International Journal of Fuzzy Logic and Intelligent Systems, vol. 5, no. 3, September 2005 pp. 230-245

Gesture based Input Device: An All Inertial Approach

Wook Chang*, Won Chul Bang*, Eun Seok Choi*, Jing Yang*, Sung Jung Cho*, Joon Kee Cho*, Jong Koo Oh* and Dong Yoon Kim*

*** Interaction LAB, Samsung Advanced Institute of Technology**

Abstract

In this paper, we develop a gesture-based input device equipped with accelerometers and gyroscopes. The sensors measure the inertial measurements, i.e., accelerations and angular velocities produced by the movement of the system when a user is inputting gestures on a plane surface or in a 3D space. The gyroscope measurements are integrated to give orientation of the device and consequently used to compensate the accelerations. The compensated accelerations are doubly integrated to yield the position of the device. With this approach, a user's gesture input trajectories can be recovered without any external sensors. Three versions of motion tracking algorithms are provided to cope with wide spectrum of applications. Then, a Bayesian network based recognition system processes the recovered trajectories to identify the gesture class. Experimental results convincingly show the feasibility and effectiveness of the proposed gesture input device. In order to show practical use of the proposed input method, we implemented a prototype system, which is a gesture-based remote controller (MagicWand).

Key Words : Input device, gesture input system, accelerometer, gyroscope, inertial navigation system, mobile phone, mobile device, remote controller

1. Introduction

This work is part of a project for the development of an inertial sensor based gesture input device involving Samsung Advanced Institute of Technology (SAIT).

A large number of computing devices found today have increasing computing power, all the while becoming extremely miniaturized even to be worn or embedded in the environment. Due to their totally different appearances compared with PCs, the wide spread WIMP (Windows Icon Menu Pointer) interface is no longer valid for those systems. Especially, input devices are one of major bottlenecks for users to fully enjoy their high performance but small sized computing devices. For example, manufacturers of TV with internet connectivity provide users with wireless keyboard mouse pairs or button rich remote controllers as input devices, both of which are still uncomfortable for most customers.

There have been fruitful research and industrial works in the design of efficient input devices for post PC devices. To cite a few, tangible input devices, voice input devices, gesture input devices, handwriting input devices, and so forth. However, no pervasive and popular input devices are available yet and it is pointed out in the literature [1] that a single input method cannot cover diverse needs of users, which leads to research on multimodal input systems.

As an attempt to reach the goal of providing users with comfortable but easy to use input device, the authors developed an input device using inertial sensing technology to control information appliances. Gestures, which are defined as stylized motions that convey meaning [2], are adopted in many input systems today due to their intuitiveness [3]. Among various kinds of gesture input devices [4], vision based gesture input devices are popular in research fields. These systems are composed of several cameras installed in the environment, an image processing unit and a recognition system. Another interesting approach is to adopt electromyogram (EMG) technology to recognize user's gestures [4]. For small sized computing systems such as wearable computers, however, the camera based gesture input devices cannot be viable solutions because they require large external installation which cannot be carried with users.

Simply speaking, the role of cameras in most gesture based input devices is to extract the motion of a user which may fall under the category of motion tracking. Very similar to the field of input systems, motion tracking also deploys a lot of available technologies; mechanical sensing, inertial sensing, acoustic sensing, magnetic sensing, optical sensing, and radio/microwave sensing [5].

Motion tracking systems can be categorized as external reference types and self contained types. External referenced types use some external reference sources or digitizing surfaces to track the position of an object, whose portability is limited due to additional devices for referencing. To the contrary, the self contained types do not require additional devices and their usability and comfortability is superior to that of the external referenced types. Motion tracker technologies using optical sensors are the most successful and commercially available devices to make self contained input devices but their functionality is limited by the fact that the movement range should be within the line of sight of the detector (camera). In addition, they also require a surface to detect the reflected light.

Unlike above mentioned motion tracking technologies, the

Manuscript received Aug. 2, 2005; revised Sep. 7, 2005.

inertial sensing approach is quite unique from the view point that it can track the motion of a system without external electromagnetic signals. Therefore, the system, which is called inertial navigation system (INS), does not suffer from signal coverage problem which is common in external referenced systems such as global positioning system (GPS). The idea of building input devices using inertial sensors such as accelerometers and gyroscopes has long been investigated and reported in the research papers [6-22] or patents [23-32].

At the earliest stage of the research, the acceleration measurements were mainly used for signature verification [6, 7, 13-15, 23, 29]. In [12], a pointing device which can be used in 3D space was discussed. Strickland et al. [17] and Kobayashi et al. [10] applied inertial measurement units to head tracker systems. In [22, 23], a data glove system was constructed using accelerometers only. The application of inertial sensors to mice was also considered in [18]. For a pen type input device, Ishikawa et al. [11] suggested to use accelerometers to recover the handwriting trajectory. Verplaetse [34] discussed possibility of the pen type input device equipped with inertial measurement units but its realization was far from current state of art. Several patents [35-37] claimed input devices with various combination of accelerometers and angular rate sensors. In [19], the vibration caused by the friction between a writing surface and a pen tip was analyzed to give the direction information of the pen tip. Miyagawa et al. [9, 38] reported a inertial pen system that senses tri axis accelerations and angular velocities. More recently, Cheok et al. [8] proposed an accelerometer based pen system for wearable computers.

Reviewing patents on input devices adopting inertial sensing technology, [24] is considered to be the most fundamental patent, whose predecessor is from IBM [27]. The pen systems considered in the both patents use accelerometer and the main difference is their applications where [27] is for signature verification and [24] is for drawing and writing. In [28], a pen type input device with two accelerometers and one gyroscope and related algorithms has been presented. Baron Inc. [29] suggested to use accelerometers to identify users' signature and positioning sensors such as an ultrasonic sensor or an electro magnetic coil to detect the position of the pen tip. Ricoh Inc. [26] claimed a device that senses six degree of freedom (DOF) inertial measurements and related algorithms. Japanese patent [30] uses accelerometers and ultrasonic sensor together to recover the handwriting motion. In [39-41], 3D input devices, which adopt different configurations of motion sensors including accelerometers, angular rate sensors, and so forth, have been claimed whose major application fields include pointing control of PCs and simple character recognition. It should be noticed that 6 DOF inertial measurements, i.e., tri axis accelerations and angular rates are required to recover the perfect 3D information of a moving object in the free space. However, except for [9, 26, 38], most works have utilized less number of sensors and can be used for only gesture input, pointing, signature verification, etc.

Although there have been fruitful research works and patents on input devices with inertial sensing technology, they have not been seriously deployed in commercially available input devices [24]. One reason is that integration of inertial sensors in the allowed size for target systems with high accuracy has been very difficult if not impossible. However, this problem is being gradually relieved as low cost MEMS inertial sensors become commercially available in the market. Another barrier to make a useful inertial input devices is the unbounded growth of error inherent in INS. In aerospace applications, GPS is used to correct the errors of INS but is not feasible for small sized systems such as a gesture input device due to its size and battery life problems.

Motivated by the above observations, this paper aims at developing a low cost, small sized inertial input device which can be used without external reference systems. More specifically, we first synthesize inertial measurement unit (IMU) with commercially available MEMS type accelerometers and gyroscopes, based on which a gesture input device is developed. In addition to the theoretically perfect 6 DOF inertial navigation algorithm, the authors also develop alternative inertial navigation methods with less number of sensors. To alleviate the problem of unbounded increase of error in INS, we adopt the slightly modified version of zero velocity update (ZUPT) [42, 43] fit for our application. The obtained 3D position information is then recognized by Bayesian network pattern recognizer. In this stage, the obtained 3D position information of motion is projected to 2D space for the pattern recognizer module to maximize its recognition performance.

The main contribution of this paper is to establish the systematic design procedure of a gesture based input device by using inertial sensing technology. The proposed schemes and techniques carried out in this work can be used to large spectrum of applications such as 3D pointing, motion tracking, and so forth. Our experimental results convincingly show that the proposed gesture input device provides a satisfactory performance. We should also note that the implementation is quite simple but successful to show a promising way of commercialization of inertial input devices. Finally, we provide an illustrative prototype system, which is applied for remote control systems, to show the commercial feasibility of the proposed technique.

The rest of this paper is organized as follows. Section 2 introduces the brief overview of gesture based input devices and related technologies, followed by a description of the motion tracking algorithms 3 and recognition algorithm 4. A fruitful analysis and discussions based on real experiments is proven in Section 5 whose commercial applicability is provided by implementing two prototype systems 6. Finally, a conclusion of this paper is given in Section 7

2. Overview of the proposed gesture based input device

By adopting inertial sensing technology, the proposed system is intended to provide users with very simple and comfortable way of interaction with their own machines with little background information required to use it.

To achieve this goal, the proposed system senses a user's hand gestures and recognizes it, which is probably used to generate appropriate command signals for controlling external devices. Fig. 1 presents the overview of the proposed gesture input system. The most important components are the inertial measurement unit, the motion tracking algorithm, the Bayesian network based recognizer. The inertial measurement unit measures tri-axis acceleration and angular rate measurements relative to a fixed reference system. The motion tracking algorithm converts the measurements to actual 3D position (X, Y, Z) and attitude information (roll, pitch, yaw). The Bayesian network based recognizer is trained based on data collected from the system to decide actual meaning conveyed by the user's gestures. The structure of the Bayesian network based recognizer is flexible enough to use various kinds of signals such as the raw inertial measurements, positional information, attitude information, etc.

2.1. Hardware system overview

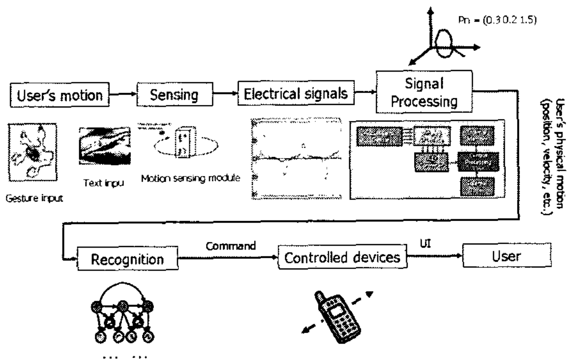


Fig. 1 Overview of the proposed gesture input device

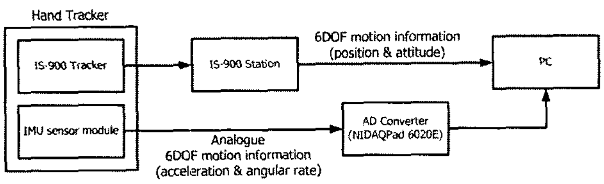


Fig. 2 Hardware configuration

Figure 2 shows the hardware configuration of the whole system. In order to qualitatively verify the performance of the proposed system, the InterSense (http://www.isense.com) IS 900 motion tracking system is adopted. The IS 900 motion tracking system utilizes the combination of ultrasonic and inertial sensors to give full 6 DOF motion information, whose main component is a tracker and stations (tracked devices).

