MFI2003 in Tokyo, Japan, pp.287-292 (2003)

A method of personal positioning based on sensor data fusion of wearable camera and self-contained sensors

Masakatsu Kourogi† Takeshi Kurata†‡

†Intelligent Systems Institute, National Institute of Advanced Industrial Science and Technology (AIST)
1–1–1 Umezono, Tsukuba Chuo Dai 2, Tsukuba, Ibaraki, 305–8568, JAPAN

‡Human Interface Techonology Laboratory, University of Washington

E-mail: m.kourogi@aist.go.jp, kurata@ieee.org

Abstract

In this paper, we propose a method of personal positioning that combines images taken from a wearable camera with data from self-contained sensors attached to the user through a Kalman filter as a data integration mechanism. The proposed method estimates the user's position and direction by image registration between the input images from the camera and a set of images captured at known positions and directions beforehand as a database. It updates the estimation of the user's position and direction with pedestrian dead-reckoning by detecting walking behavior of the user and by estimating the heading direction of the body with the self-contained sensors.

1 Introduction

In recent years, wearable computing [2] has become more feasible due to rapid progress in low battery consumption and compact, light-weight computers and their I/O peripherals such as displays, cameras and sensors. To provide suitable applications to the wearable computers, it is essential for the computers to understand the context of users all the time.

One of the most important contexts to be determined by the computers is the user's location and orientation relative to the environment. By acquiring the context, for example, it is possible to create Augmented Reality (AR) systems that overlay annotative information on a live image from the user's viewpoint [3][4]. This paper focuses on pedestrian users walking both indoors and outdoors, and presents a novel method of real-time personal positioning for them.

Previous researches have dealt with personal positioning from several approaches. In one study using images from a wearable camera as an information source, Aoki et al. proposed a method of personal positioning based on dynamic programming matching between a time sequence of color histograms of successive images and a database of sequences captured and registered beforehand [6]. In general, methods of personal positioning exploiting images have the advantage that no surrounding infrastructure is required. However, image-based methods require a database for registration covering the entire location where the user can go, and it is very time-consuming to construct such a database and thus it is difficult to use the method for a large scale environment.

On the other hand, many researchers proposed multi-sensor based methods aiming at human-specific behaviors for personal positioning. Golding et al. proposed an indoor navigation method based on a statistical analysis of multi-sensor data such as accelerometers, magnetometers, thermometer and light sensor [7]. However, since this method takes time to acquire a statistically meaningful result from multi-sensor data, it cannot easily be used for real-time applications. Lee et al. proposed an incremental positioning method based on detection of human walking behavior by accelerometers and walking direction by magnetometers[11]. Although it estimates relative displacement of the user, it can be used for a limited range for which paths of displacement are preregistered. It also requires absolute positioning information from an external source and a mechanism of data integration between the relative displacement and the absolute positioning, which are not realized by the method.

Some of the previous works proposed the combination of dead-reckoning (DR) of human walking and a method of absolute positioning. For example, Levi et al. from Point Research Inc. [8] proposed a personal positioning system using the Global Positioning System (GPS) as an absolute locating method, and pedometer with accelerometers and a magnetic compass as a relative displacement positioning method [8][9]. However, when used in an indoor environment, the GPS signals are easily blocked by building structures and the earth's magnetic field can be disturbed by electronic devices such as cathode ray tubes (CRT) and copier machines and thus the magnetic compass becomes vulnerable to disturbance.

2 The proposed method

We propose a method of personal positioning by utilizing four sensing modules as described below.

1. A wearable camera attached to the user's head:

The images taken by the camera are tested if matched with pre-registered images stored in a database. As a result, the direction of the user's head and the absolute position are acquired.

2. An attitude sensor attached to the user's head: This sensor estimates the absolute attitude of the user's head (i.e. viewpoint) with a drifting effect.

3. Accelerometers attached to the user's torso: The user's walking steps are recognized and measured.

4. Magnetometers, gyro-sensors and inclinometers attached to the user's torso: The absolute attitude of the user's torso are acquired. By combining module 3, a DR module that estimates relative displacement based on the user's walking behavior can be constructed.

The arrangement of these sensing modules is shown in Figure 1.

