

High-precision optical position measurement in indoor environments*

A method for locating people by capturing multiple optical markers with a mobile device

Ralf Vandenhouten¹ and Ole Wendlandt¹

¹ Wildau University of Applied Sciences, Hochschulring 1, 15745 Wildau, Germany

Abstract. Indoor localization is still a challenge today. Conventional methods based on satellite navigation, like GPS, can only be used outdoors. On the other hand, radio signal architectures, like WiFi or NFC, require high cost infrastructure to be installed inside the building. In collaboration with ixellence GmbH, the research group for telematics at Wildau University of Applied Sciences is developing a new inexpensive solution based on the camera of a mobile device. This approach only requires the availability of multiple identifiable landmarks that can be installed cost-efficiently in an indoor environment. Our investigations show that this method is very accurate and can produce repeatable, robust results.



Figure 1. Indoor navigation with a tablet using the proposed method

1 Indoor localization today and related problems

Modern mobile devices typically have a sensor for satellite navigation, mostly GPS. This allows spatial outdoor tracking with an accuracy of one or a few meters, depending on the conditions of the signal transmission. Since this type of detection is based on the reception of the very weak satellite signals that are available only outside of buildings, an indoor localization is not possible.

Therefore, other methods have been developed for indoor positioning, based on the fact that the mobile device receives a number of signals from a special infrastructure installed in the building (e.g. WiFi or infrared transmitters) and uses them for calculating its location, e.g. by triangulation. Apart from the relatively high infrastructure costs, this approach has further disadvantages. Firstly, the accuracy is affected due to reflections of the signals on walls and objects in the room. Using calibration methods these

errors can indeed be partially compensated, but such methods are complex and expensive. Secondly, the positioning using radio signals is not very robust, as they can be disrupted in an unpredictable way related to moving objects (such as people, furniture, shopping carts). For the above reasons, indoor positioning methods based on radio infrastructure could not prevail in broad so far.

The presented method in this article dispenses with an elaborate infrastructure and uses - besides ordinary mobile devices - only simple and inexpensive landmarks and 2D codes. The application field for such indoor positioning ranges from navigation solutions for large buildings or localized advertising at the point-of-sale through to location-based value-added services, offered to the user on the mobile device depending on his position [1].

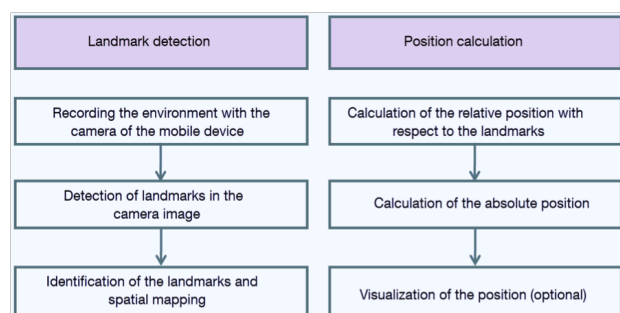


Figure 2. Processing steps of localization

2 Architecture of our optical indoor localization

The basic architecture of the system consists of the landmark infrastructure, a mobile device such as a tablet or

*This work was published in *Optik & Photonik* 10 (3), 2015.

mobile phone that will be located, and an optional server for data storage. The building in which to place the localization system must be equipped with landmarks. Landmarks can be two-dimensional (e.g. a print on a wall) or three-dimensional (e.g. sphere, cube, cuboid, hexagonal shapes). In our first approach we use datamatrix codes on a two dimensional robust cardboard with a colored rectangle. For reading datamatrix codes ixellence GmbH provides a quick (real time) and reliable decoder (ixMAT) for mobile devices which is optimized for long distance scanning. In general the landmarks should be optically prominent (e.g. cubes with colored edges) so that they can be identified easily. Each landmark has to be positioned in a way that it is clearly visible from as many positions in the area as possible. If a landmark has several outer surfaces facing in different directions of the area, it is advantageous if each of these surfaces has an individual marking in order to be distinguishable, and so the viewing direction can be determined reliably. The size, the number and spatial density of landmarks to be installed depend on the site area and the desired spatial resolution of the location.

The localization itself is done in two steps (Fig. 2). One step is the image processing part that detects the landmarks in the digital image that is received from the camera of the mobile device. The second step is computing the position of the device and consequently of the person using it. For this the research group for telematics at Wildau University of Applied Science developed an algorithm that processes both steps in real time on the mobile device. The relative position and orientation of the mobile device with respect to the landmarks is determined based on the size of the visible landmarks, their positions in the image and their perspective distortion by estimating the camera pose from 3D-2D point correspondences. From the relative position of the mobile device with respect to the landmarks and the known absolute position of these landmarks as well as their orientation the absolute position and orientation of the mobile device can be calculated. The absolute positions of the landmarks and their orientation can be stored either in the mobile device itself or on a server with a database to which the mobile device can connect wirelessly and download required data.

