Augmenting Indoor Inertial Tracking with Polarized Light

Zhao Tian[†1], Yu-Lin Wei[†2], Wei-Nin Chang[2], Xi Xiong[1],
Changxi Zheng[3], Hsin-Mu Tsai[2], Kate Ching-Ju Lin[4], and Xia Zhou[1]

[1]Dartmouth College, [2]National Taiwan University, [3]Columbia University, [4]National Chiao Tung University

[†] Co-primary author

tianzhao@cs.dartmouth.edu,r03922027@csie.ntu.edu.tw,r05922097@ntu.edu.tw,xi@cs.dartmouth.edu,cxz@cs.columbia.edu,hsinmu@csie.ntu.edu.tw,katelin@cs.nctu.edu.tw,xia.zhou@dartmouth.edu

ABSTRACT

Inertial measurement unit (IMU) has long suffered from the problem of integration drift, where sensor noises accumulate quickly and cause fast-growing tracking errors. Existing methods for calibrating IMU tracking either require human in the loop, or need energy-consuming cameras, or suffer from coarse tracking granularity. We propose to augment indoor inertial tracking by reusing existing indoor luminaries to project a static light polarization pattern in the space. This pattern is imperceptible to human eyes and yet through a polarizer, it becomes detectable by a color sensor, and thus can serve as fine-grained optical landmarks that constrain and correct IMU's integration drift and boost tracking accuracy. Exploiting the birefringence optical property of transparent tapes – a low-cost and easily-accessible material – we realize the polarization pattern by simply adding to existing light cover a thin polarizer film with transparent tape stripes glued atop. When fusing with IMU sensor signals, the light pattern enables robust, accurate and low-power motion tracking. Meanwhile, our approach entails low deployment overhead by reusing existing lighting infrastructure without needing an active modulation unit. We build a prototype of our light cover and the sensing unit using off-the-shelf components. Experiments show 4.3 cm median error for 2D tracking and 10 cm for 3D tracking, as well as its robustness in diverse settings.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **Computer systems organization → Sensors and actuators**;

KEYWORDS

Inertial tracking, light polarization, particle filter

ACM Reference Format:
Zhao Tian[†1], Yu-Lin Wei[†2], Wei-Nin Chang[2], Xi Xiong[1], Changxi Zheng[3], Hsin-Mu Tsai[2], Kate Ching-Ju Lin[4], and Xia Zhou[1]. 2018. Augmenting Indoor Inertial Tracking with Polarized Light. In *MobiSys '18: The 16th Annual International Conference on Mobile Systems, Applications,* and Services, June 10–15, 2018, Munich, Germany. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3210240.3210340

1 INTRODUCTION

Inertial measurement unit (IMU) has been widely used to track object motion. Comprising accelerometer, gyroscope, and compass, IMU measures object's acceleration, angular velocity, and orientation. As a low-cost, small, and low-power unit, IMU is common in mobile devices (e.g., smartphones, smart watches, drones [5, 26]) and home appliances [2].

Despite its popularity, IMU has a long-standing problem of integration drift [46]. As IMU sensors measure acceleration and angular velocity, one must *integrate* these sensor signals to compute object's movement trajectory. The integration contains sensor errors from bias (nonzero output when the sensor is at rest), electrical and thermal-mechanical noise, as well as other random noises due to calibration or temperature [46, 64]. These errors accumulate quickly, resulting into a fast growth of the location error [54, 60, 75].

Active research has examined approaches to addressing IMU's drift problem. Traditional approaches focus on individual sensor calibration, either requiring laboratory equipments [45] or demanding repeated human efforts [16, 24]. Recent studies explore fusion of IMU with other sensors and leverage external landmarks with absolute locations. However, these methods either require human in the loop (e.g., footsteps used in the pedestrian dead reckoning) [30, 40, 64], or rely on cameras [31, 36] that are typically power-hungry, or consider only outdoor scenarios with GPS signals and outdoor landmarks available [19], or leverage coarse-grained landmarks (e.g., landmarks based on Wi-Fi signal strength or building structures) and thus are unable to achieve centimeter-level accuracy reliably [34, 54, 60, 67].

In this paper, we address above problems using fine-grained landmarks provided by ubiquitous lights. Specifically, we consider reusing indoor luminaries (e.g., LED or fluorescent lights) to cast static, imperceptible *light polarization patterns*. We create such patterns using *a low-cost polarizer and a birefringent [8] film (e.g., everyday transparent tapes)* attached to existing light cover/diffuser. These patterns are imperceptible to naked eyes [32] and yet detectable by low-cost color sensors[1] viewing through a second polarizer. The patterns alone cannot determine object's precise location; however, they provide fine-grained (e.g., centimeter-level) landmarks to constrain the integration drift of the IMU collocated with color sensors. With a thin, low-cost layer atop existing light cover,

---

[1]A color sensor consists of a set of blue-filtered, yellow-filtered, and red-filtered photodiodes, which sense the intensity of incoming light on red (R), green (G), and blue (B) channels.
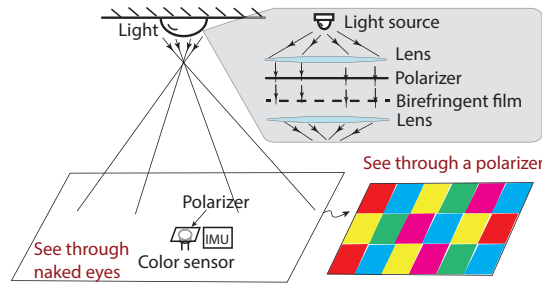
Figure 1: Creating an imperceptible grid pattern of light polarization using a polarizer and birefringent film (e.g., transparent tapes). The light polarization varies across cells, manifested in different colors observed by a color sensor through a polarizer as it moves across cells. The observations can constrain the drift of a collocated IMU.

our approach reuses existing lighting infrastructure without requiring any active modulation unit, thus entailing low deployment overhead and cost. Our approach will benefit indoor object tracking, including robotics applications (e.g., autonomous navigation, robotic collaboration), and HCI applications (e.g., tracking hand controllers or handset in VR), where cm-level accuracy is desired.

Figure 1 illustrates the system setup. Light is first collimated after a thin lens (e.g., a Fresnel lens [3]). It then passes a polarizer (e.g., a 0.18-mm thin film [7], <$1) that allows light rays of a certain polarization direction to pass. The polarized light then passes a birefringent film (e.g., transparent tape stripes arranged in a pattern) that alters the light polarization to create a spatial pattern (e.g., a grid pattern). Through another lens (e.g., Fresnel lens), the pattern is projected to the space. Color sensors covered with polarizer films will perceive various colors as it moves across cells and these cell-crossing events can augment IMU tracking. We use point light source for illustration and will discuss other luminary types in §8.

Fusing light and IMU sensors has three benefits: **(1) High accuracy**: With wavelengths in nanometers, visible light can generate landmarks with fine granularity. These optical landmarks best complement IMU to address its drift error and achieve high tracking accuracy; **(2) Reliability**: While light provides precision, it can be easily blocked by other objects, rendering light signals unavailable. IMU sensors, on the other hand, steadily provide sensor readings that can compensate for losses of light signals upon occasional blockage, rendering the system robust against occasional blockage in tracking; **(3) Low power and portability**: Both IMU and color sensors are cheap, small, and low-power. They can be easily embedded in everyday objects for them to self-track their locations.

To enable effective fusion, we face three main challenges. The first two challenges concern the design and detection of the polarization pattern. Two practical factors affect the sensing and detection of the pattern. First, the orientation of the object (and thus sensors) also affects the color perceived by color sensors. Changes in object orientation alter the direction of its polarizer and, thus, the perceived color at the sensor, making it difficult to identify color changes from crossing cells. Second, a small portion of ambient light can still be sensed by the color sensor and affect its perceived color. It is particularly problematic when ambient light is colored, spatially nonuniform, and time-varying. The last challenge is on the fusion of light pattern and IMU signals. The detected pattern

(i.e., cell crossings) does not indicate the specific coordinates for calibration and can be occasionally unavailable due to blockage. Also, as a 2D pattern, it introduces ambiguity in determining object's vertical position. Tracking object's 3D location is nontrivial.

We address these challenges via following design elements. *First*, using a grid pattern as the starting point, we carefully design the pattern to ensure sharp color contrast between adjacent cells under different orientations and facilitate pattern detection at color sensors. We also arrange the pattern among consecutive cells to mitigate the ambiguity of mapping an observed color sequence to a movement trajectory. *Second*, to ensure robust detection of the pattern, we propose a dual-sensor design that examines the *differential color values* of two collocated color sensors to remove ambient light interference. We also propose an empirical model to estimate the impact of object orientation on the resulting color values sensed by sensors. It allows us to correctly identify changes associated with crossing cells. *Finally*, we model the data fusion problem using a Bayesian inference framework. We apply the sequential Monte Carlo (i.e., particle filter) algorithm [14] to update the probability distribution of the object location based on observed differential color values and IMU data. We also extract features from differential color values to extend the fusion algorithm to 3D tracking.

We build a prototype of our proposed light cover layer, as well as the sensing component with collocated color sensors and IMU, all using off-the-shelf hardware. We experiment our system in an indoor lab and compare our results to Vicon [9]. We also conduct a user study with 33 participants to examine the impact of our projected pattern on illumination. Our key findings are as below:

- Light polarization pattern effectively corrects IMU drift errors, achieving 4.3-cm median position error for 2D tracking and 10-cm for 3D tracking, where tracking error is bounded by the projected cell size.
- Fusing light and IMU signals, the system is robust against object's orientation change, diverse ambient light conditions, and occasional losses of light signals.
- The current data fusion (implemented in C) outputs a tracking result in 7.18 ms, supporting potential tracking rates up to 140 Hz.
- Our light cover layer does not affect the comfort and uniformity of illumination and participants do not perceive projected polarization patterns.

2 RATIONALE AND CHALLENGES

Most common IMUs today are produced by MEMS (micro-machined electromechanical systems) technology[2]. Advances in manufacturing silicon chips have made MEMS IMUs low-cost (e.g., $1.5–20), small, and lightweight [58, 63]. MEMS inertial sensors, however, suffer from larger sensor noises/errors. Our experiments with latest MEMs IMUs [4] reveal that while most errors are deterministic and can be calibrated in advance or in repeated recalibration, the stochastic error is still time-variant and movement-dependent, and thus difficult to be removed by any calibration process [59]. We need to seek external landmarks that can provide periodical recalibration to constrain IMU's drift errors.