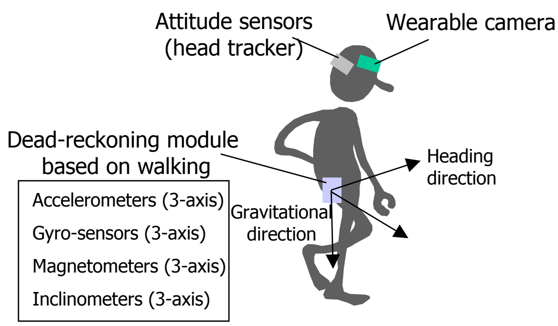

Figure 1: The arrangement of sensing modules attached to the user.

In the proposed method, multiple output data from these modules are acquired, and fed into the Kalman filter framework and are integrated into a single statistically consistent result with its uncertainty as described in the following section.

2.1 Data fusion framework with Kalman filter

The method integrates data from the wearable camera and the multiple sensors within the Kalman filter into a single output result with a particular uncertainty. The state vector x_k at a discrete time k is defined as:

$$x_k = [x_k, y_k, \alpha_k, \beta_k, \gamma_k, \psi_k, \theta_k, \phi_k]^T, \tag{1}$$

where (x_k, y_k) is the position of the user, $(\alpha_k, \beta_k, \gamma_k)$ is the absolute attitude of the user's head, and $(\psi_k, \theta_k, \phi_k)$ is the absolute attitude of the user's torso. In this paper, attitude vector is notated by Euler angle (yaw, pitch and roll). In the framework of Kalman filter of linear system with a control input vector, the state vector is updated by the equation:

$$x_{k+1} = f(x_k) + D_k(x_k)u_k + G_k w_k, \tag{2}$$

where f is a state transition function, w_k is the systematic noise vector and u_k is the control input vector. We apply the dead-reckoning based on human walking to the framework by regarding the control input vector as a walking distance u_k, $f(x_k) = x_k$, and $D_k(x_k) = [\cos\psi_k, \sin\psi_k, 0, .., 0]^T$. Because the systematic noise vector w_k is generated by the dead-reckoning estimation, we can assume that $G_k = I$.

The filter equation predicts the optimal estimation of the state vector $\hat{x}_k$ by:

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}) + D_k(\hat{x}_{k|k})u_k, \tag{3}$$

where $\hat{x}_{k+1|k}$ is the state vector predicted from the state vector $\hat{x}_{k|k}$ at the previous discrete time.

After the measurement process, $\hat{x}_k$ is updated by the following equation:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k[\tilde{y}_k - h(\hat{x}_{k|k-1})], \tag{4}$$

where K_k is Kalman gain, $\tilde{y}_k$ is measurement vector, and h is the prediction function of measurement vector from the state vector. We set the measurement vector $\tilde{y}_k$ acquired from the wearable camera and the multiple sensors as:

$$\tilde{y}_k = [\hat{x}_k, \hat{y}_k, \hat{\alpha}_k, \hat{\beta}_k, \hat{\gamma}_k, \hat{\psi}_k, \hat{\theta}_k, \hat{\phi}_k]^T. \tag{5}$$

Therefore, $h(\hat{x}_k) = \hat{x}_k$. In the Kalman filter framework, the measurement vector and the state vector satisfy:

$$\tilde{y}_k = h(x_k) + v_k, \tag{6}$$

where v_k is measurement noise vector.

Kalman gain K_k is updated by the following equation:

$$K_k = \hat{\Sigma}_{k|k-1}H^T(H\hat{\Sigma}_{k|k-1}H^T + \Sigma_{vk})^{-1}, \tag{7}$$

where $\hat{\Sigma}_k$ is the covariance matrix of the optimal estimation of the state vector $\hat{x}_k$, and Σ_{vk} is the covariance matrix of the measurement noise vector. By the definition of the prediction function $h(x)$, $H = I$.

The covariance equation in the Kalman filter framework updates the covariance matrix $\hat{\Sigma}_k$ by the following equations:

$$\hat{\Sigma}_{k|k} = \hat{\Sigma}_{k|k-1} - K_k H\hat{\Sigma}_{k|k-1}, \tag{8}$$

$$\hat{\Sigma}_{k+1|k} = F\hat{\Sigma}_{k|k}F^T + G_k\Sigma_{wk}G_k^T, \tag{9}$$

where Σ_{wk} is the covariance matrix of systematic noise vector w_k, and by the definition of the state transition function $f(x)$, $F = I$.