We have also implemented an in house tracked device which consists of IS 900 station and the main sensor module as shown in Fig. 3. Motion information given by IS 900 motion tracking system is connected to an RS232C interface on a PC.

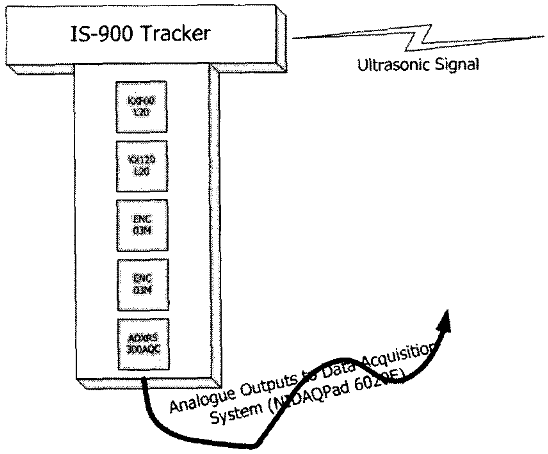


Fig. 3 Tracked device composed of a IS 900 station and the sensor module

A data acquisition program written in Visual C++ was used at the PC to receive and save the data from the sensor module and the IS 900 motion tracker to the hard disk. Data sampling rates for the sensor module and the motion tracker are set to 100 Hz. In addition, it guides users through a predefined sequence of gestures. The data is stored in the structured database and used to evaluate the performance of the system.

The sensor module in the system used in this study is composed of a dual axis accelerometer, a single z axis accelerometer, two single axis gyroscopes, and a single z axis gyroscope, which are all installed in 11.7 X 2 X 0.26 (cm) printed circuit board (PCB).

As shown in Fig. 2, the accelerometers and gyroscopes, which are sometimes called inertial sensors collectively, are arranged so that tri-axis accelerations and angular rates can be measured. The resulting configuration gives inertial measurement unit (IMU). In this configuration, the gyroscope measures the change of the angular rate of the system about its main axis of rotation and the accelerometer senses linear acceleration as well as gravity along its sensing axis. By combining signals from both sensors, one can compute the 3D motion relative to a fixed inertial frame.

On the presented system, following inertial sensors are packaged. The characteristics of each unit is described in Tables 1 - 4.

- A Kionix dual axis accelerometer (KX120 L20) with range of $\pm 2g$, sensitivity of 1000 mV/g and 2.5 V of zero g offset voltage for the measurement of the specific forces resolved along the x and y axes of the system. The physical value is calculated as follows:

$$f_{KX120-L20} = -\frac{V_{KX120-L20}-1.5}{0.6}9.8\,(m/s^2) \qquad (1)$$

- A Kionix single axis accelerometer (KXF00 L20) with range of $\pm 2g$, sensitivity of $1000mV/g$ and $2.5V$ of zero g offset voltage for the measurement of the specific force resolved along the z axis of the system. The physical value is calculated as follows:

$$f_{KXF00-L20} = \frac{V_{KXF00-L20} - 1.5}{0.6} 9.8\,(m/s^2) \quad (2)$$

- Two single axis gyroscope (Murata ENC 03M) with range of $\pm 300 deg/s$, sensitivity of 0.67 *mV/deg/sec* and offset voltage $1.35V$ for the measurement of angular velocities along the x and z axes. The physical value is calculated as follows:

$$\omega_{ENC-03M} = \frac{V_{ENC-03M} - 1.35}{K_{ENC-03M}} \frac{\pi}{180}\,(deg/s) \quad (3)$$

where $K_{ENC-03M} = 3.33 \times 0.67 \times 10^{-3}$.

- A single axis gyroscope (ADI ADXRS300AQC) with range of $\pm 300 deg/s$, sensitivity of $5mV/deg/s$ and offset voltage $2.5\,V$ for the measurement of angular velocity along the y axis. The physical value is calculated as follows:

$$\omega_{ADXRS300AQC} = \frac{V_{ADXRS300AQC} - 1.35}{K_{ADXRS300AQC}} \frac{\pi}{180}\,(deg/s) \quad (4)$$

where $K_{ADXRS300AQC} = 3.0 \times 10^{-3}$.

Table 1 Characteristics of the KX120 L20 accelerometer

Parameters	Units	Specifications
Range	g	±2.0
Sensitivity	mV/g	1000
Offset vs. Temperature	mV	±100
Sensitivity error	%	±2.0 typical
Resolution	mg	0.1 to 0.3
Electrical		
Input supply voltage	V	5.0 ± 0.25
Input supply current	mA	6.5 typical
Package	mm	16-pin SOIC Over-modeled Plastic

Table 2. Characteristics of the KXF00 L20 accelerometer

Parameters	Units	Specifications
Range	g	±2.0
Sensitivity	mV/g	1000
Offset vs. Temperature	mV	±150
Sensitivity error	%	±2.0 typical
Resolution	mg	0.1 to 0.3
Electrical		
Input supply voltage	V	5.0 ± 0.25
Input supply current	mA	6.5 typical
Package	mm	16-pin SOIC Over-modeled Plastic

Table 3. Characteristics of the ENC03 M gyroscope

Parameters	Units	Specifications
Range	deg./sec.	±300
Sensitivity	mV/deg./sec.	0.67
Offset vs. Temperature	%	± 20 at -5 to 75 ° C
Electrical		
Input supply voltage	V	2.7 to 5.25
Input supply current	mA	5

Table 4. Characteristics of the ADXRS300AQC accelerometer

Parameters	Units	Specifications
Range	deg./sec.	±300
Sensitivity	mV/deg./sec.	5 ± 0.4
Electrical		
Input supply voltage	V	5.0 ± 0.25
Input supply current	mA	6.0 typical
Package	mm	16-pin SOIC Over-modeled Plastic

The data output by sensors are analogue voltage output whose range is between 0 and 5 volts. The data acquisition board to collect data from the developed sensor module is National Instruments' NI DAQPad6020E. It has 16 analogue inputs with 12 bit resolution, two analogue outputs with 12 bit resolution, 8 digital input output lines and two 24 bit counter/timers.

2.2. Software system overview

Since we have intended to collect large amount of data from diverse types of users and apply various motion tracking and recognition algorithms, we developed an integrated software platform to evaluate the data. The platform is called GPIEE (Gesture input device Integrated Experiment Environment). The integrated system gets the hand motion data files and configuration files as input and produces the quantitative evaluation and recognition results as its output. In order to recognize the hand motion trajectory, several modules work corporately. First, the trajectory estimation module (Trajectory Estimator) converts raw inertial sensor signals into 2D trajectories. In this module, quantitative analysis is also performed. Then, the training module (Trainer) identifies the optimal structure and parameters for the Bayesian network recognizer. With the trained Bayesian network recognizer, output of the trajectory estimation module is used as input to recognition module (Recognizer). Finally, the recognition analysis module (Result Analyzer) calculates the statistics of recognition performance.

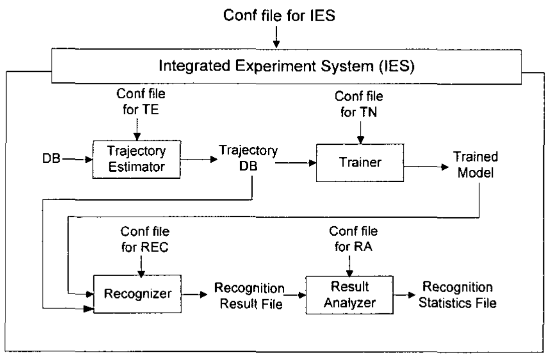


Fig. 4. Overview of GPIEE

3. Sensing Motions Produced by User's Gesture

The first step to realize the concept of the proposed system is to identify physical motion properties, i.e., 3D position and orientation, shown in a fixed inertial frame. In this work, we adopt INS theory since it gives the relative position and attitude of the system without external reference sources whose requirements are quite suitable for the concept of the proposed system. In the following, the principle of operations of inertial sensing technology is briefly introduced.

Fig. 5 shows the system in the free space. The navigation frame *n* represented by the orthogonal axes *X*, *Y*, and *Z* is the coordinate frame with respect to which the location of the system needs to be estimated. The body frame b is attached to the system, whose x, y, and z axes are aligned with those of the accelerometers. In order to compute the trajectories produced by the handwriting movement in the free space from the acceleration, three axes acceleration measurements and Euler angles (roll, pitch, and yaw) are required, by which the acceleration measurements are transformed into the accelerations in the navigation frame. The actual trajectories are obtained after double integration process. The governing INS equations of the 3D positioning used in this paper can be expressed as follows:

$$\begin{aligned}
\dot{P}_n &= V_n \\
\dot{V}_n &= C_b^n A_b - G \\
\dot{\psi} &= \frac{\omega_{by}\sin\phi + \omega_{bz}\cos\phi}{\cos\theta} \\
\dot{\theta} &= \omega_{by}\cos\phi - \omega_{bz}\sin\phi \\
\dot{\phi} &= \omega_{bx} + (\omega_{by}\sin\phi + \omega_{bz}\cos\phi)\tan\theta
\end{aligned} \tag{5}$$

where the subscript n denotes the navigation frame and b denotes the body frame, $A_b = [A_{bx}\ A_{by}\ A_{bz}]^T$ is the measured acceleration along the axes of the object, A_{bx}, A_{by}, and A_{bz} are the component of the acceleration vector in each axis, $P_n = [P_{nx}\ P_{ny}\ P_{nz}]^T$ is the position vector of the moving object (the origin of the body frame) in the inertial frame and $V_n = [V_{nx}\ V_{ny}\ V_{nz}]^T$ is the rate of change of P_n, i. e., velocity, G is the constant gravity vector shown along the z axis, (ω_{bx}, ω_{by}, ω_{bz}) are body frame inertial angular rate vector resolved along the object axis, (ϕ, $theata$, ψ) = (roll, pitch, yaw) are Euler angles. Here, the matrix $C_b^n = C_n^{bT}$ refers to the direction cosine matrix which describes the rotational relationship between the navigation frame and the body frame and the function of the three Euler angles. The detailed mathematical description of the direction cosine matrix is given in (6).

$$C_n^b = \begin{bmatrix} c\psi c\theta & s\psi c\theta & -s\theta \\ -s\psi c\phi + c\psi s\theta s\phi & c\psi c\phi + s\psi s\theta s\phi & c\theta s\phi \\ s\psi s\theta + c\psi s\theta c\phi & -c\psi s\phi + s\psi s\theta c\phi & c\theta c\phi \end{bmatrix} \tag{6}$$

where c denotes cos and s denotes sin.

