COLLABORATIVE ENVIRONMENT LEARNING: THE KEY TO LOCALIZATION OF SOLDIERS IN URBAN ENVIRONMENTS

Shahram Moafipoor¹, Dr. Lydia Bock¹, Dr. Jeffrey A. Fayman¹, Thomas J. MacDonald², Dr. Gerry Mader¹, Michael Strong¹

¹Geodetics Inc., La Jolla, California
²MacroVision, Reading, Massachusetts

Key Words: Multi-sensor integration, Collaborative navigation, Environment learning.

ABSTRACT: Several navigation technologies exist, which can facilitate the generation of Time Space Positioning Information (TSPI) in urban environments. These include GPS, image-based localization, radio-based localization and Dead Reckoning. This paper first presents a basic overview of these techniques including advantages and limitations of each. We present an approach to localization in urban environments, based on environment learning and collaborative navigation using multiple homogeneous and non-homogeneous localization technologies, fused to form a multi-sensor system.

1. INTRODUCTION

The U.S. military has an immediate need for a robust Joint Urban Operations (JUO) testing and training capability. An important component of such a capability is accurate Time Space Positioning Information (TSPI) for combat soldiers operating in urban environments. However, urban environments present substantial physical occlusions and electromagnetic anomalies restricting the use of GPS traditionally used for localization. Urban environments make tracking, identification, and localization of friendly and hostile forces difficult. Further, soldiers maneuver in these environments with a variety of dynamics adding complexity to the problem.

It is becoming increasingly apparent that no-single navigation technology will emerge in the near term to fulfill the most central requirements in urban environments (Fax and Volk, 2007). Recent research has focused on novel integration of existing sensors and technologies to satisfy the requirements. In this context, the generation of accurate TSPI in modern urban combat requires careful selection of navigation sensors and localization approaches. In this paper we provide a basic overview of several techniques useful for generating TSPI in urban environments, highlighting the advantages and limitations of each. Next, we present a multi-sensor fusion architectural framework and present two new technical concepts: collaborative navigation and environment learning, which are added to the architectural framework. This new infrastructure is likely to address the immediate concerns of military urban navigation requirements.

2. NAVIGATION TECHNOLOGIES

Several navigation techniques exist, predominantly in research or prototypes, which combine a variety of sensors to tracking a soldier operating in urban environments. These techniques can be divided into four main areas, as listed in Table 1.
<table>
<thead>
<tr>
<th>Positioning system</th>
<th>Technique</th>
<th>Required sensors</th>
<th>Measurements</th>
</tr>
</thead>
</table>
| GPS               | - Reference network GPS  
- Assisted GPS (A-GPS)  
- High Sensitive GPS (HS-GPS) | - Transmitter  
- Receiver  
- Antenna | - Position  
- Velocity  
- Time |
| Image-based       | - Feature matching  
- Map-aided using laser range | - Optical image  
- Laser image | - Range, attitude  
- Intensity |
| Radio-based       | - Short-range communication  
- Ultra-Wideband (UWB) | - Tag, reader  
- Server | - Time  
- Direction |
| Dead Reckoning (DR)| - Point-to-point navigation  
- Locomotion-based navigation | - MEMS IMU  
- Barometer altitude  
- Magnetometer  
- Step switch | - Acceleration  
- Attitude  
- Locomotion  
- Step event |

Table 1: Approaches to localization.

In what follows, we provide a basic overview of the techniques listed in Table 1, highlighting the advantages and limitations of each.

2.1. GPS Navigation

High-accuracy TSPI, predominantly provided by GPS, is one of the key enablers of precision and net-centric warfare. However, GPS is not effective in electromagnetically and physically impeded environments or in environments where GPS is significantly degraded or not available, yet Test and Evaluation as well as training operations have become more focused specifically on these types of environments.

