Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication

Gyula Simon, *Member, IEEE*, Gergely Zachár, and Gergely Vakulya

***Abstract*— A novel indoor localization system is presented, where LED beacons are utilized to determine the position of the target sensor, including a camera, an inclinometer, and a magnetometer. The beacons, which can be a part of the existing lighting infrastructure, transmit their identifiers for long distances using visible light communication techniques. The sensor is able to sense and detect the high-frequency (flicker free) code by properly undersampling the transmitted signal. The localization is performed using novel geometric and consensus-based techniques, which tolerate well measurement inaccuracies and sporadic outliers. The performance of the system is analyzed using simulations and real measurements. According to large-scale tests in realistic environments, the accuracy of the proposed system is in the low decimeter range.**

***Index Terms*— Accelerometers, cameras, estimation, magnetometers, position measurement, sensor fusion.**

I. Introduction

WHILE GPS mostly solved the problem of outdoor localization, indoor localization is still a challenging research area. Most of today's low-cost systems use WiFi signals to provide a few meters of accuracy, with the great advantage of low (zero) deployment cost [1], [2]. For higher accuracy, other approaches are actively researched, utilizing e.g., ultra-wideband (UWB) radio time of flight (TOF) [3], sound or ultrasound [4], [5], laser scanners [6], inertial sensors [7], or various optical solutions [8], [9].

Since LED lighting becomes more widespread, the utilization of such existing infrastructure for localization is a plausible idea. Using ceiling landmarks, especially light sources, was considered also advantageous in earlier research since they are easy to detect and their positions do not change in time [10].

In this paper we propose Lookup, a novel indoor localization system utilizing LED-based beacon infrastructure and a sensor containing a camera, an inclinometer, and a digital compass. The accuracy of the proposed system, according to tests, is comparable with the best of today's technologies.

In the proposed system LED beacons are utilized, which transmit their identifiers using visible light communication (VLC). The coding utilizes high frequency modulation of the light source, providing flicker-free operation, thus beacons can be part of the lighting system. The high frequency codes are detected by the camera. Properly designed codes allow undersampling of the transmitted signals, thus conventional cameras can be utilized. The detection is aided by auxiliary sensors, providing the inclination and optionally the heading of the camera. From the detected beacon positions a novel consensus-based localization method determines the location of the camera with high accuracy.

Extensive error analysis will be presented using measurements and simulations, to characterize the operation of the sensors and the localization method. Results of real tests, conducted in realistic scenarios, will be presented. The performance of the system will be compared to that of other systems, utilizing today's emerging technologies.

The novelties of the proposed solutions are the following.

1) Novel VLC-based beaconing and detection scheme is proposed, which can be utilized in existing lighting infrastructure and require only inexpensive sensors.
2) Novel consensus-based localization method is proposed, which tolerates well measurement errors and outliers.
3) Extensive error analysis is provided for the implemented system, utilizing real and realistic measurement scenarios.

In Section II, related work is reviewed. Section III introduces the proposed localization system and Section IV includes the error analysis. In Section V, the performance of the system is evaluated, using real measurements. In Section VI, the proposed system is compared to other indoor localization techniques. Section VII concludes the paper.

II. Background and Previous Work

A. Indoor Localization Systems

Nowadays, there are numerous approaches for indoor localization, including radio signals, light, sound waves, and magnetic fields. One of the most common methods utilizes the propagation characteristics of the radio waves. The measured received signal strength (RSS) correlates with the distance between the transmitter and the receiver. Devices can be a part of a readily available infrastructure e.g., WiFi, Bluetooth, or proprietary radios. Localization is possible with multiple devices based on the known location of the anchor devices and the propagation model [2]. Other solutions use fingerprints or maps, which are constructed from previous RSS measurements at several known positions [1]. The fusion of the two basic methods is also widely used. The accuracy of the position estimation is often enhanced by additional motion-based methods for short distances.

Manuscript received August 12, 2016; revised February 10, 2017; accepted March 28, 2017. Date of publication June 12, 2017; date of current version August 9, 2017. This work was supported by Széchenyi 2020 under Project EFOP-3.6.1-16-2016-00015. The Associate Editor coordinating the review process was Dr. Antonios Tsourdos. *(Corresponding author: Gyula Simon).*

The authors are with the Department of Computer Science and Systems Technology, University of Pannonia, 8200 Veszprém, Hungary (e-mail: simon@dcs.uni-pannon.hu; zachar@dcs.uni-pannon.hu; vakulya@dcs.uni-pannon.hu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2017.2707878

Motion-based localization (e.g., pedestrian dead reckoning) is possible by calculating the covered distance and orientation using various sensors, e.g., accelerometer, magnetometer, and gyroscope, which are available in most smartphones. These techniques are relatively accurate on small distances, but due to the accumulating errors, are often supported with other, usually RSS-based localization methods [7], [11].

Localization methods are often based on the measurements of signal propagation times. The distance between two devices can be calculated by measuring the TOF, or the time difference of arrival (TDOA) of signals (e.g., radio waves and sound) in various mediums (e.g. air and concrete). The most common applications based on radio waves are utilizing UWB signals for ranging [3], [16]. A novel ranging technique using radio phase difference measurements and super-resolution techniques was recently proposed [17]. Ultrasonic signal-based solutions are also available, requiring much simpler hardware solutions, due to the slower propagation speed of sound [5]. Using quadrature detectors and parabolic interpolation sub-centimeter ranging accuracy is achieved [18]. Uncommon applications are also present, e.g., measuring TDOA of vibrations in concrete, which can be the base for positioning in buildings [19]. For the fusion of TOF/TDOA measurements various trilateration techniques are proposed. The analytical approach was recently extended by optimal selection of the utilized anchor nodes [20].

Magnetic fields are also used for localization purposes. These fields are not influenced by humans or furniture in contrast to the radio waves. One possible application is to create an oscillating magnetic field and measure the magnetic resonant coupling. With multiple emitters the receiver position can be calculated [21]. Other solutions, which do not require anchors, are utilizing magnetic signal strength maps from *a priori* measurements [22].

The latest 3-D sensor technology (e.g., laser scanners and 3-D cameras) enables the creation of real time 3-D point cloud representation of the environment. Localization is based on fitting the measured model to a previously created reference model [6].

Light-based approaches were also proposed, using multiple light sources as anchors, commonly installed as a part of the lighting infrastructure. The receivers (light sensors) measure the modulated light's intensity and from the modulation also identify the beacons. From the light intensity, the range is estimated and the location is calculated using trilateration [9], [23].

Image processing is widely used for positioning purposes. Several possibilities are present from passive, discrete position localization (e.g., quick response-codes on important places) to active positioning, which is utilizing light sources as anchors [8], [10]. VLC is often deeply integrated part of these systems.

B. VLC in Localization

In most localization systems known anchor positions are used to calculate the current position of the moving device. In systems that are based on light (either imaging or light sensors), it is necessary to distinguish between the light sources, used as anchors. Therefore anchor nodes often transmit their identification code, using VLC. This is enabled by current LED technology, where instantaneous switching ON and OFF the light at high frequency is possible, providing a convenient method for encoding information. Note that while in VLC systems, high bandwidth is often a requirement, in localization systems, only small amount of information is to be transmitted.

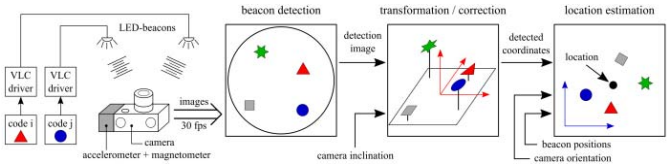


Fig. 1. Block diagram of Lookup.

Some of the VLC methods modulate the light in different frequencies—similar to radio transmitters—allowing data coding. This can be done e.g., with binary frequency shift keying, utilizing two different frequencies [9], [23]. Note that channel collisions may also be present. The receiver side utilizes only one photo-detector and samples the light intensity signal. The actual IDs can be retrieved with continuously analyzing the frequency spectrum. In these systems the ranging information (distances between the light sources and the receiver) is calculated based on the magnitudes of the associated frequencies and on the propagation model.

VLC systems are utilized with cameras (e.g., mobile phones) as well. Most of these methods are based on the rolling shutter effect: the image sensors do not capture the entire image at the same time, but rather scan the rows after each other. Due to this operation mode different pixels are sampled at different time instants, thus the image of the modulated light source contains alternating dark and light bands, which can be used to determine the transmitted code [8]. Note that this solution converts the temporal information to spatial information, thus a single image is enough to detect the transmitted code. The disadvantage of the rolling shutter solutions is that relatively large image is required to correctly decode the message; thus, either large-sized light sources must be used or the camera must be close to the source.