In addition, the initial roll and pitch angle obtained by using

$$\psi = tan^{-1}\frac{-A_{by}}{-A_{bz}} \tag{7}$$

$$\theta = sin^{-1}\frac{A_{bx}}{g} = tan^{-1}\frac{A_{bx}}{\sqrt{A_{by}^2 + A_{bz}^2}} \tag{8}$$

Theoretically speaking, the system can measure its three dimensional movement with signals from attached inertial sensors only. Unfortunately, the described inertial sensing scheme cannot be applied directly since there is an unbounded growth of error. To clarify this point, let us consider the simplest case where an object is moving in the one dimensional space. Then, the governing INS equation can be described by

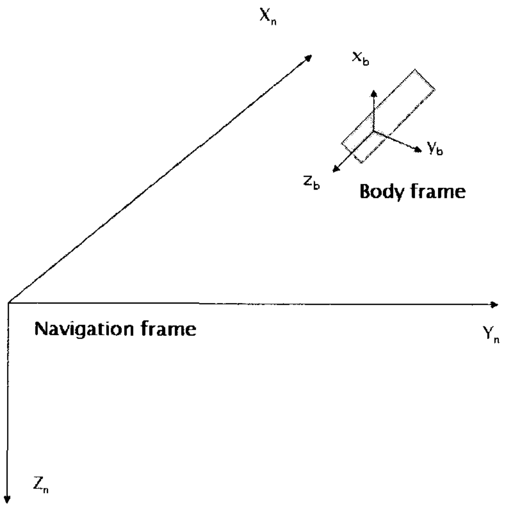


Fig. 5. Vehicle and Tangent Plane Coordinate Systems

$$p = p_0 + v_0 t + \frac{1}{2}\tilde{a}t^2 \tag{9}$$

where p is the position, p_0 is the initial position, v_0 is the initial velocity, $\tilde{a}$ is the acceleration measurement and t is time. Therefore, if there is a small error in the measured acceleration, the final positional error will unlimitedly grow with proportional to t^2. Since INS applied to compute 3D positioning involves more inertial sensors and integration steps, the resultant error is much more severe.

In the field of aerospace or robotics, this problem is usually remedied by taking use of additional positioning sensors, which cannot be used to the system since it lacks those additional sensors.

Another problem, which is specific to the system, is that it is a small commercial device. In the sense of size and battery life time, more the number of sensors increases, more size and battery power are required. Therefore, it is definitely necessary to investigate various combination of inertial sensor configurations which can provide enough performance for the proposed system.

With those problems in mind, the aim of the remaining part of this section is to suggest simple but effective signal processing methods to limit or avoid the positioning errors of the presented system with various sensor configurations. Specifically, we present some of possible sensor configurations and corresponding motion sensing methods. Before going further, it should be noticed that some of motion sensing algorithms presented in this section are applied for patents [44-47].

3.1. Method 1: Recovery of user's handwriting motions based on full 6 DOF inertial measurement

The INS typically employs additional aiding sensors such as GPS to guarantee accuracy for long term operations. If there are no additional navigation sensors, pure INS algorithm (5) is utilized to calculate the position, velocity, and attitude of the object in the free space. The most popular way of removing errors without aiding sensors is zero velocity update

(ZUPT) [42, 43, 48-50] or zero velocity compensation [51]. In the ZVC scheme, the navigation system regularly stops at a certain position. Since the system is completely stationary, one can correct navigation errors using the observed velocity errors at the ZVC point. In the correction, an optimal estimation procedure such Kalman filtering is generally utilized, which is not feasible for the proposed system due to the computational burden to implement Kalman filter. Therefore, we implement a simplified version of ZVC on the proposed system.

Figure 6 illustrates the concept of the proposed algorithm. The method described here employs the fact that the velocity should be zero at each stationary point. At the stationary point, calculated velocity of the proposed system and the hypothesized velocity, i.e., zero velocity is compared to correct the all the velocity history computed by the presented system.

Let us first assume that all error sources of the INS are collectively accumulated in the calculated acceleration in the navigation frame by $\epsilon_{A_n}(t) = [\epsilon_{A_{nx}}(t)\ \epsilon_{A_{ny}}(t)\ \epsilon_{A_{nz}}(t)]^T$. If one identifies the starting and ending instants t_1 and t_2 of the system, known variables are $V_1(t) = 0$ and $V_2(t) = 0$.

If we model the errors in the calculated inertial frame accelerations as constants i.e., $\epsilon_{A_n}(t) = \epsilon_{A_n} t$, then we have

$$\epsilon_{A_n} = \frac{\widehat{V_n}(t_2)}{t_2 - t_1} \tag{10}$$

where $\epsilon_{A_n} = [\epsilon_{A_{nx}}\ \epsilon_{A_{ny}}\ \epsilon_{A_{nz}}]^T$ is the constant error vector in each axis of the inertial frame and $V_n = [V_{nx}\ V_{ny}\ V_{nz}]^T$ is the computed velocity in the inertial frame using (5). It should be noted that $\widehat{V_n}(t_1) = 0$.

Analogously, if we model the acceleration errors as $\epsilon_{A_n}(t) = \epsilon_{A_n} t$, then parameters of the error model can be identified as

$$\epsilon_{A_n} = \frac{\widehat{V_n}(t_2)}{(t_2 - t_1)^2} \tag{11}$$

In either case, the histories of velocity and position are simply corrected by

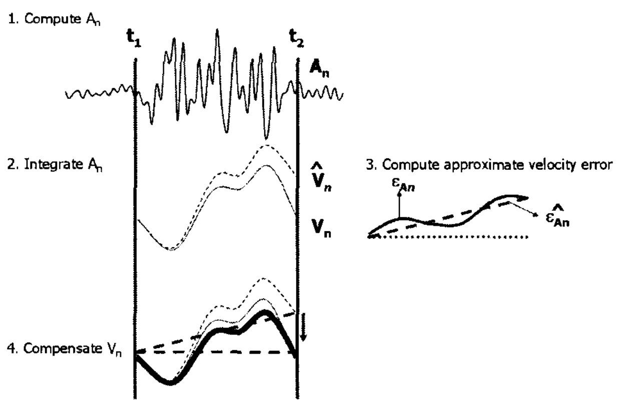


Fig. 6. Concept of the simplified ZVC

$$\begin{aligned} \widehat{V_{n,c}} &= \widehat{V_n}(t) - \int_{t_1}^{t} \epsilon_{A_n}(\tau) dtau \\ \widehat{P_{n,c}} &= \widehat{P_n}(t) - \int_{t_1}^{\tau} \int_{t_1}^{t} \epsilon_{A_n}(\tau) d\tau \end{aligned} \tag{12}$$

In order to apply the above described algorithm, the instants when the system is stationary should be accurately detected. In other words, we need an algorithm which extracts only valid motion interval of signals from the whole interval of measured signals. In this paper, the existence of the motion of the system is detected by checking the root of squared value of sum of angular rates [52].

$$m(t) = \begin{cases} 1 & \text{if } \sqrt{\omega_{bx}(\tau)^2 + \omega_{by}(\tau)^2 + \omega_{bz}(\tau)^2} \geq m_{th}, \\ 0 & otherwise \end{cases} \tag{13}$$

where $m(t)$ is the flag to denote the motion of the system at time t and m_{th} and t_{th} are a priori determined thresholds.

The measure (13) can be experientially justified by the fact that human arm movements usually involve translational motion and angular motion, which are measured by accelerometers and gyroscopes, respectively. In addition, we prefer gyroscope signals to accelerometer signals since they are less susceptible to the change of gravity.

The recovered motion information is three dimensional. Rather than construct a recognition system to process three dimensional motion information, we transform the motion information into the two dimensional information. This approach has an advantage that current available recognition systems, which assumes two dimension positional information as their inputs, can be used without modification. In addition, most users try to write on an imaginary two dimensional space although they are asked to make gestures in the three dimensional space [53, 54]. The algorithm proceeds as follows:

1) Solve the following matrix equation for α and β

$$\begin{bmatrix} \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i y_i \\ \sum_{i=1}^{m} x_i y_i & \sum_{i=1}^{m} y_i^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} z_i x_i \\ \sum_{i=1}^{m} y_i z_i \end{bmatrix} \tag{14}$$

where $P_i = (x_i, y_i, z_i)$, $i = 1, 2, \cdots, m$ are points of the recovered three dimensional trajectory and m is the number of points in the trajectory.

2) Determine a plane which has minimal z directional distance from points of trajectories

$$ax + by + cz = 0 \tag{15}$$

where $a = \alpha$, $b = \beta$, and $c = -1$.

3) Compute projected points $P_{i.p} = (x_{i.p}, y_{i.p}, z_{i.p})$ on the imaginary writing plane by

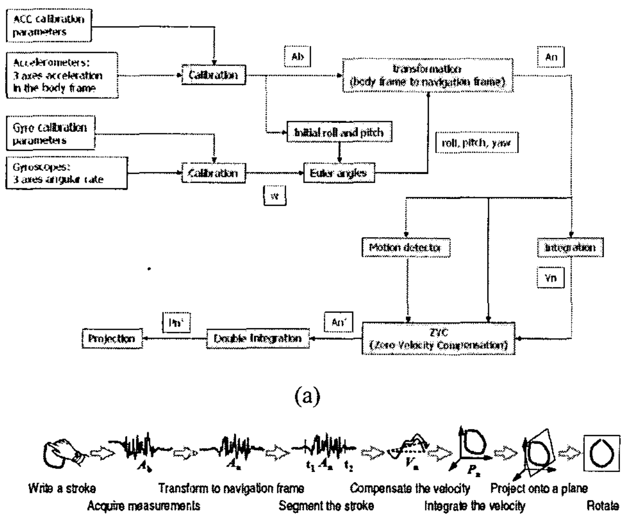
$$x_{i.p} = x_i - ka, \quad y_{i.p} = y_i - kb, \quad \text{and} \quad z_{i.p} = z_i - kc \tag{16}$$

where $k = \dfrac{ax_i + by_i + cz_i}{a^2 + b^2 + c^2}$

4) Compute 2D transformed points $P_{i.2D}$ by

$$P_{i,2D} = (x_{i,p}cos\theta + y_{i,p}sin\theta sin\phi + z_{i,p}sin\theta cos\phi,$$
$$y_{i,p}cos\phi - z_{i,p}sin\phi, 0)$$

where
$\phi = arctan2(-b, -c)$
$\theta = arctan2(a, \sqrt{b^2 + c^2})$
or
$\phi = arctan2(b, c)$
$\theta = arctan2(-a, \sqrt{b^2 + c^2})$
Figure 7 illustrates the proposed method.



