A Novel Inertial-Visual Heading Determination System for Wheeled Mobile Robots

Wenjun Lv, Yu Kang, Senior Member, IEEE, Yun-Bo Zhao, Senior Member, IEEE, Yu-Ping Wu, and Wei Xing Zheng, Fellow, IEEE

Abstract—Finding an alternative way to replace the magnetic compass to determine the robot heading angle indoor is always a challenge in the robotics society. This paper proposes a structurally simple yet efficient non-magnetic heading determination system, which can be used in the planar indoor environment with abundant ferro- and electro-magnetic interferences, by the combination of gyroscope and vision. The gyroscope is utilized to perceive the yaw rate, while a downward-looking camera is used to capture the pre-laid auxiliary strips to acquire the absolute angle of the robot heading. Due to the existence of pseudo measurement, varying noise statistical characteristics, and asynchronization between state propagation and measurement, the existing Kalman filters cannot be applied to fuse the gyroscopic and visual data. Therefore, a novel fusion algorithm named pseudo-measurement-resistant adaptive asynchronous Kalman filter is proposed, which is experimentally verified to be efficient in the environment with various interferences.

Index Terms—Heading angle, magnetic compass, Kalman filter, gyroscope, vision.

I. INTRODUCTION

The recent decades have witnessed the development of wheeled mobile robots and their applications in military, rescue, service, transport, warehousing, etc. One of the fundamental problems in mobile robotics is mobility which enables robots to move from location to location [1]. To realize the location-to-location transferability, the robot should be able to efficiently control its wheel rotational rates to track the planned route [2]. During the tracking control, the robot should accurately obtain its pose (i.e., the 2-dimensional coordinate and heading angle) in real time, thus forming a close-loop system [3]. The localization problem has already been intensively investigated, while there is still room in the study of heading determination system (HDS), which will be discussed in the paper.

The robot localization can be achieved by resorting to various measures like satellite, WiFi, vision, lidar, but the sensor to perceive the robot heading is almost unique, that is, magnetic compass [4]. The magnetic compass is usually realized by a magnetometer which perceives the magnetic intensity along different axes, thus it is able to provide the heading angle relative to the north. Apart from the magnetic compass, the heading angle can be determined by a localization system. In outdoor scenarios, the dual-antenna GPS is able to calculate the robot heading by processing the baseline between the main antenna and the second one. The similar heading determination system can be implemented by WiFi localization in the indoor environments. As another non-magnetic compass, one can use the vision, whether on-board or off-board, to capture the environmental or robotic markers and then determine the robot heading angle from the capture images. The state-of-the-art work related to the mentioned three HDSs follows.

A. Related Work

The magnetic compass is generally considered to originate in the ancient China as early as the 206 B.C., so it has been used as a device to find the direction since the ancient times [5]. By perceiving the geomagnetic intensity, the magnetic compass could calculate the direction with respect to the north. However, in the environment with a large body of ferro-and electro-materials, the distorted magnetic field gives rise to the inaccuracy and unreliability in magnetic compass. In order to eliminate the errors rendered by hard- and soft-iron distortion magnetic field, we can calibrate the magnetometer in advance, by means of the vector compensation or ellipsoidal calibration [6]. The time-varying and unpredictable magnetic interference, however, limits the real-time application in the heading measurement of a mobile robot. A preferred solution is integrating a gyroscope which is complementary to the magnetic compass. In [7], the regular and irregular noises of a magnetometer are filtered by a fuzzy-based compensator and the Kalman filter, respectively. Due to the insensitivity to magnetic interference, the gyroscope can be used to isolate the anomalies in the magnetometer readings, thus avoiding the pollution to heading estimation [8]. More work can be found in [9] and [10]. The aforementioned magnetometer-involved magnetic HDSs are usually used in the scenarios without serious geomagnetic distortion. However, in some scenarios,
especially the industrial environment, magnetic compass is extremely unwelcome and thus should be avoided [1].

Apart from the magnetic compass, the non-magnetic HDS has attracted more and more attention in the current decade. The location-to-heading method implemented by GPS is called GPS compass which is frequently used in outdoor vehicles [11]. The vehicle is desired to be moving as straightly as possible when using GPS compass to determine its heading, otherwise a large body of errors occur; and therefore, such a method is not reliable if the vehicle moves irregularly [12]. Meanwhile, the non-line-of-sight communication may degrade the localization system, and further reduce the accuracy of GPS compass. Another promising way to obtain the heading information is vision. With a dedicated coloured marker sticking on the robot roof, one can use the off-board vision to monitor the robot and extract the robot pose from the captured image sequence [13]–[15]. However, the method fails if the robot runs out of the visual field or a table blocks the line of sight. On the contrary, the on-board vision utilizes a camera directly-mounted on the robot to capture the landmarks distributed in the environment; and then, the robot location and heading could be worked out. The quick response codes are the most frequently used landmarks [16], which has received a great achievement in the warehousing robots [17].