---

[2]High-end IMUs (e.g., optical sensors) with high accuracy are expensive [17], bulky, and thus used in a limited range of applications.

To this end, we turn our attentions to light patterns given the ubiquity of lights and light's small wavelength enabling patterns in fine granularity. We next describe our rationale in designing light patterns to best augment IMU tracking, followed by the challenges to realize light-based augmentation.
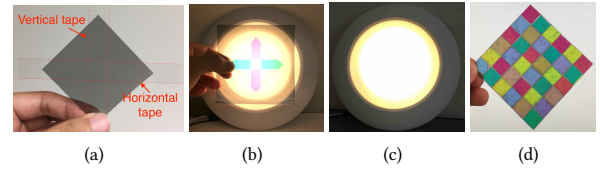
2.1 Rationale

We seek light patterns satisfying following goals: (1) The pattern should offer clear optical landmarks that can be detected by low-cost photodiodes reliably; (2) The pattern can be easily created reusing existing luminaries for minimal deployment overhead and cost; (3) The pattern should not affect the luminary's ability of illumination; ideally it should be imperceptible to naked eyes.

The above goals lead us to consider patterns in *light polarization*. Polarization is the oscillating direction of electric fields as light waves travel. It is classified into *linear polarization*, where electric fields oscillate in a single direction, and *elliptical polarization*, where the fields rotate in a plane based on light phase. Most natural or artificial light sources emit unpolarized light containing an equal mixture of all polarization directions. Linearly or elliptically polarized light can be generated by a linear or elliptical polarizer[3], which allows light of certain polarization direction to pass.

Polarization patterns offer two benefits: (1) Unlike light intensity or color, polarization is inherently imperceptible to naked eyes [32]. Thus, creating polarization patterns does not affect light's illumination; (2) Since ambient light is typically unpolarized, polarization patterns are potentially more robust against ambient light interference, in comparison to patterns on light intensity. The remaining question is whether we can create light polarization patterns in a low-cost manner by reusing existing luminaries.

The answer is yes. Unlike prior works [21, 62, 68] that demand an LCD shutter to dynamically change light polarization, we consider a static, passive polarization pattern that can be generated without adding any active modulation unit. In particular, we exploit birefringence [8], an optical property of an optically anisotropic material whose refractive index depends on the polarization and propagation direction of incoming light. The difference of the refractive indexes in different directions changes the polarization of the incoming light depending on the light wavelength (i.e., color). As a result, as a polarized white light (i.e., mixture of light in different wavelengths) passes the material, light rays with different wavelengths (i.e., colors) are dispersed to different polarization directions. This dispersion of light polarization is imperceptible to naked eyes. Yet through a second polarizer, a color beam is visible, which has the polarization direction aligning with that of the second polarizer.

Interestingly, everyday transparent tapes are made of optically anisotropic material [57]. Figure 2(a) shows a thin polarizer film [7] with two transparent tape stripes glued horizontally and vertically. We attach the polarizer to an existing LED light cover, where the surface without the transparent tapes faces the light cover. Viewing the light cover through another polarizer, we see various colors in regions with transparent tape stripes (Figure 2(b)). The color differences indicate that *the thickness* and *orientation* of transparent tape layers both affect the amount of polarization dispersion. This

---

[3]Our current system uses linear polarizers to generate polarized light. Thus the term *polarizer* in the rest of the paper refers to linear polarizer.

(a) (b) (c) (d)

Figure 2: A polarizer [7] with two transparent tape stripes glued atop (a). Attaching this film to a light cover and viewing it through another polarizer, we see various colors at regions with different transparent tape configurations (b). The colors, however, are invisible without a polarizer (c). Viewed through a second polarizer, (d) shows a grid color pattern (invisible without the second polarizer) created by sophisticated arrangement of transparent tape layers.

color pattern, however, is invisible to naked eyes without the second polarizer (Figure 2(c)).

Our results validate that a static polarization pattern can be created using only a cheap polarizer with transparent tapes glued atop. By varying the thickness and orientation of the tape layer across different regions on the polarizer, we can create a spatial polarization pattern. In this work, we consider grid patterns as the starting point and will discuss more advanced patterns in §8. Figure 2(d) shows an example, where the color of each cell is achieved by a configuration of the tape layer thickness (i.e., the number of tape layers) and tape orientation. Attaching this film of polarizer and transparent tapes to a light cover, we can add a thin Fresnel lens to project the pattern to the 3D space. Then for an object equipped with a color sensor with a polarizer atop, the color sensor can sense color changes as the object moves across cells. Such color changes (i.e., crossings of cell boundaries) serve as optical landmarks to constrain and correct drift errors of a collocated IMU.

2.2 Practical Challenges

To realize the augmentation using light polarization patterns, we face the following practical challenges.

First, in addition to cell crossings, orientation changes of the tracked object also affect the color perceived by the color sensor on the object, thus interfering with the sensing of the polarization pattern. This is because as the orientation of the polarizer atop the color sensor changes, it allows light with a different polarization direction to pass, leading to light rays with different wavelength. Figure 3 shows an example pattern viewed through a second polarizer in three orientations ($45°$, $90°$, $135°$), where the same pattern exhibits different colors. Additionally, certain orientations (e.g., $45°$) lead to colors with lower saturation, rendering the color detection prone to errors. These above factors impose challenges in both the design and detection of polarization patterns: the design of the color/polarization pattern should ensure high-contrast colors that can be robustly detected under diverse orientations, while the detection of the pattern needs to filter out the color changes caused by orientation change.

The second challenge is to deal with interference from ambient light, which contains light rays in all polarization directions. Although polarization patterns are more robust against ambient light interference than patterns in light intensity, part of ambient light – the light rays with polarization directions aligning with that of the polarizer atop the color sensor – is still sensed by the color sensor and thus can interfere with sensing the polarization pattern
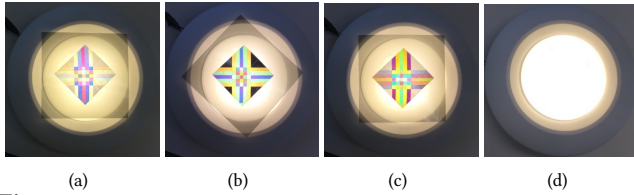
Figure 3: Different colors observed through a second polarizer with three orientations: 45°(a), 90°(b), 135°(c). The pattern is invisible without the second polarizer (d).

from our luminary, especially when ambient light is colored. Furthermore, ambient light can vary spatially and temporally. Hence fingerprinting ambient light condition entails expensive overhead: we need to exhaustively measure all locations and we cannot use an one-time measurement of ambient light in the beginning to filter its influence in later measurements.

Finally, efficient fusion of the detected light pattern and IMU signals is challenging, given that the optical landmarks only indicate crossings of a boundary at x or y direction, providing no information on the specific coordinates of the crossing locations. Thus, the data fusion algorithm needs to best exploit such information to correct IMU tracking errors. Also, since the polarization pattern is 2D, it leads to ambiguity in determining the object's vertical position. Extending the augmentation to enable 3D tracking remains difficult. Furthermore, light signals can be occasionally unavailable due to the blockage of other objects or holes in light coverage in the 3D space. The fusion algorithm needs to be robust against such losses of optical landmarks.

Next, we describe our three main system components to address these challenges.

3 DESIGNING POLARIZATION PATTERNS

The first component is designing the polarization pattern to ease its detection by a color sensor under diverse settings (e.g., orientation, ambient light). Again take the grid pattern (Figure 2(d)) as an example. Ideally, one can assign a unique color to each cell, so that a perceived color can narrow down object's location to a single cell for calibrating IMU. This method, however, requires a sufficient number of colors. Given that the color range is fixed, more colors implies smaller differences among colors, making color detection prone to sensor noises and ambient light interference. Thus, to ensure robust pattern detection, we reuse a few distinctive colors across cells to create a polarization/color pattern. The color reuse, however, creates ambiguity in mapping a perceived color change to the actual cell boundary that the object is crossing. To balance this tradeoff, we next describe how to select transparent tape configurations that lead to distinctive colors, and how to assign tape configurations to cells. Here a tape configuration is defined as a setting of the orientation and thickness of the tape layer for a cell.

3.1 Selecting Tape Configurations

We first select a set of transparent tape configurations whose resulting colors can be reliably differentiated by a color sensor in different orientations and ambient light conditions. We evaluate the distinctiveness of a color based on light intensity values perceived

by the color sensor on R, G, B channels. Instead of directly examining the differences of RGB values, which are dependent on the light intensity and thus affected by the distance to the luminary, we convert the RGB values to the HSV (Hue, Saturation, Value) color space [55]. Here hue is the attribute of a color and defined as the degree to which a stimulus can be discernible to red, green, blue and yellow; saturation describes the purity of the hue; and value represents the brightness or light intensity. We consider hue and saturation values to evaluate the differences of two perceived colors, so that color detection is independent of changes in overall light intensity caused by distance changes during object's movement.

To select tape configurations that are robust against sensor's orientation change, we exhaustively test various tape configurations by varying the number of tape layers and orientation. For each tape configuration, we compute the corresponding hue and saturation using measurements of a color sensor in three orientations (i.e., 0°, 45°, 90°). We then plot these data points in the hue-saturation (HS) plane, where hue is arranged in a radial slice and saturation is the distance to the origin. Data points of the same tape configuration form a cluster. We identify the most separated clusters based on their Euclidean distances. Thus, colors resulted from these configurations are well separated. We skip clusters close to the origin (i.e., with low saturation) to ensure robust color detection, as our experiments show that colors in low saturation (< 0.25) are vulnerable to sensor noises and lead to less accurate hue measurements.

3.2 Assigning Tape Configurations

With the selected set of tape configurations, the next step is to assign them to cells. An effective assignment should meet two goals: (1) ensuring robust detection of crossing cell boundaries and (2) reducing the ambiguity in mapping a sequence of observed colors to a movement trajectory to facilitate the calibration of IMU tracking. To achieve the first goal, we assign tape configurations resulting into colors with higher contrast to adjacent cells, so that the crossing of a cell boundary produces a sharp change in the perceived color, which can be more reliably detected even with sensor noises. To achieve the second goal, we judiciously arrange the configurations in various consecutive cells, so that a movement trajectory in a particular direction can lead to a unique sequence of observed colors. Figure 4(a)-(d) show example shapes of local regions with three consecutive cells. Figure 4(e) is an assignment that ensures unique color sequences for each of these three-cell shapes in Figure 4(a)-(d). These local region shapes and length can be adapted based on object's movement characteristics.