III. Lookup System

A. System Overview

The operation of the proposed system is shown in Fig. 1. The beacons are LED lamps, which can possibly be part of the existing lighting infrastructure. The coordinates of the beacons are known in the world coordinate system *CW*. The LED lamps' illumination is controlled by the VLC drivers, providing modulated light sources, each transmitting its unique code. The modulation utilizes high frequency, which is not visible for human eyes (flicker-free). The lamps are detected by a fisheye camera, which looks upward. Notice the important fact that in this case the image center corresponds to the ceiling position under which the camera is placed. The localization goal is equivalent to finding the position of the camera, i.e., the image center in *CW*.

The video stream of the observed ceiling is used to detect and identify beacons. Since the modulation of the beacons has

TABLE I

MAIN PROPERTIES OF THE LOCATION ESTIMATION METHODS. ACC: ACCELEROMETER, MAG: DIGITAL COMPASS, AND H: ALTITUDE OF THE CAMERA

Beacons	Sensor	Additional req.	Estimation principle	Accuracy
1	ACC, MAG	H	distance, orientation	poor
2	ACC	(H)	distance	medium
≥3	ACC	-	angle	high

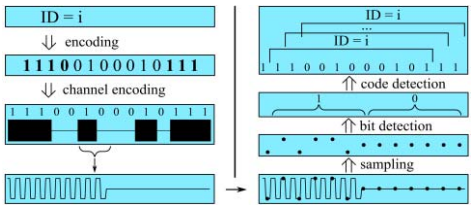


Fig. 2. Beacon encoding and detection.

much higher frequency than the available sampling frequency of conventional cameras, the received beacon signals are undersampled. The beacon modulation is designed so that the undersampled signal can efficiently be used for the detection and identification of beacons. The details of beacon detection will be described in Section III-B.

The detected beacon coordinates on the fisheye image are transformed to the orthographic camera coordinate system *CC* (shown by red axes in Fig. 1), the origin of which is the image center, and its axis z is parallel with that of *CW*. Since the camera may not be looking exactly upward, the camera inclination, measured by the 3-D accelerometer, is used to compensate for the effect. In this processing step filtering and masking are also performed to compensate for detection errors and inaccuracies, as will be described in Section III-C.

The location estimation utilizes the detected and estimated beacon coordinates in *CC*, and the known beacon coordinates in *CW*. Three different location estimation methods are used, depending on the number of detected beacons. In case of one detected beacon the camera orientation, measured by a digital compass, is also utilized. When two beacons are present, the beacon positions are used to calculate the location. When three or more beacons are available, only the angles between the detected beacons are used. The main features of the algorithms, including the utilized sensors, additional requirements, the estimation principle, and the accuracy, are summarized in Table I, and the details will be described in Section III-D.

B. Beacon Detection

The proposed infrastructure-based localization system implements LED beacons and the corresponding camera-based detection algorithm. In the system (see Fig. 1), beacons transmit their unique IDs encoded as blinking patterns, similar to other VLC systems. From the video stream of the camera, the IDs and the detected positions of the beacon nodes can be recovered on the receiver side. A layer-based model is used to transmit and receive beacon IDs, as shown in Fig. 2. First, the ID is encoded into a continuously repeated bit pattern containing n_{CODE} bits. The code segments are preceded with a header of length of n_{HEAD} (in Fig. 2 the header is 1110). Notice that the header bit pattern is not allowed in any other position of the code. Channel encoding generates the appropriate blinking pattern and the power level settings for the LED. During bits with value 1 the LED is driven with full power and blinked with 50% duty cycle and frequency f_{LED}. The zero bits are encoded with constant half-bright light, as shown at the lowest layer of Fig. 2.

On the receiver side the LED beacons are observed with a camera. Regular cameras cannot provide sampling rates to fulfill the Nyquist criterion of the high-speed light intensity change of the beacons, required for flicker-free operation (typical camera frame rates are around 30 fps); thus; undersampling is proposed. With high-speed shutter mode (where shutter time is much smaller than the sampling period) the camera takes short samples (frames), using its normal sampling frequency f_S. With this method, the undersampled alias frequency can be detected: for bit 0, there is no significant change between samples, while for bit 1, the undersampled signal oscillates.

After adaptive thresholding of the frames, for each pixel a bit stream containing oscillating and constant segments is gained. The segments, corresponding to bits, are then identified with a state machine-based detection algorithm, which examines whether the samples are oscillating (bit 1) or not (bit 0), and also measures the length of the pattern (to detect multiple consecutive 1's or 0's). For each pixel the detected bit series is stored.

The code detection algorithm continuously tries to match the last $n_M = n_{\mathrm{HEAD}} + n_{\mathrm{CODE}}$ bits to all possible IDs used in the system, thus all identifiers are detected in parallel. For this a table of length 2^{n_M} is utilized, which is indexed by the last n_M detected bits, to provide fast ID decoding. Note that the code detector can provide the ID in any bit position of the code, due to the cyclic property of the codes, as illustrated in Fig. 2. Thus, the algorithm produces a set of matching pixels for every corresponding ID, and for each matching pixel the number of detections is also stored. (The latter is used for detection quality measurement, as described in Section III-C3.) The set of matching pixels is further processed: connected regions (blobs) are selected, and for each blob the centroid (x_F, y_F) of the blob and the sum of the number of detections TD is calculated.

The final result of the beacon detection is a set of triplets $\{(\mathrm{ID}, (x_F, y_F), \mathrm{TD})\}$, each triplet containing the detected code, the detected coordinates of the beacon, and the total number of detections. Notice that ideally the set contains one triplet for each visible beacon, but in real cases incorrect multiple detections may also be present.

In the proposed system $n_{\mathrm{CODE}} = 7$ and $n_{\mathrm{HEAD}} = 4$ was used with $f_{\mathrm{LED}} = 165$ Hz and $f_S = 30$ Hz, resulting 15 Hz alias frequency. The bit lengths were set to approximately 150 ms, thus approximately $t_{\mathrm{ID}} = 1.7$ s is necessary to transmit and detect an ID. We used redundancy to provide more robust detection: the detector was operated for 4 s. The implemented detector algorithm, running on an ordinary laptop, was able to process the 30 fps 1080p video stream in real time.

Because of the considerable amount of detection time the system is to be used with static camera. Now the effects of the

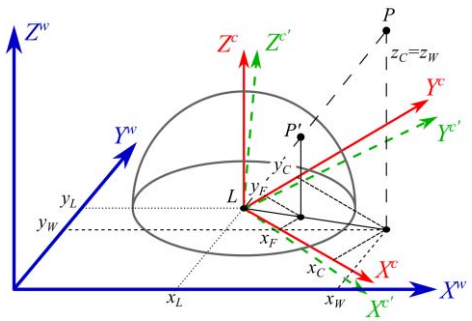


Fig. 3. Coordinate systems and transformations of a detected beacon.

possible camera movement on the detection is discussed and an upper bound is provided for the maximum allowed camera speed. For a successful detection of a beacon, at least some pixels must be decoded, which implies that some common pixels must be present in all of the images of the beacon, while the ID is transmitted (in our case $t_{\mathrm{ID}} = 1.7$ s). Thus, the image of the beacon on the first and last frame in this interval must be overlapping. To fulfill this requirement, the relative speed between the camera and the beacon must be smaller than $d_{\mathrm{LAMP}}/t_{\mathrm{ID}}$, where d_{LAMP} is the diameter of the lamp. For example, with $d_{\mathrm{LAMP}} = 5$ cm the upper bound for the camera speed is around 3 cm/s, which is impractical for real dynamic applications. Thus, the method is proposed for static cameras only.

C. Beacon Preprocessing

1) Position Transformation: The goal of the position transformation step is to restore the orthographic projection of the real objects onto plane xy, from the heavily distorted image of the fisheye camera.

The utilized coordinate systems and their relationships are shown in Fig. 3. The location of beacon P is (x_w, y_w, z_w) in the world coordinate system CW, denoted by the blue axes. The unknown location L is the center of the camera, with unknown coordinates (x_L, y_L, z_L) in CW. Let us assume that the altitude z_L of the camera is known. (We will relax this constrain when multiple beacons are available.) For sake of simplicity and without loss of generality, let us choose CW so that $z_L = 0$, as shown in Fig. 3.