**B. Contribution**

In this paper, we propose a novel non-magnetic way to determine the robot heading angle, which does not suffer from ferro- and electro-magnetic interferences. The system is illustrated in Fig. 1. The floor of the area that the robot operates should be laid with some auxiliary strips in advance. All strips are parallel to the $X$-axis and equally-spaced. The strips’ colour is optional, but complementary to the floor colour in hue. The gyroscope is mounted on the robot with its axis perpendicular to the floor, thus perceiving the pure yaw rate of the robot. An on-board downward-looking camera captures the floor images. Because all the auxiliary strips are parallel to the $X$-axis, the relative angle between the robot heading and $X$-axis could be extracted from the floor images. By adjusting the camera pose, the area of visual field could be enlarged; and therefore, it can be guaranteed that the vision could capture one auxiliary strips at least.

In order to estimate the robot heading, we first establish the relationship from the angle of the captured strips in the image framework to the heading angle in the world framework, which enable the non-magnetic way to perceive the absolute angle of the robot heading. However, the fusion of gyroscopic yaw rate and visual heading measurement is not a straightforward work by applying the Kalman filter, which is due to the following issues: 1) Because of the unknown direction of the auxiliary strips, the visual heading measurement outputs two values, a real measurement and a pseudo one. 2) Because of the camera shake and illumination variation, the noise variances, particularly the measurement noise variance, change slowly. 3) Because the visual heading measurement is time-consuming and resource-intensive, the gyroscopic state propagation proceeds at a higher rate than the visual measurement, which causes the asynchronization issue. Although many modified Kalman filters have been proposed (e.g., robust Kalman filter [18], switch Kalman filter [19], adaptive Kalman filter [20]), there dose not exist such a Kalman filter capable of solving the aforementioned three issues, to the best of our knowledge. Therefore, a novel fusion algorithm named pseudo-measurement-resistant adaptive asynchronous Kalman filter is proposed in this paper.

In the rest of the paper, the models of the gyroscopic measurement of yaw rate and the visual measurement of robot heading will be presented in Section II, and the proposed pseudo-measurement-resistant adaptive asynchronous Kalman filter applying to the inertial-visual fusion will be stated in Section III. In Section IV, the real-world experiment is exhibited to verify the effectiveness of the inertial-visual heading determination system. The paper is concluded in Section V.

**II. MODELLING OF SENSORS**

**A. Gyroscopic Measurement of Yaw Rate**

Gyroscope is a device sensing the rotational rate of a rigid body. If considering all error items, the model could be very complicated, so the model precision and complexity should be compromised. Because the systematic uncertainties could be pre-calibrated by experiments, only the non-systematic items will be considered to establish a feasible gyroscopic error model [21]. A large body of experiments have demonstrated
that the angular random walk (ARW) and angular rate random walk (ARRW) affect the gyroscopic accuracy significantly [22]. The ARW reflects the characteristics of white noise in the angular rate, so it can be seen as a fast-changing noise. Similarly, the ARRW reflects the characteristics of white noise in the angular acceleration, so it is a slow-changing noise. According to the above analysis, we have the model that relates the real yaw rate to the gyroscopic readings, that is,

$$\delta_{g,k} = \delta_k + b_{g,k} + n_{g,k},$$

(1)

where $\delta_k$ denotes the robot yaw rate, $\delta_{g,k}$ denotes the gyroscopic readings, $b_{g,k}$ and $n_{g,k}$ denote the ARRW and ARW error, respectively, all at sampling point $k$. The fast-changing noise $n_{g,k}$ is modelled as Gaussian noise, i.e., $n_{g,k} \sim \mathcal{N}(0, Q_n)$. Meanwhile, the slow-changing noise $b_{g,k}$ is modelled as a one-order Markov-Gaussian stochastic process, that is,

$$b_{g,k+1} = \gamma b_{g,k} + w_{g,k},$$

(2)

where $w_{g,k} \sim \mathcal{N}(0, Q_{w})$ is Gaussian noise, $\gamma \in [0, 1]$ denotes the correlation coefficient which indicates the correlation between two successive states. The correlation gets stronger with $\gamma$ getting larger. Because the ARRW error changes very slowly, we often set $\gamma = 1$ empirically.

**B. Visual Measurement of Heading**

The visual heading measurement is implemented by the monocular vision with the aid of floor auxiliary strips. The process contains the following steps:

1) **Image Preprocess:** This step aims to extract the central lines of the auxiliary strips. First, the coloured floor image is converted to a binary image with reserving the auxiliary strips, by using thresholding segmentation. Second, the binary image is further processed by using mathematical morphology (e.g., hole filling, skeleton extraction, burring algorithm), thus obtaining the central lines of the auxiliary strips.

2) **Line Detection:** This step aims to acquire the parameters of the auxiliary strips in the $U$-$V$ framework, by using Hough transform and perspective transform. The auxiliary strips are parameterized by

$$u \cos \rho + v \sin \rho = r,$$

(3)

where $r \in [r_{\min}, r_{\max}]$ denotes the vertical distance from the origin of the $U$-$V$ framework to the strip’s central line, and $\rho \in [-\pi/2, \pi/2]$ denotes the angle from the $U$-axis to the vertical line. For a $\mu \times \nu$ image, we have $r_{\max} = -r_{\min} = \sqrt{\mu^2 + \nu^2}$. The parameters $(r, \rho)$ can be obtained by using Hough transform.