Specifically, we formulate our assignment problem as follows. We assume a grid pattern with $M \times N$ cells, L tape configurations $C = \{c_1, c_2, \ldots, c_L\}$ (recall that each configuration is specific arrangement of the tape layer thickness and orientation), function $d(c_1, c_2)$ evaluating the difference between configuration c_1 and c_2 (e.g., the Euclidean distance on the HS plane), and the set T of types of local regions (e.g., $T = \{t_1, t_2, t_3, t_4\}$ in Figure 4) where we aim to assign unique configuration sequences. Let A denote an assignment, where $A(i, j) \in C, i \in [1, M], j \in [1, N]$, and the multiset of all the configuration sequences for a type of local region $t \in T$ under assignment A is $S(A, t)$. The optimal assignment $A^\star$ maximizes the minimal difference of tape configurations in any two adjacent cells,

(a) t_1 (b) t_2 (c) t_3 (d) t_4 (e)
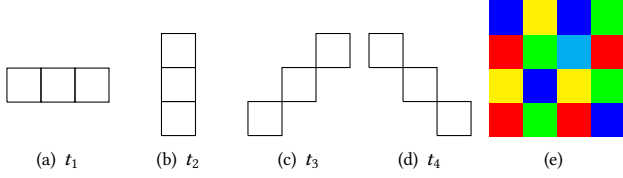
Figure 4: (a)-(d) are example shapes of local regions with three consecutive cells. (e) is an assignment where any three consecutive cells with one of the shapes in (a)-(d) lead to a unique sequence of tape configurations, where a tape configuration is represented as a color.

ALGORITHM 1: findAssignment

Input: M: # of rows, N: # of columns, $C = \{c_1, \ldots, c_L\}$: tape
         configurations, $pairs$: usable configuration pairs for adjacent cells,
         $T = \{t_1, \ldots, t_K\}$: local regions, r: completed # of rows
Output: A: feasible assignment or null
if $r < 0$ **then**
    | **return** null
else if $r = 0$ **then**
    | **for** $i \leftarrow 1$ **to** M **do**
    | | $map[i] \leftarrow \{\}$
    | | **for** $j \leftarrow 1$ **to** K **do**
    | | | $pattern[i][j] \leftarrow \{\}$ // assigned t_j from row 1 to i
    | | **end**
    | **end**
else if $r = M$ **then**
    | **return** map
$map[r + 1] \leftarrow next(map[r + 1], map[r], C, pairs)$
while $map[r + 1] \neq \{\}$ **and** $map[1 : r + 1]$ *has repeated patterns* **do**
    | $map[r + 1] \leftarrow next(map[r + 1], map[r], C, pairs)$
end
if $map[r + 1] = \{\}$ **then**
    | $findAssignment(M, N, C, pairs, T, r - 1)$
end
else
    | **for** $i \leftarrow 1$ **to** K **do**
    | | $pattern[r + 1][i] \leftarrow pattern[r][i] \cup$
    | | $\{$sequences added by $map[r + 1]$ for $t_i\}$
    | **end**
    | $findAssignment(M, N, C, pairs, T, r + 1)$
end

---

while avoiding repetition of configuration sequences among each local region in T. Thus, we model the problem as:

$$\text{Maximize:} \quad \min_{\substack{|i_1 - i_2| = 1, 1 \leq i_1, i_2 \leq M \\ \text{or } |j_1 - j_2| = 1, 1 \leq j_1, j_2 \leq N}} d(A(i_1, j_1), A(i_2, j_2)) \quad (1a)$$

$$\text{subj. to:} \quad \forall t \in T, S(A, t) \text{ has no repeated elements.} \quad (1b)$$

To solve this optimization problem, we first sort all configuration pairs in C based on their pairwise differences evaluated by $d(\cdot)$. We then iteratively search for a feasible assignment. In each iteration i, only the top-i most distinct configuration pairs are assigned to adjacent cells and we use a backtracking algorithm (Algorithm 1) to search through possible assignments and examine whether they satisfy the constraint in Eq. (1b). We keep increasing i until finding a feasible solution. Since configuration pairs are sorted in the descending order of their pairwise differences, the first output assignment is the optimal solution.

Given that this is a combinatorial optimization problem, our algorithm entails exponential complexity. However, since the assignment is computed only once offline, running time is less a concern for a relatively small number of cells. In our implementation with Python, it takes 0.1 second to find an assignment for 49 cells. For a larger number of cells, we will explore approximate algorithms to speed up computation. We leave it to future work.

We also note that feasible solutions may not exist when the local regions in T have a short length (i.e., small number of cells), because short configuration sequences more likely repeat given a fixed number of configurations. The problem can be mitigated by increasing the length of each local region where longer sequences have more combinations to differentiate different trajectories. The side effect is that it introduces a longer bootstrap latency to output the tracking locations. Thus, we should carefully configure the local region length to strike the best balance.

4 DETECTING POLARIZATION PATTERNS

With the polarization pattern projected in the space, the second component is to detect the pattern reliably. The detection boils down to detecting the light polarization state at current location, which manifests as a color observed by a color sensor in a given orientation. Here the color sensor views through a polarizer atop sensor's photodiode array area and monitors light intensity values on R, G, B channels. Key challenges come from dealing with ambient light interference and object's orientation changes, as they both affect the color perceived by the sensor (§2.2). We address these challenges via two design elements elaborated as below.

Dual-Sensor Design. To address ambient light interference, we propose a dual-sensor design, where two collocated color sensors, each with a separate polarizer atop the photodiode array area, simultaneously monitor light intensity values on three color channels. More importantly, two sensors' polarizers are arranged in different orientations: one in $0°$ and the other in $90°$ relative to the object's orientation. Since ambient light is typically unpolarized, containing equal mixture of light waves in different polarization directions, equal amount of ambient light – light waves with polarization direction that aligns with the orientation of a color sensor's polarizer – will be sensed by each color sensor. Thus, we can remove ambient light interference by taking the differential between two sensors' values on each color channel. Let R_1, G_1, B_1 and R_2, G_2, B_2 denote the intensity value observed by each sensor on each color channel, respectively. Then *differential color values* can be written as:

$$\Delta R = R_1 - R_2, \Delta G = G_1 - G_2, \Delta B = B_1 - B_2. \quad (2)$$

To infer light polarization state (i.e., color) in current location, we follow the same formulas that convert RGB values to HSV color space [55], by replacing the absolute R, G, B values with their corresponding differential values in Eq. (2). Note that these differential values can be negative. Existing formulas for calculating hue still apply but the saturation calculation faces problems, as negative values can lead to negative saturation and change the saturation scale. To address the problem, we slightly revise the calculation of saturation S' by adjusting the denominator of the original formula:

$$S' = \frac{\max(\Delta R, \Delta G, \Delta B) - \min(\Delta R, \Delta G, \Delta B)}{2 \times \max(|\Delta R|, |\Delta G|, |\Delta B|)}. \quad (3)$$

The new calculation guarantees a positive value within $[0, 1]$, and most importantly, preserves the orthogonality with hue and value because they are independent. We use the hue and saturation derived from differential color values to determine the best set of tape configurations (§3.1) and to evaluate the difference (distance on the HS-plane) between two tape configurations ($d(c_1, c_2)$ function in §3.2). With the above process, our implementation ends up with selecting five tape configurations.

Estimating Impact of Orientation Change. To deal with object's orientation change, we leverage the compass sensor[4] in IMU to monitor object's orientation. Upon any change in orientation, we estimate its impact on the differential values (Eq. (2)) observed by color sensors. Such estimates allow us to correctly identify the light polarization state at current location and thus recognize color changes caused by crossing cell boundaries.

To estimate the impact of orientation change, a simple method is to exhaustively test all orientations at each cell and measure these values. This method, however, is time-consuming and does not scale to different environments. Instead, we propose an empirical model that characterizes the relationship between the object orientation and the light intensity value of each color channel, given a light polarization state. With this model, we can estimate the differential color values of the two color sensors, given object's orientation and the cell the object is located in. Comparing these estimated differential values to the actual observed values, we can then identify whether the object has moved into another cell. This empirical model will be later integrated in the data fusion algorithm (§5) to evaluate whether a candidate location matches the observed differential color values.

Our modeling starts from the recognition that the transparent tape layer turns the incoming linearly polarized light into *elliptically polarized* light, as our extensive measurements show. We hypothesize that it is due to the fact that transparent tapes are made of *uniaxial* materials, the simplest type of birefringence materials, where a single direction (termed as the *fast axis*) dominates the optical anisotropy while its perpendicular axis is referred to as the *slow axis*. Because of the differences of refractive index in these two axises, the phases of light on these axises can be retarded. As a result, a light traversing tapes becomes two linearly polarized waves with phase difference of $90°$, which are then combined as an elliptically polarized wave.

In optics, elliptically polarized light can be described as below using Jones vector [35]:

$$\mathbf{J_{ellip}} = \begin{bmatrix} A \\ B \pm Ci \end{bmatrix}, \qquad (4)$$

where A, $B \pm Ci$ represent the amplitude and phase of the electric field in the x and y direction, respectively. The light intensity is proportional to $A^2 + (B \pm Ci)^2$. Based on Jones calculus [25], when an elliptically polarized light passes through a linear polarizer with orientation θ, the polarization state of the resulting light is described

as

$$\mathbf{E}(\theta) = \mathbf{J_{polar}} \cdot \mathbf{J_{ellip}} = \begin{bmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{bmatrix} \begin{bmatrix} A \\ B \pm Ci \end{bmatrix}, \quad (5)$$

where $\mathbf{J_{polar}}$ is the Jones matrix of a rotated linear polarizer [25]. Then the resulting light intensity $I(\theta)$ is:

$$I(\theta) = \mathbf{E}^{\mathsf{T}}(\theta)\mathbf{E}(\theta) = a_1 \cos^2(\psi - \theta) + a_2 \sin^2(\psi - \theta) + a_3, \quad (6)$$

where $\psi = \text{atan2}(B, A)$, i.e., the direction of the fast axis polarization, a_1 and a_2 are terms combining the amplitude and phase of the elliptically polarized light, and a_3 is noise.

