurate such a low-cost setup really is, all sensors were evaluated in several scenarios, representing a variety of conditions. Furthermore, the standalone sensor tracking was used in a simple GIS visualization application.

4.1 GPS Evaluation

We started by comparing the sensor in our cube to a conventional mobile GPS receiver. A walk through a park was tracked with the uBlox receiver on the sensor cube and an LG G5 smartphone for comparison. As shown in Figure 3, the smartphone drifts significantly and is more than five meters off most of the time, demonstrating that smartphone sensors are not very accurate.

To test the accuracy of the uBlox receiver, it is compared with a high-precision DGPS receiver, the Novatel OEMv2. For representing the results, GPS coordinates were converted into UTM. In the Figures, X refers to *east* and Y refers to *north*. Both receivers were placed side by side with an approximate distance between antennas of about 2 cm. The reference GPS position was taken from the Novatel receiver. All accuracy measurements are calculated with respect to the final measurement of the Novatel receiver (the last measurement in the record).

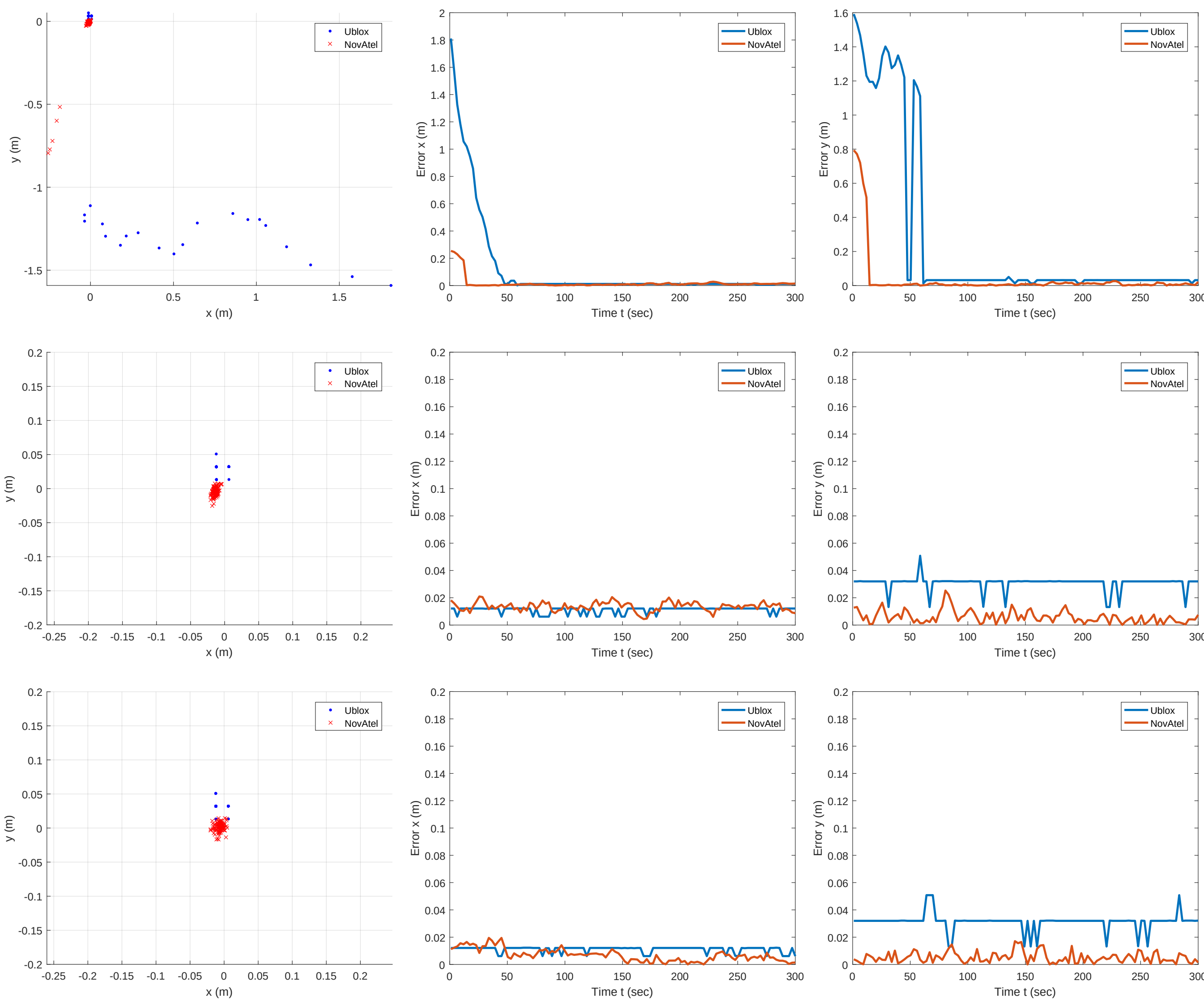
For this test, a wooden platform with open sky view was chosen as location. Due to the absence of any obstacles and full view onto the hemisphere, the GPS positioning was tested under optimal conditions. The first row in Fig.5 shows the first five minutes of the position estimation procedure, the second and third row represent time intervals from 5-10 and 10-15 minutes, respectively. The Novatel receiver already starts with an error smaller than one meter and achieves centimeter-level accuracy after about 15 seconds. The uBlox receiver starts with quite a large error of more then 2 meters overall, but reaches centimeter-level accuracy within the first minute. The deviation map of the first five minutes shows the position approximation of both devices, revealing that both receivers are performing reasonably well. The maximum error of the uBlox receiver is about 5 cm, with a slightly greater error in Y dimension and a maximum error of 2 cm compared to the Novatel receiver. Also, the maximum resolution of 2 cm for the uBlox can be seen in Figures 5 (a), while the Novatel receiver has a resolution under one centimeter. However, the uBlox receiver can keep up with the Novatel receiver, as long as a constant accuracy of less than 2 cm is sufficient for a particular use case. This test exhibits the accurate positioning in best conditions, which is more than sufficient for initial positioning of hybrid tracking. This level of precision can be considered as high-precision positioning in open space environments for sensor-only tracking.

