Survey of Accuracy Improvement Approaches for Tightly Coupled ToA/IMU Personal Indoor Navigation System

Vladimir Maximov  
LLC RTLS  
Moscow, Russia  
Email: vmaximov@rtlsnet.ru

Oleg Tabarovsky  
LLC RTLS  
Moscow, Russia  
Email: olg@rtlsnet.ru

Abstract—In this work we present the personal indoor navigation system based on ranges measurements with Time of Arrival (ToA) principle and Inertial Measurement Unit (IMU). Survey of accuracy improvement approaches including monocular camera Simultaneous Localization an Mapping (SLAM) and WiFi SLAM is provided. Provided experimental results show that integration of data from various navigation systems with different physical principles can increase the accuracy and robustness of the overall solution.

Index Terms—PDR; IMU; ToA; Monocular SLAM; Vector Field SLAM;

I. INTRODUCTION

Indoor real-time locating systems (RTLS) are spread widely nowadays, and use various physical layers - from RF to acoustic and infrared. RTLS system, developed in our company employs RF ToA range measurements between mobile receivers - tags, and stationary base stations - anchors. It can provide steady solution with 1 meter accuracy 80% of a time. But it is known that indoor RF range-measuring systems are suffering from NLOS measurements, another limitation comes from measurements update rate that is about 1 Hz. In order to provide more smooth and robust updates, other sources of navigation information should be used. For this work we choose only those sources, that are relatively autonomous - inertial, visual, field strengths, with more autonomous and ubiquitous navigator in mind. This paper has following structure: first part describes tightly coupled PDR navigator, second part is devoted to inertially augmented monocular SLAM, and third part gives some results on WiFi SLAM.

II. TIGHTLY COUPLED TOA/IMU NAVIGATOR

A. Pedestrian Dead Reckoning

There is a lot of reports on pedestrian navigation systems that use inertial sensors - from foot-mounted, with full 3D starupdown INS [1] to 2D strapdown INS attached to a pedestrian body [2]. In our system we use practical solution, where pedestrian dead reckoning (PDR) navigator uses velocity \( \dot{V}_b \), that is determined by estimating step frequency using accelerometer signals. Angle \( \dot{\Psi} \) defines the rotation of body frame with respect to navigation frame. Fig. 1 shows frames used for navigation system, where: \( X, Y \) - navigation frame (n-frame) axes, \( X_p, Y_p \) - pedestrian frame (p-frame) axes, \( X_b, Y_b \) - body frame (b-frames) axes, \( \dot{\Psi} \) - heading angle, \( \delta \dot{\Psi} \) - heading angle misalignment. Fig. 2 shows the functional diagram of tightly-coupled ToA/IMU navigator, where: \( \ddot{a}_b \) - acceleration vector in b-frame; \( \ddot{m}_b \) - magnetic field vector in b-frame; \( \dot{\Psi}_{AHRS} \) - AHRS heading angle; \( V_p \) - velocity in p-frame; \( \delta \dot{V}_p \) - estimated p-frame velocity error; \( \delta \dot{\Psi} \) - estimated heading error; \( \ddot{R}_n \) - n-frame coordinates; \( \delta \ddot{R} \) - estimated n-frame coordinates error; \( \ddot{R}_n \) - corrected coordinates in n-frame; \( \dot{R}_{CSS} \) - measurements of ranges by RF chirp spread spectrum (CSS) ToA system. Velocity absolute value is calculated by multiplying inverse of counter value by experimentally estimated scale-factor:

\[
V_p = \frac{S_v}{S_{cnt}}
\]

(1)
where $S_{cnt}$ - counter value; $S_v$ - step scale-factor. Heading angle is estimated by AHRS (attitude and heading reference system) that fuses data from inertial sensors and vector magnetometer in 15-state Kalman filter.

### B. PDR Errors Correction

Pedestrian velocity and heading angle can be used to calculate coordinates in n-frame, drift will inevitably occur. Drift is caused by magnetic field disturbances that are highly probable indoors; velocity derived form step frequency is also prone to various errors, ranging from false step detection to person-varying scale factor in (1). Tightly-coupled ToA/IMU prone to various errors, ranging from false step detection to probable indoors; velocity derived form step frequency is also drift is caused by magnetic field disturbances that are highly out short-time magnetic disturbances, long-time disturbances still pose the problem. It is known that monocular camera SLAM algorithm can provide information on camera attitude, velocities and coordinates [3], [4]. Typically 30 frames per second (fps) camera rate is used to make smooth tracking possible. In order to make algorithm more suitable for mobile platforms, data from gyroscopes was used on prediction step, and accelerometer data on correction step - so it was possible to reduce camera frame rate to 10 fps. Monocular SLAM augmented with inertial data is based on EKF with dynamically changing state vector. State vector $\bar{x}$ includes camera state $\bar{x}_c$ and number of features states $\bar{x}_f$. Two different modes of monocular SLAM were tested - compass mode - where only the attitude of the camera being estimated, and full 6-D mode - where attitude is estimated alongside with coordinates and velocities of the camera in starter frame:

$$\bar{x} = [\bar{x}_c \quad \bar{x}_f \ldots \bar{x}_f]^T$$  

#### 1) System model: It can be easily shown that linearized PDR errors dynamic equation can be written as:

$$\dot{\delta \tilde{R}} = \begin{bmatrix} -V^v_x \\
V^v_y \\ 
\cos \Psi \\
\sin \Psi
\end{bmatrix} \delta \Psi + \begin{bmatrix}-V^v_x \\
V^v_y \\
\cos \Psi \\
\sin \Psi
\end{bmatrix} \delta S_v$$

where $V^v_x, V^v_y$ - velocity components in n-frame.