We use the above equation to estimate light intensity of each color channel perceived by a color sensor and thus the differential value between two sensors. Take the red color channel as an example. Assuming the orientation of the two sensors' polarizers are θ and $\theta + 90°$, we can infer $R_1 = I(\theta)$ and $R_2 = I(\theta + 90°)$. Based on Eq. (5), we can compute the estimated $\widetilde{\Delta R}$ as below, with omission of detailed derivations:

$$\widetilde{\Delta R}(\theta) = b \cdot \cos(2(\psi - \theta)) \qquad (7)$$

where $b = a_1 - a_2$. The same model applies to other two color channels. To find the best-fit parameters $\Theta = \{b, \psi\}$ for each color channel, we calibrate the model using a small number (e.g., 30 per tape configuration) of sample measurements. Specifically, we place the two sensors (each with its polarizer film) collocated with IMU in each cell and then randomly rotate them. We collect sensor data on three color channels to compute the differential value on each color channel and record the corresponding orientation value θ reported by compass. We apply gradient descent to identify the optimal parameters that lead to estimates best matching the measurements. Again, taking the red color channel as an example, its best-fit parameters are:

$$\Theta = \underset{\Theta'}{\arg\min} \, (\widetilde{\Delta R}_{\Theta'}(\theta) - \Delta R(\theta))^2 . \qquad (8)$$

The parameters for other color channels are determined similarly. Note that since the differential color values do not contain ambient light, for a given luminary, the calibration for each tape configuration needs to be conducted only once and is reusable for different lights in any lighting condition. Also, given the sampling rates of IMU and color sensors, the calibration can be done fairly quickly (within 2 minutes in our experiments). Thus, the overall overhead of calibration is low.

5 FUSING IMU AND LIGHT DATA

With differential color values sensed by color sensors, the last component fuses these observations with IMU data and outputs object location. Estimation theory [53] in general provides a theoretical framework for fusing different types of sensory data on the basis of Bayesian inference. The goal is to estimate the evolution of system's *internal state* (e.g., object location) based on noisy *observations* (e.g., differential color values). In our context, the nonlinearity of the state evolution renders classic Kalman filter inapplicable. Thus, we apply sequential Monte Carlo (or particle filter) [14] method to solve the fusion problem. It has been proven successful in prior localization systems [67].

At high level, particle filter is a sampling-based approach. It estimates the probability density function (PDF) of the system state

using a set of particles (i.e., samples of the state). Each particle is assigned with a weight that represents the probability of that particle being sampled from the PDF. At each time step k, we sample a set of possible states (particles) based on the current estimated probability density. We then advance (or *predict*) these particles to the next time step $k + 1$ to form a new set of particles based on an evolution model, which takes into account current IMU measurements. Based on new observations (e.g., differential color values) at time step $k + 1$, we then *update* particle weights and thus estimate the probability density function. The new density function is used to re-sample particles and move on to the next time step. We next elaborate on our model and functions used for sampling particles and updating particle weights.

Problem Model. We model our problem as follows:
(1) Hidden State: The hidden state vector $s_k = (x_k, v_k, q_k)$ at time k contains object's 3D coordinate x_k, its velocity v_k in the global frame of reference (i.e., with respect to the coordinate system of a room), and its 3D orientation q_k represented as the quaternion that takes the global frame onto the target frame.
(2) Observations: The observation c_k is the observed differential color values at time k. The relationship between the observations and state is described by the following conditioning probability:

$$p(c_k | s_k) \sim \mathcal{N}(C(x_k, q_k), \Sigma_c), \qquad (9)$$

where $p(c_k | s_k)$ is the probability of observing c_k conditioning on hidden state s_k. We estimate it by assuming c_k is normally distributed with mean of the estimated differential color values from our empirical model $C(\cdot)$ as described in §4, with Σ_c as the covariance matrix of modeling errors.
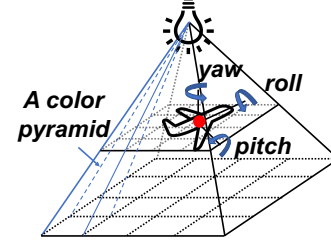*(3) Evolution Model:* As IMU sensors measure acceleration a_k (converted to the global frame) and angular velocity ω'_k (measured in the target frame), s_k evolves as below [20]:

$$v_k = v_{k-1} + a_k \Delta t + n_k^v,$$
$$x_k = x_{k-1} + \frac{1}{2}(v_{k-1} + v_k)\Delta t + n_k^x, \qquad (10)$$
$$q_k = q_{k-1} + \frac{1}{2} q_{k-1} \omega'_k \Delta t + n_k^q,$$

where n_k's are the process noises: $(n_k^v, n_k^x, n_k^q) \sim \mathcal{N}(0, \Sigma_s)$. We assume all the noises are independent, which means the covariance matrix Σ_s is diagonal. In our system, q_k is computed by the internal fusion algorithm of IMU. We simply add some noise to the orientation output from IMU.
(4) Initial State: The initial distribution $p(s_0 | c_0) \equiv p(s_0)$ is known, where we assume that initially (1) target's coordinates x_k are uniformly distributed in the space; (2) the velocity v_k is normally distributed with zero mean; (3) the orientation q_k is normally distributed with mean obtained from the orientation reported by IMU.

Predict and Update. With the above model, the tracking problem is to estimate state s_k at time k given all the observations $c_{0:k}$ up to time k along a trajectory, by constructing the PDF $p(s_k | c_{0:k})$. Since deriving the actual PDF is analytically intractable in our problem[5], we use particle filter to approximate the PDF with a set

---

[5]Although we assume Gaussian process and measurement noise, the state distribution conditioned on all previous observations is not necessarily Gaussian, because the observation model (Eq. 9) and the evolution model (Eq. 10) are nonlinear.

Figure 5: Color pattern alone cannot differentiate object's vertical position within the pyramid from the luminary to a cell.

of N_s particles $\{(s_k^i, w_k^i)\}_{i=1}^{N_s}$, where w_k^i is the weight of particle s_k^i, $\sum_i w_k^i = 1$. Then, the discrete approximation to the true posterior probability density is

$$p(s_k | c_{0:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(s_k - s_k^i), \qquad (11)$$

where $\delta(\cdot)$ is the multi-dimensional Dirac delta function that only returns 1 when $s_k = s_k^i$.

We predict the particles of the new state k based on the sampling importance resampling (SIR) filter [29]. We advance existing particles $\{(s_{k-1}^i, w_{k-1}^i)\}_{i=1}^{N_s}$ at time $k - 1$ to time k by drawing samples from the importance distribution

$$s_k^i \sim p(s_k | s_{k-1}^i). \qquad (12)$$

Specifically, we first draw samples from the process noise $\mathcal{N}(0, \Sigma_s)$ and then advance particles according to the evolution model in Eq. (10). With new observation c_k at time k, we then update particle weights based on the observation model (Eq. (9)) so that particles conforming to the new observation are favored (update stage):

$$w_k^i \propto p(c_k | s_k^i). \qquad (13)$$

A common problem of the SIR particle filter is the degeneracy phenomenon, where after a few iterations, all but one particle will have a negligible weight, making the approximation of the PDF inaccurate. The effect of degeneracy can be mitigated by resampling, where we generate samples from the discrete approximation so that $\Pr\{s_k^{i*} = s_k^i\} = w_k^i$ and the weights are reset to $w_k^i = 1/N_s$.

After each iteration, we identify the cell in which the sum of particles' weights is maximal. Then object location is estimated as the weighted mean of the particles in this cell.

3D Tracking. In 3D tracking, we face the challenge that the color pattern alone is 2D and thus introduces ambiguity in determining the vertical position. As shown in Figure 5, when evaluating the light polarization state at a cell only based on color (hue and saturation), the pattern is the same within a pyramid that connects a cell on the floor and the luminary on the ceiling. Although the shrinking of cell sizes as height increases leads to cell-crossing events when object moves vertically, such events are rare, providing limited opportunities to calibrate IMU data.

To improve the accuracy in 3D tracking, we extract a feature related to light intensity, in addition to color, based on the differential color values. Since ambient light is already removed from the differential color values, this method is robust against ambient light interference. In particular, when converting the differential color values to the HSV space, we derive the *value* as the channel

(red, green, or blue) whose absolute value is maximum. Because the received light intensity is determined by the incident angle, irradiant angle, and distance from the light source to the light sensor, we apply the following model in that channel to predict the *value*. According to the Lambertian model [37], the received light intensity is $I_r = I_0 \cos^m(\phi) \cos(\beta)/d^2$, where ϕ is the irradiant angle, β is the incident angle, and d is the distance from the light source to the color sensor. I_0 and m are constants that are related only to the lamp[6]. In each pyramid, the irradiant angle is approximately a constant because the pyramid is very narrow. Thus, $I_0 \cos^m(\phi)$ is a constant for every pyramid that can be estimated a priori. Conditioning on the state of a particle, including the 3D position and orientation of the color sensors, we can calculate β and d. Therefore, by adding *value* in c_k, in addition to hue and saturation, we allow the vertical position to be calibrated using the observed *value*, which improves 3D tracking accuracy.

6 PROTOTYPE IMPLEMENTATION

Sensing Unit. Our sensing unit prototype consists of four components: (1) two color sensors (AMS-TAOS TCS3472) with polarizer films. Each sensor (2 mm × 2.4 mm in size) contains a photodiode array and consumes 2.04-mW active power from our measurements. Its integration time is configurable (2.4 ms – 700 ms) and we set it to 50 ms, resulting into 20-Hz sampling rate of the optical signals. Light intensities of the red (R), green (G), and blue (B) channels are streamed to the micro-controller via I^2C bus. We cover sensor's photodiode array area with a 1 cm × 1 cm polarizer film [7] so that the sensor perceives light polarization states as colors. The color sensors in our current prototype are saturated when light intensity is above approximately 2000 lux; (2) a 9-degree-of-freedom (9-DoF) IMU (BNO055 [4]), which integrates a MEMS accelerometer, magnetometer, and gyroscope into a single die with an ARM Cortex-M0 based processor to fuse all sensor data. The IMU outputs linear acceleration and quaternions at 100 Hz; (3) a micro-controller (Raspberry Pi Zero W). We choose Raspberry Pi for its flexibility and ease of programming. We host the color sensors and IMU on a shield board and plug the shield board into the Raspberry Pi for a compact look (Figure 6(a)); (4) an external 500 mAh lithium-ion polymer battery to power the whole system. For devices with built-in IMU, micro-controller, and internal power source, we can reuse these components on the device and only add two color sensors, further reducing the power, weight, and size of our sensing unit.