4.2 IMU Evaluation

Schall *et al*. [9] point out that orientation sensors can be extremely susceptible to magnetic disturbances. Especially the magnetic compass is vulnerable, which is a serious problem for accurate orientation estimation. Therefore, compass accuracy was tested. To provide an exact ground truth of the bearing values for the compass validation, the setup makes use of UTM (Universal Transverse Mercator) coordinate system, which has a north-directed Y axis. Using precisely measured reference points (error < 2 cm), shown on the left of Fig. 4, and a setup to provide an exact alignment from one to another point, the validation can be done using traditional surveying. The alignment system consists of two tripods, one with a prism and one with the sensor cube combined with a laser range finder to align to the prism. This setup is shown on the right of Fig. 4.

The compass bearing was tracked once per second to check how accurate and precise the compass is. The first measurements were made from Point IF-15. The calculated ground truth angle and the compass measurements are shown in Tab. 2. The obtained error seems to be consistent and decreasing with higher distance between

Figure 4: Environment setup for the compass validation. Left: Map of all reference Points. Middle: Reflective prism on a tripod. Right: Laser range finder mounted on the sensor cube.



(a) Deviation Map (b) Error X (c) Error Y

Figure 5: Positioning comparison of uBlox and Novatel receiver for the ideal operation test. First row: Minute 0-5, Second row: Minute 5-10, Third row: Minute 10-15.