3) **Perspective Transform:** Because the camera’s optical axis may not be vertical to the floor, we should project the captured floor image onto the floor plane. Define the perspective transformation function

$$(u, v) = \mathcal{P} \{(u, v)\},$$

(4)

where $(u, v)$ is a point in the $U$-$V$ framework, while $(\tilde{u}, \tilde{v})$ in the $\bar{U}$-$\bar{V}$ framework. Now we are in the position to convert $(r, \rho)$-lines to $(\tilde{r}, \tilde{\rho})$-lines in the $\bar{U}$-$\bar{V}$ framework. Find two points on the $(r, \rho)$-line, and calculate

$$(\bar{u}_1, \bar{v}_1) = \mathcal{P} \{(u_1, v_1)\},$$

(5a)

$$(\bar{u}_2, \bar{v}_2) = \mathcal{P} \{(u_2, v_2)\},$$

(5b)

where $(\bar{u}_1, \bar{v}_1)$ and $(\bar{u}_2, \bar{v}_2)$ are two points on the $(\tilde{r}, \tilde{\rho})$-line. Substituting them to (3) yields

$$\bar{u}_1 \cos \tilde{\rho} + \bar{v}_1 \sin \tilde{\rho} = \tilde{r},$$

(6a)

$$\bar{u}_2 \cos \tilde{\rho} + \bar{v}_2 \sin \tilde{\rho} = \tilde{r},$$

(6b)

and consequently, we have

$$\tilde{\rho} = \begin{cases} \arctan(\frac{\bar{u}_1 - \bar{u}_2}{\bar{v}_1 - \bar{v}_2}) & \text{when } \bar{v}_1 \neq \bar{v}_2, \\ -\frac{\pi}{2} & \text{when } \bar{v}_1 = \bar{v}_2, \end{cases}$$

(7)

which realizes the conversion from the $U$-$V$ framework to $\bar{U}$-$\bar{V}$ one.

4) **Heading Acquirement:** This step aims to transform $\tilde{\rho}$ to $\theta$. We analyze the situations that the robot heading $\theta$ lies in $(0, \pi/2), (\pi/2, \pi), (\pi, 3\pi/2)$, and $(3\pi/2, 2\pi)$, respectively. Situation 1 (i.e., $(0, \pi/2)$) is illustrated in Fig. 2. It is observed that the camera captures two parallel auxiliary strips which are represented by $(\bar{r}_1, \bar{\rho}_1)$ and $(\bar{r}_2, \bar{\rho}_2)$. Since all auxiliary strips are parallel to the $X$-axis, $\bar{\rho}_1$ equals $\bar{\rho}_2$ if without considering the measurement error; and therefore, it is sufficient to infer $\theta$ with the aid of only one auxiliary strip. Select one of the $(\bar{r}_1, \bar{\rho}_1)$ if there are multiple auxiliary strips, and ignore the subscript of $\bar{\rho}$. It is easy to derive

$$\theta_{(c1)} = -\tilde{\rho},$$

(8)
where $\theta_{(c1)} \in \left(0, \frac{\pi}{2}\right]$. Furthermore, for Situation 2 to 4, we have

$$
\begin{align*}
\theta_{(c2)} &= -\bar{\rho} + \pi, \\
\theta_{(c3)} &= -\bar{\rho} + \pi, \\
\theta_{(c4)} &= -\bar{\rho},
\end{align*}
$$

(9a) (9b) (9c)

where $\theta_{(c2)} \in \left(\frac{\pi}{2}, \pi\right]$, $\theta_{(c3)} \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right]$, and $\theta_{(c4)} \in \left[\frac{3\pi}{2}, 2\pi\right]$. Combining (8), (9a), (9b), and (9c) together yields

$$
\theta \in \{\theta_{(c1)}, \theta_{(c2)}, \theta_{(c3)}, \theta_{(c4)}\} = \{-\bar{\rho}, -\bar{\rho} + \pi\},
$$

(10)

and consequently, we have

$$
\theta \in \begin{cases} 
\{-\bar{\rho}, -\bar{\rho} + \pi\}, & \text{if } \bar{\rho} \in \left[-\frac{\pi}{2}, 0\right), \\
\{2\pi - \bar{\rho}, -\bar{\rho} + \pi\}, & \text{if } \bar{\rho} \in \left[0, \frac{\pi}{2}\right),
\end{cases}
$$

(11)

which can be used to derive $\theta$ from $\bar{\rho}$.

As seen from the above discussion, the visual heading measurement is time-consuming and resource-intensive, so it cannot synchronize with the gyroscopic heading estimation. As $k = 1, 2, \cdots$, we define sampling point $t_k \in \{1, 2, \cdots\}$ where $\ell = 1, 2, \cdots$. Define $\bar{\rho}$ at sampling point $t_k$ by $\bar{\rho}_{t_k}$, and the measurement of $\bar{\rho}_{t_k}$ by $\bar{\rho}_{c,t_k}$. Since the measurement contains the additive noise, we have

$$
\bar{\rho}_{c,t_k} = \bar{\rho}_{t_k} + n_{c,t_k},
$$

(12)

where $n_{c,t_k}$ denotes the measurement noise. There are many factors causing the measurement noise, such as environmental illumination, misalignment, camera shake, so the noise statistical characteristics may change with variation of these factors. For example, when the robot enters the area without sufficient light, the performance of image preprocess may degrade, thus causing a heavier noise. Most frequently, the variation is reflected in the noise variance, so we assume $n_{c,t_k} \sim \mathcal{N}(0, R_{t_k})$.