(a)

(b)

Fig. 7. Illustration of the overall algorithm

3.2. Method 2: Recovery of user's handwriting motions based on acceleration measurements only [55]

The motion tracking methods presented in the previous sections require 6 DOF inertial measurements, which need totally five sensors in our study. Some real applications, however, cannot provide enough space to install all of the sensors. For example, if one tries to apply the proposed system to mobile phones, he/she may find that it is very difficult to locate enough space to install the five sensors. In this case, it is reasonable to develop another simple but effective algorithm to operate with least number of sensors. In this section, we propose a motion tracking algorithm with three acceleration measurements only. With this method, the number of required sensors is only two. The block diagram of the algorithm is depicted in Fig. 8. In the overall algorithm in Fig. 8, the estimation part of the rotation angles (roll ϕ, pitch θ, and yaw ψ) is different from the algorithm in Fig. 7. The other parts of the algorithm in Fig. 8 are identical with them in Fig. 7.

Empirically, we have found that the rotation angles (roll ϕ, pitch θ, and yaw ψ) do not change seriously during the gesture input period. Based on the empirical observation, we propose a simple linear approximation method to estimate the rotation angles with accelerometer only.

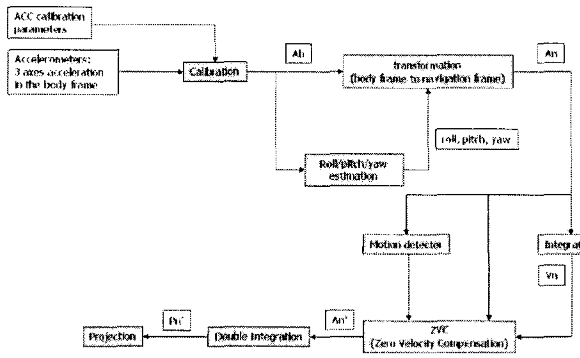


Fig. 8. Block diagram of Method 2.

By (7) and (8), we can calculate roll and pitch when the system does not move. If we assume that the system does not move before and after gesture input, it is possible to estimate roll and pitch angle by the following linear approximation method:

$$\phi(t) = at + b \tag{18}$$
$$\theta(t) = ct + d \tag{19}$$

where the constant a, b, c and d is given as

$$a = \frac{\phi(t_2) - \phi(t_1)}{t_2 - t_1}; b = -at_1 + \phi(t_1) \tag{20}$$

$$c = \frac{\theta(t_2) - \theta(t_1)}{t_2 - t_1} d = -ct_1 + \theta(t_1) \tag{21}$$

respectively. Using (18) and (19) we calculate roll ϕ and pitch θ in the interval of gesture input. From the experiments, we found that the yaw ψ angle did not affect the accuracy of the motion tracking algorithm severely. Therefore, we fixed $\psi = 0$. The estimation algorithms are summarized as follows:

$$\phi(t) = \begin{cases} \phi_1; t_0 \le t < t_1 \\ at + b; t_1 \le t < t_2 \\ \phi_2; t_2 \le t < t_e \end{cases}$$
$$\theta(t) = \begin{cases} \theta_1; t_0 \le t < t_1 \\ ct + d; t_1 \le t < t_2 \\ \theta_2; t_2 \le t < t_e \end{cases}$$
$$\psi(t) = 0; t_0 \le t < t_e$$

where constants ϕ_1, ϕ_2, θ_1 and θ_2 are in (20) and (21), and t_0 and t_1 are the starting and ending times of data acquisition, respectively. We assume that there is no motion when t is in $t_0 \le t < t_1$ and $t_2 \le t < t_f$.

Now we are in the position of applying INS equation (22) since we have three acceleration measurements and three Euler angles. It should be noted that governing equations for Euler angles in (5) are removed in (22).

$$\begin{aligned} \dot{P}_n &= V_n \\ \dot{V}_n &= C_b^n A_b - G \end{aligned} \tag{22}$$

3.3. Method 3: Recovery of user's handwriting motions based on two angular rate information

The presented algorithms in sections 3.1 and 3.2 recover the real three dimensional movements of the system. However, if we confine our interest within translating the user's motion into meaningful commands, i.e., gesture recognition, then the positional information is not the only source for the purpose.

The system can provide many sets of information regarding three dimensional movements such as acceleration and angular rates shown in the body frame and attitude, acceleration, velocity and position shown in the navigation frame, etc. The signals should satisfy a variety of requirements for the recognition. For example, features of signals from sensors should not vary from person to person for a specific gesture (low distinctiveness) while they provide good separability among different gestures (high intra class variation).

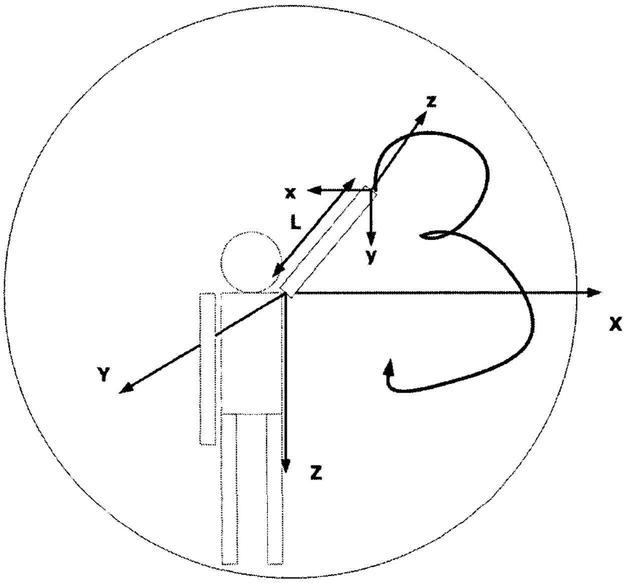
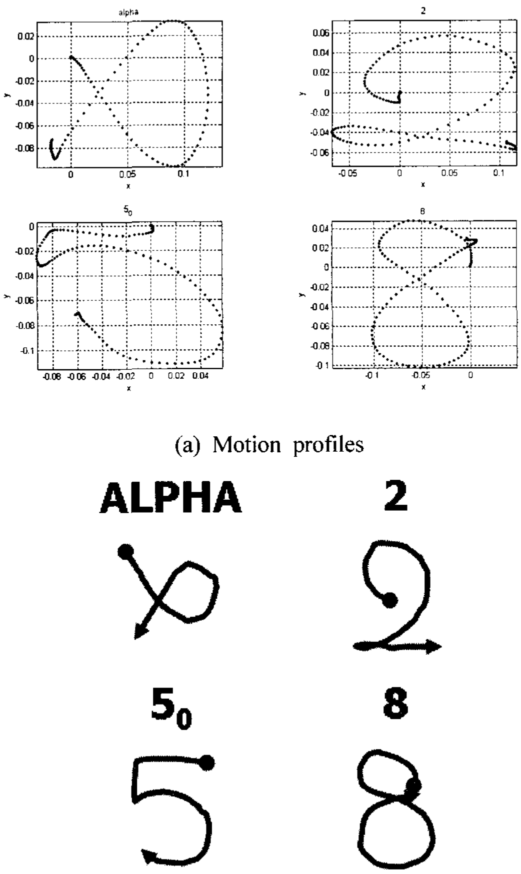


Fig. 9. Simplified kinematics of human arm movement



(a) Motion profiles

(b) Reference gestures

Fig. 10. Exemplary motion profiles by Method 3.

Among various physical motion quantities, we select rotation information as a good candidate gesture recognition since most of human arm movements consist of circular motions [56]. Thus, as shown in Fig. 9. In the figure, we know that rotational movement is dominant while z directional movement in the body frame is almost negligible. Therefore, we can roughly map the x and y axis angular rates to virtual 2D space as follows:

$$P_x = \int_0^t \omega_{bx(\tau)} d\tau \qquad (23)$$
$$P_y = \int_0^t \omega_{by(\tau)} d\tau$$

Figure 10 shows some exemplary motion profiles obtained by (23) with reference gesture set to help readers to verify the qualitative performance of the proposed scheme. In spite of less physical motion information, it is enough for a recognizer to classify different gestures. Detailed discussions of the performance of the proposed method will be given in the later of this paper.

4. Recognition

After the construction of motion information, the actual recognition process is applied to it, which is based on Bayesian network approach [57-59] for its high performance and robustness to random disturbances. It can more explicitly model basic strokes and their relationship compared to the conventional approaches such as template matching or hidden Markov model. In this framework, Bayesian networks [60] are adopted to probabilistically represent all the components and relationships in gestures.

In order to implement Bayesian network based recognizer, effective Bayesian network models of gestures are required. A gesture can be viewed as an aggregation of strokes which represents a nearly straight part of the whole gesture trace. Analogously, a stroke is regarded as a structural combination of points which merely contain two dimensional position information. Fig. 11 illustrates components of a gesture.

Bayesian network models relationship between components as well as components themselves. A point is represented by a two dimensional Gaussian distribution model, which corresponds to a single node in the Bayesian network gesture model. The stroke model is then constructed to reflect the

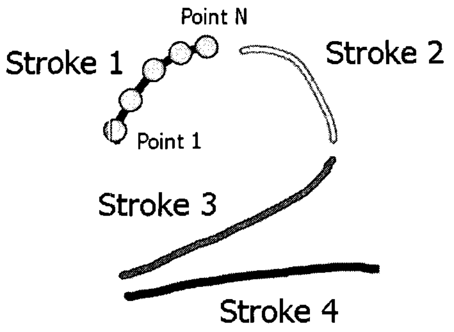
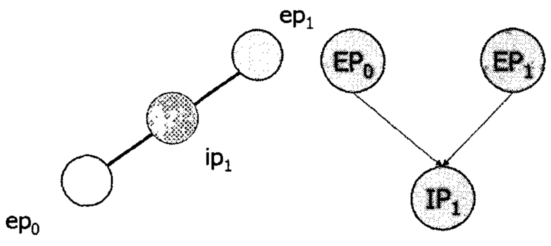


Fig. 11. Possible stroke segmentation for the number 2

point models and their inter relationships, i.e., within stroke relationships (WSRs). WSRs describe the dependency of a mid point from two specific points. Figure 12 illustrates the concept of WSRs and the corresponding Bayesian network. In Fig. 12, a stroke consists of two points ep_0 and ep_1. A mid point ip_1 is added and the resulting Bayesian network model is constructed as shown in Fig. 12 (b).



(a) (b)

Fig. 12. An example of WSR and its model using a Bayesian network

One can construct a Bayesian network model for a stroke as shown in Fig. 13 by recursively adding mid points and constructing WSRs until covariances of newly added point models are smaller than the predefined threshold.