If the camera is looking upward, the camera coordinates systems CF and CC, both denoted by the same red axes in Fig. 3, are centered at L and axis Z^W is parallel to Z^C. Note that the camera orientation, i.e., the angle between X^W and X^C, is not known and when required, is estimated by the system.

The light from P intersects the unity sphere, placed around L, in P'. The fisheye camera in CF detects the orthographic projection of P' into plane xy as image coordinates (x_F, y_F). The z coordinate of P' on the unity sphere is $z_F = (1 - x_F^2 - y_F^2)^{1/2}$. Since the z coordinate z_w of P is known, using the resize factor

$$s = z_w/z_F \tag{1}$$

the coordinates of P in CC can be computed as follows:

$$P_C = (x_C, y_C) = (x_F s, y_F s). \tag{2}$$

Notes:

Real cameras do not implement the exact orthographic projection of P', illustrated in Fig. 3. Camera distortions, however, can be measured and compensated for. In the proposed system, the camera was calibrated by the OCamCalib toolbox [14]. For the details of camera distortion model and the calibration method, see [15]. The calibrated camera approximates the ideal orthographic projection well. The effect of remaining error will be analyzed in Section IV-A.

The camera is assumed to look exactly upward, but in practice it may be tilted. To compensate for this effect, the inclination is measured by a 3-D accelerometer, as described in the following section.

2) Inclination Correction: If the camera is tilted, the beacon detection is performed in coordinate system *CF'* (shown in green in Fig. 3). The coordinates $(x_{F'}, y_{F'}, z_{F'})$, measured in *CF'*, must be transformed to CF to obtain (x_F, y_F, z_F).

The 3-D accelerometer (type BMA180) is attached to the camera, thus the axes of the camera and the accelerometer are aligned. The accelerometer detects the direction of gravitational force, which is identical to $-Z^W$, and this vector is measured as $n = (n_x, n_y, n_z)$ in CC', where $|n| = 1$ unity vector in the common camera-accelerometer coordinate system.

To compensate for the inclination, a transformation T_{inc} is required, which transforms axis $-Z^{C'}$ into vector n (and thus $Z^{C'}$ into Z^C). Using e.g., a modified version of Rodriguez's formula [12], the transformation has the following form:

$$T_{\mathrm{inc}} = \begin{bmatrix} c\alpha & -s\alpha & 0 \\ s\alpha & c\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c\beta & 0 & -s\beta \\ 0 & 1 & 0 \\ s\beta & 0 & c\beta \end{bmatrix} \begin{bmatrix} c\alpha & s\alpha & 0 \\ -s\alpha & c\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

where $sx \equiv \sin x$, $cx \equiv \cos x$, $\alpha = \mathrm{atan2}(n_y, n_x)$, $\beta = \mathrm{atan2}(R_{xy}, n_z)$, and $R_{xy} = (n_x^2 + n_y^2)^{1/2}$.

Thus if the image coordinates in the tilted coordinate system *CF'* are (x'_F, y'_F, z'_F) then the corrected coordinates (x_F, y_F, z_F) in the ideal CF can be obtained as

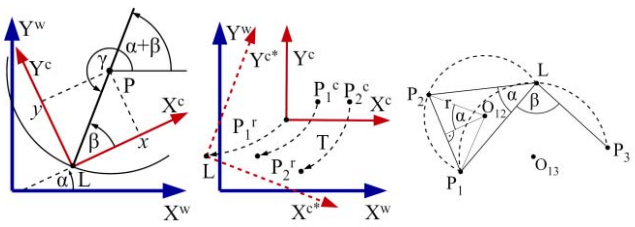
$$\begin{bmatrix} x_F \\ y_F \\ z_F \end{bmatrix} = T_{\mathrm{inc}} \begin{bmatrix} x'_F \\ y'_F \\ z'_F \end{bmatrix}. \tag{4}$$

3) Filtering of Detections: The detected beacon positions are filtered using the detection quality index (DQI) and a consistency check, as follows.

The quality of the detection is quantified using the size of the image of the detected beacon, and the total number of detections. If the number of transmitted bits is $N_B > N_M$ then the DQI for detection i is the following:

$$\mathrm{DQI}(i) = \frac{1}{N_B - N_M + 1}\mathrm{TD}(i). \tag{5}$$

Detections with low DQI are ignored, using a threshold $\mathrm{DQI}_{\mathrm{MIN}}$, thus sporadic false detections are efficiently filtered out (in practice $\mathrm{DQI}_{\mathrm{MIN}}$ is chosen in the range of 5–7). If for a beacon multiple detections are present, the one with the highest DQI is kept, the others are ignored.



(a) (b) (c)

Fig. 4. Position estimation using (a) one beacon, (b) two beacons, and (c) three or more beacons.

Since distances in *CC* and *CW* under ideal circumstances are the same, this fact can also be used to detect bad detections (e.g., reflected images). For a pair of beacons let the detections in *CC* be $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, while the corresponding beacon positions in *CW* be $P_1^r = (x_1^r, y_1^r, z_1^r)$ and $P_2^r = (x_2^r, y_2^r, z_2^r)$. Using

$$d_{12} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{6}$$

$$d_{12}^r = \sqrt{(x_1^r - x_2^r)^2 + (y_1^r - y_2^r)^2} \tag{7}$$

scaling factor r_{12} is defined as follows:

$$r_{12} = \frac{d_{12}^r}{d_{12}^r}. \tag{8}$$

Ideally $r_{12} = 1$, and any deviation from the ideal value indicates detection inaccuracy/error. In the proposed system, a simple binary quality index is used

$$q_{i,j} = \left\{ \begin{array}{ll} 1 & |1 - r_{12}| < q_{\text{lim}} \\ 0 & \text{otherwise} \end{array} \right\}. \tag{9}$$

Quality value $q_{i,j} = 0$ shows that the detection pair is not consistent, i.e., at least one of the detections is bad. Thus when multiple beacon detections are present, the localization method, discussed in Section III-D3, will only use detection pairs with nonzero quality index.

Parameter q_{lim} in (9) reflects the allowed detection inaccuracy. Small detection inaccuracies are tolerated, but erroneous detections, e.g., reflections from a shiny surface, are filtered out with high probability. Notice that when the altitude of the camera is not accurately known, q_{lim} must reflect the inaccuracy of the camera altitude as well. In the proposed system $q_{\text{lim}} = 0.1$ was used.

D. Position Estimation

In the system a large number of beacons may be deployed but only a subset of them may be visible due to obstacles. The system utilizes the actually visible (i.e., successfully detected) beacons, and depending on their number different localization methods are utilized, as illustrated in Fig. 4.

1) Position Estimation Using One Beacon: Let us denote the detected beacon position in *CC* by $P = (x, y)$, and the corresponding real beacon positions in *CW* by $P^r = (x^r, y^r, z^r)$, as shown in Fig. 4(a). The camera orientation α, with respect to *CW*, is measured by the digital compass. The distance between the camera and the beacon is

$$D = \sqrt{x^2 + y^2} \tag{10}$$

and the measured beacon direction in *CC* is

$$\beta = \text{atan2}(y, x). \tag{11}$$

The camera location in *CW* is on a circle centered at (x^r, y^r) with radius D. From P the direction of the camera is $\gamma = \pi + \beta + \alpha$, as shown in Fig. 4(a). Thus, the location estimate is the following:

$$\begin{bmatrix} x_L \\ y_L \end{bmatrix} = \begin{bmatrix} x^r \\ y^r \end{bmatrix} + D \begin{bmatrix} \cos\gamma \\ \sin\gamma \end{bmatrix}. \tag{12}$$

2) Position Estimation Using Two Beacons: Let us denote the detected positions of the beacons in *CC* by $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, while the corresponding beacon positions in *CW* by $P_1^r = (x_1^r, y_1^r, z_1^r)$ and $P_2^r = (x_2^r, y_2^r, z_2^r)$.