When the result of image matching from the wearable camera (that is, sensing module 1) is available, the user's position $(\hat{x}_k, \hat{y}_k)$ and the absolute attitude of the user's head $(\hat{\alpha}_k, \hat{\beta}_k, \hat{\gamma}_k)$ are given as a part of the measurement vector. The sensing modules 2, 3 and 4 give the absolute attitude of the user's head $(\hat{\alpha}_k, \hat{\beta}_k, \hat{\gamma}_k)$ and the absolute attitude of the user's torso $(\hat{\psi}_k, \hat{\theta}_k, \hat{\phi}_k)$ as measurements.

The measurement noise vector v_k, in the Kalman filter framework, is the discrepancy vector between the measurement vector predicted by the function $h(x)$ and that obtained from real observations by the multiple sensors. The proposed method uses the state vector updated by the dead-reckoning from the previous discrete time as a prediction of the next measurement vector. Thus, the components of the covariance matrix of the absolute attitude incrementally grow because of the drifting effect of the attitude sensor attached to the user's head. On the other hand, the components of the covariance matrix of the absolute attitude output $(\hat{\psi}_k, \hat{\theta}_k, \hat{\phi}_k)$ should be below a constant and sufficiently small since its estimation uses the gravitational direction and the magnetic vector as the absolute reference.

At the stage of updating the state vector, the accumulation of systematic noise caused by the dead-reckoning based on equation (3) makes the components of the position of the covariance matrix of the optimal estimation of the state vector incrementally grow. Because of the drifting effect of the attitude sensor attached to the head, the components of the absolute attitude of the user's head of the covariance matrix of the measurement noise increase. On the other hand, the components of the absolute attitude of the user's torso should not grow since the components of the measurement noise vector remain constant. When the position and the attitude of the user's head are acquired from the image registration (within sensing module 1), the accumulated components of the position and the absolute attitude of the user's head are thought to shrink below a certain level of accuracy provided by the image registration.

Figure 2 shows the system diagram of the Kalman filter described in this section.

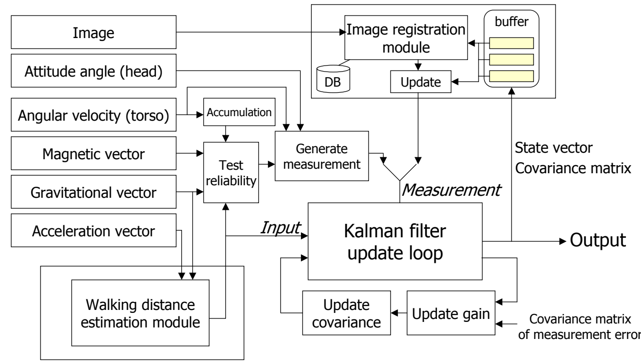

Figure 2: System diagram of the Kalman filter.

2.2 Estimation of position and attitude of head based on image registration

We use a set of images captured at known positions and posture of the camera as an image database. The current position and the attitude of the wearable camera that is attached to the user's head are acquired by the image registration between live images from the camera and images in the database. As for the method of image registration, we use the method of estimating global affine parameters between frames described in the literature [5]. After the affine parameters are acquired, the normalized cross correlation of image brightness between the aligned frames with the parameters is computed and tested as to whether the correlation value exceeds the pre-defined threshold. If it passes the test, the image registration is deemed to be successful. When the image registration successively matches an input frame and one of the frames in the database, the position and the posture of the registered frame are given as measurements to the update loop of the Kalman filter described in Section 2.1.

Since the computational cost of image registration is expensive and the image database covers the target area at a sparse density in practice, it is essential to precisely judge the necessity of image registration when an input frame is given at a discrete time k. The proposed method determines the necessity by the optimal estimation of the position and the absolute attitude of the user's head and their components of the covariance matrix $\hat{\Sigma}_k$, which correspond to the uncertainty of the estimation. That is, the method initiates the image registration if any frame is registered in the database whose position and posture are covered by the uncertainty range defined by the optimal estimation and its covariance.