The next step is to build a gesture model by merging stroke models and so called inter stroke relationships (ISRs), which describe writing sequence of strokes and dependencies

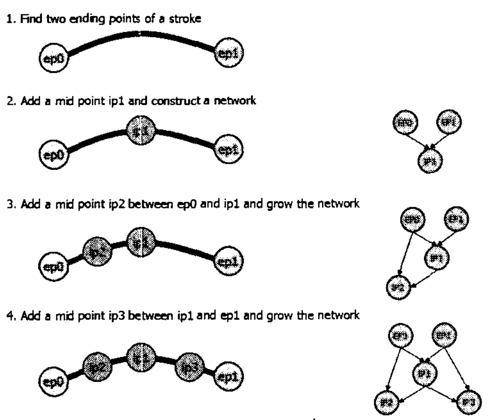


Fig. 13. Recursive construction of a stroke model

among end points of strokes. Fig. 14 shows the Bayesian network based gesture model with four strokes and stroke recursion depth of two, where EP_{i-1} and EP_i $(i = 1, 2, \cdots, 4)$ denote the beginning and ending point models for i th stroke, $IP_{i,j}$ are internal point models for ith stroke, arcs between EP_i's represent ISRs and incoming arcs to $IP_{i,j}$ describe WSRs.

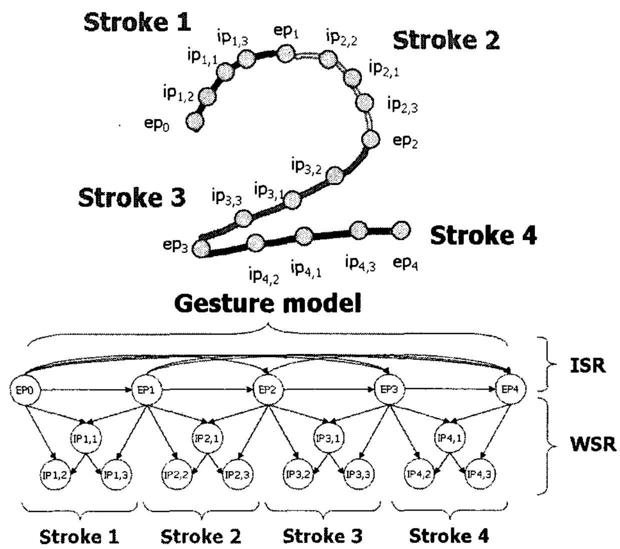


Fig. 14. Gesture model with 4 strokes and stroke recursion depth of 2

Generally speaking, the training of the above described Bayesian network involves two phases: structure determination and parameter identification. The former is related to finding a suitable structure of the Bayesian network, a proper recursion depth of stroke models and the number of strokes for each character. The latter is concerned with the adjustment of network parameters such as conditional probability parameters. In this paper, the structure of the Bayesian network is determined by a designer's a priori knowledge and parameters are identified by using the well known expectation and maximization (EM) algorithm.

After training, each gesture class m has its own gesture model BN_m $(m = 1, 2, \cdots, N)$, where N is the number of gestures. Two dimensional trajectory given by the presented motion tracking algorithms is fed to the gesture model with the form of sequence of points $O(1), \cdots, O(T)$. Then the recognition problem is to find the gesture model BN^* which produces the highest model likelihood for all possible stroke segmentations as follows:

$$\begin{aligned} BN^* = argmax_m P(BN_m | O(1), \cdots, O(T)) = \\ argmax_m P(BN_m) P(O(1), \cdots, O(T) | BN_m) \end{aligned} \tag{24}$$

where

$$P(O(1),\cdots,O(T)|BN_m) = \sum_{\gamma \in \Gamma} \prod_{i=1}^{N} P(EP_i = O(t_i)|O(t_0),\cdots,O(t_{i-1})) \prod_{i=1}^{N} \prod_{j=1}^{2^d-1} P(IP_{i,j} = ip_{i,j}(O(t_{i-1},t_i))|pa(IP_{i,j}))$$

and $\gamma = (t_0,\cdots,t_N)$, $t_0 = 1 < t_1 < \cdots < t_N = T$ is a stroke segmentation instance, Γ is the whole set of the possible stroke segmentations, $O(t_i,t_j) = (O(t_i), O(t_{i+1}),\cdots,O(t_j))$, N is the number of segments, d is the recursion depth and $pa(IP_{i,j})$ is the configuration of parent nodes from which arcs come to IP_i.

5. Experiments

In the actual experiment, 7 users are guided to write twenty four gestures defined in gesture set shown in Fig. 15. Each gesture is drawn twelve times. Therefore, we have 240 gestures per each user.

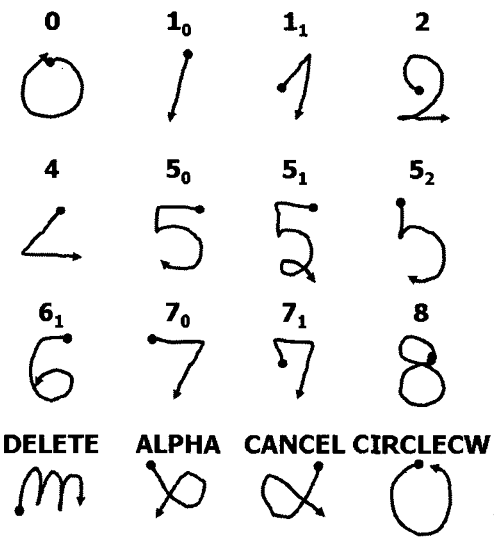


Fig. 15. Gesture set

Figure 16 illustrates the experiment procedure. As shown in the figure, user first inputs predefined set of gestures, signals of which are captured by data acquisition system and stored in database. The database is qualitatively and quantitatively analyzed by GPIEE and the result is stored in the analysis database.

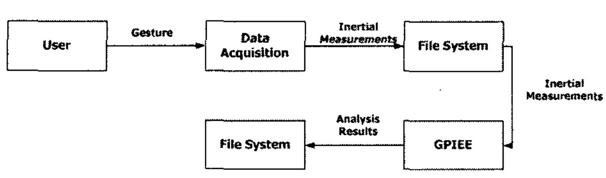


Fig. 16. Schematic diagram of the experimental environment involving a user

5.1. Method 1: Recovery of user's handwriting motions based on full 6 DOF inertial measurement

Typical position profiles estimated from the proposed system with ZVC, are displayed in Fig. 17 along with reference positions. Root mean of squared errors (RMSEs) of position estimation with ZVC for X, Y, and Z axis are 0.15175 (m), 0.09602 (m) and 0.04885 (m). Figure 18 shows three dimensional view of the recovered trajectories. Tables 5 - 6 shows RMSEs for each gesture and each user.

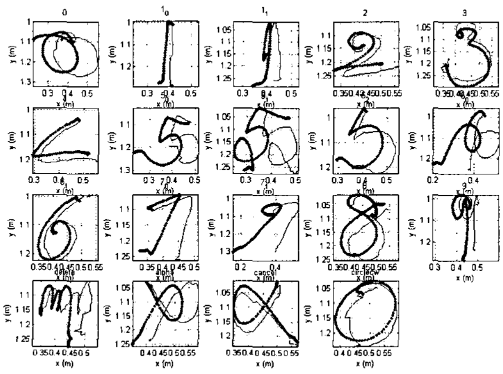


Fig. 17. Comparison of the estimated 2D trajectories by the proposed method (thick) and the reference 2D trajectories (solid)

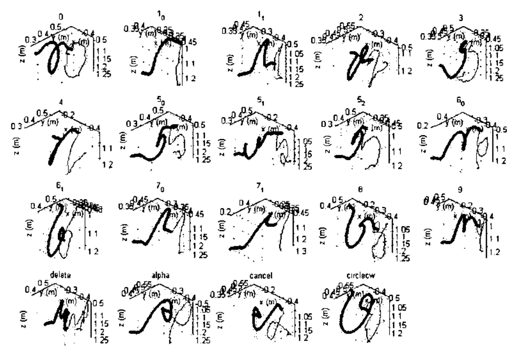


Fig. 18. Comparison of the estimated 3D trajectories by the proposed method (thick) and the reference 3D trajectories (solid)

Table 5. RMSE of x axis for the motion tracking method with full 6 DOF inertial measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.09297	0.13795	0.14604	0.05204	0.05384	0.03437	0.16825	0.09792
	1_0	0.05110	0.13450	0.12579	0.09917	0.07605	0.11113	0.15855	0.10804
	1_1	0.09132	0.13657	0.15144	0.10132	0.05070	0.63000	0.22553	0.19813
	2	0.12524	0.18132	0.09802	0.10861	0.05873	0.10203	0.36052	0.14778
	3	0.14009	0.17622	0.14932	0.11406	0.07538	0.11191	0.40204	0.16700
	4	0.05727	0.12877	0.10601	0.09778	0.05102	0.10556	0.23301	0.11135
	5_0	0.15205	0.18718	0.21433	0.11591	0.07611	0.13940	0.35302	0.17686
	5_1	0.22447	0.23934	0.26118	0.20301	0.11214	0.20507	0.52723	0.25320
	5_2	0.12009	0.14812	0.11929	0.09599	0.05986	0.10125	0.27330	0.13113
	6_0	0.09515	0.17870	0.18016	0.06538	0.10423	0.13861	0.30461	0.15241
	6_1	0.05944	0.11002	0.06592	0.10358	0.06886	0.04413	0.10735	0.07990
	7_0	0.09438	0.12676	0.13017	0.17506	0.04338	0.10606	0.30715	0.14042
	7_1	0.12931	0.15606	0.17380	0.20366	0.05486	0.09168	0.27120	0.15437
	8	0.10970	0.16400	0.16782	0.05792	0.06571	0.06609	0.27022	0.12878
	9	0.16995	0.17951	0.21088	0.16852	0.05017	0.13496	0.36249	0.18235
	delete	0.21041	0.16093	0.25966	0.26892	0.13370	0.12600	0.53919	0.24269
	alpha	0.06969	0.15803	0.15754	0.10597	0.05495	0.10896	0.25691	0.13029
	cancel	0.09999	0.15837	0.17900	0.14717	0.04374	0.38651	0.27738	0.18460
	circlecw	0.09047	0.13009	0.08690	0.05493	0.05505	0.04042	0.21493	0.09611
	Average	0.11490	0.15750	0.15701	0.12310	0.06781	0.14653	0.29542	0.15175