Notice that *CC* provides orthographic projections of real objects to plane xy, in such a way that coordinates in *CC* can be transformed to the real xy coordinates using transformation T, which contains only translation, rotation, and—in real cases—some scaling, as shown in Fig. 4(b). T can be constructed in the following way:

Let us calculate scaling factor r_{12}, according to (8), and let us define transformation T_1 as a scaling with factor r_{12}; (notice that ideally $r_{12} = 1$, thus T_1 compensates for possible detection inaccuracies)

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = r_{12} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{13}$$

Let $\varphi = a\text{tan2}(y_2 - y_1, x_2 - x_1)$, $\varphi^r = a\text{tan2}\left(y_2^r - y_1^r, x_2^r - x_1^r\right)$, and $\Delta\varphi = \varphi^r - \varphi$. Transformation T_2 is a rotation with $\Delta\varphi$, which compensates for the unknown camera orientation

$$T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\Delta\varphi & -\sin\Delta\varphi \\ \sin\Delta\varphi & \cos\Delta\varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{14}$$

Let $\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} x_1^r \\ y_1^r \end{bmatrix} - T_2\left(T_1\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right)\right)$ and T_3 be the following translation (compensating for the unknown camera location):

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \tag{15}$$

Transformation T is defined as

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = T_3\left(T_2\left(T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)\right)\right). \tag{16}$$

Since the unknown location is the camera position in *CW*, the location estimate is

$$\begin{bmatrix} x_L \\ y_L \end{bmatrix} = T(C) \tag{17}$$

where C is the center position of the camera in *CW* [in Fig. 3 $C = (0, 0)$].

Notice that it was assumed that the altitude of the camera is known (i.e., $z_L = 0$). If z_L is not accurately known but the altitudes of the beacons are the same, then scaling factor s in (1) will be inaccurate, but will be the same for both beacons. Thus coordinates P_1 and P_2 will be transformed by the same scaling factor, which is compensated for by T_1. Thus, the camera altitude is not required for the positioning in this case. However, if the beacon altitudes are different, the camera altitude is still required (see Table I).

3) Position Estimation Using Three or More Beacons: Let us denote the detected positions of the beacons in CC by $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, and $P_3 = (x_3, y_3)$, while the corresponding beacon positions in CW are $P_1^r = (x_1^r, y_1^r, z_1^r)$, $P_2^r = (x_2^r, y_2^r, z_2^r)$, and $P_3^r = (x_3^r, y_3^r, z_3^r)$. Let us choose a pair of detected beacons, e.g., P_1 and P_2, as shown in Fig. 4(c). Let us measure angle $\alpha_{12} = P_1 L P_2$, which is the viewing angle between beacons at real positions P_1^r and P_2^r, observed from the unknown location L. Thus, L is on a circular arc corresponding to chord $P_1^r P_2^r$, with inscribed angle α_{12}, as shown in Fig. 4(c). The center O_{12} of the circular arc can be calculated as follows:

$$O_{12} = \boldsymbol{P}_1^r + \boldsymbol{\Delta P}/2 + m\boldsymbol{v_o} \tag{18}$$

where $\boldsymbol{\Delta P} = \boldsymbol{P}_2^r - \boldsymbol{P}_1^r$, $m = |\boldsymbol{\Delta P}|/(2\tan\alpha_{12})$, $\boldsymbol{P}_1^r = [x_1^r, y_1^r]^T$, $\boldsymbol{P}_2^r = [x_2^r, y_2^r]^T$, and

$$\boldsymbol{v_0} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\Delta P}/|\boldsymbol{\Delta P}|. \tag{19}$$

The radius r_{12} of the circular arc is the following:

$$r_{12} = \frac{1}{2}|\boldsymbol{\Delta P}|\sqrt{1 + 1/(\tan\alpha_{12})^2}. \tag{20}$$

Let us denote the circular arc with center (18) and radius (20), created from viewing angle α_{12} of detections P_1 and P_2, and the corresponding beacon locations P_1^r and P_2^r, by $\Theta\ (\alpha_{12}, P_1^r, P_2^r)$.

Using another pair of detected beacons [e.g., P_1 and P_3 on Fig. 4(c)], another arc $\Theta\ (\alpha_{13}, P_1^r, P_3^r)$ can be defined, on which L is located. The intersection of the arcs gives the location estimate. Note that the circular arc, calculated from P_2 and P_3 [not shown in Fig. 4(c)], intersects at the same point.

When more than three beacons are available, the process can be extended and a redundant, more robust solution can be created, by defining circular arcs for every possible pair of beacons. Ideally all of these arcs intersect in one point, corresponding to the camera location L. In the presence of measurement inaccuracies and errors, however, the arcs may produce several intersections. Conventional methods (e.g., least squares), are sensitive to outliers and thus one large measurement error, e.g., detection of a reflected beacon, may result in large estimation error, even if several accurate measurements are available. In the proposed system, a consensus-based fusion is utilized, which takes into account the possible detection errors and tolerates well sporadic large measurement errors as well. Consensus-based approaches were successfully utilized earlier in TDOA-based localization systems [24].

If a beacon is detected at $(x_{F,i}, y_{F,i})$ and the maximal detection error is D pixels, then the correct detection for the beacon may be anywhere inside of a circle with radius D and center $(x_{F,i}, y_{F,i})$ in CF. The maximum angle detection error $\Delta\alpha_i$ for point $x_{F,i}, y_{F,i}$ is the following:

$$\Delta\alpha_i = \arcsin\left(D/\sqrt{x_{F,i}^2 + y_{F,i}^2}\right). \tag{21}$$

Since view angles in CF and CC are equal, the maximum detection error $\Delta\alpha_{ij}$ of view angle α_{ij} is the following:

$$\Delta\alpha_{ij} = \Delta\alpha_i + \Delta\alpha_j \tag{22}$$

thus the view angle, given the maximum detection error D, can be between $\alpha_{12,\mathrm{MIN}}$ and $\alpha_{12,\mathrm{MAX}}$

$$\begin{aligned} \alpha_{12,\mathrm{MIN}} &= \alpha_{12} - \Delta\alpha_{ij} \\ \alpha_{12,\mathrm{MAX}} &= \alpha_{12} + \Delta\alpha_{ij}. \end{aligned} \tag{23}$$

Let us create the set of possible solutions for a pair of detection, assuming maximum detection error D

$$\Omega_{i,j} = \bigcup_{\alpha_{12,\mathrm{MIN}} \le \alpha \le \alpha_{12,\mathrm{MAX}}} \Theta(\alpha, P_i^r, P_j^r). \tag{24}$$

Let us define $f_{i,j}(x, y)$, as follows:

$$f_{i,j}(x, y) = \left\{ \begin{array}{ll} 1 & \text{if}(x, y) \in \Omega_{i,j} \\ 0 & \text{otherwise} \end{array} \right\}. \tag{25}$$

For a position (x, y) in CW, the consensus index is defined as follows:

$$\lambda\ (x, y) = \sum_{\forall i,j,i \ne j} q_{i,j} f_{i,j}(x, y). \tag{26}$$

Thus $\lambda\ (x, y)$ is the number of detected beacon pairs supporting the hypothesis that the unknown location is (x, y), given the maximum detection error D. The location estimate is the following:

$$(x_L, y_L) = \arg\max \lambda(x, y). \tag{27}$$

If (27) provides multiple points, their centroid is used as location estimate.

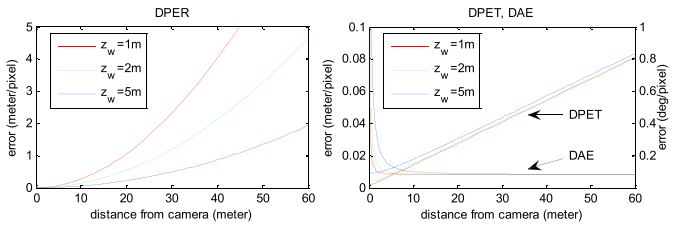
In practice a fast estimation of (27) is utilized as follows. The solution is searched for on a grid with resolution ΔG. For sake of uniformity, let us refer to the grid points by coordinates $(\bar{x}, \bar{y})$, where the coordinates are discrete values in CW, with resolution of ΔG. Initially each point in the grid is marked by 0. For a pair of beacons, the maximum and minimum view angles are calculated from detected coordinates $(x_{F,i}, y_{F,i})$ and $(x_{F,j}, y_{F,j})$, using (21)–(23). Using a series of angles $\alpha_k = \alpha_{12,\mathrm{MIN}} + k\Delta\alpha$, with step size $\Delta\alpha = (\alpha_{12,\mathrm{MAX}} - \alpha_{12,\mathrm{MIN}})/N_\alpha$ and $k = 0, \ldots, N_\alpha$, a series of circular arcs $\Theta_k = \Theta\ (\alpha_k, P_i^r, P_j^r)$ are defined, using (18) and (20). Each arc Θ_k is approximated with N_C equidistant points $p_1, p_2, \ldots, p_{N_C}$, and for each p_i the closest grid point is marked by 1. Thus for each pair of detections the marked grid contains the discrete approximation $\bar{f}_{i,j}(\bar{x}, \bar{y})$ of $f_{i,j}(x, y)$.