2.3 Estimation of walking distance

2.3.1 Detection of walking behavior by acceleration

Since human walking behavior generates a typical pattern of acceleration in the vertical and horizontal directions, the accelerometers attached to the torso (which can be nearly regarded as the center of gravity of the human body) can be used to detect the behavior. Figure 3 shows a typical pattern of acceleration in the vertical and horizontal directions caused by walking behavior.

A simple peak detection of acceleration in the vertical direction to recognize the behavior, however, can easily fail if the user turns around at the same po-

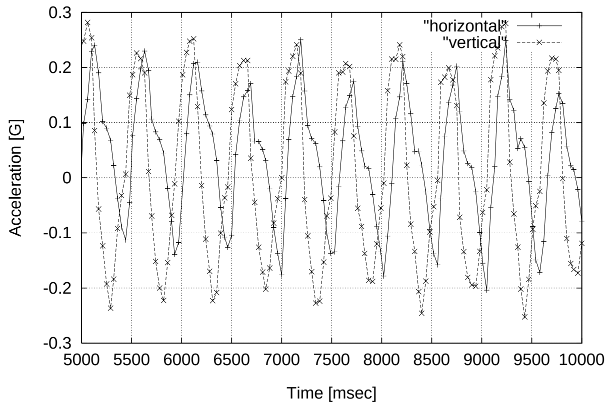

Figure 3: Example of acceleration caused by walking behavior.

sition, for example. The proposed method simultaneously observes the acceleration in the horizontal direction in addition to the peak detection in the vertical direction. The method integrates the acceleration in the horizontal direction during the half cycle time of unit of walking behavior before the point of peak detection of acceleration in the vertical direction. If the integration value exceeds a threshold, it regards that the walking behavior to be detected.

2.3.2 Estimation of walking distance

The method gives the standard step length (= height $\times 0.45$) as an estimation of walking distance if walking behavior is detected and gives 0 otherwise. Because the detection of walking behavior is accurate but the length of footsteps actually varies, the systematic noise covariance is set to be large if detected, and is set to be small otherwise considering the case of false detection.

2.4 Estimation of the absolute attitude of the user's torso

2.4.1 Estimation of the absolute attitude by magnetometers and inclinometers

In principle, the absolute attitude can be estimated from the output of the inclinometers which sense the gravitational acceleration vector and the magnetometers which sense the earth's magnetic vector. However, the earth's magnetic vector cannot easily to be stably sensed by the magnetometers especially in an indoor environment because electronic devices and building structures disturb the earth's magnetic field. Therefore, we utilize the fact that the dip angle (Figure 4) between the horizontal plane and the earth's magnetic vector is uniquely determined by the latitude and longitude in the absence of disturbance.

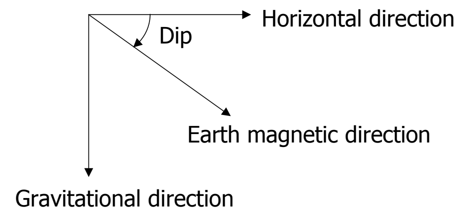

Figure 4: The dip angle.

The method evaluates the reliability of the magnetic field sensed by the magnetometers as an estimation of the earth's magnetic vector, by comparing the pre-determined dip angle and the angle between the horizontal plane and the magnetic vector. However, a disturbed magnetic field can generate a magnetic vector whose angle of depression accidentally matches the dip angle.

Therefore, the method successively compares the two angles when the user is in motion, and regards the magnetic field as reliable if the two angles match continuously. This successive comparison increases the reliability of the test for the magnetic field sensed by the sensors. The method detects walking behavior as described in Section 2.3 to judge if the user is in motion. Figure 5 shows an example of the time sequence of angle between the horizontal plane sensed by the inclinometers and the magnetic vector sensed by the magnetometers while the user is walking in an indoor environment. The angle should be equal to the dip angle if no magnetic disturbance is present in the environment. The result shows that the range of time when the magnetic vector can be reliable is limited.