Light Cover Layer. We fabricate a light cover using two Fresnel lenses [3] (12-cm focal length, 0.15-cm center thickness) and a 0.18-mm thin polarizer film [7] with transparent tape (HOME360™) stripes applied atop. The polarizer and tapes are sandwiched between Fresnel lenses. We add the cover to a lamp (CREE XHP70 chip, 7 mm× 7 mm in size) on the ceiling at 2.6-m height (Figure 6(b)).

To create the tape pattern, we cut tapes into pieces and paste them atop the polarizer film (6 cm × 6 cm) based on the assignment of tape configurations generated by Algorithm 1. To enhance the saturation of resulting colors, we arrange the orientations of the first and last tape layers so that they differ by 45°. This arrangement

disperses the polarization of light rays of different wavelengths to the largest extent, thus increasing the color purity in a polarization direction. We employ two lenses to produce sharp cell edges (Figure 1 and 6(b)). The first lens collimates the light before the polarizer and tape, resulting in the same incident angles impending all cells at the first polarizer[7], while the second lens projects the light pattern to a larger area. We create a grid pattern with 6×6 cells. Each cell is 1 cm × 1 cm, projected to the floor as a 20 cm × 20 cm cell (Figure 6(c)). The cell size is constrained by our manual tape pasting/cutting. When manufactured with proper machinery, the thickness and orientation of the tape layer can be controlled in a finer granularity (at millimeter-level), producing much smaller cells to improve tracking accuracy. At the same time, by selecting and arranging the lenses properly, we can project the light pattern to a much larger area without losing tracking accuracy, if we maintain the cell size on the floor through smaller tape cells on the lamp.

7 EVALUATION

Using our prototype, we examine our method based on both system performance and user perception of the projected pattern. We evaluate system performance on tracking accuracy, robustness (i.e., dealing with light blockage, ambient light, and orientation change), tracking latency, and power consumption. A user study is conducted to examine the impact of the polarization pattern on illumination.

Experimental Setup. We conduct our experiments in a lab setting (Figure 7), where we mount our light cover in front of a luminary on the ceiling (2.6-m height) and cast a polarization pattern in a 1.2 m × 1.2 m area on the floor. The color pattern is invisible via naked eyes (Figure 6(d)). We observe gray lines at cell boundaries due to artifacts of our tape pasting. We will discuss eliminating these imperfections in §8.

We examine two types of application scenarios: (1) indoor navigation of autonomous robots on a 2D plane, where we add the sensing unit to a robotic car [6]; and (2) tracking on-body devices in the 3D space (e.g., hand controller in virtual reality), where a user holds our sensing unit in hand while freely gesturing. In both scenarios, the Raspberry Pi of our sensing unit streams IMU and light data to a laptop, which then runs the data fusion algorithm (§5) to output object location [9]. We calibrate our empirical model (Eq. (8)) using data collected under no ambient light, where we turn off all other artificial lights and close the windows.

To gather the ground truth, we deploy the Vicon motion capture system [9] by setting up four Vicon MX-F40 cameras around our experiment area. We attach a retro-reflective marker to the target object and Vicon cameras track marker's 3D location at 100-Hz rate. The mean absolute positioning error of Vicon is less than 2 mm for dynamic tracking [43]. Given that the tracking rates of Vicon and our system are different, we synchronize the results by an anchor frame where the color sensor and Vicon marker are blocked at the same time. We then associate each tracking output from our

---

[6]If we know the irradiance pattern of the light source from the lampshade customization and the power of the light, we do not need train the Lambertian model to obtain those parameters.

[7]If light strikes the first polarizer and birefringent film with nonuniform incident angles, cells with the same tape configuration can have different polarization states (i.e., different colors perceived by the color sensor).

[8]With optimized lens to collimate light from the light source, all components (lenses, polarizer, transparent tapes) can be closely stacked, resulting into overall thickness less than 1 cm.

[9]We did not run the fusion algorithm on Raspberry Pi because of its sufficient RAM.

| (a) Sensing unit | (b) Light cover layer | (c) w/ second polarizer | (d) w/o second polarizer |
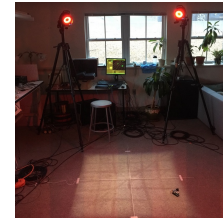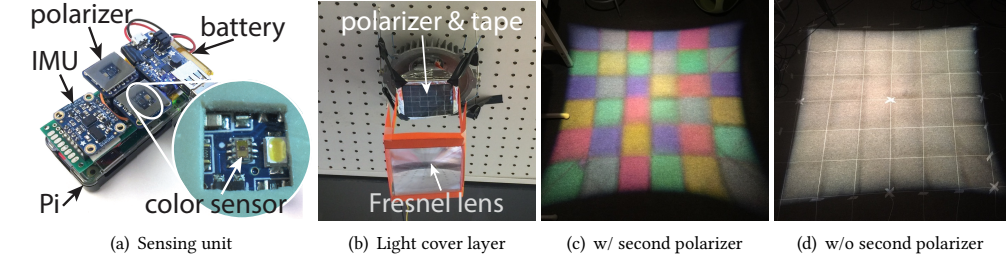
Figure 6: Our prototype of the sensing unit (a) and light cover layer (b). As the first lens does not perfectly collimate light, we have to place the projection lens 12 cm (its focal length) away from the polarizer and tapes using a 3D-printed holder[8]. (c) shows the projected color pattern when we place a second polarizer at the light cover to make the color visible. The pattern otherwise is invisible (d). We place white thin lines along cell boundaries to ease experiments.

Figure 7: Experiment setup with four VICON cameras to gather the ground truth.

| (a) Trajectory | (b) Error over time | (c) CDF |
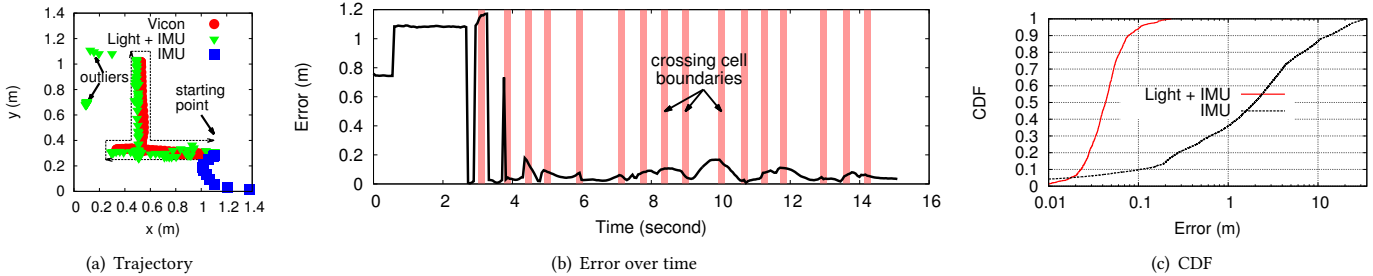
Figure 8: Tracking accuracy in comparison to IMU tracking alone. The locations inferred from our system closely match those from Vicon, except a few outliers (a). These outliers occur during bootstrapping before the object crosses any cell boundary and the error is quickly corrected once two or three boundary-crossing events occur (b). Overall the tracking error after bootstrapping is 5.2 cm in median, with 11 cm as the 90th percentile.

system with the Vicon output that has the closest timestamp. The tracking error is then calculated as the distance between the inferred coordinate as described in §5 and that from Vicon. Experiments are conducted during the day by default, where ambient light consists of sunlight and red light from Vicon. Overall it is measured as 178 lux at the center of the experiment area. We examine more ambient light settings in §7.2.

7.1 Tracking Accuracy

2D. In 2D tracking, the robotic cart freely moves in our experiment area for 2 minutes for 10 rounds. In addition to our method and Vicon, we also include the results by only integrating IMU data, which serve as a baseline to examine the efficacy of light-based augmentation. In the IMU-alone case, we assume the object's initial displacement and velocity are known as priors, while our method has no prior knowledge of the initial location. For both methods, we ignore acceleration data below 0.1 m/s^2 to discard small-magnitude acceleration due to environmental factors (e.g., vibrations).

We plot the tracking results from our system, Vicon, as well as IMU-alone tracking in the 2D plane (Figure 8(a)), where the arrow of each point indicates the direction of velocity. Our main observation is that the trajectory inferred by our system closely matches that of Vicon, except a few outliers in the beginning. In comparison, the trajectory integrated from IMU data alone deviates quickly from the actual trajectory and cannot be calibrated later (beyond the figure border and thus not shown). The result demonstrates the efficacy of calibration provided by the light polarization pattern.

To further examine the outliers of our tracking results and the evolution of the tracking error, we plot the time series of tracking

| (a) Bootstrapping stage | (b) After crossing cells |
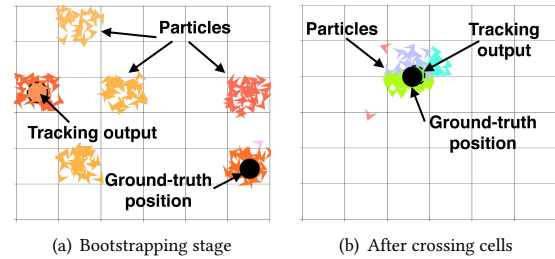
Figure 9: Particle filter in 2D tracking: particles are triangles, with arrows as moving directions and colors as predicted colors. For clean illustration, we show only 200 particles uniformly sampled from 1024. Vicon output is the black circle while ours is a circle with color based on Eq (2). Particles are in multiple cells in bootstrapping (a) and quickly converge to the correct cell after a few boundary-crossing events (b).

errors in Figure 8(b), where we mark in red bars the time points when a boundary-crossing event occurs. After the bootstrapping stage, the valleys of the tracking error are not always perfectly aligned with cell-crossings, because only cell-crossings with strong color contrast decrease the error dramatically. In some rare cases where the colors of adjacent cells are less contrastive, the cell-crossing does not reduce the error as much. We also plot the state of the particles in the 2D plane in the beginning of the tracking, and after a boundary-crossing event occurs (Figure 9). We observe that these outliers are from the bootstrap stage before any cell-crossing events occur. Observing a single polarization state cannot narrow down object's location to a single cell, leading to higher location errors (Figure 9(a)). However, the error is quickly corrected once two or three boundary-crossings occur (Figure 9(b)).