Finally, we obtain the visual heading measurement $\Theta_{c,t_k}$, which is formulated as

$$
\Theta_{c,t_k} = \begin{cases} 
\{-\bar{\rho}_{c,t_k}, -\bar{\rho}_{c,t_k} + \pi\}, & \text{if } \bar{\rho}_{c,t_k} \in \left[-\frac{\pi}{2}, 0\right), \\
\{2\pi - \bar{\rho}_{c,t_k}, -\bar{\rho}_{c,t_k} + \pi\}, & \text{if } \bar{\rho}_{c,t_k} \in \left[0, \frac{\pi}{2}\right),
\end{cases}
$$

(13)

which can be rewritten as

$$
\Theta_{c,t_k} = \begin{cases} 
\{\theta_{t_k} + n_{c,t_k}, \theta_{t_k} + \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[0, \pi\right), \\
\{\theta_{t_k} + n_{c,t_k}, \theta_{t_k} - \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[\pi, 2\pi\right),
\end{cases}
$$

(14)

It is observed that $\Theta_{c,t_k}$ contains two measurements, a real one $\Theta_{c,t_k}^{(1)}$ equalling $\theta_{t_k} + n_{c,t_k}$, and a pseudo one $\Theta_{c,t_k}^{(2)}$ equalling $\theta_{t_k} + \pi + n_{c,t_k}$. Hence, the robot heading estimation system must have the ability to recognize the pseudo-measurement and isolate it.

### III. INERTIAL-VISUAL FUSION

#### A. Problem Formulation

The inertial-visual fusion problem is based on

$$
\begin{align*}
\theta_{k+1} &= \theta_k + T \delta_k, \\
\delta_{g,k} &= \delta_k + b_{g,k} + n_{g,k}, \\
b_{g,k+1} &= b_{g,k} + w_{g,k}, \\
\Theta_{c,t_k} &= \left\{ \begin{array}{ll}
\{\theta_{t_k} + n_{c,t_k}, \theta_{t_k} + \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[0, \pi\right), \\
\{\theta_{t_k} + n_{c,t_k}, \theta_{t_k} - \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[\pi, 2\pi\right),
\end{array} \right.
\end{align*}
$$

(15a) (15b) (15c)

(15d)

Modifying (15a) to (15d) by state augmentation yields

$$
\begin{align*}
\vartheta_k &= F \vartheta_{k-1} + B \delta_{k-1} + C w_{k-1}, \\
\vartheta_{c,t_k} &= \left\{ \begin{array}{ll}
\{H \theta_{t_k} + n_{c,t_k}, H \theta_{t_k} + \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[0, \pi\right), \\
\{H \theta_{t_k} + n_{c,t_k}, H \theta_{t_k} - \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[\pi, 2\pi\right),
\end{array} \right.
\end{align*}
$$

(16a) (16b) (16c)

where

$$
\begin{align*}
\vartheta_k &= \begin{bmatrix} \theta_k \\ b_{g,k} \end{bmatrix}, \\
w_k &= \begin{bmatrix} n_{g,k} \\ w_{g,k} \end{bmatrix} \sim \mathcal{N}(0, Q), \\
Q &= \begin{bmatrix} Q_n & 0 \\ 0 & Q_w \end{bmatrix}, \\
F &= \begin{bmatrix} 1 - T & T \\ 0 & 1 - T \end{bmatrix}, \\
B &= \begin{bmatrix} T \\ 0 \end{bmatrix}, \\
C &= \begin{bmatrix} -T & 0 \\ 0 & 1 \end{bmatrix}, \\
H &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}.
\end{align*}
$$

It is observed that the measurement and state propagation are not in synchronization when $t_k \neq t_\ell$. This multi-rate problem is caused by the resource-intensive visual heading measurement. The visual heading measurement is accomplished by image acquisition and the subsequent image processing operations, so it costs much more time than the yaw rate measurement which is accomplished by a simple wired data communication with the gyroscope. For the purpose of reducing resource consumption, we may decrease the frequency of the heading estimation correction by visual heading measurement.