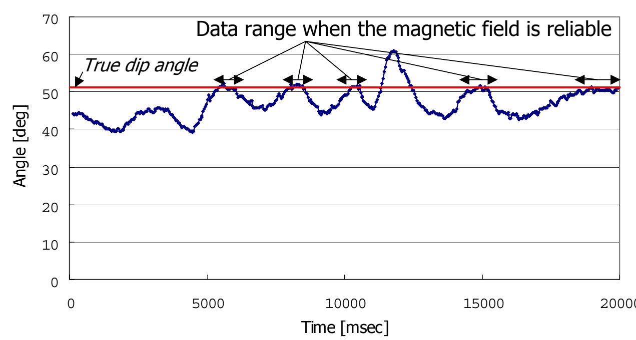

Figure 5: An example of the time sequence of angle between the horizontal plane and the magnetic vector.

Ladetto et al. proposed a method of canceling the effect of local and temporary disturbance of the magnetic field by using low-cost magnetometers and gyrosensors. Our method, however, is capable of actively locating the area where the magnetic field is reliable to estimate the earth's magnetic vector and thus can be applied to an environment where the disturbance of the magnetic field affects a large area. The method also uses low-cost gyro-sensors to cover the area where the magnetic vector is disturbed by integration of the angular velocity vector sensed by the gyro-sensors as described in the following section.

2.4.2 Estimation of the absolute attitude by accumulation of angular velocity

The method uses the absolute attitude computed from the reliable earth's magnetic vector and gravitational vector as a reference and updates it by integrating the angular velocity vector $(\omega_x, \omega_y, \omega_z)$ sensed by the gyro-sensors while the magnetic field is unreliable. Since the output of gyro-sensors always contains a drift component, the component of the covariance matrix of the measurement noise vector should grow incrementally without adjustment by the magnetic vector.

3 Experiments

3.1 Implementation

3.1.1 Hardware

In this research, we use InterTrax^2 from InterSense Inc. as the attitude sensor attached to the user's head. As for accelerometers, gyro-sensors, inclinometers and magnetometers attached to the user's torso, we use 3DM-G from MicroStrain Inc. that integrates these sensors into a package and enables us to acquire sensor data with a time stamp of acquisition via a serial port. In the package of sensors, ADXL202JE (3-axis) from Analog Devices Inc. are used for accelerometers, ENC-03JA (3-axis) from Murata Manufacturing Co. are used for gyro-sensors, HMC1052 (2-axis) and HMC1051Z (1-axis) from Honeywell Co. are used for magnetometers, and inclinometers are implemented by combining the accelerometers and the gyro-sensors. These sensors are very small and available at low cost compared to the sensors integrated in the Inertial Measurement Unit (IMU). We use a web image server (YOKOGAWA, IPV) for image capturing and use a card-sized PC (CPU: Mobile Pentium-III 700MHz, OS: Linux-2.4) from Cell Computing Inc. as the wearable PC. We use Clip-On Display from MicroOptical Inc. for the head-worn display.

3.1.2 System architecture

We have developed a prototype wearable system with the wearable visual interface *Weavy* [1]. Figure 6 shows the diagram of our prototype of personal positioning system named *Wyvern* composed of the hardware described in Section 3.1.1. All software processing of our method is done on the wearable side. The image database and map information are stored in the database server on the infrastructure side and are designed to be fetched from the wearable side via a wireless network compliant with IEEE 802.11b. The system is designed not to halt even if the wireless network becomes unreachable within an area where the fetched image database and map information are available in the cache storage on the wearable side.

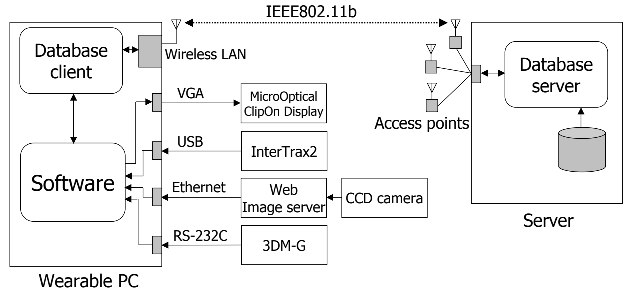

Figure 6: The diagram of the personal positioning system.