(a) Error over time   (b) CDF
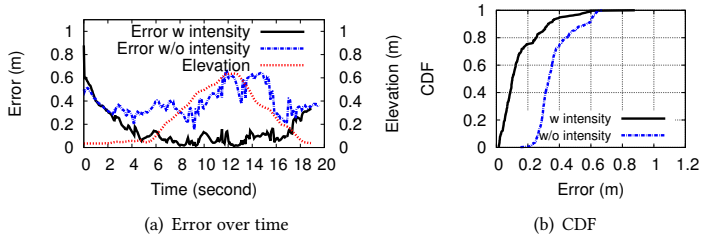
Figure 10: Accuracy of 3D tracking with and without adding the feature related to light intensity.

Figure 11: Tracking robustness against object's orientation change and light blockage.

(a) Hue error   (b) Saturation error   (c) Value error

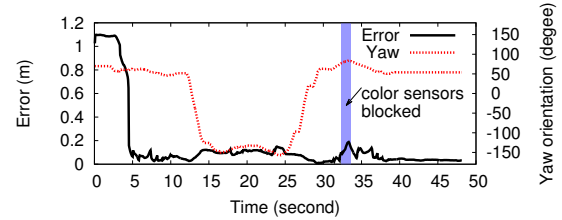**Figure 12: Impact of ambient light on empirical model accuracy.**

From Figure 9, we also observe that tracking errors are bounded within 20 cm, which is the cell size on the floor. This is because from the observed light pattern and device orientation, the system knows that the object does not move out of a particular cell, and thus effectively bounds the tracking error, even in the worst-case scenario when the object stays stationary for long. Furthermore, the tracking errors grow in between the boundary-crossing events, because only IMU data are used to infer object location. The error is quickly corrected upon each boundary-crossing, which explains the fluctuation of tracking errors over time. Overall, as we plot the CDF of tracking errors in Figure 8(c) that excludes the bootstrap stage, the median tracking error is 4.3 cm and the 90th percentile error is 8.1 cm. In comparison, the error of IMU-tracking is unbounded.

3D. In 3D tracking, a user holds our sensing unit in hand while freely gesturing in the 3D space. We plot the time series of the tracking error in Figure 10(a) along with the elevation change over time. We also include the results when we remove the feature related to light intensity from the data fusion algorithm and use only the boundary-crossing events for calibration. This allows us to examine the efficacy of adding the light intensity feature for 3D tracking. Figure 10(b) further compares the CDFs of errors for these two cases. We make two observations. First, 3D tracking error is slightly higher than 2D given the limited calibration opportunities offered by vertical boundary-crossings. The larger errors occur mainly in the beginning of the movement when the system is still accumulating observations from color sensors. The error gradually decreases as more calibration events occur. Second, adding the feature related to light intensity greatly drives down the tracking error, where the median error decreases from 33 cm to 10 cm. Note that the intensity feature is currently calculated based on the predicted differential color values using training data collected on the floor. With a few sample measurements at other heights for training, 3D tracking accuracy can be further improved.

7.2 Tracking Robustness

We next conduct experiments to examine system robustness against object's orientation change, light blockage, and ambient light. We use 2D tracking as an example and the result holds in 3D scenarios.

Object Orientation. We start with examining the tracking accuracy as object changes its orientation during movement. We conduct experiments with the robotic cart moving on the floor, where the cart gradually varies its orientation (yaw). In Figure 11, we plot the tracking error and its yaw orientation over time. We observe that the tracking error remains within 15 cm as the orientation changes. It indicates the efficacy of our empirical model in §4 in predicting the impact of orientation change. It allows the data fusion algorithm
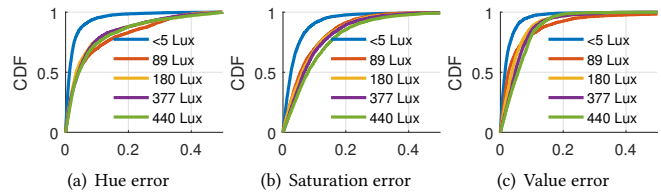
to correctly identify color changes due to boundary crossings and calibrate IMU tracking. We will elaborate on the accuracy of our empirical model later in Figure 12.

Light Blockage. We next examine how our system deals with occasional losses of light signals, which can be caused by either blockage from other objects or entering coverage holes in the 3D space. We experiment with the robotic cart moving on the floor, where we occasionally block the two color sensors for 1.3 seconds (Figure 11). We mark the blockage duration as a blue bar. We observe that the tracking error starts to increase during this period, as no light pattern is observed for calibrating IMU tracking. Since the particle filter method is unable to calibrate the position in the update stage, it outputs the prediction using the IMU data, leading to increasing error. Once light signal becomes available, the system quickly corrects the tracking error based on newly observed light pattern, indicating that the system gracefully handles losses of light signals. The re-initialization is required only when the blockage period is too long (e.g., > 5 seconds). In this case, the system needs to re-enter the bootstrapping stage, which often takes two or three cell-crossing events to recover tracking.

Ambient Light. Finally we examine the impact of ambient light. We conduct experiments under different ambient light conditions: daytime with sunlight (178 lux) and nighttime with fluorescent lights (232 lux) close to our luminary. In both cases, the red light from Vicon also contributes to ambient light interference and is measured 20 lux at the center of our experiment area. Ambient light is unpolarized in our experiments. We observe similar tracking performance in all these conditions. It demonstrates that our dual sensor design effectively mitigates the interference from unpolarized ambient light. We omit the figure in the interest of space.

We also examine the impact of ambient light on the accuracy of our empirical model (§4). Given that the data used to calibrate the model are collected under one lighting condition, we aim to examine its prediction accuracy in different light conditions. To do so, we collect the differential color values using two color sensors in five ambient light settings: under 5 lux with all lights turned off, 89, 180, 377, and 440 lux with different florescent light turned on respectively. We randomly rotate color sensors in different cells.
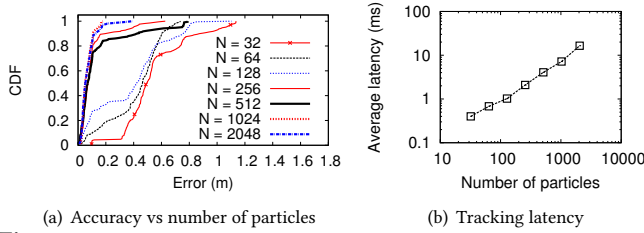
(a) Accuracy vs number of particles   (b) Tracking latency

Figure 13: Tracking accuracy with different numbers of particles.

We then calculate the hue, saturation (Eq (3)), and *value* (used in 3D tracking) based on measured differential color values in these five settings. We compare them to that calculated based on predicted differential color values from our calibrated model. We plot the accuracy of the predicted hue, saturation, and *value* in Figure 12, where *value* is scaled based on the maximum. We observe that prediction accuracy is the highest under the darkest condition (< 5 lux), as it is the closest to the setting where calibration data were collected. Ambient light difference affects prediction accuracy because the polarization of the ambient light is not completely uniform. Across other settings, however, the difference of prediction accuracy is negligible. The result demonstrates the accuracy of the calibrated model in predicting differential color values in other different lighting conditions.

7.3 Practical Considerations

Tracking Latency. The tracking latency of our system depends on the number of particles in the particle filter algorithm (Figure 13(b)). More particles represent the posterior state distribution more accurately and yet also entail higher time complexity. To identify the particle number achieving the best tradeoff, we run the particle filter algorithm offline with different numbers of particles and plot the CDFs of tracking errors in Figure 13(a). We observe that the accuracy saturates once the number of particles reaches 1024. Thus, we use 1024 particles in our experiments. The average tracking latency for 1024 particles is 7.18 ms with our current C implementation of the particle filter algorithm. It can potentially support tracking rates up to 140 Hz.

Impact on Illumination. Finally, we examine user's perception on the projected polarization pattern via a user study. We set up two lights on the ceiling, one without our light cover and the other with our cover. We mount the lights behind the ceiling tiles so that participants cannot differentiate the lights based on the look. We cut two holes on the tiles to allow light rays to pass through and illuminate two areas (A, B) on a table beneath. We invite 33 participants (9 females, 22 – 37 years old). Each participant stands or sits around the table for 30 seconds and then fills in a questionnaire with following questions: (1) *Do you feel the light in location A (B) comfortable?* (2) *Do you think the light evenly distributed in location A (B)?* (3) *Do feel the light intensities/colors in two locations are different?* Participants rate their answers from 1 (worst) to 5 (best), not informed of the study purpose.

Our results show that the mean comfort score of location A (without our light cover) and B (with our light cover) is 4.03 and 4.15, respectively, and the mean score on light uniformity of location A and B is 3.75 and 3.87, respectively. The differences between scores of location A and B are negligible, verifying that the effect

of our light cover layer is negligible. The mean scores on the light intensity and color difference between two locations are 4 and 3.5, respectively[10]. The color difference is partly contributed by surrounding objects in the environment. Overall, no participant perceives any color pattern. The results confirm that our light cover layer does not affect normal illumination.

8 DISCUSSIONS

Advanced Polarization Patterns. Our current study considers a grid polarization pattern for its simplicity. Other types of patterns are worth future explorations as we move forward. For instance, a hexagon cell shares boundaries with more (6) cells than a cell in grid pattern (4). Thus a boundary crossing better differentiates moving direction. Another possibility is to generate a pattern based on a De Bruijn sequence [27]. It allows a color sequence of visited cells to be unique within a large pattern. Additionally, to generate patterns in finer granularity, we plan to explore more advanced manufacturing methods to fabricate the birefringent film. It will further boost tracking accuracy as the maximal tracking error is bounded by the cell size. Our key design components are general and apply to other types of patterns.