Table 6. RMSE of y axis for the motion tracking method with full 6 DOF inertial measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.09687	0.11617	0.11262	0.07624	0.06128	0.06352	0.20850	0.10503
	1_0	0.02764	0.04653	0.04003	0.04335	0.06580	0.02948	0.07280	0.04652
	1_1	0.02474	0.06031	0.03816	0.06379	0.04110	0.54366	0.12255	0.12776
	2	0.10649	0.07334	0.06678	0.05892	0.03975	0.07312	0.18314	0.08594
	3	0.03982	0.09351	0.05154	0.05399	0.10711	0.07520	0.19665	0.08826
	4	0.09736	0.09996	0.08075	0.06187	0.02776	0.08689	0.14195	0.08522
	5_0	0.10678	0.13974	0.08916	0.06863	0.11399	0.04277	0.21224	0.11047
	5_1	0.21510	0.15410	0.09744	0.13587	0.12874	0.09628	0.30535	0.16184
	5_2	0.09811	0.09826	0.06501	0.03856	0.09571	0.06743	0.14069	0.08625
	6_0	0.11626	0.14350	0.14474	0.11850	0.04353	0.08466	0.23325	0.12635
	6_1	0.05735	0.05860	0.04248	0.04636	0.04009	0.06659	0.08358	0.05643
	7_0	0.03625	0.07234	0.04456	0.06111	0.06891	0.09340	0.13286	0.07278
	7_1	0.04894	0.09577	0.04294	0.10035	0.05011	0.04261	0.13081	0.07307
	8	0.09803	0.12280	0.07907	0.05985	0.05330	0.04393	0.22382	0.09726
	9	0.10593	0.11402	0.08763	0.10670	0.04670	0.04264	0.24444	0.10687
	delete	0.08304	0.10453	0.09981	0.10045	0.20686	0.20405	0.16532	0.13772
	alpha	0.03910	0.09437	0.11402	0.10780	0.03472	0.06502	0.17636	0.08948
	cancel	0.09162	0.08174	0.10110	0.08890	0.06320	0.19181	0.12126	0.10566
	circlecw	0.04332	0.07120	0.03740	0.04068	0.07324	0.06025	0.10354	0.06137
	Average	0.08067	0.09688	0.07554	0.07510	0.07168	0.10386	0.16837	0.09602

Table 7. RMSE of z axis for the motion tracking method with full 6 DOF inertial measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.02151	0.03963	0.03917	0.02668	0.02332	0.02610	0.02370	0.02859
	1_0	0.02172	0.01385	0.08377	0.02027	0.03316	0.12394	0.01321	0.04428
	1_1	0.02292	0.03647	0.05672	0.02670	0.02716	0.11887	0.01747	0.04376
	2	0.03502	0.02806	0.08272	0.01805	0.02392	0.09274	0.02197	0.04321
	3	0.07147	0.03187	0.06704	0.02523	0.02196	0.06676	0.03896	0.04618
	4	0.03822	0.03102	0.08084	0.02497	0.02153	0.11983	0.01738	0.04769
	5_0	0.09797	0.04623	0.06270	0.02539	0.02989	0.06268	0.04055	0.05220
	5_1	0.14083	0.03545	0.08949	0.03085	0.02912	0.12151	0.02837	0.06795
	5_2	0.06945	0.01845	0.04091	0.01713	0.03865	0.06975	0.02356	0.03970
	6_0	0.11202	0.03101	0.06878	0.04187	0.03153	0.06592	0.01944	0.05294
	6_1	0.03564	0.03359	0.14057	0.02536	0.02153	0.10429	0.02331	0.05490
	7_0	0.02530	0.02533	0.04533	0.03331	0.01143	0.09636	0.01893	0.03657
	7_1	0.05150	0.03430	0.08191	0.02618	0.02159	0.10891	0.02111	0.04936
	8	0.03209	0.03842	0.06577	0.01980	0.01751	0.05514	0.04445	0.03903
	9	0.05258	0.02624	0.06690	0.02862	0.02577	0.12247	0.03232	0.05070
	delete	0.17847	0.04384	0.10339	0.04719	0.11332	0.11227	0.08687	0.09791
	alpha	0.07659	0.03114	0.07827	0.02614	0.04241	0.09304	0.01950	0.05244
	cancel	0.07609	0.03820	0.06985	0.02360	0.02489	0.10503	0.01480	0.05035
	circlecw	0.03377	0.02236	0.03091	0.02124	0.03011	0.04275	0.03128	0.03034
	Average	0.06280	0.03187	0.07132	0.02677	0.03099	0.08991	0.02827	0.04885

Next, we apply recognition system to the recovered trajectories. This test can be viewed as a qualitative analysis of the accuracy of the recovered trajectories. The generalized recognition rate is 75.9 %.

5.2. Method 2: Recovery of user's handwriting motions based on acceleration measurements only

Typical position profiles estimated from the proposed system with ZVC, are displayed in Fig. 17 along with reference positions. Root mean of squared errors (RMSEs) of position estimation with ZVC for X, Y, and Z axis are 0.40820 (m), 0.19998 (m) and 0.29663 (m). Figure 18 shows three dimensional view of the recovered trajectories. Tables 5 ~ 6 shows RMSEs for each gesture and each user.

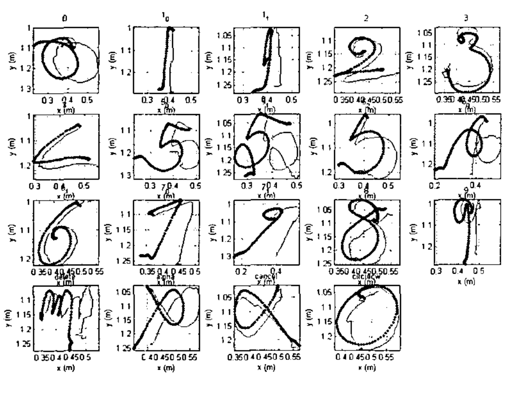


Fig. 19. Comparison of the estimated 2D trajectories by the proposed method (thick) and the reference 2D trajectories (solid)

Next, we apply recognition system to the recovered trajectories. This test can be viewed as a qualitative analysis of the accuracy of the recovered trajectories. The generalized recognition rate is 67.8 %.

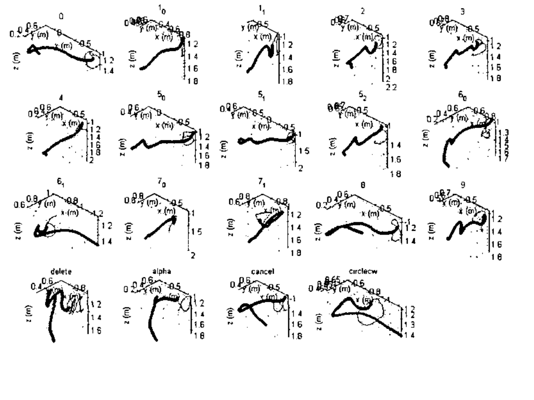


Fig. 20. Comparison of the estimated 3D trajectories by the proposed method (thick) and the reference 3D trajectories (solid)

Table 8. RMSE of x axis for the motion tracking method with 3 axis acceleration measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.09297	0.13795	0.14604	0.05204	0.05384	0.03437	0.16825	0.09792
	1_0	0.05110	0.13450	0.12579	0.09917	0.07605	0.11113	0.15855	0.10804
	1_1	0.09132	0.13657	0.15144	0.10132	0.05070	0.63000	0.22553	0.19813
	2	0.12524	0.18132	0.09802	0.10861	0.05873	0.10203	0.36052	0.14778
	3	0.14009	0.17622	0.14932	0.11406	0.07538	0.11191	0.40204	0.16700
	4	0.05727	0.12877	0.10601	0.09778	0.05102	0.10556	0.23301	0.11135
	5_0	0.15205	0.18718	0.21433	0.11591	0.07611	0.13940	0.35302	0.17686
	5_1	0.22447	0.23934	0.26118	0.20301	0.11214	0.20507	0.52723	0.25320
	5_2	0.12009	0.14812	0.11929	0.09599	0.05986	0.10125	0.27330	0.13113
	6_0	0.09515	0.17870	0.18016	0.06538	0.10423	0.13861	0.30461	0.15241
	6_1	0.05944	0.11002	0.06592	0.10358	0.06886	0.04413	0.10735	0.07990
	7_0	0.09438	0.12676	0.13017	0.17506	0.04338	0.10606	0.30715	0.14042
	7_1	0.12931	0.15606	0.17380	0.20366	0.05486	0.09168	0.27120	0.15437
	8	0.10970	0.16400	0.16782	0.05792	0.06571	0.06609	0.27022	0.12878
	9	0.16995	0.17951	0.21088	0.16852	0.05017	0.13496	0.36249	0.18235
	delete	0.21041	0.16093	0.25966	0.26892	0.13370	0.12600	0.53919	0.24269
	alpha	0.06969	0.15803	0.15754	0.10597	0.05495	0.10896	0.25691	0.13029
	cancel	0.09999	0.15837	0.17900	0.14717	0.04374	0.38651	0.27738	0.18460
	circlecw	0.09047	0.13009	0.08690	0.05493	0.05505	0.04042	0.21493	0.09611
	Average	0.11490	0.15750	0.15701	0.12310	0.06781	0.14653	0.29542	0.15175

Table 9. RMSE of y axis for the motion tracking method with 3 axis acceleration measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.12653	0.10639	0.30825	0.26917	0.17391	0.19636	0.15437	0.19071
	1_0	0.02688	0.04222	0.12149	0.11082	0.07771	0.08544	0.07013	0.07638
	1_1	0.04262	0.12552	0.13690	0.14515	0.11452	0.56379	0.06474	0.17046
	2	0.24079	0.11847	0.17213	0.20532	0.12431	0.17717	0.19371	0.17599
	3	0.28667	0.11690	0.17151	0.23413	0.18876	0.25600	0.16311	0.20244
	4	0.17505	0.10530	0.27890	0.15119	0.08025	0.20790	0.16889	0.16678
	5_0	0.43141	0.17582	0.40207	0.27323	0.33403	0.18111	0.28853	0.29803
	5_1	0.27169	0.17542	0.72191	0.27071	0.39097	0.33734	0.44578	0.37343
	5_2	0.34097	0.07742	0.22880	0.17990	0.15952	0.20147	0.15957	0.19252
	6_0	0.13575	0.16776	0.28025	0.17194	0.22664	0.26236	0.19966	0.20634
	6_1	0.30432	0.14310	0.20912	0.21688	0.16879	0.16317	0.14543	0.19297
	7_0	0.08399	0.11586	0.12472	0.15268	0.13788	0.12848	0.11764	0.12304
	7_1	0.12652	0.13652	0.17817	0.13488	0.16192	0.10574	0.08751	0.13304
	8	0.24555	0.15917	0.48383	0.33123	0.33917	0.26084	0.46342	0.32617
	9	0.11449	0.14373	0.33701	0.14821	0.26198	0.24612	0.25839	0.21571
	delete	0.13135	0.10963	0.35171	0.24634	0.25881	0.23057	0.27932	0.22968
	alpha	0.08409	0.12024	0.30160	0.15309	0.17277	0.24179	0.14518	0.17411
	cancel	0.15857	0.10152	0.28450	0.22156	0.22448	0.21190	0.18477	0.19819
	circlecw	0.15816	0.08314	0.17741	0.17008	0.14120	0.18169	0.16364	0.15362
	Average	0.18345	0.12232	0.27738	0.19929	0.19672	0.22312	0.19757	0.19998

Table 10. RMSE of z axis for the motion tracking method with 3 axis acceleration measurements