The synchronized model of (16a) and (16c) is

$$
\begin{align*}
\vartheta_{t_{k+1}} &= \mathcal{F}_t \vartheta_{t_k} + d_{t_k} + \omega_{t_k}, \\
\vartheta_{c,t_k} &= \left\{ \begin{array}{ll}
\{H \theta_{t_k} + n_{c,t_k}, H \theta_{t_k} + \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[0, \pi\right), \\
\{H \theta_{t_k} + n_{c,t_k}, H \theta_{t_k} - \pi + n_{c,t_k}\}, & \text{if } \theta_{t_k} \in \left[\pi, 2\pi\right),
\end{array} \right.
\end{align*}
$$

(17a) (17b) (17c)

where

$$
\mathcal{F}_t = \mathcal{F}^{n_t} = \begin{bmatrix} 1 & -\zeta_t T \\ 0 & 1 \end{bmatrix}, \\
\zeta_t = t_{k+1} - t_\ell, \\
d_{t_k} = \sum_{i=t_k}^{t_{k+1}-1} F^{t_{k+1}-i-1} B \delta_i, \\
\omega_{t_k} = \sum_{i=t_k}^{t_{k+1}-1} F^{t_{k+1}-i-1} C w_i.
$$

It is a 2-dimensional Gaussian noise with the covariance $Q = \sum_{i=t_k}^{t_{k+1}-1} F^{t_{k+1}-i-1} C Q(C^{T} F^{t_{k+1}-i-1})$.

The problem is to develop a inertial-visual heading estimation algorithm based on the gyroscopic readings $\{\delta_k\}$ and visual measurements $\{\vartheta_{c,t_k}\}$. It cannot be easily solved by using a standard Kalman filter, because of the following three issues: the existence of pseudo measurement, varying noise statistical characteristics, and asynchronization between state propagation and measurement.
B. Estimation Algorithm

The estimation algorithm contains 6 steps: initialization, a-priori estimation, visual measurement trigger, real measurement recognition, a-posterior estimation, and noise covariance estimation.

1) Initialization: The initial a-posterior estimation \( \hat{\theta}_0 \) should be determined at first, where \( \hat{\theta}_0 = [\theta_0, \dot{\theta}_0, 0]' \). The initial heading \( \hat{\theta}_0 \) is measured artificially, and the initial gyroscopic bias is obtained by averaging the gyroscopic readings in stationary state. The initial a-posterior error correlation \( \hat{P}_0 \) is initialized by a diagonal matrix with appropriate values.

2) A-Priori Estimation: Suppose that \( \hat{\theta}_{k-1} \) has been obtained, then the a-priori estimates are achieved by

\[
\begin{align}
\hat{\theta}_k &= F \hat{\theta}_{k-1} + B \delta_{k-1}, \\
\hat{P}_k &= F \hat{P}_{k-1} F' + C Q C' ,
\end{align}
\]

where \( \hat{\theta}_k \) and \( \hat{\theta}_k \) denotes the a-priori and a-posterior estimations of \( \theta_k \), respectively. \( \hat{P}_k \) and \( \hat{P}_k \) denotes the a-priori and a-posterior estimation error covariances at sampling point \( k \), respectively.

3) Visual Measurement Trigger: The gyroscopic bias (i.e., ARRW) is included in the augmented state as shown in (16a), so it can be estimated and removed in the a-priori estimation process, which means that the gyroscopic heading estimation is relatively accurate in short term. Hence, to reduce the computational burden, the a-priori estimation is not necessary to be corrected at each sampling point, until the error variance of heading estimation gets larger than a given bound. The visual measurement is triggered at

\[
\hat{P}_k < P \text{ and } H(F \hat{P}_k F' + C Q C') H' + R < (\frac{\pi}{2\beta})^2
\]

(19)

where \( \beta \) is an integer not less than 3. The function of the first inequation in (19) is to limit the heading estimation error by the upper bound \( \hat{P} \), while the second inequation prevents the real measurement at the next sampling point from being unrecognizable. The discussion of the recognizable condition of real measurement can be found in Appendix A. According to [23], \( \hat{P}_t \) and \( \hat{P}_c \) converge to time-invariant matrices for a completely-observable and completely-controllable system. Therefore, the event-based trigger designed above will trigger the visual measurement in equal intervals of time after convergence.

4) Real Measurement Recognition: If the visual heading measurement is triggered at \( \ell_i \), the real measurement should be recognized. Calculate the \( \theta_{c,k} \) and \( \theta_{c,k}^{(2)} \) based innovations \( \phi^{(1)} \) and \( \phi^{(2)} \) by

\[
\begin{align}
\phi^{(1)} &= \theta_{c,k}^{(1)} - H \hat{\theta}_k, \\
\phi^{(2)} &= \theta_{c,k}^{(2)} - H \hat{\theta}_k,
\end{align}
\]

(20)

where \( \theta_{c,k}^{(1)} \) and \( \theta_{c,k}^{(2)} \) are the elements in \( \Theta_{c,k} \). The real measurement can be picked out by

\[
\theta_{c,k}^{(b)} = \begin{cases} 
\theta_{c,k}^{(1)}, & \text{if } |\phi^{(1)}| < |\phi^{(2)}|, \\
\theta_{c,k}^{(2)}, & \text{otherwise},
\end{cases}
\]

(21)

where \( \theta_{c,k}^{(b)} \) denotes the recognized real measurement. In Appendix A, we show the recognizable condition of the real measurement, in which \( \theta_{c,k}^{(b)} = \theta_{c,k}^{(1)} \) holds in an extreme high probability.