Pattern Imperfections. With our current prototype, we observe imperfections in the projected pattern, including gray boundary lines, subtle differences of cell sizes, and slight pattern distortion near coverage corners. These imperfections are due to multiple factors, such as our imperfect manual cutting and application of transparent tapes, and artifact of our lenses. They can be accounted for via a quick one-time calibration, where we collect information (e.g., take a photo) of the projected pattern and build a mapping to translate any 3D coordinate to a cell. Moreover, pattern imperfections can also occur because of the gradual aging and distortion of the fresnel lenses in close range of the heat dissipated from the LED. As an example, the fresnel lens in our prototype is made of optical PVC with maximum operating temperature around 60°C [12]. We observe lens distortion after 5-month use. This problem, however, can be mitigated by using lenses made of more heat-resistant materials. Overall the problem of pattern imperfections is not fundamental to our approach and can be eliminated using more advanced photonic devices and materials in mass production.

Diverse Luminary Types. While our current prototype is for a point light source, we can extend our design to diverse luminary types (e.g., light panels, elongated fluorescent lights, and recessed lights), by adapting the shape of the first lens after the light source (Figure 1). As long as light rays passing the first lens are collimated, the other layers (i.e, polarizer, birefringent film, lens for projection) of the light cover remain the same. Optimizing lens shape for collimating incoming light rays has been studied in optical lens design [22, 51]. We will explore these methods to design light covers for other types of luminaries.

Extending Coverage. With a single luminary, our current experiments are within a limited area. To expand the sensing coverage, we will examine scenarios with multiple lights that jointly

---

[10]Although the low-cost polarizer (42% transmission ratio) in our light cover degrades light intensity, Fresnel lens concentrates light into a smaller region, which compensates for the degradation from polarizer, rendering area B slightly brighter than A.

generate a spatial polarization pattern covering a large area. Key challenges arise on possible overlaps of adjacent lights' coverage regions, where cells in the overlap region contain mixed light polarization states. To mitigate the overlap, we will consider regulating the coverage shape of each light. This can be achieved by optimizing the shape of the projection lens, a well-studied problem in computer graphics [74] offering solutions for our test. If coverage regions of multiple lights overlap significantly, it is not necessary to augment each luminary. Instead, we can select to augment a subset of luminaries that collectively cover the space of interest. If small gaps remain between lights, it is similar to scenarios with temporary light blockage and thus can be addressed by sensor fusion with IMU. Furthermore, when extending to multiple lights, our system only needs 3D positions of luminaries as the input, without the need to map a floor or fingerprint each cell.

Sensor Orientation.   In our current experiment, we have placed the sensing unit atop the tracking object where the color sensors are kept facing the ceiling. In scenarios of integrating our sensing units with wearable devices (e.g., smart watches), sensor orientation can vary depending on the body movement (e.g., wrist rotation, arm movement). As a result, light from the augmented luminary may occasionally be outside sensors' field of view and thus not perceived by the sensors. To support various sensor orientations, we will consider adding multiple pairs of color sensors each facing a different direction. It increases the sensing unit's angle diversity and reduces the likelihood of not perceiving the light pattern. The overhead of additional light sensors is small given their small size (e.g., 2 mm × 2.4 mm) and low power consumption (e.g., 2 mW).

9 RELATED WORK

IMU Calibration.   Prior studies have revealed IMU's serious drift in orientation and location estimation [63, 64, 75]. Prior methods calibrate IMU with or without high-end equipments. In particular, [45] uses accelerometers to calibrate gyroscope, while [44] leverages a thermal chamber to reduce the impact of the thermal drift. Other approaches [16, 24] combine the three-axis effect of physical signals such as local gravity, earth's rotation and the magnetic filed without the assistance of high-end equipments. They, however, require user involvement to move sensors and thus are not applicable in robotic applications (e.g., autonomous drones).

Recent research then focuses on calibrating IMU using known locations as landmarks, such as GPS signals, outdoor landmarks [19], footsteps [30, 40, 64], Wi-Fi signal strengths [23, 54, 60], maps [1, 18, 39, 47, 65, 69], or visual features [31, 36]. Our approach differs in that we study finer-grained light patterns as landmarks. Though some studies [28, 34, 67] have considered fusing light with IMU for indoor localization, they exploit coarse light features and target 2D localization. For example, [67] leverages the rise of light intensity as a user approaches a light source to calibrate IMU, achieving a mean location error of 38 – 74 cm. Our approach proposes a novel light pattern to significantly improve the localization resolution (cm-level) and enable 3D tracking.

Light-based Localization.   Active research has studied the use of visible or infrared light for indoor localization. Some modulate lights to broadcast anchor signals, while the others exploit inherent

light features. The former uses photodiodes or cameras to detect anchor signals, estimate distances to each LED, and infer their locations [13, 38, 41, 48, 66, 68]. Among them, [68] leverages polarized lens to create coarse-grained light patterns, [13] generates OFDM-based anchor signals, and [66] covers infrared lights with a rotating lampshade to emit coded signals. [21, 62] design *active* polarization patterns for 3D tracking and long-range communications. [42] enables a single light of LED arrays to blink a unique frequency at each pixel. Those systems however all need considerable costs of modifying the lighting infrastructure. The latter exploits ambient light features for coarse-grained (room-level) localization [15, 28, 50]. Later designs [49, 71, 72, 76] identify physical features of individual lights to further reduce localization errors. Overall, light-based localization alone is vulnerable to light blockage and ambient light, and thus unable to achieve reliable localization in practice. Our design combines light and IMU to embrace the precision from light and reliability from IMU.

Camera-based Tracking.   Optical motion capturing systems, such as OptiTrack [10], Vicon [9], and PhaseSpace [11], track retro-reflective markers using an array of customized cameras. They are widely used to collect the ground truth to evaluate other tracking technologies with inertial sensors [56], vision [70], or RF [61]. Despite millimeter-level accuracy [10, 43], those systems are expensive ($600−$6000 per camera) and require special perceptible lighting. In comparison, our method entails lower cost and reuses existing luminaries. Markerless optical motion capturing technologies employ cameras to reconstruct the skeleton, head pose, and/or facial expression using vision algorithms [33, 52, 73]. Those algorithms require a large dataset (hundreds of thousands of training images) to robustly track users in different body shapes and sizes in varied poses [33, 52, 73]. Our method does not require training to track various objects.

10 CONCLUSION

We studied augmenting IMU tracking with well-designed, static light polarization patterns that are created by covering an existing luminary with a thin cheap polarizer with specially-arranged transparent tapes. The imperceptible pattern is detectable by color sensors viewing through another polarizer and thus constrains a collocated IMU's drift errors. We investigated the design and detection of the light pattern, as well as the fusion algorithm with IMU data. Our prototype experiments demonstrated the tracking accuracy and robustness of our approach.

11 ACKNOWLEDGMENTS

REFERENCES

[1] [n. d.]. APFiLoc: An Infrastructure-Free Indoor Localization Method Fusing Smartphone Inertial Sensors, Landmarks and Map Information. ([n. d.]).

[2] 2015. Review: iRobot Roomba 980. https://spectrum.ieee.org/automaton/robotics/home-robots/review-irobot-roomba-980. (2015).

[3] 2017. 2.3" x 2.3", 1.5" Focal Length, Fresnel Lens. https://www.edmundoptics.com/optics/optical-lenses/fresnel-lenses/fresnel-lenses/43024/. (2017).

[4] 2017. BNO055 Smart Hubs and ASSNs. https://www.bosch-sensortec.com/bst/products/all_products/bno055. (2017).

[5] 2017. Crazyflie 2.0. https://www.bitcraze.io/crazyflie-2/. (2017).

[6] 2017. Elegoo EL-KIT-012 UNO Project Smart Robot Car Kit. https://www.elegoo.com/product/arduinocarv3-0/. (2017).

[7] 2017. Linear polarizer, 3Dlens.com. https://www.3dlens.com/shop/linearpolarizer.php. (2017).

[8] 2017. Optical Birefringence. http://www.olympusmicro.com/primer/lightandcolor/birefringence.html. (2017).

[9] 2017. Vicon Motion Capture System. https://www.vicon.com/. (2017).

[10] 2018. OptiTrack. http://optitrack.com/support/faq/general.html. (2018).

[11] 2018. PhaseSpace Impulse X2E Motion Capture System. http://phasespace.com/x2e-motion-capture/. (2018).

[12] 2018. Polyvinyl chloride. https://en.wikipedia.org/wiki/Polyvinyl_chloride. (2018).

[13] Mohammadreza Aminikashani, Wenjun Gu, and Mohsen Kavehrad. 2015. Indoor positioning in high speed ofdm visible light communications. *arXiv preprint arXiv:1505.01811* (2015).

[14] M Sanjeev Arulampalam, Simon Maskell, Neil Gordon, and Tim Clapp. 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing* 50, 2 (2002), 174–188.

[15] Martin Azizyan, Ionut Constandache, and Romit Roy Choudhury. 2009. SurroundSense: Mobile Phone Localization via Ambience Fingerprinting. In *Proc. of MobiCom*.

[16] E. R. Bachmann, Xiaoping Yun, D. McKinney, R. B. McGhee, and M. J. Zyda. 2003. Design and implementation of MARG sensors for 3-DOF orientation measurement of rigid bodies. In *IEEE International Conference on Robotics and Automation*.

[17] Neil M. Barbour. 2004. Inertial Navigation Sensors. In *NATO RTO Lecture Series 232 "Advances in Navigation Sensors and Integration Technology"*.

[18] S. Beauregard, Widyawan, and M. Klepal. 2008. Indoor PDR performance enhancement using minimal map information and particle filters. In *IEEE/ION Position, Location and Navigation Symposium*.

[19] Cheng Bo, Xiang-Yang Li, Taeho Jung, Xufei Mao, Yue Tao, and Lan Yao. 2013. SmartLoc: Push the Limit of the Inertial Sensor Based Metropolitan Localization Using Smartphone. In *Proc. of MobiCom*.

[20] Michael Boyle. 2017. The Integration of Angular Velocity. *Advances in Applied Clifford Algebras* 27, 3 (01 Sep 2017), 2345–2374. https://doi.org/10.1007/s00006-017-0793-z

[21] Chun-Ling Chan, Hsin-Mu Tsai, and Kate Ching-Ju Lin. 2017. POLI: Long-Range Visible Light Communications Using Polarized Light Intensity Modulation. In *Proc. of MobiSys*.

[22] Chen Chen and Xiaohui Zhang. 2014. Design of optical system for collimating the light of an LED uniformly. *J. Opt. Soc. Am. A* 31, 5 (May 2014), 1118–1125.