Similar to (26), the discrete consensus function on the grid is calculated as follows:

$$\bar{\lambda}(\bar{x}, \bar{y}) = \sum_{\forall i,j,i \ne j} q_{i,j} \bar{f}_{i,j}(\bar{x}, \bar{y}). \tag{28}$$

The centroid of the coordinates $(\bar{x}, \bar{y})$, corresponding to the maximum value of $\bar{\lambda}(\bar{x}, \bar{y})$, results the location estimate.

Since the location estimate is the position where the highest number of beacon detections are in consensus, the fusion can successfully eliminate false detections (e.g., reflections), provided that there are enough accurate detections available.



(a) (b)

Fig. 5. Beacon detection errors on the image. (a) Radial error component. (b) Tangential error component and angle detection error.

IV. Error Analysis

A. Detection Error Sources

The main detection error sources of the Lookup system include imaging errors, resulting in inaccurate detection of the beacon on the image, inaccurate inclination correction, and compass error.

The various error sources result in error of the detected beacon position [detected position error (DPE)], denoted by Δx_C and Δy_C, and also error in the detected beacon viewing angle [detected angle error (DAE)], denoted by $\Delta\beta$. The DPE may significantly differ in radial and tangential direction, thus in relevant cases we will refer to separate radial DPE (DPER) and tangential DPE (DPET), denoted by Δr and Δt, respectively.

1) Image Detection Error: First, the radial effect of incorrect detection of the beacon on the fisheye image is studied. Since the camera is circularly symmetric, the effect will be studied using distances from the center of the camera. In *CF*, this distance is p pixels, while in *CC* the corresponding distance of the beacon from the camera is $r_C = (x_C^2 + y_C^2)^{1/2}$. Let us define the mapping between the pixel coordinate p and the elevation $\vartheta = \text{atan2}(z_w, r_C)$ of the beacon as $\vartheta = f(p)$. (The mapping f is measured during the calibration process.) Since $r_C(p, z_w) = z_w \cot f(p)$, the DPER, caused by 1 pixel of radial detection error is the following:

$$\Delta r_C^{\text{im}}(p, z_w) = z_w(\cot f(p+1) - \cot f(p)). \tag{29}$$

One pixel of tangential detection error corresponds to the following DPET:

$$\Delta t_C^{\text{im}}(p, z_w) = \frac{2\pi r_C(p, z_w)}{2\pi p} = \frac{r_C(p, z_w)}{p} = \frac{r_C}{f^{-1}(\text{arccot}\frac{r_C}{z_w})}. \tag{30}$$

The DAE resulted by 1 pixel of tangential error can be calculated as follows:

$$\Delta\beta^{\text{im}} = \text{arccot}\frac{p}{1} = \text{arccot} f^{-1}\left(\text{arccot}\frac{r_C}{z_w}\right). \tag{31}$$

If the image detection contains N_p pixels of error, for small N_p good estimates of the resulting DPER, DPET, and DAE are $N_p\Delta r_C$, $N_p\Delta t_C$, and $N_p\Delta\beta$, respectively.

In Fig. 5(a), Δr_C is shown as a function of r_C, for different values of z_w. It is clearly visible that higher beacon positions result smaller DPER. It is also visible that for large beacon distances the DPER can be significant. The situation is much better for DPET, as shown in Fig. 5(b): for reasonable distances, the DPET is in the range of a few centimeters. Notice that in case of DPET lower beacon positions result in smaller error, but the dependence is not significant. The DAE, also shown in Fig. 5(b), becomes small when the beacon is not very close to the camera. Notice that although the angle error is higher for small distances, in such cases the detected image is large and thus subpixel detection accuracy is achievable.

The camera used in our tests was an inexpensive web camera with a simple plastic lens. The calibration process provided good fit with subpixel accuracy at the center of the camera, but toward the edge the average error was significantly larger (3–4 pixels).

2) Inclination Error: To evaluate the effect of imperfect inclination compensation, we will assume that the camera is tilted by angle $\Delta\vartheta$, and it is not compensated for by the inclination correction described in Section III-C2.

The fisheye coordinates of the detected beacon are $P' = (x_F, y_F, z_F) = (x_F, r_x\cos\vartheta, r_x\sin\vartheta)$, where $r_x = (y_F^2 + z_F^2)^{1/2} = (1 - x_F^2)^{1/2}$ and $\vartheta = \text{atan2}(z_F, y_F)$ are the distance and elevation of P' from axis X^C (see Fig. 3). The corresponding coordinates in CC, according to (2) are

$$x_C = sx_F = \frac{x_F z_w}{r_x \sin\vartheta} \tag{32}$$

$$y_C = sy_F = z_w \,\text{ctg}\vartheta. \tag{33}$$

Since the camera can be considered circularly symmetric, without loss of generality we can assume that it is tilted around axis X^C, and the detected elevation ϑ is changed by $\Delta\vartheta$. Using

$$\frac{dx_C}{d\vartheta} = \frac{-x_F z_w\cos\vartheta}{r_x\sin^2\vartheta} = -x_C\,\text{ctg}\vartheta = \frac{-x_C y_C}{z_w}$$
$$\frac{dy_C}{d\vartheta} = \frac{-z_w}{\sin^2\vartheta} = \frac{-z_w}{\sin^2(\text{atan2}(z_w, y_C))} \tag{34}$$

the DPE for small $\Delta\vartheta$ can be estimated as follows:

$$\left(\Delta x_C^{\text{tilt}}, \Delta y_C^{\text{tilt}}\right) = \left(\frac{-x_C y_c}{z_w}\Delta\vartheta, \frac{-z_w}{\sin^2(\text{atan2}(z_w, y_C))}\Delta\vartheta\right). \tag{35}$$

Notice that in general in (34) y_C means the distance of the beacon from the tilting axis, and x_C means the distance between the camera and the beacon's projection to the tilting axis. Similarly, in (35) Δx_C^{tilt} and Δy_C^{tilt} are the detection errors in parallel and perpendicular to the tilting axis, respectively.

Since the detection angle of a beacon is

$$\beta = \arctan\frac{y_F}{x_F} = \arctan\frac{r_x\cos\vartheta}{x_F} \tag{36}$$

its derivative with respect to the tilting axis is the following:

$$\frac{d\beta}{d\vartheta} = \frac{-x_F r_x\sin\vartheta}{x_F^2 + r^2\cos^2\vartheta} = \frac{-x_F z_F}{x_F^2 + y_F^2}. \tag{37}$$

Thus, the DAE for small tilting angle $\Delta\vartheta$ can be estimated in the following way:

$$\Delta\beta^{\text{tilt}} = -\frac{x_F z_F}{x_F^2 + y_F^2}\Delta\vartheta. \tag{38}$$

TABLE II
COMPASS INDOOR DETECTION ERRORS $\Delta\phi_C$

Altitude	10 cm	40 cm	150 cm
Mean	8°	6°	6°
Max	27°	12°	12°

TABLE III
ERROR SOURCES AND THEIR MAXIMUM VALUES USED IN THE MONTE CARLO SIMULATION

Source	Image detection	Inclination	Compass
Max	$N_p = 3$ pixels	$\Delta\vartheta$ =2°	$\Delta\phi_C = 12°$

The residual tilting error $\Delta\vartheta$ of the system, after the inclination correction, was measured. In the measurements the device was tilted by angles from 1° to 10° and the remaining tilting error was measured. Without and with sensor calibration the maximum tilting error was 1.8° and 0.7°, respectively.

3) Compass Error: The angle error $\Delta\phi_C$ of the compass causes DPET in the following way:

$$\Delta t^{\text{comp}} = \sqrt{x_C^2 + y_C^2}\tan\Delta\phi_C. \tag{39}$$

Thus for small $\Delta\phi_C$, DPET is proportional with $\Delta\phi_C$ and the distance from the beacon. According to measurement results, $\Delta\phi_C$ depends largely on the surrounding environment. In an office room of size 6 m × 11 m, 25 measurement points were used in a grid. At each point, eight measurements were performed, setting the compass to directions corresponding to the principal points (N, NE, E, etc.). The measurements were repeated at three different altitudes. The mean and maximum errors are shown in Table II. Close to the floor (10 cm) the error was significantly higher with large errors at certain points (probably due to the high metallic content of the reinforced concrete floor), while at the two elevated positions the error was almost the same. The error also showed large variance at various positions, e.g., it was significantly larger at the proximity of a metallic whiteboard in the room.