5) A-Posterior Estimation: If the visual measurement is not triggered, which means that the state propagation and measurement are not synchronized, then the a-posterior estimates equals the a-priori estimates, that is,

\[
\begin{align}
\hat{\theta}_k &= \hat{\theta}_k, \\
\hat{P}_k &= \hat{P}_k,
\end{align}
\]

(22a, 22b)

otherwise, the a-posterior estimate is given by

\[
\begin{align}
K_k &= \hat{P}_k H'[H \hat{P}_k H' + \hat{R}_k]^{-1}, \\
\hat{\theta}_k &= \hat{\theta}_k + K_k [\theta_{c,k}^{(b)} - H \hat{\theta}_k], \\
\hat{P}_k &= [I_2 - K_k H] \hat{P}_k,
\end{align}
\]

(23a, 23b, 23c)

where \( K_k \) denotes the filtering gain at time \( k \), and \( I_2 \) denotes the 2-dimensional unit matrix.

6) Noise Covariance Estimation: Define another time s-tamp \( T_{\ell} \) where \( \tau \) is set as 0 initially. When \( \varsigma_\ell = \varsigma_{\ell-1} \) and \( \varsigma_{\ell-3} = \varsigma_{\ell-4} \), we increase \( \tau \) by 1, calculate \( \nabla \Theta_{c,t}^{(s)} = \theta_{c,t}^{(s)} - 2 \theta_{c,t-1}^{(s)} + \theta_{c,t-2}^{(s)} - T \delta_{g,t-1} + T \delta_{g,t-2} \) where \( \delta_{g,t-1} = \sum_{i=t-1}^{\varsigma_{t-1}} \delta_{g,i} \), and assign the value of \( \nabla \Theta_{c,t-1}^{(s)} \nabla \Theta_{c,t-2}^{(s)} \) to \( T_{\ell} \). Finally, the measurement noise variance is estimated by

\[
\hat{R}_{\ell} = \frac{1}{n(T_{\ell} - T_{\ell-1})} (T_{\ell} - T_{\ell-1})
\]

(24)

where \( \tau > n \). At the starting phase, there may not be a sufficient number of \( T_{\ell} \), i.e., \( \tau < n \), thus the measurement noise variance can be \( \hat{R}_{\ell} = \frac{1}{\tau} \sum_{i=1}^{\tau} T_i \). The derivation of (24) can be found in Appendix B. Furthermore, it is noted that \( R_{\ell} \) should not be introduced to the filter at each \( T_{\ell} \) for the reason that a varying \( R_{\ell} \) leads to a varying \( \varsigma_\ell \), and thus causing a smaller number of \( T_{\ell} \).

IV. EXPERIMENTAL VERIFICATION

A. Experiment Setup

As shown in Fig. 3, the experimental robot (Turtlebot3 Burger) is mounted with a camera looking downward to the...
floor which is covered with red parallel auxiliary strips. All these auxiliary strips are equally spaced by 0.5 meter, and parallel to the X-axis. Other sensors, like gyroscopes, are embedded on the circuit board. The key information of the sensors is shown in Table I.

The implementation of the proposed system is not demanding to have expensive dedicated gyroscope or camera. As an example, we use a customer-grade MEMS gyroscope, the type of which is MPU6050 (manufactured by InvenSense, less than 1 US dollar per chip), to perceive the yaw rate. Additionally, we use a general-purpose full-colour camera, the resolution of which is 640×480 (less than 10 US dollar per piece), to capture the floor image. All sensors work at 10 Hz.

B. Evaluation of the Proposed HDS

1) Pseudo Measurement Isolation: The estimation results of the proposed HDS within the first 500 sampling points are exhibited in Fig. 4. The initial setup follows: \( \theta_0 = [90.5, 1.7]' \), \( Q = \text{diag}(0.38^2, 0.05^2) \), \( R = 1.4^2 \). It is observed that two measurements (the real one and pseudo one) exist, but the pseudo measurement can be isolated correctly all the time; and consequently, the heading estimation tracks the truth with high accuracy. As shown in Fig. 4a, the visual heading measurement is triggered by 25 times during the 500 sampling points, and the heading estimation achieves the smallest root mean square error (RMSE) of 0.98 deg. As shown in Fig. 4b and Fig. 4c, the period without visual correction gets longer with \( \tilde{P} \) increasing, thus causing larger RMSEs. It can also be observed, however, that the pseudo measurements are always outside the range of innovation boundary \(-3(H \tilde{P}_t H' + R_{tt})^{1/2}, +3(H \tilde{P}_t H' + R_{tt})^{1/2}\), while the real ones inside, provided that the visual heading measurement is triggered in time. Finally, observing the visual measurements (blue crossings), their intervals get larger and remain fixed over time for the filtering convergence.

2) Noise Variance Estimation: At the sampling point of about 11000, we dim the light and loose the screws that fix the downward-looking camera, thus causing an increase of the noise variance of the visual heading measurement. The initial \( \tilde{P} = 2^2, R = 1.4^2 \), and the sliding window width of \( R_{tt} \) equals 500. If the noise variance \( R \) is not adjusted, then the heading estimation RMSE is 2.71 deg. After introducing the noise variance estimator, the RMSE is reduced to 1.57 deg. As shown in Fig. 5, the estimated \( R \) converges to 2 before the environmental variation, and goes larger soon afterwards. It is also interesting to observe that the visual heading measurement is triggered more and more frequently until \( R_{tt} \) converges.