It gives following system matrix $F$ for discrete EKF:

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\
-V^v_0 \Delta t & 1 & 0 & \cos \Psi \Delta t \\
V^v_0 \Delta t & 0 & 1 & \sin \Psi \Delta t \\
0 & 0 & 0 & 1
\end{bmatrix}$$

where $\Delta t$ - sampling period.

#### 2) Measurement model: Range measurement, delivered by CSS ToA system can be written as a function of current position $x, y$ and known base stations coordinates $X_i, Y_i$:

$$r_i = h(x) = \sqrt{(X_i - x)^2 + (Y_i - y)^2}$$

Measurement vector $z$ is formed as a difference between predicted and measured ranges:

$$z = [r_1^{PDR} - r_1^{CSS} \ldots r_n^{PDR} - r_n^{CSS}]^T$$

#### C. Filtering and experimental results

As system and measurement models are defined, the standard set of discrete Joseph-form EKF equations was applied to get the estimated errors values. Experimental setup included CSS tag paired with custom-built inertial module. Data was acquired in a typical office environment. Inertial sensors data was sampled with 20 Hz rate together with 1 Hz CSS ToA ranges. It can be seen from Fig. 3 and 4 that tightly coupled system can effectively cope with various misalignments and step scale-factors, continuously adapting to the pedestrian. Variations of misalignment angle plots on Fig. 3 are due to external magnetic distortions.

![Fig. 3. Misalignment angle estimation](image)

![Fig. 4. Step scale factor multiplier estimation](image)

### III. USING MONOCULAR SLAM FOR PDR ACCURACY IMPROVEMENT

Tightly integrated scheme described above, can effectively fuse the data, compensate PDR errors and smooth CSS ToA ranges measurements. But it also has several drawbacks, and reliance on magnetic heading is one of them. While AHRS combines the gyroscope and magnetometer data and can filter out short-time magnetic disturbances, long-time disturbances still pose the problem. It is known that monocular camera SLAM algorithm can provide information on camera attitude, velocities and coordinates [3], [4]. Typically 30 frames per second (fps) camera rate is used to make smooth tracking possible. In order to make algorithm more suitable for mobile platforms, data from gyroscopes was used on prediction step, and accelerometer data on correction step - so it was possible to reduce camera frame rate to 10 fps. Monocular SLAM augmented with inertial data is based on EKF with dynamically changing state vector. State vector $\bar{x}$ includes camera state $\bar{x}_c$ and number of features states $\bar{x}_f$. Two different modes of monocular SLAM were tested - compass mode - where only the attitude of the camera being estimated, and full 6-D mode - where attitude is estimated alongside with coordinates and velocities of the camera in starter frame:
State vector for the 6-D mode:
\[
\bar{x}_c = [\bar{R}_c \ \bar{q} \ \delta \bar{\omega} \ \bar{v}]^T \quad \bar{x}_f = [\bar{R}_f \ \theta \ \phi \ \rho]^T
\]  
(8)
where \(\bar{q}\) - camera attitude quaternion, \(\bar{R}_c\) - camera coordinates vector in starter frame, \(\bar{v}\) - camera velocity vector in starter frame, \(\delta \bar{\omega}\) - gyroscopes biases vector, \(\bar{R}_f, \theta, \phi, \rho\) - standard inverse depth parametrization model of visual features [4]. For the compass mode \(x_c\) will contain only \(\bar{q}\) and \(\delta \bar{\omega}\) terms, and \(x_f\) - only \(\theta\) and \(\phi\).

A. System model

System state model for 6-D case can be written as:
\[
\begin{bmatrix}
\dot{\bar{R}}_c \\
\dot{\bar{q}} \\
\delta \dot{\bar{\omega}} \\
\dot{\bar{v}}
\end{bmatrix} = f(\bar{x}, \dot{\bar{x}}) = \begin{bmatrix}
\bar{v} \\
1/2\Omega \bar{q}
\end{bmatrix}
\]  
(9)
where \(\Omega\) - is a a following skew-symmetric matrix:
\[
\Omega = \begin{bmatrix}
0 & -\omega_z + \delta \omega_x & -\omega_y + \delta \omega_z & -\omega_z + \delta \omega_z \\
\omega_x - \delta \omega_x & 0 & -\omega_z + \delta \omega_z & -\omega_y + \delta \omega_y \\
\omega_y - \delta \omega_y & -\omega_x + \delta \omega_z & 0 & -\omega_x + \delta \omega_x \\
\omega_z - \delta \omega_z & \omega_y - \delta \omega_y & \omega_x - \delta \omega_x & 0
\end{bmatrix}
\]
Gyroscope biases and camera velocity are modeled as random walk processes with correspondent noises \(\bar{v}_\omega\) and \(\bar{v}_v\). For the compass case only \(\bar{q}\) and \(\delta \bar{\omega}\) components of the model are used [5].