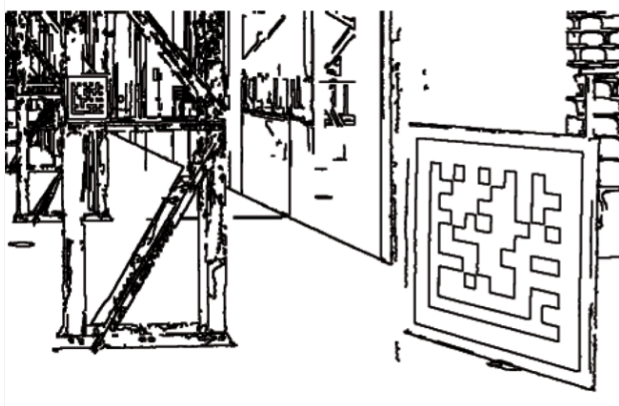


Figure 3. Detection of edges with Canny algorithm

3 Landmark detection

The layout of our landmark contains a datamatrix code with an identifier for the position of the landmark in the database and a surrounding rectangle.

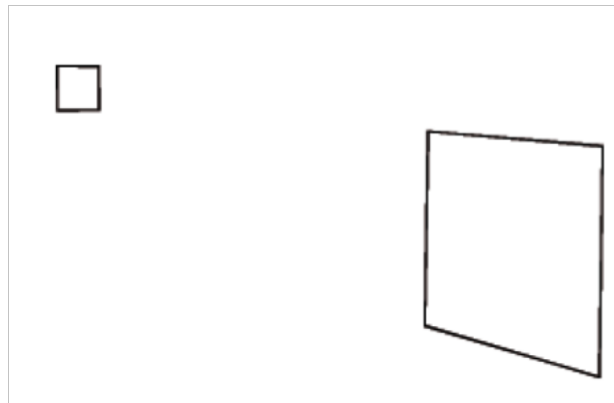


Figure 4. Candidates of landmark images

The detection process of the landmarks in the camera image is done in a two-stage procedure. In the first stage all rectangles are detected. Because of the morphologic characteristics rectangles are easy to detect in a camera image of an environment. The rectangles further allow computing the transformation matrix to get an undistorted rectangle that may contain a datamatrix barcode. The computation of the undistorted rectangle is a pre-processing step that simplifies and increases robustness of further image processing steps (e.g. decoding of the datamatrix label). The detection of rectangles is implemented with an edge detection filter on a grey-image using the well-known Canny algorithm [2]. The result of the Canny algorithm is a binary image containing only two color values, black and white. White pixel values belong to an edge (Fig. 3). Within the binary image, we are looking for those points that approximate closed polygons with only four edges using topological structural analysis and curve approximation. From this polygon approximation we get a list of sub-images that may contain possible landmarks – see Fig. 4 for positions of possible landmarks in the source image.

The last pre-processing step is dewarping the perspective of each sub-image of a landmark candidate. The perspective of the landmark is transformed to a central perspective to make it orthogonal to the camera viewpoint, like the camera would be positioned directly in front of the landmark. As mentioned, this transformation increases the robustness of the decoding of the datamatrix code within the rectangle.

The second stage deals with the decoding of the datamatrix codes in our landmarks. All possible landmarks are processed by the ixMAT library [3, 4] from ixellence GmbH. As a result we get the identifier of the landmark in our database where the known absolute position is stored and additionally the position in the camera image. False positive landmarks (rectangles containing no barcode) are filtered out, when the decoder cannot identify them as

datamatrix codes. The decoding is parallelized for speeding up the real time detection.

4 Position calculation

The position calculation can be performed now. The inputs are the absolute 3D positions of the landmarks stored in a database and the relative (2D) positions of the detected landmarks computed in the preceding step. We calculate the position of the camera (i.e. the user position) from the 2D-3D correspondence of the absolute landmark positions in the real world and the relative landmark positions in the source image. The calculation of the 2D-3D correspondence is called a Perspective-n-Point (PnP) problem. With solving the PnP problem the position and orientation of a calibrated camera can be determined. We use the iterative Levenberg-Marquardt algorithm [5] that provides an accurate solution to the PnP-Problem. The result of the computation process is a translation vector (t) and a rotation matrix (R). With t and R it is now possible to transform points from 2D space (R^2) to 3D space (R^3) and vice versa. This is based on the pinhole camera model described in Hartley and Zissermann [6]. The 3D space is the coordinate system of the real world with the known absolute positions of the landmarks. The 2D space is a plane (i.e. the image plane) within the camera coordinate system (R^3). The distance between the image plane and the camera position is called focal length. The focal length belongs to the set of intrinsic camera parameters and is determined by a pre-processed calibration procedure for the camera of the mobile device.