One promising technique in GPS navigation is based on the Epoch-by-Epoch® (EBE) processing (Bock et al., 2004). EBE technology performs Precise Instantaneous Network (“PIN”) positioning utilizing data from one or more GPS reference receivers and the soldiers GPS receiver to produce a rigorous network solution at each measurement epoch. Unlike traditional RTK approaches, there is no need for re-initialization immediately following loss-of-lock problems such as occurs when GPS satellites are occulted from the antenna’s view. This feature of instantaneous integer ambiguity resolution is of utmost importance when trying to position a soldier operating in environments which disrupt GPS reception such as in foliage and structures in urban environments.

Another approach to accelerate GPS position fixes in urban environments is using Assisted GPS (A-GPS) capable receivers. In urban environments, there is always the restriction of downloading almanac and ephemeris data from the GPS constellation. A-GPS receivers can assist this shortcoming using a wireless reference network, which communicates with the GPS receiver and transfers the required data (LaMance et al., 2002). Similarly, when GPS signals are weak, extra processing power is necessary to integrate weak signals to the point where they can be used to provide TSPI data. High Sensitivity GPS (HS-GPS) receivers use large banks of correlators and digital signal processing techniques to search GPS signals very quickly (Lachapelle et al., 2006).
The GPS modernization program, including the development of HS-GPS or A-GPS, will enable operation with much weaker signals (even indoors), and has shown significant improvement in the past few years; however, there are still situations where even these techniques cannot provide a sufficiently accurate position fix within an acceptable time interval. It is expected that soldiers in high multipath or extremely weak signal environments may experience low positioning accuracy and/or long delays in achieving a position fix. Even when the above mentioned strategies are fully implemented to provide the soldiers with a gracefully degrading position fix service, the position fix will eventually become unavailable.

2.2. Image-Based Navigation

Another method of navigation is based on image processing, (e.g., Moafipoor, 2006; Koch and Teller, 2008). This approach recognizes the surrounding environment through feature matching and tracking. An area map is built by the system using this information, which is used to locate the soldier as he navigates the environment. The general algorithms use the geometry of images and feature tracking over consecutive images. As an example, Figure 1 shows a sequence of images taken by the soldier with image overlap sufficient to track features and reconstruct the trajectory. In this algorithm, when GPS data is available, top-left corner in Figure 1, the captured images are geo-referenced. When the soldier leaves the GPS coverage area, the new captured images that overlap with the geo-referenced images are used to estimate the soldier’s position. The sub problems in the transition from one image to the next are: image registration, feature extraction, feature tracking, mosaic generation, and detection of targets and moving features in the subsequent images. Applying this algorithm to a sequence of overlapping images can reconstruct the navigation solution along that strip. Image resolution is the crucial key to improving the performance of this technique; however, by increasing the resolution, the computational burden of the process for real-time performance is increased.

![Figure 1](image)

Figure 1: Image-based navigation using feature tracking; the user started moving from an outdoor to an indoor environment; in the figures several features of interest are extracted and tracked from one image to the next, facilitated the position fix.

Another type of image-based tracking technology is based on laser ranging, which provides range measurements to dynamic or static targets (Heikkilä, 2005; Farley and Chapman, 2008). The laser scanner performs well in environments with many and apparent geometrical features, where it can measure distances from several meters (with mm-level accuracy) to a few hundreds of meters (with dm-level accuracy). The laser scanner also has the capability for online landmark-based map building and simultaneous utilization of the generated map to constrain the errors in the Inertial Measurement Unit (IMU) (Guivant et al., 2000).
Another promising technique in this category is based on the integration of a wide field of view passive 2D imager with an active laser radar 3D imaging (Naikal et al., 2009). The goal is to create tools that support cooperative navigation enabling across distributive imaging sensors to share data in meaningful ways, and to examine the impact of their navigation solution.

2.3. Radio-Based Navigation

The third navigation technique is radio-based navigation designed for a pre-instrumented environment using active and passive tracking technology. Several radio-based positioning techniques are available today on the market (Yoneki and Bacon, 2006). According to our application, they can be categorized into two systems: short-range wireless, and Ultra-Wideband (UWB).