B. Measurement model

Monocular SLAM measurement model with pin-hole camera was augmented with accelerometer data, which gives the system an ability to keep the local vertical. Normalized acceleration vector, measured in body frame is transformed to navigation frame, where it is compared with normalized local gravity vector:
\[
\| \bar{a}_n \| = C_b^n (\bar{q}) \| \bar{a}_b \|
\]  
(10)
where \(C_b\) - direction cosines matrix; \(\|\bar{a}_b\|\) - normalized acceleration vector in b-frame; \(\|\bar{a}_n\|\) - estimated acceleration vector in n-frame; Acceleration measurement is gated by measured acceleration absolute value - only undisturbed acceleration vectors are used. Ranges from CSS ToA system were also added to the measurement model for the 6-D case.

C. Experimental results

The experimental setup included BeagleBone board with 320x240 web-camera, custom-built AHRS module, CSS ToA tag. Data from camera was sampled at 10 fps rate, inertial data at 100 Hz rate, and CSS ToA data at 1 Hz rate.

1) Compass mode: Pedestrian heading angle can be evaluated with range-only system while pedestrian is moving, but it is ambiguous for the stationary case. Magnetic field disturbances, especially in industrial areas will finally distort the AHRS readings also. To address this problem, ceiling-looking camera was used as a source of heading angle information. Two plots of Fig. 5 show heading angle behavior in gyro-only mode (magnetic correction was switched off) - \(\Psi_w\) and heading angle \(\Psi_{SLAM}\), provided by monocular SLAM, augmented with inertial data. It is clear that monocular SLAM helps to eliminate heading drift considerably.

2) 6-D mode: For this mode pedestrian was equipped with forward-looking hand-held camera, AHRS and CSS ToA tag. Standard office environment proved to be very difficult place for pedestrian monocular SLAM, as landmarks set is changing fast and features base can quickly grow more then a 100, nevertheless, for short periods of time, while pedestrian is located in the same room, monocular SLAM can provide good support to PDR navigator. Fig. 7 shows the estimated trajectory.
IV. WiFi SLAM

Received signal strength information (RSSI) of WiFi signals is another ubiquitous source of navigation information. Additional reason to take that information into account is that WiFi as well as inertial sensors and monocular camera are at the core of any modern mobile platform such as smartphone. RSSI has long been used for navigation purposes with the fingerprinting approach as a basic one. One of the approaches, that enables to collect the RSSI data without need for fingerprinting, is the approach, called Vector Field SLAM (VFSLAM) [6]. Being the standard EKF-SLAM with dynamically changing state vector, that consists of vehicle part and a number of features parts, it is original in a way of representing RSSI surface. Features represent the RSSI levels for the regular square grid corners. RSSI value inside the current grid is calculated by bilinear interpolation.

A. WiFi SLAM simulation and test results

Matlab simulation of VFSLAM was performed to evaluate it’s effectiveness in well-controlled environment. Gaussian processes [7] were used to approximate random surfaces to simulate signal strengths for three base stations. Grid size was chosen to be 1m. Fig. 8 shows the result of simulation. Green and blue and red surfaces represent reference fields, while same color markers correspond to field values, estimated by VFSLAM algorithm. It can bee seen that estimated values are pretty close to reference values, so the next step is to evaluate the algorithm with real-life data.

To evaluate performance of VFSLAM in real life situation, data from PDR and WiFi RSS collected in office environment was fused by VFSLAM with following state vector:

$$\bar{x} = [x \ y \ \Psi \ m_1 \ \ldots \ m_n]^T$$ (11)

where $m_1 \ldots m_n$ - RSSI levels at regular grid corners.

Pedestrian velocity and angular rate around vertical axis (derived from AHRS and corrected with gyroscope bias estimation) are used to propagate the state forward, while WiFi RSS measurements, taken with 0.5Hz rate served as the only external information. Grid size was chosen to be 5 meters. Fig. 9 shows two calculated pedestrian paths, and it can be easily seen that even for rough 5-meter grid, WiFi VFSLAM delivers much robust results than PDR alone.

V. Conclusion

In this paper authors propose practical approach to indoor pedestrian navigation which combines several different sources of navigation information. Survey of accuracy improvement approaches is provided, with emphasis to usefulness of the system on mobile smartphone-like platform. There is a lot of open questions remain - both from practical and from theoretical standpoint - how to optimally fuse different frameworks - WiFi, visual and inertial, switch from one mode of navigation to the other, and so on. In our ongoing research activity we use BeagleBone as the prototype platform, and the next step will be the integration of all surveyed approaches onto this compact platform to get fully-functioning prototype.

REFERENCES