Between 2D space and 3D space, the following coordinate transformation equation can be used:

$$\begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}. \quad (1)$$

(x,y,f) is the position on the image plane, i.e. camera coordinate system, which is computed with the rotation matrix R , the 3D space point (X,Y,Z) and the translation vector (t_x,t_y,t_z) . The equation above can also be written as follows:

$$P_{camera} = R \cdot P_{world} + t. \quad (2)$$

Now we want to determine the camera position in 3D space. We know the position of the camera in the camera coordinate system which is exactly the centre C :

$$P_{camera} = C = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3)$$

This can be translated to the 3D space with the relation derived from eq. 2 above:

$$-P_{world} = R^{-1} \cdot t. \quad (4)$$

The result $-P_{world}$ is the absolute camera position in 3D space with the coordinates (X, Y, Z) . These coordinates

can be displayed to the user or visualized graphically as his position in the building (see Fig. 1).

The results of our validation procedure in the laboratory of computer vision are relatively robust. We compared different positions in 3D space measured manually with the results of our indoor location method. The relative error for the x- and y-coordinates was about 3.5 – 3.8%, which corresponds to an absolute deviation of 10-12 cm.

5 Conclusion and future work

We have implemented our method in a prototype running on an Android tablet device with a dual core A15 ARM processor. With this kind of processor type it is possible to obtain a frame rate of 10 frames per seconds. Compared to other approaches, the method is very precise and robust with respect to changes in the environment. A major advantage of this technique lies in the fact that – apart from the inexpensive landmarks – it does not need any additional components or equipment but can be implemented by simply using commercially available smartphones. This ensures low investment costs and makes the method interesting for the growing market of indoor tracking solutions. Our future work will be focused on including additional sensors of the mobile device like acceleration and / or gyroscope sensors. These sensors can be used to improve the accuracy of the position. Furthermore, a tracking component with a dead reckoning component could be realized. With these technologies the mobile device can be tracked without permanently having landmarks visible from the camera. Beyond determination of the absolute position other content can be shown on the display of the device. Key point for that is the integration of location-based content for the user (e.g. advertising in a supermarket). A next step could be the integration of augmented reality scenarios to the user. Augmented reality can enrich the real live images of the camera with additional information on the user interface.

References

- [1] D. Mietzner, R. Vandenhoueten, C. Schultz, *Neue Ansätze in der Indoor-Navigation - Echter Mehrwert für den stationären Handel?* (Shaker, Berlin, 2014), pp. 49–62, Forum Geo-Bau, ISBN 978-3-8440-2643-6
- [2] J. Canny, IEEE Trans. on Pattern Analysis and Machine Intelligence **8**, 679 (1986)
- [3] *The mobile phone as a mobile scanner* (2015), <http://www.ixellence.com/ixmat>
- [4] R. Vandenhoueten, M. Selz, *Identification and tracking of goods with the mobile phone*, in *2007 International Symposium on Logistics and Industrial Informatics* (IEEE, 2007), pp. 25–29, ISBN 978-1-4244-1440-6
- [5] D. Marquardt, Journal of the Society for Industrial and Applied Mathematics **11**, 431 (1963)

- [6] R. Hartley, A. Zisserman, *Multiple view geometry in computer vision* (University Press, Cambridge, UK, 2003)

Authors

Prof. Dr. Ralf Vandenhousten is vice president for research and business relations and head of the department of telematics at the Wildau University of Applied Sciences. His research focuses on software development for telematics applications, mobile computing and image processing. He has founded several companies and developed software tools for signal processing and time series

analysis. He received his doctoral degree at the Technical University Aachen (RWTH) in 1997 and is a professor at Wildau University of Applied Sciences since 2000.

Ole Wendlandt works as a scientific assistant in the research group for telematics at the Wildau University of Applied Sciences. Research topics are plug-in-oriented software engineering, image processing and medical data processing. He received his Master of Engineering in 2010.

Contact: Prof. Dr. Ralf Vandenhousten, Department of Telematics, Wildau University of Applied Sciences, Hochschulring 1, 15745 Wildau, Germany. E-Mail: ralf.vandenhousten@th-wildau.de