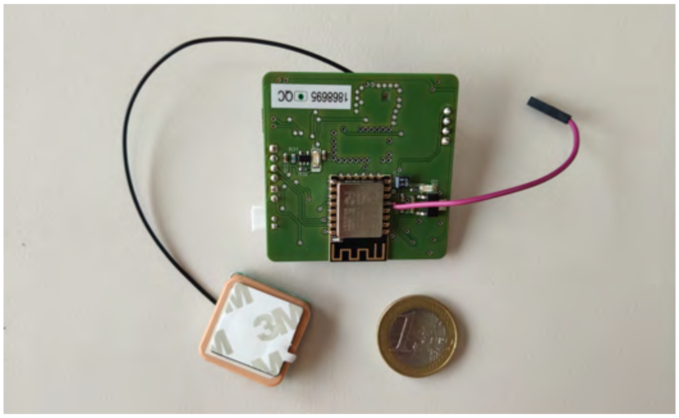


Figure 6: The first circuit board prototype with a size of 5×5 cm and a small passive GPS antenna attached to it.

15 to	ground truth	compass
03	344.7685	343.437
04	280.8829	280.312
14	186.2871	185.812
16	96.085371	96.437
17	85.162329	85.562
18	66.898923	67.312

Table 2: The obtained compass values of the Yoctopuce 3D and the ground truth northing angle calculated in UTM coordinate system.

the reference points. It can be assumed that small distances induce a larger systematic error from surveying with the prism.

The compass error was smaller than 0.5 degree overall, which is quite accurate and sufficient for AR. Hwoever, this test was taken without any metallic objects in the vicinity. Cars and steel-reinforced concrete leads to more outliers; making it necessary to rely on vision for correction in dense urban environments.

4.3 Visualization Attempts

For a first application, GIS data are used to visualize buildings within the hardware tracking setup. The sensor cube was mounted onto a Microsoft Surface Pro 3. On the right of Fig. 1, some visualizations with high-precision localization are depicted. Sensor-only tracking is prone to noise, especially affecting compass readings. Orientation error leads to larger registration errors for far away virtual objects. Most of the time, tracking seem to be quite accurate in semi-dense urban areas. Simple vision tracking approaches, like keyframe tracking, would suffice to build a robust hybrid tracking system. The interested reader is referred to the work of Liu *et al*. [15] for an approach towards achieving robust tracking.

5 Conclusion and Future Work

The sensor cube provides accurate and robust 6DOF localization. It demonstrates that inexpensive sensors can keep up with much more expensive high-precision setups used in previous approaches. However, sensor tracking alone cannot guarantee robustness at all time. Especially in urban areas, satellite shadowing and reflections from buildings effect GPS accuracy. Similarly, the magnetometer is easily affected by close-by metallic objects, such as cars.

For these reasons, we built a hybrid tracking system to combine the advantages of vision and sensor tracking. The hybrid approach already provides reasonable additional value to the sensor setup. Nevertheless, the current SLAM setup is not working in real-time on mobile devices yet and is under construction.

Future steps include the investigation of suitable vision tracking extensions, whether using the current SLAM approach or a different one, *e.g.,* the use of edge matching. The goal is to create a system which is accurate and precise in most everyday situations. Moreover, the sensor cube's size is reduced to approximately 5×5 cm by designing an integrated circuit board, which is already in progress and the first prototype is exhibit in figure 6. However, this first iteration of the circuit still has some construction errors, which have to be fixed. Nevertheless, this prototype already shows the compactness we wanted to go for. Another design challenge will be to find a suitable antenna of a minimum size while still providing reasonable GPS signals for highly accurate positioning. Finally, a calibration routine for registration of the the cube to a discrete camera would be of great practical value for applications.

References

[1] B. Jiang, U. Neumann and S. You. A robust hybrid tracking system for outdoor augmented reality. *Int. Symp. on Virtual Reality*, pages 266–275, 2004.

[2] C. Arth, C. Pirchheim, J. Ventura, D. Schmalstieg and V. Lepetit. Instant Outdoor Localization and SLAM Initialization from 2.5D Maps. *IEEE Trans Vis Comp Graph*, pages 1309–1318, 2015.