B. Localization Error

The localization algorithm utilizes detections with possible inaccuracies, resulting in inaccurate location estimates. In addition to the detection errors, the localization error also depends on the beacon deployment and the actual camera position. Since the dependence on the relative beacon/target location is complex, including effects of geometric dilution of precision (GDOP) [25], especially when several beacons are utilized, the localization error is estimated using Monte Carlo simulations. Notice that the simulation also covers error resulting from the finite resolution of the search grid. The possible detection errors are modeled as random variables with uniform distribution between zero and a maximum value. The error sources utilized in the tests, along with their ranges are shown in Table III. Notice that the error ranges were selected to reflect the measured typical error ranges. The four test

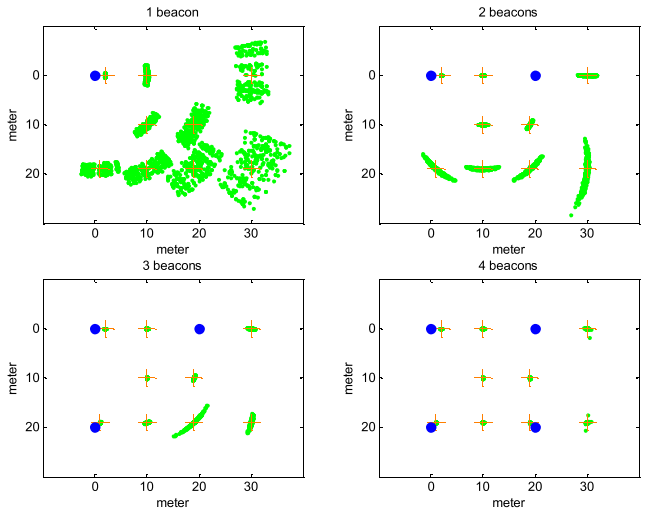


Fig. 6. Monte Carlo simulation results for 1, 2, 3, and 4 beacons.

setups, using 1, 2, 3, and 4 beacons, are shown in Fig. 6. The beacons, denoted by blue dots, were placed at positions (0 m, 0 m, 3.5 m), (0 m, 20 m, 4.5 m), (20 m, 0 m, 3 m), and (20 m, 20 m, 2.5 m), modeling a realistic large scale environment, as shown in Fig. 6. Nine test positions were selected both inside and outside of the sensor area, shown by red crosses. During the simulation, for each test positions 300 simulations were performed, with random image detection, inclination, and compass errors. The heading of the inclination error was also chosen randomly.

The test results are shown in Fig. 6. Each localization result is represented by a green dot, forming an error cloud. In the 1-beacon case the localization error is increasing with the distance from the beacon. The radial component is due to DPER, while the tangential component is dominated by the compass error.

The 2-beacon localization has moderate error when the target is between the two beacons, the error being smaller when the target is close to one of the beacons. As the distance from the beacon axes increases the localization error also increases. As shown in Fig. 6, the error cloud is spread along the corresponding viewing circles.

In the 3-beacon case the error is low in the area surrounded by the three sensors, but is significant around position (20, 20 m). This large error is caused is by GDOP: the viewing circles are almost parallel here and their intersections vary greatly even in the presence of small detection errors.

When four beacons are present the GDOP is much less in the area surrounded by the sensors, providing a uniformly small localization error.

V. EVALUATION

A. Sensors

The equipment used in the performance evaluation contains the following sensors: The main sensor of the localization system is an inexpensive USB web camera with a 1/2.7" OV2710 sensor. The sensor can provide a video stream with a resolution of 1920 × 1080 pixels at 30 fps. The camera is equipped with a fixed plastic fisheye lens, providing a

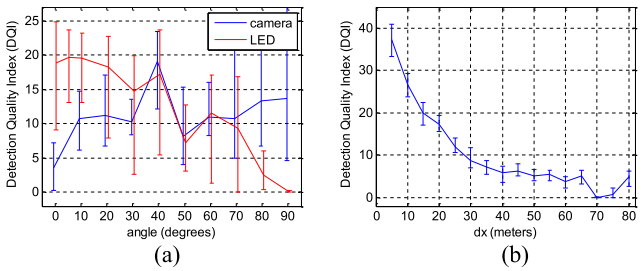


Fig. 7. Beacon detection performance. (a) DQI versus LED and camera angles. (b) DQI versus distance.

180° viewing angle. The tilt of the equipment is detected by a Bosch BMA180 three-axis accelerometer, commonly used in smartphones. The resolution of the sensor is 12 bit, providing a sensitivity of 2.5 mm/s^2. The compass contains a Freescale MAG3110 three-axis magnetometer, with full-scale range of $\pm 1000 \mu T$ and sensitivity of $0.1 \mu T$.

B. Beacon Detection

The performance of the beacon detection subsystem was tested. Let us denote the LED angle (the angle between the LED-camera axis and the normal vector of LED's plane) by ψ_{LED} and the camera angle (the angle between the LED-camera axis and the focal axis of the camera) by ψ_{CAM}. In the first test the performance versus the LED angle was examined. The distance d between the camera and the LED was 20 m, ψ_{CAM} was 45° and ψ_{LED} was varied between 0° and 90°, at each angle 50 measurements were taken. The mean of the DQI (5) is shown in Fig. 7(a), along with the minimum and maximum values. According to measurement results, the performance somewhat degrades above 45°, and severely degrades above 70°. This is inline with the 120° viewing angle (corresponding to $\psi_{\text{LED}} = 60°$) of the device.

In the second test, the camera angle was varied (with $\psi_{\text{LED}} = 45°$ and $d = 20$ m), and the results are shown in Fig. 7(a). The detection properties do not significantly depend on ψ_{CAM}. Interestingly, at the edge of the image ($\psi_{\text{CAM}} = 90°$) the detection is even better than at the center, partly probably due to the poor quality of the plastic camera lens: images at the edge are smeared, providing more pixels for detection.

The third test examines the performance as a function of distance between the camera and the beacon. The test setup reflects a realistic operational scenario: the LED was mounted at position (0 m, 0 m, 3 m), and the camera was moved along the x-axis from 0 to 80 m; thus, the camera coordinates were $(dx, 0, 0)$. The camera was looking upward, while the LED was pointing toward point (3 m, 0 m, 0 m). With this setup ψ_{LED} was changing between $-45°$ and $+45°$, while ψ_{CAM} was changing from 0° to 90°, as dx increased. According to the results, shown in Fig. 7(b), reliable detection can be performed until 50–60 m.

C. Indoor Localization

The performance of the system was tested in a gymnasium with size approx. 30 m × 17 m. Four beacons were installed at

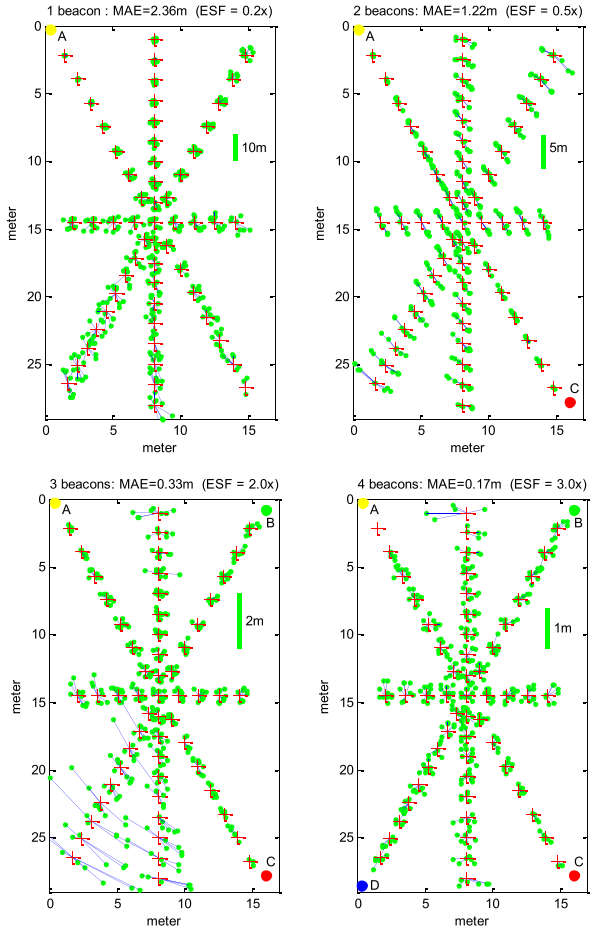


Fig. 8. Results of the localization test for 1-,2-,3-, and 4-beacon cases. The error vectors were scaled by ESF for better visibility. Green bars also show the error scale.