Short range wireless communication is, by any measure, the fastest growing segment of the wirelesscommunications industry. However, it is very difficult to provide TSPI information using technologies such as WLAN, RFID, ZigBee (IEEE802.15.4), Bluetooth, and WI-FI (Rantakokko et al., 2007). These systems usually require pre-installed infrastructure in the operational environment. However, they may also be deployed and pinpointed during an operation. The result is a navigation network consists of a number of sensors, i.e. transmitters and receivers, spread across geographical area; each sensor has wireless communication capabilities and sufficient intelligence for networking of data. The reading range of these networks is less than 15m in the proximity of each node, and cm-level accuracy requires dense deployment of sensors (Parikh et al., 2004).

The UWB signals, created by a direct-sequence spread-spectrum (DSSS) have been used in ranging systems for many years. UWB transmits the radio frequency, RF, in a series of narrow pulses that reach their destination within very short time, i.e., a few ns. The UWB spreads these pulses over a very large spectral bandwidth; typically from 3 to 10 GHz (Pittet et al., 2008). Because of these attributes, i.e., low transmit power, large bandwidth, low frequency (which improves the penetration of the radio signal through walls, etc), and short pulse duration, UWB is considered as one of the most accurate and highly regarded technologies for location estimates. Several types of observables based on RF transmissions have been used for indoor localization. These include 1) the angle of arrival (AOA) indicating the direction from which the signal was received, 2) the received signal strength (RSS), 3) the time of arrival (TOA) of the received signal, and 4) time-difference-of-arrival (TDOA) of the received signals for the estimation of distance to mobile user (Bouet and Santos, 2008). However, users in high multipath environments may encounter low positioning accuracy and/or long delays in achieving a position fix. These factors have a more significant impact on the AOA/RSS observables than on the TOA/TDOA, and consequently these observables will not be considered for this effort. The range of UWB is about 10m, which can be extended by increasing the transmitter power and antenna gain.

2.4. Dead Reckoning Navigation

Dead Reckoning (DR) navigation is a relative measurement approach, the fundamental idea of which is to integrate incremental motion information over time. Starting from a known position,
successive position displacements, derived in the form of changes in step length (SL) and step direction (SD), are accumulated. DR systems are usually designed on the basis of a range of self-contained sensors including MEMS IMU (3-axis accelerometers, gyros), magnetometers, barometric altimeters, and step-sensors. These sensors are integrated into a sensor package suite, carried by the soldier, usually on a backpack or on the feet.

The SL is defined as the distance between two successive points of heel-ground contact made during walking. The primary sensor enabling easy determination of step events is a set of step-sensors, located in the shoe soles, used to sense impact, i.e., the instances when the operator’s shoes hit the ground (Toth et al., 2008). They can also be used as a reliable indicator of whether the operator is in motion or at rest. In many DR systems, the SL is simply approximated by an average distance for each user, because in the range of existing self-contained sensors, SL cannot be measured directly. With shoe-mounted placement, it is possible to measure SL using zero-velocity updates (ZUPT) (Ojeda and Borenstein, 2007). However, this schema has two shortcomings: first, the shoe direction is difficult to measure and is also unstable. Second, it is more difficult to engage the lever-arm in the process because of frequent changes as range of movement increases. The proper approach, applied here, is to use human body as navigation sensor to facilitate SL prediction (Grejner-Brzezinska et al., 2008).

Instead of SL, the SD can directly be measured using the magnetometer compass. The magnetometer compass operates based on sensing and measuring the Earth’s magnetic field; therefore it is sensitive to any ferrous materials close to the sensor and can be easily degraded by uncompensated magnetic anomalies, requiring a substantial calibration effort. To improve the reliability of heading determination in urban environment, a high quality gyro should be added to the system configuration. The current MEMS-gyro technology has not yet represented stand-alone the performance accurate enough for providing heading; requiring aiding and constant calibration. However, recently developed researches appear to provide the possibility of meeting the difficult tactical grade (e.g., Waters et al., 2002; Panhorst et al., 2006).

The main limitation of DR navigation is that by propagation of the DR position, uncertainty grows with time unbounded, because of the errors in estimation of SL and SD, respectively. So a reliable navigator must be protected by a mechanism of controlling the accumulation of errors. For this purpose, independent and external reference information in terms of position, direction, or velocity should be periodically assigned to the system. Images (e.g., optical cameras or laser scanner images) and maps are two typical sources of this information, which can represent references (landmark) points. Simultaneous Localization and Mapping (SLAM) algorithm is a landmark-based navigation system that has the capability for online map building and simultaneously utilization of the generated map to bound the errors in the DR system (Guivant, et al., 2000).