Fig. 4: Estimation results of the proposed HDS. The black curves (−) stand for the truth values, the red curves (−−) stand for the estimates, the blue crossings (+) stand for the measurements, and the gray areas stand for the innovation boundaries.

Fig. 5: Estimation results of the proposed HDS with noise variance adjustment. (a) RMSE = 2.57 deg. (b) RMSE = 1.56 deg. (c) RMSE = 0.98 deg.

if \( R \) increases and the HDS works at a low correction rate, then the heading estimation may be invalid due to the confusion

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyroscope</td>
<td>MPU6050, range: ±250 deg/s, initial ZERO° tolerance: ±5 deg/s, total RMS° noise: 0.1 deg/s.</td>
</tr>
<tr>
<td>Gyroscope</td>
<td>ADXRS453, range: ±400 deg/s, accuracy: ±0.4 deg/s.</td>
</tr>
<tr>
<td>Magnetometer</td>
<td>AK8963, range: ±4800 uT.</td>
</tr>
<tr>
<td>Camera</td>
<td>R.MONCAM, normal lens; resolution: 480×640; rate: 30 fps.</td>
</tr>
</tbody>
</table>

* Zero-Rate Output; ** Root Mean Square.
in distinguishing the real and pseudo measurements. In our design, however, an increasing $R$ leads to a more frequent visual correction, thus accelerating the convergence of $\hat{R}_{t}^{\beta}$.

C. Comparative Study

A magnetic HDS is composed of a gyroscope and a magnetic compass (usually a magnetometer), and outputs the heading estimation by fusing the readings of the two complementary sensors. In [8], an advanced gyroscope-magnetometer integrated algorithm is proposed. The algorithm could recognize and isolate the ferro- or electro-interference-induced anomalies in magnetic compass. The comparative results are shown in Fig. 6. In the case without interference, the magnetic compass performs in a stable and accurate manner as the up-left subfigure exhibited. As the down-left subfigure exhibited, if the interference incurs a sudden variation to the magnetic compass, then these anomalies could be recognized accurately and isolated, thus the heading estimation is not affected. However, the bias is small or turns to larger in a gradual way, the anomalies cannot be recognized. Because the Kalman filter tracks the measurements, the unrecognized anomalies give rise to the failure of magnetic heading estimation, exactly as the up- and down-right subfigures exhibited. In conclusion, the ferro- and electro-magnetic interferences cause the potential performance instability to a magnetic HDS, but do not affect the proposed non-magnetic one.

V. CONCLUSION

In this paper, a structurally simple yet efficient non-magnetic heading determination system has been developed, which can be used in the planar indoor environment with abundant ferro- and electro-magnetic interferences, by the combination of gyroscope and vision. As the experiments shown, such a visual-inertial heading determination system, coupled with the proposed pseudo-measurement-resistant adaptive asynchronous Kalman filter, could isolate the pseudo measurement accurately, and adapt to the environmental variation by estimating the measurement noise variance. Furthermore, as a non-magnetic way to determine the robot heading, the developed system provides an effective alternative solution in the environments with abundant magnetic interferences.

APPENDIX A

Define the innovations based on $\theta_{c,t}^{(1)}$ and $\theta_{c,t}^{(2)}$ by

$$\phi_{t}^{(1)} = \theta_{c,t}^{(1)} - H \hat{\theta}_{t},$$

$$\phi_{t}^{(2)} = \theta_{c,t}^{(2)} - H \hat{\theta}_{t} = \pm \pi + \phi_{t}^{(1)},$$

which are Gaussian noises. $\phi_{t}^{(1)}$ is a Gaussian white noise, i.e., $\phi_{t}^{(1)} \sim \mathcal{N}(0, H \hat{P}_{H} H' + R_{t})$, and $\phi_{t}^{(2)}$ is a coloured one, i.e., $\phi_{t}^{(2)} \sim \mathcal{N}(\pm \pi, H \hat{P}_{H} H' + R_{t})$. Obviously, $\phi_{t}^{(1)}$ and $\phi_{t}^{(2)}$ are not independently and identically distributed, but differ by an offset $\pm \pi$. If it is desired that $|\phi_{t}^{(1)}| < |\phi_{t}^{(2)}| = |\pm \pi + \phi_{t}^{(1)}|$ holds, then $|\phi_{t}^{(1)}| \leq \frac{\pi}{2}$ is expected. Since $\phi_{t}^{(1)} \sim \mathcal{N}(0, H \hat{P}_{H} H' + R_{t})$, $|\phi_{t}^{(1)}| \leq \frac{\pi}{2}$ cannot hold always, but can hold in an extremely high probability. For example, if $\frac{\pi}{2} \geq 3 \sqrt{H \hat{P}_{H} H' + R_{t}}$, then $\mathcal{P}(|\phi_{t}^{(1)}| \leq \frac{\pi}{2}) \geq 99.73\%$, which means that the real measurement derived innovation is greater than the pseudo one in a probability of 99.73%. The probability can be increased by increasing the multiple of $\sqrt{H \hat{P}_{H} H' + R_{t}}$. Therefore, the recognizable condition is $\frac{\pi}{2} \geq \beta \sqrt{H \hat{P}_{H} H' + R_{t}}$, where $\beta > 0$ is a parameter which is generally set as an integer greater than 3.