the following positions: $P_A = (0.26 \text{ m}, 0.4 \text{ m}, 3.05 \text{ m})$, $P_B = (16.0 \text{ m}, 0.8 \text{ m}, 4.3 \text{ m})$, $P_C = (16.0 \text{ m}, 27.76 \text{ m}, 4.3 \text{ m})$, and $P_D = (0.3 \text{ m}, 28.5 \text{ m}, 3.1 \text{ m})$, as shown in Fig. 8. During the tests 49 test positions were marked on the floor (shown by red crosses in Fig. 8) and each position was measured by a laser distance meter. In each test position ten measurements were conducted, each with random camera orientation. The measurements, using all beacons, were stored, and in the offline evaluation phase subsets (A), (A,B), (A,B,C), and (A,B,C,D) were used to test the 1-, 2-, 3-, and 4-beacon cases, respectively. The results are shown in Fig. 8 as follows. The error vector for each estimated location was calculated and, for better visibility, was multiplied by an error scaling factor (ESF), then was plotted in Fig. 8 with a green dot, connected to the true location with a blue line. The value of ESF and the mean absolute error (MAE) are also shown in Fig. 8.

The MAE was 2.36 m in the 1-beacon case. As shown in Fig. 8, the error is moderate when the sensor is close to the beacon, but both the radial error component (due to incorrect distance estimation) and the tangential error component (mainly due to incorrect orientation measurement) are increasing as the distance from the beacon increases, with maximum error of 9.7 m. When two beacons were available, the MAE was 1.22 m. Fig. 8 shows, the error is large when the sensor is far from both sensors, as was predicted in Section IV-B. In the 3-beacon case the MAE decreased to

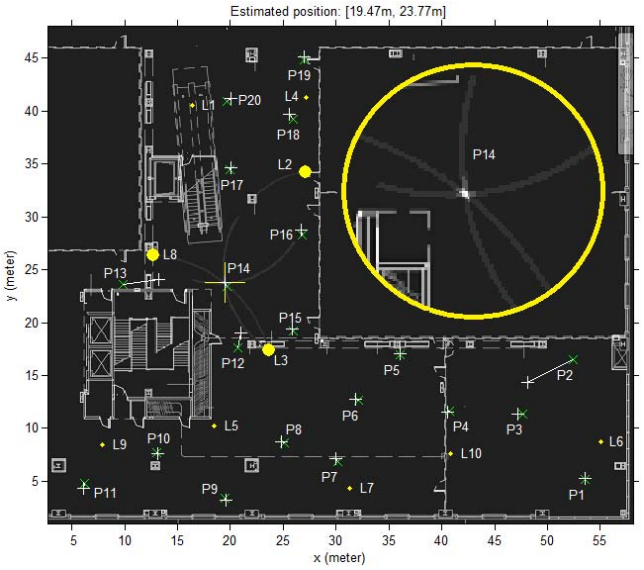


Fig. 9. Lookup evaluation display. L1–L10 (yellow dots) are the deployed beacons, P1–P20 (green x's) are the reference points. Detected positions are shown with white crosses, connected to the true location. The display shows the detection of P14, where three beacons were in line of sight (shown with larger dots). The enlarged insert shows the consensus function around the true location, lighter colors meaning higher values.

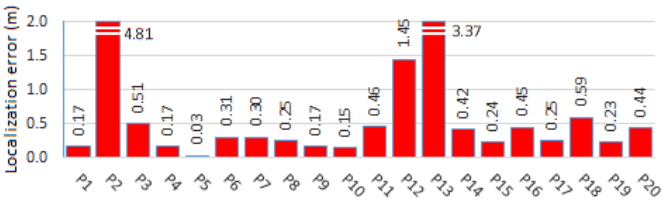


Fig. 10. Lookup localization errors for reference points P1–P20.

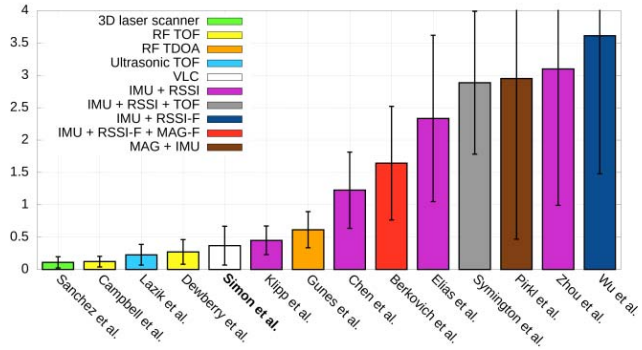


Fig. 11. Ninetieth percentile error and variance of the localization systems used in the comparison.

0.33 m, but here systematically the lower-left region has larger error, due to the GDOP: all three circular arcs are close to parallel in this region, resulting large localization errors even in case of small detection errors. Four beacons sufficiently covered the area resulting MAE = 0.17 m.

VI. Comparison

The performance of the system is compared to that of several other systems, utilizing various up-to-date technologies. All the systems in the comparison participated in the Microsoft Indoor Localization Contest in 2005, thus their performances can realistically be compared [26]. The contest was organized in a large conference building containing several walls, pillars, staircases, see the map in Fig. 9. The systems were evaluated using 20 reference points (shown by P1–P20 in Fig. 9), the location of which was not known in advance to the participants. At each reference point one measurement was performed by each system, and the location estimates were recorded. After the contest the organizers published the reference positions and the measurement results of each participant at [26], which will be used here for comparison.

In the contest the deployment of maximum ten beacons was allowed. The beacon positions of Lookup are represented by L1–L10 in Fig. 9. The estimated positions are shown by white crosses, connected to the true position. The localization errors are shown for each test position in Fig. 10. There are three large outliers (P2, P12, P13), caused by both software errors and installation errors (mismatched beacon IDs). The rest of the estimates are around or below 0.5 m. Since many systems participating in the contest encountered the same problems, in the comparision the 90th percentile accuracy will be used.

For comparison 14 representative systems were selected from the contest, as shown in Fig. 11, where the mean localization error and the variance for each system is shown, using the best 90% of the measurements. The short descriptions of the systems can be found in [26], identified by the the first author's name, as shown in Fig. 11.

Sanchez *et al.* [6] used 3-D laser scanners to build the real-time map of the surrounding area; this map was fitted to a reference map, with average error of 11 cm. Systems based on radio propagation time measurement also provided excellent results: Campbell *et al.* and Huseth *et al.* [16] used UWB radios for RF TOF ranging, producing location estimates with error of 12 and 27 cm, respectively. Instead of TOF, Gunes *et al*. measured TDOA of radio signals, allowing localization with average error of 61 cm. The system of Lazik *et al.* [5] utilized ultrasonic TOF ranging, with 23 cm of average error. The proposed Lookup system (Simon *et al.*) had average error of 37 cm.

Several systems utilized dead reckoning, using inertial measurement units (IMUs), and increased the accuracy by fusion of various received signal strength indicator (RSSI)-based methods, with average error ranging from 0.5 to 3.5 m. Among these systems Klipp *et al.* produced the highest accuracy (with 45 cm average error), using IMU placed on the shoe; here the RSSI measurements were used only for establishing the initial position. The systems of Chen *et al.* [11] and Elias *et al.* utilized RSSI based distance measurements to correct the dead reckoning results. Zou *et al.* utilized upgraded WiFi access points to receive RSSI of mobile devices; the location estimate was improved by IMU data. Symington *et al.* fused dead reckoning estimates with RSSI and RF TOF measurements. The system of Wu *et al.* utilized RSSI fingerprinting, while Berkovich *et al.* used both RSSI and magnetic fingerprinting, to improve the accuracy of dead reckoning. In system of Pirkl and Lukowicz [21] oscillating magnetic field was used for localization, improved by IMU data, providing 2.9 m of average error.