3. NAVIGATION STRATEGY

In previous sections we provided an overview of several navigation technologies. Table 2 below provides an evaluation of these technologies in terms of limitation, advantage, coverage, operation time, and accuracy.
As shown in Table 2, each localization approach has its own advantages as well as limitations. A robust solution is increasingly pointing to an integrated system using combinations of these techniques rather than any single technique. Substantial research and conceptual work have been conducted in recent years to develop reliable and ubiquitous multi-sensor navigation system for military personnel (Fax and Volk, 2007; Filjar, et al., 2008). The main advantage with multi-sensor integration is in increasing redundancy, integrity, and availability, as well as robust estimation of the TSPI parameters. In this work, we adopt a multi-sensor fusion “layered” architecture which abstracts sensors, modules and module fusion techniques. This architecture is illustrated in Figure 2 below.

The first layer, sensor hardware layer, represents sensors required for each technique, also shown in Table 1 (last column). The sensor abstraction layer takes advantage of advances in sensor technologies without having to re-invent the wheel on integration for each advance. The third layer is the positioning module layer. The primary modules are GPS, DR, and RF-ranging systems. The GPS module is utilized when GPS is available. It is mainly intended to perform two tasks: first, providing TSPI data in outdoor environments, and second, calibrating the self-contained sensors. Once the self-contained sensors are calibrated, two interconnected procedures are performed as a byproduct of the outdoor navigation solution: first, training a knowledge-based system to support the human locomotion modeling and predict the SL/SD parameters for DR module (Grejner-Brzezinska et al., 2007), and second, deploying reference stations for radio-based (RF-ranging) infrastructure. Later, in the absence of GPS signals, the last two modules are augmented to provide TSPI data. Together, these modules have the potential to deliver accurate TSPI information to the soldier but for a limited time due to drift in internal components and environmental effects. In order to expand the performance duration, the system has to be augmented with non-inertial data, such as optical image, laser range, or map data. These modules can provide absolute position information by which the accumulated errors in the DR module can be removed or bounded.

<table>
<thead>
<tr>
<th>Position module</th>
<th>Limitation</th>
<th>Advantages</th>
<th>Coverage range</th>
<th>Operation time</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS</td>
<td>- Line-of-sight - Cycle slip - Jamming</td>
<td>Absolute navigation information</td>
<td>City 0-10 km</td>
<td>Unlimited</td>
<td>~ cm</td>
</tr>
<tr>
<td>Optical image</td>
<td>- Require images overlap - Image orientation</td>
<td>Target identification</td>
<td>Building 1-100 m</td>
<td>1-10 minutes</td>
<td>~ dm</td>
</tr>
<tr>
<td>Laser image</td>
<td>- Feature extraction - Feature association</td>
<td>Learning environment</td>
<td>Building 1-30 m</td>
<td>1-15 minutes</td>
<td>~ cm</td>
</tr>
<tr>
<td>Radio-based</td>
<td>- Multipath - Pre-install - Time synchronization</td>
<td>Provide unique identification</td>
<td>Building 10-20 m</td>
<td>Restricted to location</td>
<td>Variable ~ dm-km</td>
</tr>
<tr>
<td>DR</td>
<td>- Error growth unbounded - Stringent environmental requirement</td>
<td>Self-contained performance</td>
<td>Several blocks 0-2 km</td>
<td>1-10 minutes</td>
<td>~ dm</td>
</tr>
</tbody>
</table>

Table 2: Characteristics of different military position modules
The forth layer of the architecture, the Control Module layer is designed to check the consistency between the modules and increases the redundancy of much of the information required to constrain the soldier navigation. A quality assurance/quality control (QA/QC) mechanism is used in this layer to monitor their integrity during navigation in real time and to test methodologies that identify any inconsistency or mis-specification in their measurements (Moafipoor et al., 2008). If there is any mis-specification in the DR/RF measurements, a redundancy check is made, and a suitable adjustment is assigned to the measurements. The estimation problem involves the determination of the extent of system failure and engages the back-up plan, i.e., image, map data, to compensate for the error caused by the failure.