3.1.3 Experimental results

Figure 7 shows a result of the user's trajectory estimated by the proposed method in an indoor environment. In this result, the estimated position is adjusted by the image registration at the five locations marked by "o". The result shows that the method can robustly estimate the user's position in the presence of electronic office appliances such copier machines and computer displays (CRT) and building structures which disturb the magnetic field.

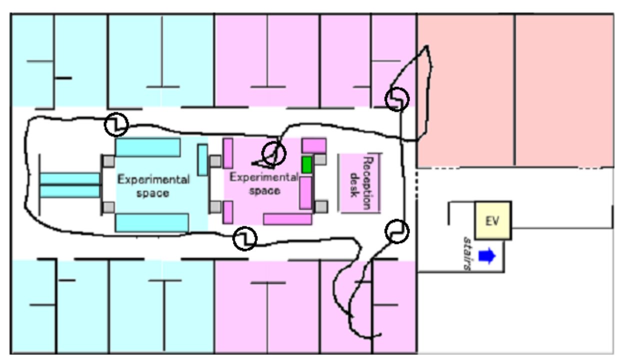

Figure 7: A result of estimation of trajectory.

Figure 8 shows examples of overlaying annotative information on live video frames based on the estimated results of the position and viewing direction of the user.

Figure 8: Examples of annotation overlay.

4 Conclusion

In this paper, we proposed a method of personal positioning for the pedestrians. We implemented the method on a prototype wearable system and found that it worked well in real-world scenes in both indoor and outdoor environments.

As future works, we need to improve the method of estimating walking steps. We used the fixed value for the length. And the estimation of the absolute attitude of the user's head is likely to degrade due to the drifting effect of the attitude sensors attached to the head, and thus the method requires the user to turn the viewing direction to a pre-registered scene in the database for the image registration to cancel accumulated error. We need to consider a natural method for prompting the user to view the scene.

References

[1] *Weavy: Wearable Visual Interfaces*, `http://www.is.aist.go.jp/weavy/`

[2] Steve Mann, "Wearable Computing: A First Step Toward Personal Imaging," *Computer*, Vol. 30, No. 2, pp. 25–32, 1997.

[3] Steven Feiner, Blair MacIntyre, Tobias Höllerer and Anthony Webster, "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment," in *Proc. Int'l Symp. on Wearable Computers (ISWC99)*, pp. 74–81, 1999.

[4] Masakatsu Kourogi, Takeshi Kurata and Katsuhiko Sakaue, "A Panorama-based Method of Personal Positioning and Orientation and Its Real-time Applications for Wearable Computers," in *Proc. Int'l Symp. on Wearable Computers (ISWC2001)*, pp. 107–114, 2001.

[5] Masakatsu Kourogi, Takeshi Kurata, Jun'ichi Hoshino and Yoichi Muraoka, "Real-time image mosaicing from a video sequence," in *Proc. Int'l Conference on Image Processing (ICIP99)*, Vol. 4, pp. 133–137, 1999.

[6] Hisashi Aoki, Bernt Schiele and Alex Pentland, "Realtime Personal Positioning System for Wearable Computers," in *Proc. Int'l Symp. on Wearable Computers (ISWC99)*, pp. 37–43, 1999.

[7] Andrew Golding and Neal Lesh, "Indoor Navigation Using a Diverse Set of Cheap, Wearable Sensors," in *Proc. Int'l Symp. on Wearable Computers (ISWC99)*, pp. 29–36, 1999.

[8] Point Research Corporation, `http://www.pointresearch.com/`

[9] Robert Levi and Thomas Judd, "Dead Reckoning Navigational System Using Accelerometer to Measure Foot Impacts," U.S. Patent Number 5,583,776, 1996.

[10] Thomas Judd, "A Personal Dead Reckoning Module," in *Proc. the Institute of Navigation, GPS'97*, pp. 169–170, 1997.]

[11] Seon-Woo Lee and Kenji Mase, "Incremental Motion-Based Location Recognition," in *Proc. International Symposium on Wearable Computers (ISWC2001)*, pp. 123–130, 2001.

[12] Quentin Ladetto and Bertrand Merminod, "Digital Magnetic Compass and Gyroscope Integration for Pedestrian Navigation," in *9th St. Petersburg International Conference on Integrated Navigation Systems*, 2002.