The best result was provided by a system utilizing 3-D laser scanner, which requires no special infrastructure but an expensive localization device is required (the price of the sensor is several thousand dollars). Methods based on radio or ultrasonic TOF ranging followed closely the best result. These methods require the installation of special infrastructure (with price ∼$100 per beacon), and the tracked device is also inexpensive (∼$100). The proposed solution uses an inexpensive sensor unit (∼$50) and requires some modification of the existing lighting infrastructure (or stand-alone beacons may be used with cost around $50/beacon), and also had error in low decimeter range, which may potentially be improved by utilizing cameras of better quality. Dead reckoning-based solutions, supported by RSSI, potentially have error in the meter range, although with special sensors (e.g., applied to the shoes) the error can be as low as 0.5 m. The advantage of these solutions is that they potentially require no extra infrastructure (the existing WiFi network may be utilized), and the inertial sensor unit's cost is as low as ∼$50 (or for lower-end solutions sensors integrated in smart-phones can be utilized); but some of these solutions require initial setup. The solution based on oscillating magnetic field had error in the meter range, with estimated cost of ∼$50 for both the sensor units and the beacons.

VII. Conclusion

A novel indoor localization system was proposed, which utilizes LED lights as beacons and VLC to identify beacons. The sensor contains a camera with fisheye lens, and auxiliary sensors (accelerometer and magnetometer). The proposed system is potentially able to utilize existing lighting infrastructure with minimal modification: the LED beacons transmit high-frequency beacon identifiers, which are detected from high distances using the camera, by undersampling the transmitted code. The system utilizes different algorithms for one, two, and higher number of visible beacons. For one and two beacons geometric solutions are utilized, while the algorithm for higher number of beacons is based on a consensus-based approach, tolerating well bad detections and providing high accuracy. The performance of the system was analyzed by simulations and also real measurements were performed. The performance of the proposed system was compared to today's emerging indoor localization technologies.

The proposed method is able to locate a static camera with accuracy of a few decimeters in large areas, depending on the number and relative locations of the beacons. With higher number of beacons the accuracy increases, but the system is able to provide reasonable location estimates with one or two beacons only. The accuracy of the system is enough for person and asset localization applications, but for more demanding applications (e.g., robot control) the accuracy should be improved with careful beacon placement. Current work includes the extension of the detection algorithm to handle moving sensors, broadening the possible application fields. Also more sophisticated beacons are under development, which eliminate the possible stroboscopic effect of the modulated LEDs.

References

[1] K. Chintalapudi, A. P. Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2010, pp. 173–184.

[2] H. Lim, L. C. Kung, J. C. Hou, and H. Luo, "Zero-configuration, robust indoor localization: Theory and experimentation," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006, pp. 1–12.

[3] L. Zwirello, T. Schipper, M. Jalilvand, and T. Zwick, "Realization limits of impulse-based localization system for large-scale indoor applications," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 39–51, Jan. 2015.

[4] J. C. Prieto *et al.*, "Performance evaluation of 3D-LOCUS advanced acoustic LPS," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2385–2395, Aug. 2009.

[5] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Ultrasonic time synchronization and ranging on smartphones," in *Proc. IEEE Real-Time Embedded Technol. Appl. Symp. (RTAS)*, May 2015, pp. 108–118.

[6] C. Sánchez, P. Taddei, S. Ceriani, E. Wolfart, and V. Sequeira, ""Localization and tracking in known large environments using portable real-time 3D sensors," *Comput. Vis. Image Understand.*, vol. 149, pp. 197–208, Aug. 2016.

[7] A. Colombo, D. Fontanelli, D. Macii, and L. Palopoli, "Flexible indoor localization and tracking based on a wearable platform and sensor data fusion," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 4, pp. 864–876, Apr. 2014.

[8] Y. S. Kuo, P. Pannuto, K. J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, New York, NY, USA, 2014, pp. 447–458.

[9] L. Li, P. Hu, C. Peng, G. Shen, and F. Zhao, "Epsilon: A visible light based positioning system," in *Proc. 11th USENIX Conf. Netw. Syst. Design Implement. (NSDI)*, Seattle, WA, USA, 2014, pp. 331–343.

[10] H. Wang, H. Yu, and L. Kong, "Ceiling light landmarks based localization and motion control for a mobile robot," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, London, U.K., Apr. 2007, pp. 285–290.

[11] Z. Chen, Q. Zhu, H. Jiang, and Y. C. Soh, "Indoor localization using smartphone sensors and iBeacons," in *Proc. IEEE 10th Conf. Ind. Electron. Appl. (ICIEA)*, Auckland, NZ, USA, Jun. 2015, pp. 1723–1728.

[12] E. Kovács, "Rotation about an arbitrary axis and reflection through an arbitrary plane," *Ann. Math. et Inf.*, vol. 40, pp. 175–186, 2012. [Online]. Available: http://ami.ektf.hu/uploads/papers/finalpdf/AMI_40_from175to186.pdf

[13] G. F. Carballo and P. W. Fieguth, "Probabilistic cost functions for network flow phase unwrapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2192–2201, Sep. 2000.

[14] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Beijing, China, Oct. 2006, pp. 5695–5701.

[15] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A flexible technique for accurate omnidirectional camera calibration and structure from motion," in *Proc. IEEE Int. Conf. Comput. Vis. Syst. (ICVS)*, New York, NY, USA, Jan. 2006, p. 45.

[16] S. Huseth, B. S. Dewberry, and R. McCrosky, "Pulsed-RF ultrawideband ranging for the GLANSER GPS-denied emergency responder navigation system," in *Proc. Int. Tech. Meeting Inst. Navigat.*, Jan. 2011, pp. 389–396.

[17] O. Oshiga, S. Severi, and G. T. F. de Abreu, "Superresolution multipoint ranging with optimized sampling via orthogonally designed Golomb rulers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 267–282, Jan. 2016.

[18] C. Medina, J. Segura, and Á. De la Torre, "Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy," *Sensors*, vol. 13, no. 3, pp. 3501–3526, Mar. 2013.

[19] M. Mirshekari, S. Pan, A. Bannis, Y. P. M. Lam, P. Zhang, and H. Y. Noh, "Step-level person localization through sparse sensing of structural vibration," in *Proc. 14th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2015, pp. 376–377.

[20] P. Cotera, M. Velazquez, D. Cruz, L. Medina, and M. Bandala, "Indoor robot positioning using an enhanced trilateration algorithm," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 3, p. 110, Jun. 2016, doi: 10.5772/63246.

[21] G. Pirkl and P. Lukowicz, "Resonant magnetic coupling indoor localization system," in *Proc. ACM Conf. Pervasive Ubiquitous Comput. Adjunct Publication (UbiComp Adjunct)*, 2013, pp. 59–62.

[22] M. Frassl, M. Angermann, M. Lichtenstern, P. Robertson, B. J. Julian, and M. Doniec, "Magnetic maps of indoor environments for precise localization of legged and non-legged locomotion," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Tokyo, Japan, Nov. 2013, pp. 913–920.

[23] P. Hu, L. Li, C. Peng, G. Shen, and F. Zhao, "Pharos: Enable physical analytics through visible light based indoor localization," in *Proc. 12th ACM Workshop Hot Topics Netw. (HotNets)*, 2013, Art. no. 5.
[24] A. Ledeczi *et al.*, "Countersniper system for urban warfare," *ACM Trans. Sensor Netw.*, vol. 1, no. 2, pp. 153–177, Nov. 2005.
[25] R. B. Langley, "Dilution of precision," *GPS World*, vol. 10, no. 5, pp. 52–59, May 1999.
[26] Microsoft Indoor Localization Competition—IPSN 2015. [Online]. Available: https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2015/

Gyula Simon (M'04) received the M.Sc. and Ph.D. degrees in electrical engineering from the Budapest University of Technology, Budapest, Hungary, in 1991 and 1998, respectively.

He is currently an Associate Professor with the University of Pannonia, Veszprém, Hungary. His current research interests include adaptive signal processing, localization, and sensor networks.

Gergely Zachár received the M.Sc. degree in information technology from the University of Pannonia, Veszprém, Hungary, in 2014, where he is currently pursuing the Ph.D. degree in information science and technology.

His current research interests include sensor networking and embedded systems.

Gergely Vakulya received the M.Sc. and Ph.D. degrees in information technology from the University of Pannonia, Veszprém, Hungary, in 2008 and 2015, respectively.

He is currently an Assistant Professor with the University of Pannonia. His current research interests include sensor networks and embedded systems.