The fifth layer, the Collaborative Module layer, implements collaborative navigation, explained in the next Section. The final layer, the Fusion Module layer, is designed to implement different multi-sensor fusion techniques such as the centralized Extended Kalman Filter (EKF), Particle Filter, or Voting techniques (Fayman et al., 1999), enabling us to construct reliable TSPI from a multitude of navigation modules.

4. COLLABORATIVE NAVIGATION

Collaborative navigation is an approach based in which independent users have their own complementary navigation modules, but as a team, they share their common resources and cooperate in order to address navigation goals for the whole team (Fox et al., 2008). Collaborative strategies can be implemented in many ways, e.g. (Vydhyanathan et al., 2007; Cui and Cao, 2008). In our design, collaborative navigation relies on resources coming from...
dismounted soldiers, radio-based networks, and map-data of the scene created in an off-line process from optical/laser image data. The use of shared resources provides advantages over single resource navigation systems in terms of efficiency, tolerance to possible failures, and capability of merging varied information, thus compensating for sensor uncertainty. Another advantage of this strategy is that different soldiers can be equipped with different navigation performance levels, enabling the reduction of the total cost of navigation for the whole team. For example, if the line of sight to satellites is blocked for a single soldier, through collaboration, the satellite information can be provided by other soldiers. These soldiers may then transmit their own positions and estimate the range to the soldier lacking GPS visibility. In this context, soldiers outside the building act similarly to GPS satellites (dynamic pseudolites) (Rantakokko et al., 2007). The common scenario for collaborative navigation is possible if: 1) some units have access to open sky and have GPS solution or other stated reference information, 2) communication between soldiers can be established through radio communications, and 3) their relative range-to-range vectors can be measured.

Another technical concept, which can enhance the collaborative navigation performance, is based on learning environment. This technique is called collaborative environment learning.

5. COLLABORATIVE ENVIRONMENT LEARNING NAVIGATION

In light of the stringent requirements for providing TSPI for periods of up to several hours, with high-accuracy in the absence of GPS, a collaborative architecture was proposed. The complementary to this architecture is environment learning. Here, ‘learning’ implies the ability to identify certain spatial features of the environment, e.g., interest points, corners, walls to some extent, so that a map can be created from these features. Because of complexity of real-time map building, this task can be performed in off-line mode, separated from current navigation operation. Figure 3 shows the algorithm proposed for real-time map building and environment learning.

![Figure 3: Collaborative environment learning architecture.](image)

Referring to Figure 3, it is assumed that a group of soldiers with different sensor systems, as discussed in Figure 2, exchange information between themselves and control command. Next, data in terms of range and bearing are collected by these units and transferred to the control commander. Once enough data is gathered, a map is built, and then, it is transferred back to the personnel. This feedback routine can also be updated while the navigation is in progress. This task can also be undertaken by a separate team, called navigation crew. This team is responsible to not only learn the ongoing environments, but also deploy the radio-based network structure and pinpoint them.
6. CONCLUSION

Modern urban combat requires the execution of highly developed cognitive process in terms of selection of navigation and corresponding infrastructure in order to provide accurate and reliable TSPI. The current target is to provide TSPI data for periods of up to one hour, with an accuracy of better than 1m in the urban environments. Meanwhile, it is increasingly apparent that no single navigation technology will emerge in the near term to capable of providing accurate location/navigation data to meet our purpose. In this paper, a new navigation architecture is proposed. This architecture is based on a collaborative structure with several layers, including, multi-sensor integration, collaborative navigation, and environment learning. It seems that the navigation in urban environments needs more leverage on learning environment, through enforcing restrictions on movement. This routine in integrating with the collaborative navigation can provide information about the state of a soldier without the need of a prior infrastructure, such as GPS, ground beacons, or a preloaded map in urban environments.

REFERNECE