APPENDIX B

The difference series of $\theta_{c,t}^{(i)}$ contain $w_{g,k}$, $n_{g,k}$ and $n_{c,t}$, which enables the estimation of noise variances. Consider the
following equations
\[ \theta_{t_{\ell+1}} = \theta_{t_{\ell}} + T \sum_{i=t_{\ell}}^{t_{\ell+1}-1} (\delta_{g,i} - b_{g,i} - n_{g,i}), \]  
(26a)
\[ \theta_{c,t_{\ell+1}}^{(b)} = \theta_{t_{\ell+1}} + n_{c,t_{\ell+1}}, \]  
(26b)
\[ \theta_{c,t_{\ell}}^{(b)} = \theta_{t_{\ell}} + n_{c,t_{\ell}}, \]  
(26c)
and define \( \bar{g}_{g,t_{\ell}} = \sum_{i=t_{\ell}}^{t_{\ell+1}-1} \delta_{g,i} \), \( \bar{b}_{g,t_{\ell}} = \sum_{i=t_{\ell}}^{t_{\ell+1}-1} b_{g,i} \), and \( \bar{n}_{g,t_{\ell}} = \sum_{i=t_{\ell}}^{t_{\ell+1}-1} n_{g,i} \). Subtracting (26b) from (26c) yields
\[ \theta_{c,t_{\ell+1}}^{(b)} - \theta_{c,t_{\ell}}^{(b)} = \theta_{t_{\ell+1}} + n_{c,t_{\ell+1}} - \theta_{t_{\ell}} - n_{c,t_{\ell}}, \]  
(27)
which includes the gyroscopic and visual error terms. For removing \( \bar{b}_{g,t_{\ell}} \), the backward difference operator \( \nabla \) is introduced. If \( \varsigma_{t_{\ell}} = \varsigma_{t_{\ell}-1} \), then we have
\[ \nabla \Theta_{c,t_{\ell+1}}^{(b)} = \Theta_{c,t_{\ell+1}}^{(b)} - \Theta_{c,t_{\ell}}^{(b)} = \nabla \left( \theta_{t_{\ell+1}} - 2 \theta_{t_{\ell}} + \theta_{t_{\ell-1}} \right) = n_{c,t_{\ell+1}} - 2 n_{c,t_{\ell}} - n_{c,t_{\ell-1}}, \]  
(28)
where \( n_{c,t_{\ell+1}} = \sum_{i=t_{\ell}}^{t_{\ell+1}-1} n_{c,i} \). It is observed that (29) is the sum of three moving average stochastic processes \( \bar{w}_{g,t_{\ell+1}}, \bar{w}_{g,t_{\ell}}, \) and \( n_{c,t_{\ell+1}} - 2 n_{c,t_{\ell}} - n_{c,t_{\ell-1}} \), which are expressed in the left- and right-hand sides of (29) yields
\[ A = \mathbb{E} \left[ \nabla \Theta_{c,t_{\ell}}^{(b)} \cdot \nabla \Theta_{c,t_{\ell-2}}^{(b)} \right] = R, \]  
(30)
where \( R \) denotes the mathematical expectation. Because \( \nabla \Theta_{c,t_{\ell}}^{(b)} \) only holds at \( \varsigma_{t_{\ell}} = \varsigma_{t_{\ell}-1} \), \( \nabla \Theta_{c,t_{\ell}}^{(b)} \), and \( \nabla \Theta_{c,t_{\ell-2}}^{(b)} \) only exists at \( \varsigma_{t_{\ell}} = \varsigma_{t_{\ell}-1} \) and \( \varsigma_{t_{\ell}-3} = \varsigma_{t_{\ell}-4} \). Define another time stamp \( t_{\ell} = \{t_{\ell}, \ell = 1, 2, \ldots \} \), where \( \tau = 1, 2, \ldots \). When \( \varsigma_{t_{\ell}} = \varsigma_{t_{\ell}-1} \) and \( \varsigma_{t_{\ell}-3} = \varsigma_{t_{\ell}-4} \), we increase \( \tau \) by 1, assign the value of \( t_{\ell} \) to \( t_{\ell+1} \), and let \( \Delta t_{\ell} = \nabla \Theta_{c,t_{\ell}}^{(b)} \cdot \nabla \Theta_{c,t_{\ell-2}}^{(b)} \). Finally, the measurement noise variance is estimated by
\[ \hat{R}_{t_{\ell}} = \frac{1}{n} \sum_{i=t_{\ell}^*}^{t_{\ell}} I_{t_{\ell}} = \frac{1}{n} \left( \sum_{i=t_{\ell}^*}^{t_{\ell}^*-1} I_{t_{\ell}} + I_{t_{\ell}} - I_{t_{\ell}^*-1} \right), \]  
(31)
when \( \tau > n \); otherwise, \( \hat{R}_{t_{\ell}} = \frac{1}{\tau} \sum_{i=t_{\ell}^*}^{t_{\ell}^*} I_{t_{\ell}} \).

REFERENCES