RMSE		User							
		User 1	User 2	User 3	User 4	User 5	User 6	User 7	Average
Character	0	0.02151	0.03963	0.03917	0.02668	0.02332	0.02610	0.02370	0.02859
	1_0	0.02172	0.01385	0.08377	0.02027	0.03316	0.12394	0.01321	0.04428
	1_1	0.02292	0.03647	0.05672	0.02670	0.02716	0.11887	0.01747	0.04376
	2	0.03502	0.02806	0.08272	0.01805	0.02392	0.09274	0.02197	0.04321
	3	0.07147	0.03187	0.06704	0.02523	0.02196	0.06676	0.03896	0.04618
	4	0.03822	0.03102	0.08084	0.02497	0.02153	0.11983	0.01738	0.04769
	5_0	0.09797	0.04623	0.06270	0.02539	0.02989	0.06268	0.04055	0.05220
	5_1	0.14083	0.03545	0.08949	0.03085	0.02912	0.12151	0.02837	0.06795
	5_2	0.06945	0.01845	0.04091	0.01713	0.03865	0.06975	0.02356	0.03970
	6_0	0.11202	0.03101	0.06878	0.04187	0.03153	0.06592	0.01944	0.05294
	6_1	0.03564	0.03359	0.14057	0.02536	0.02153	0.10429	0.02331	0.05490
	7_0	0.02530	0.02533	0.04533	0.03331	0.01143	0.09636	0.01893	0.03657
	7_1	0.05150	0.03430	0.08191	0.02618	0.02159	0.10891	0.02111	0.04936
	8	0.03209	0.03842	0.06577	0.01980	0.01751	0.05514	0.04445	0.03903
	9	0.05258	0.02624	0.06690	0.02862	0.02577	0.12247	0.03232	0.05070
	delete	0.17847	0.04384	0.10339	0.04719	0.11332	0.11227	0.08687	0.09791
	alpha	0.07659	0.03114	0.07827	0.02614	0.04241	0.09304	0.01950	0.05244
	cancel	0.07609	0.03820	0.06985	0.02360	0.02489	0.10503	0.01480	0.05035
	circlecw	0.03377	0.02236	0.03091	0.02124	0.03011	0.04275	0.03128	0.03034
	Average	0.06280	0.03187	0.07132	0.02677	0.03099	0.08991	0.02827	0.04885

Remark 1. The accuracy of the motion tracking and the recognition rate of Method 2 is quite disappointing. However, Method 2 is still appealing since accelerometers are usually cheaper and smaller than gyroscopes. Therefore, we adopted the sensor configuration of Method 2 in the commercialization process and have taken different path to recognizer user's gestures. Due to security, we briefly mention that the recognition rate is around 97.3 %.

5.3. Method 3: Recovery of user's handwriting motions based on two angular rate information

In this experiment only dual axis angular rate measurements are used to track and recognize a user's gesture. Since integration of gyroscope signals are only meaningful to recognition, we provide recognition rate only, which is 99.8 % and is acceptable for many commercial applications. Figure 21 illustrates the estimated position profiles obtained by Method 3. For the reader's convenience, corresponding reference 2D trajectories are also plotted.

Table 11 summarizes the comparison of characteristics of three proposed algorithms.

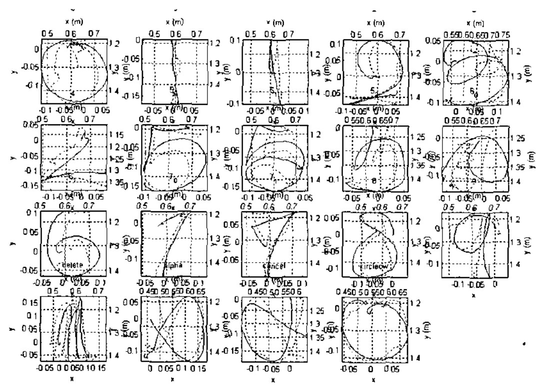


Fig. 21. Comparison of the estimated 2D trajectories by the proposed method (solid) and the reference 2D trajectories (dotted)

Table. 11. Comparison of the proposed motion tracking algorithms

	Method 1 (3A3G)	Method 2 (3A)	Method 3 (2G)
Physical motion information	Available	Available	Not Available
Recognition Rate	75.9 %	67.8 %	99.8 %
Number of sensors	5	1 or 2	2
RMSE of X axis (m)	0.15175	0.40820	Not Available
RMSE of Y axis (m)	0.09602	0.19998	Not Available
RMSE of Z axis (m)	0.04885	0.29663	Not Available

6. Application example

This section introduces a universal remote controller adopting the proposed input technology to show the possible applications of inertial sensing input systems, which is called the MagicWand [61].

Current universal remote controllers have many buttons and users usually are frustrated when they try to control multiple devices with the remote control system. In the perspective of remote controller, by replacing the current button rich remote controller with our system, users can control home appliances in very intuitive way. Besides the gesture recognition feature of the proposed input technology, the menu navigation and selection can be more comfortable since the proposed system is used as an alternative form of mouse. In addition, the text entry function of the proposed system is very useful for the high end digital TVs since many of them are equipped with internet connectivity and they need some form of text entry. Other remote controlling devices can perform same functions but our approach has an advantage in the fact that increasing function does not require increasing size of the device. Fig. 22 illustrates the conceptual usage scenario of the proposed gesture input device applied to the field of remote controller. Figure 22 and 23 illustrate the overview and configuration of the overall system.

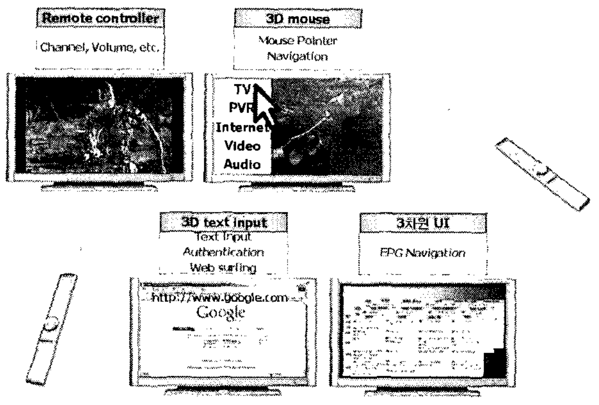


Fig. 22. Concept illustration of the gesture input device used as a universal remote controller

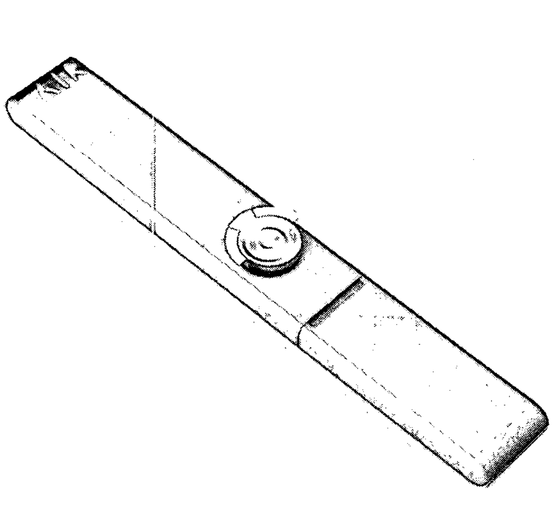


Fig. 23. MagicWand

The sensor configuration of the MagicWand is same with that used for the previous gesture based input device. The MagicWand is also equipped with a central processing unit for stand alone operation and an infrared emitter to control external devices. Figure 24 shows the system configuration of the MagicWand. Implementation of this specific example requires only good recognition rate. Therefore, we pick out motion tracking method with two gyroscopes. Furthermore, we utilize Fisher discriminant analysis (FDA) based recognizer instead of the previously presented Bayesian network based one due to inevitable memory limitation. However, the authors believe that the change of recognizer does not corrupt the fundamental concept of gesture input with all inertial approach.

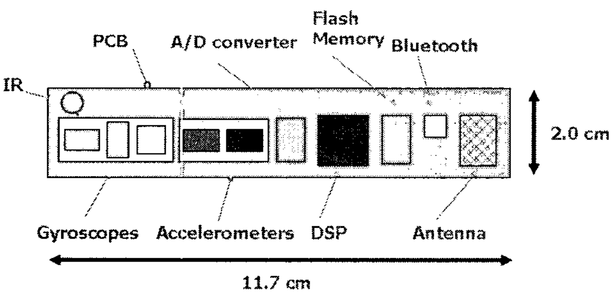


Fig. 24. Hardware configuration of the MagicWand

7. Conclusions

In this paper, we proposed a gesture based input device which is capable of capturing three dimensional motion. The system is equipped with accelerometers and gyroscopes. The sensors measure the inertial measurements, i.e., accelerations and angular velocities produced by pen's movement when a user is writing on a plane surface or three dimensional space. The gyroscope measurements are integrated to give orientation of the system and consequently used to compensate the accelerations. The compensated accelerations are doubly integrated to yield the position of the system. With this approach, a user's gesture inputs can be recovered without any external sensors. We have proposed three motion tracking algorithms to deal with various kinds of commercial applications. A Bayesian network based recognition system has been implemented for the proposed system. generate meaningful information. The evaluation of the proposed system has been conducted quantitatively and qualitatively. Position and angular accuracy of the proposed system has been adopted for the quantitative analysis while recognition rate has been utilized to verify the qualitative performance of the system. From the experiments, it has been shown that the performance of the system is acceptable, which leads to a promising way of commercialization of application of inertial sensing technologies. The proposed method is successfully implemented on an exemplary application system. Finally, it may be stated that algorithms and systems presented in this paper and other patents [62, 63] are partially applied to a commercial product SCH S310 which is developed by Samsung Electronics

References

[1] S. Keates and P. Robinson, "The use of gestures in multimodal input," in Proceedings of the third international ACM conference on Assistive technologies, Marina del Rey, California, United States, 1998, pp. 35-42.

[2] R. B. Dannenberg and D. Amon, "A gesture based user interface prototyping system," in Proceedings of the 2nd annual ACM SIGGRAPH symposium on User interface software and technology, Nov. 1989, pp. 127-132.

[3] C. G. Wolf, "Can people use gesture commands?" IBM, Tech. Rep. 86A002284, 1986.

[4] K. R. Wheeler and C. C. Jorgensen, "Gestures as input: neuroelectric joysticks and keyboards," IEEE Pervasive Computing, no. 2, pp. 56- 61, April-June 2003.

[5] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," IEEE Computer Graphics and Applications, pp. 24-38, November / December 2002.

[6] J. Morris, "Accelerometry - a technique for the measurement of human body movement," Biomechanics, pp. 729-736, 1973.

[7] N. Herbst and C. Liu, "Automatic signature verification," IBM Thomas J.Watson Research Center, Yorktown Heights, New York 10598, Tech. Rep. RC5810, Nov. 1975.

[8] A. D. Cheok, K. G. Kumar, and S. Prince, "Micro-accelerometer based hardware interfaces for wearable computer mixed reality applications," in Proceedings of ISWC2002, 2002.