[23] Zhenghua Chen, Han Zou, Hao Jiang, Qingchang Zhu, Yeng Soh, and Lihua Xie. 2015. Fusion of WiFi, Smartphone Sensors and Landmarks Using the Kalman Filter for Indoor Localization. *Sensors* 15, 1 (Jan 2015), 715–732.

[24] C. M. Cheuk, T. K. Lau, K. W. Lin, and Y. Liu. 2012. Automatic calibration for inertial measurement unit. In *International Conference on Control Automation Robotics Vision*.

[25] Edward Collett. 2005. *Field guide to polarization*. Vol. FG05. SPIE Press.

[26] DJI. [n. d.]. *Spark User Manual*.

[27] Jason Geng. 2011. Structured-light 3D surface imaging: a tutorial. *Advances in Optics and Photonics* 3, 2 (Jun 2011), 128–160.

[28] A. R. Golding and N. Lesh. 1999. Indoor navigation using a diverse set of cheap, wearable sensors. In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*.

[29] Neil J Gordon, David J Salmond, and Adrian FM Smith. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, Vol. 140. IET, 107–113.

[30] R. Harle. 2013. A Survey of Indoor Inertial Positioning Systems for Pedestrians. *IEEE Communications Surveys Tutorials* 15, 3 (2013), 1281–1293.

[31] Joel A. Hesch, Dimitrios G. Kottas, Sean L. Bowman, and Stergios I. Roumeliotis. 2014. Camera-IMU-based localization: Observability analysis and consistency improvement. *International Journal of Robotics Research* 33, 1 (2014), 182–201.

[32] Gábor Horváth and Dezsö Varjú. 2013. *Polarized light in animal vision: polarization patterns in nature*. Springer Science & Business Media.

[33] Umar Iqbal, Anton Milan, and Juergen Gall. 2017. Posetrack: Joint multi-person pose estimation and tracking. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1.

[34] Antonio R. JimÃŋnez, Francisco Zampella, and Fernando Seco. 2014. Improving Inertial Pedestrian Dead-Reckoning by Detecting Unmodified Switched-on Lamps in Buildings. *Sensors* 14, 1 (2014), 731.

[35] R. Clark Jones. 1941. A New Calculus for the Treatment of Optical Systems I. Description and Discussion of the Calculus. *J. Opt. Soc. Am.* 31, 7 (Jul 1941), 488–493. http://www.osapublishing.org/abstract.cfm?URI=josa-31-7-488

[36] Jonathan Kelly and Gaurav S. Sukhatme. 2011. Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration. *International Journal of Robotics Research* (2011), 56–79.

[37] Toshihiko Komine and Masao Nakagawa. 2004. Fundamental analysis for visible-light communication system using LED lights. *IEEE transactions on Consumer Electronics* 50, 1 (2004), 100–107.

[38] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor positioning with mobile phones and visible light. In *Proc. of MobiCom*.

[39] K. C. Lan and W. Y. Shih. 2014. Using smart-phones and floor plans for indoor location tracking. *IEEE Transactions on Human-Machine Systems* 44, 2 (April 2014), 211–221.

[40] Fan Li, Chunshui Zhao, Guanzhong Ding, Jian Gong, Chenxing Liu, and Feng Zhao. 2012. A Reliable and Accurate Indoor Localization Method Using Phone Inertial Sensors. In *Proc. of UbiComp*.

[41] Liqun Li, Pan Hu, Chunyi Peng, Guobin Shen, and Feng Zhao. 2014. Epsilon: A visible light based positioning system. In *Proc. of NSDI*.

[42] Song Liu and Tian He. 2017. SmartLight: Light-weight 3D Indoor Localization Using a Single LED Lamp. In *Proc. of SenSys*.

[43] Pierre Merriaux, Yohan Dupuis, Rémi Boutteau, Pascal Vasseur, and Xavier Savatier. 2017. A Study of Vicon System Positioning Performance. *Sensors* 17, 7 (2017), 1591.

[44] Xiaoji Niu, You Li, Hongping Zhang, Qingjiang Wang, and Yalong Ban. 2013. Fast Thermal Calibration of Low-Grade Inertial Sensors and Inertial Measurement Units. *Sensors* 13, 9 (2013), 12192–12217.

[45] A. Olivares, G. Olivares, J. M. Górriz, and J. Ramírez. 2009. High-efficiency low-cost accelerometer-aided gyroscope calibration. In *International Conference on Test and Measurement*, Vol. 1. 354–360.

[46] Shashi Poddar, Vipan Kumar, and Amod Kumar. 2016. A Comprehensive Overview of Inertial Sensor Calibration Techniques. 139 (08 2016).

[47] Jiuchao Qian, Jiabin Ma, Rendong Ying, Peilin Liu, and Ling Pei. 2013. An improved indoor localization method using smartphone inertial sensors. In *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*.

[48] Niranjini Rajagopal, Patrick Lazik, and Anthony Rowe. 2014. Visual Light Landmarks for Mobile Devices. In *Proc. of IPSN*.

[49] Julian Randall, Oliver Amft, Jürgen Bohn, and Martin Burri. 2007. LuxTrace: Indoor Positioning Using Building Illumination. *Personal Ubiquitous Comput.* 11, 6 (2007).

[50] Nishkam Ravi and Liviu Iftode. 2007. FiatLux: Fingerprinting Rooms Using Light Intensity.

[51] Patrick W. Rhodes and David L. Shealy. 1980. Refractive optical systems for irradiance redistribution of collimated radiation: their design and analysis. *Appl. Opt.* 19, 20 (Oct 1980), 3545–3553.

[52] Grégory Rogez, Philippe Weinzaepfel, and Cordelia Schmid. 2018. LCR-Net++: Multi-person 2D and 3D Pose Detection in Natural Images. *CoRR* abs/1803.00455v1 (2018).

[53] Ali H Sayed. 2003. *Fundamentals of adaptive filtering*. John Wiley & Sons.

[54] Guobin Shen, Zhuo Chen, Peichao Zhang, Thomas Moscibroda, and Yongguang Zhang. 2013. Walkie-Markie: Indoor Pathway Mapping Made Easy. In *Proc. of NSDI*.

[55] Alvy Ray Smith. 1978. Color Gamut Transform Pairs. *SIGGRAPH Comput. Graph.* 12, 3 (Aug. 1978), 12–19.

[56] Jochen Tautges, Arno Zinke, Björn Krüger, Jan Baumann, Andreas Weber, Thomas Helten, Meinard Müller, Hans-Peter Seidel, and Bernd Eberhardt. 2011. Motion reconstruction using sparse accelerometer data. *ACM Transactions on Graphics (ToG)* 30, 3 (2011), 18.

[57] Zhao Tian, Yu-Lin Wei, Xi Xiong, Wei-Nin Chang, Hsin-Mu Tsai, Kate Ching-Ju Lin, Changxi Zheng, and Xia Zhou. 2017. Position: Augmenting Inertial Tracking with Light. (2017).

[58] David Titterton and John L Weston. 2004. *Strapdown inertial navigation technology*. Vol. 17. IET.

[59] Derya Unsal and Kerim Demirbas. 2012. Estimation of deterministic and stochastic IMU error parameters. In *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*. IEEE, 862–868.

[60] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. 2012. No Need to War-drive: Unsupervised Indoor Localization. In *Proc. of MobiSys*.

[61] Jue Wang, Deepak Vasisht, and Dina Katabi. 2014. RF-IDraw: virtual touch screen in the air using RF signals. In *Proc. of SIGCOMM*, Vol. 44. ACM, 235–246.

[62] Yu-Lin Wei, Chang-Jung Huang, Hsin-Mu Tsai, and Kate Ching-Ju Lin. 2017. CELLI: Indoor Positioning Using Polarized Sweeping Light Beams. In *Proc. of MobiSys*.

[63] Oliver J Woodman. 2007. *An introduction to inertial navigation*. Technical Report. University of Cambridge, Computer Laboratory.

[64] Oliver J. Woodman. 2010. *Pedestrian localisation for indoor environments*. Ph.D. Dissertation. University of Cambridge.

[65] Zhuoling Xiao, Hongkai Wen, Andrew Markham, and Niki Trigoni. 2014. Light-weight Map Matching for Indoor Localisation Using Conditional Random Fields. In *Proc. of IPSN*.

[66] Bo Xie, Guang Tan, and Tian He. 2015. SpinLight: A High Accuracy and Robust Light Positioning System for Indoor Applications. In *Proc. of SenSys*.

[67] Qiang Xu, Rong Zheng, and Steve Hranilovic. 2015. Idyll: indoor localization using inertial and light sensors on smartphones. In *Proc. of UbiComp*.

[68] Zhice Yang et al. 2015. Wearables Can Afford: Light-weight Indoor Positioning with Visible Light. In *Proc. of MobiSys*.

[69] Zheng Yang, Chenshu Wu, and Yunhao Liu. 2012. Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention. In *Proc. of Mobi-Com*.

[70] Hashim Yasin, Umar Iqbal, Bjorn Kruger, Andreas Weber, and Juergen Gall. 2016. A dual-source approach for 3D pose estimation from a single image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4948–4956.

[71] Chi Zhang and Xinyu Zhang. 2016. LiTell: Robust Indoor Localization Using Unmodified Light Fixtures. In *Proc. of MobiCom*.

[72] Chi Zhang and Xinyu Zhang. 2017. Pulsar: Towards Ubiquitous Visible Light Localization. In *Proc. of MobiCom*.

[73] Zhengyou Zhang. 2012. Microsoft kinect sensor and its effect. *IEEE multimedia* 19, 2 (2012), 4–10.

[74] Haisen Zhao, Lin Lu, Yuan Wei, Dani Lischinski, Andrei Sharf, Daniel Cohen-Or, and Baoquan Chen. 2016. Printed perforated lampshades for continuous projective images. *ACM Transactions on Graphics (TOG)* 35, 5 (2016), 154.

[75] Pengfei Zhou, Mo Li, and Guobin Shen. 2014. Use It Free: Instantly Knowing Your Phone Attitude. In *Proc. of MobiCom*.

[76] Shilin Zhu and Xinyu Zhang. 2017. Enabling High-Precision Visible Light Localization in Today's Buildings. In *Proc. of MobiSys*.