[3] D. H. Shin and P. S. Dunston. Identification of application areas for Augmented Reality in industrial construction based on technology suitability. *Automation in Construction*, 17(7):882–894, 2008.

[4] E. Artemciukas, L. Sakalauskas and E. Zulkas. Kalman Filter for Hybrid Tracking Technique in Augmented Reality. *Elektronika ir elektrotechnika*, 22:1392–1215, 2016.

[5] E. Bostanci, N. Kanwal, S. Ehsan and A. F. Clark. User Tracking Methods for Augmented Reality. *Int. Journal of Computer Theory and Engineering*, 5:93–98, 2013.

[6] G. King, W. Piekarski, and B. Thomas. ARVino - Outdoor Augmented Reality Visualisation of Viticulture GIS Data. *Proc. ISMAR*, 2005.

[7] G. Reitmayr and T. W. Drummond. Going out: Robust Model-based Tracking for Outdoor AR. *Proc. ISMAR*, pages 109–118, 2006.

[8] G. Reitmayr and T. W. Drummond. Initialisation for Visual Tracking in Urban Environments. *Proc. ISMAR*, pages 161–172, 2007.

[9] G. Schall, A. Mulloni and G. Reitmayr. North-centred orientation tracking on mobile phones. *Proc. ISMAR*, 9:267–268, 2010.

[10] G. Schall, D. Wagner, G. Reitmayr, E. Taichmann, M. Wieser, D. Schmalstieg and B. Hofmann-Wellenhof. Global Pose Estimation using Multi-Sensor Fusion for Outdoor Augmented Reality. *Proc. ISMAR*, pages 153–162, 2009.

[11] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, D. Schmalstieg. Handheld Augmented Reality for Underground Infrastructure Visualization. *Pers. Ubiq. Comp.*, 4(13):281–291, 2009.

[12] G. Schall, S. Zollmann and G. Reitmayr. Smart vidente: advances in mobile augmented reality for interactive visualization of underground infrastructure. *Pers. Ubiq. Comp.*, page 15331549, 2012.

[13] J. Karlekar, S. Z. Zhou, W. Lu, Z. C. Loh, Y. Nakayama, and D. Hii. Positioning, tracking and mapping for outdoor augmentation. *Proc. ISMAR*, page 175184, 2010.

[14] M. Hakkarainen, C. Woodward and K. Rainio. Software Architecture for Mobile Mixed Reality and 4D BIM Interaction. *25th CIB W78 Conference*, pages 1–8, 2009.

[15] Ruyu Liu, Jianhua Zhang Shengyong, Clemens Arth. Towards SLAM-based Outdoor Localization using Poor GPS and 2.5D Building Models. *Proc. ISMAR*, 2019.

[16] Dieter Schmalstieg and Tobias Höllerer. *Augmented Reality - Principles and Practice*. Addison-Wesley Professional, jun 2016.

[17] B. Thomas, V. Demczuk, W. Piekarski, D. Hepworth, and B. Gunther. A wearable computer system with augmented reality to support terrestrial navigation. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, pages 168–171, Oct 1998.

[18] Jonathan Ventura, Clemens Arth, Gerhard Reitmayr, and Dieter Schmalstieg. Global Localization from Monocular SLAM on a Mobile Phone. *IEEE Trans Vis Comp Graph*, 20(4):531–539, mar 2014.

[19] W, Piekarski and B. H. Thomas. Tinmith-Metro: New Outdoor Techniques for Creating City Models with an Augmented Reality Wearable Computer. *Int. Symp. on Wearable Computing*, page 3135, 2001.

[20] W. T. Fong, S. K. Ong and A. Y. C. Nee. Computer Vision Centric Hybrid Tracking for Augmented Reality in Outdoor Urban Environments. *Int. Conf. on Virtual Reality Continuum and its App. in Industry*, pages 185–190, 2009.