[9] T. Miyagawa, Y. Yonezawa, K. Itoh, and M. Hashimoto, "Handwritten pattern reproduction using 3d inertial measurement of handwriting movement," Transactions of the Society of Instrument and Control Engineers, vol. 38, Jan. 2002.

[10] E. Kobayashi, K. Masamune, S. Tadokoro, M. Suzuki, T. Dohi, and D. Hashimoto, "Head mouse system using a gyro sensor - the man-machine interface for laparoscopic manipulator," in Proceedings of the 12th International Symposium and Exhibition on CAR'98. Computer Assisted Radiology and Surgery. Elsevier Science, 1998, pp. 742-746.

[11] H. Ishikawa, Y. Yonezawa, K. Itoh, and M. Hashimoto, "Extraction of character information from acceleration of handwriting pen," Transactions on the Institute of Electronics, Information and Communication Engineers D-I, vol. J76D-I, pp. 541-543, Oct. 1993.

[12] H. Nonaka and T. Da-Te, "The development of a pointing device with a cursor in inertial motion," Transactions of the Information Processing Society of Japan, vol. 31, no. 2, pp. 268-274, 1990.

[13] R. Gjone and J. Miraglia, "Pen for signature verification by acceleration and pressure," IBM Technical Disclosure Bulletin, vol. 24, p. 902, July 1981.

[14] N. M. Herbst and C. J. Liu, "Signature verification based on complete accelerometry," IBM Technical Disclosure Bulletin, pp. 4827-4829, May 1977.

[15] J. H. Morrissey, "Electronic calculator based on character recognition of input from stylus acceleration dynamics," IBM Technical Disclosure Bulletin, vol. 19, pp. 2816-2817, Dec. 1976.

[16] P. E. Jones, "Three-dimensional input device with six degree of freedom," Mechatronics, vol. 9, pp. 717-729, Oct. 1999.

[17] D. Strickland, A. Patel, C. Stovall, J. Palmer, and D. McAllister, "Self tracking of human motion for virtual reality systems," in Proceedings of SPIE - the International Society for Optical Engineering, vol. 2177, 1994, pp. 278-287.

[18] S. Lee, G. J. Nam, J. Chae, H. Kim, and A. J. Drake, "Two-dimensional position detection systems with MEMS accelerometer for mouse applications," in Proceedings of the 38th Design Automation Conference, June 2001, pp. 852-857.

[19] M. Haukijarvi and J. Yrjanajnen, "Accelerometer vibration analysis in finding out velocities," in Proceedings of EUSIPCO 2000, vol. 1, Sept. 2000, pp. 585-588.

[20] T. N. nd R. B. Yates, "A novel accelerometer for sensing pen tip motion," in Proceedings of the 4th European Conference on Smart Structures and Materials in conjunction with the 2nd International Conference on Micromechanics, Intelligent Materials and Robotics, Bristol, UK, 1998, pp. 709-714.

[21] B. Milner, "Handwriting recognition using acceleration-based motion detection," in IEE colloquium on document image processing and multimedia, London, UK, 25 March 1999, pp. 5/1-5/6.

[22] J. K. Perng, B. Fisher, S. Hollar, and K. S. J. Pister, "Acceleration sensing glove (ASG)," in Digest of paper. Third International Symposium on Wearable Computers. Los Alamos, CA, USA: IEEE Computing Society, Oct. 1999, pp. 178-180.

[23] N. M. Herbst and J. H. Morrissey, "Signature verification method and apparatus," U.S. Patent 3 983 535, Sept. 28, 1976.

[24] M. Epperson, "Autonomous computer input device and marking instrument," U.S. Patent 5 247 137, Sept. 21, 1993.

[25] Y. Sato, T. Inoue, E. Fujisawa, T. Kitaguchi, T. Furuta, N. Murata, and M. Shingyouchi, "Pen-shaped handwriting input apparatus using accelerometers and gyroscopes and an associated operational device for determining pen movement," U.S. Patent 5 981 884, Nov. 9, 1999.

[26] --, "Pen-shaped handwriting input apparatus using accelerometers and gyroscopes and an associated operational device for determining pen movement," U.S. Patent 5 902 968, May 11, 1999.

[27] T. J. Chainer, R. A. Scranton, and T. K. Worthington, "Data input pen for signature verification," U.S. Patent 4 513 437, Apr. 23, 1985.

[28] Seiko, "Seiko," Japan Patent JP H6-67 799, 1994.

[29] E. Baron, A. Prishvin, Z. Bar-Itzhak, and V. Korsensky, "Handwriting input apparatus for handwriting recognition using more than one sensing technique," U.S. Patent 5 517 579, May 14, 1996.

[30] S. Hidenori, "Pen type input device," Japan Patent Application 2000-330 717, Nov. 30, 2000.

[31] "Method and device for tracking an object," U.S. European Patent 02 071 324, Sept. 12, 2002.

[32] "Method for utilization of a gyroscopic or inertial device as a user interface mechanism for computer," U.S. European Patent 0 188 896, Nov. 22, 2002.

[33] G. J. Grimes, "Digital data entry glove interface device," U.S. Patent US 4 414 537, Nov. 8, 1983.

[34] C. Verplaetse, "Inertial proprioceptive devices: self-motion-sensing toys and tools," IBM Systems Journal, vol. 34, no. 3 & 4, pp. 639-650, 1996.

[35] K. Matsushima, "Coordinate detecting apparatus having acceleration detectors," U.S. Patent US 5 587 558, Dec. 24, 1996.

[36] B. J. Glynn, "Computer apparatus input device for three-dimensional information," U.S. Patent US 5 181 181, 1993 19, Jan.

[37] M. Tokioka, A. Tanaka, Y. Yoshimura, R. Yanagisawa, K. Kobayshi, and H. Sato, "Gesture input method and apparatus," U.S. Patent US 5 714 698, Feb. 3, 1998.

[38] T. Miyagawa, Y. Yonezawa, K. Itoh, and M. Hashimoto, "Handwriting pattern reproduction using 3D inertial measurement of handwriting movement," Transactions of the Society of Instrument and Control Engineers, vol. 38, pp. 1-8, Jan. 2002.

[39] M. LaBiche and A. Derouen, "Spatial input apparatus," U.S. Patent US 4 839 838, June 13, 1989.

[40] T. Kirabayashi, "Multidimensional signal input device," U.S. Patent US 5 329 276, July 12, 1994.

[41] C. W. C. Jr., T. Hovden, G. C. Smith, S. P. Hotelling, and M. K. Sabina, "Electronic pointing apparatus and method," U.S. Patent US 5 825 350, Oct. 20, 1998.

[42] A. Noureldin, D. Irvine-Halliday, and M. M. P. Mintchev, "Accuracy limitations of fog-based continuous measurement-while-drilling surveying instruments for horizontal wells," IEEE Transactions on Instrumentation and Measurement, vol. 51, no. 6, pp. 1177-1191, Dec. 2002.
[43] J. R. Huddle, "Trends in inertial systems technology for high accuracy AUV navigation," in Proceedings of the 1998 workshop on autonomous underwater vehicles, 1998. AUV'98, Aug. 1998, pp. 63-73.
[44] W. Chang, D. Y. Kim, K. H. Kang, E.-S. Choi, and W.-C. Bang, "Method and apparatus for inputting information spatially," Korea A 2004-0 107 144, Dec. 20, 2004.
[45] W. Chang, D. Y. Kim, K. H. K. and Eun-Seok Choi, and W.-C. Bang, "Apparatus and method for inputting information spatially," Korea A 2004-0 108 218, Dec. 23, 2004.
[46] E.-S. Choi, D. Y. Kim, W.-C. Bang, W. Chang, and K. H. Kang, "a input system based on three dimensional inertial navigation system and method for trajectory estimation thereof," Korea A 2004-0 082 128, Sept. 24, 2004.
[47] W. Chang, D. Y. Kim, K. H. Kang, E.-S. Choi, and W.-C. Bang, "Pen-shaped input device using inertial measurement units and method thereof," Korea A 2004-0 082 558, Sept. 30, 2004.
[48] J. S. A. Hepburn, J. Z. Zywiel, and D. B. Reid, "INS reset removal for enhanced post-flight performance analysis," in Position location and navigation symposium, 1990. Record. The 1990's - A decade of Excellence in the navigation sciences. IEEE PLANS'90., Las Vegas, NV, USA, Mar. 1990, pp. 226-230.
[49] W. Sohne, O. Heinze, and E. Groten, "Integrated INS/GPS system for high precision navigation applications," in IEEE 1994 Position Location and Navigation Symposium, Apr. 1994, pp. 310-313.
[50] B. Stieler, "Gyroscopic goniometry during quasistatic and kinematic conditions using a new and conventional methods for aiding," Optical gyros and their application, p. 13/1, 1999.
[51] W. C. Bang, D. Y. Kim, W. Chang, K. H. Kang, and E.-S. Choi, "Method and apparatus for compensating the acceleration error and inertial navigation system using thereof," Korea A 2004-0 107 968, Dec. 23, 2004.
[52] W.-C. Bang, D.-Y. Kim, W. Chang, K. H. Kang, and E.-S. Choi, "Motion recognition system capable of distinguishment a stroke for writing motion and method thereof," Korea A 2004-0 081 855, Sept. 23, 2004.
[53] W.-C. Bang, W. Chang, E.-S. Choi, K.-H. Kang, and D.-Y. Kim, "Design of self-contained gesture-based input device for human robot interface," Oct. 2004, to be published.
[54] W.-C. Bang, D.-Y. Kim, W. Chang, K. H. Kang, and E.-S. Choi, "Motion recognition system using a imaginary writing plane and method thereof," Korea A 2004-0 081 854, Sept. 23, 2004.
[55] E. S. Choi, W. Chang, W. C. Bang, J. Yang, S. J. Cho, J. K. Oh, J. K. Cho, and D. Y. Kim, "Development of the gyro-free handwriting input device based on inertial navigation system (ins) theory," in SICE Annual Conference in Sapporo, Hokkaido Institute of Technology, Sapporo, Japan, Aug. 2004, pp. 1176-1181.
[56] T. Sakaguchi, T. Kanamori, H. Katayose, K. Saot, and S. Inokuchi, "Human motion capture by integrating gyroscopes and accelerometers," in Proceedings of IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems, Dec. 1996, pp. 470-475.
[57] S. J. Cho and J. H. Kim, "Bayesian network modeling of hangul characters for on-line handwritten recognition," in Proc. Seventh International Conference on Document Analysis and Recognition, Edinburgh, Scotland, Aug. 2003, pp. 207-211.
[58] --, "Bayesian network modeling of strokes and their relationships for on-line handwriting recognition," Pattern Recognition, vol. 37, no. 2, pp. 253-264, Feb. 2004.
[59] S. J. Cho, J. K. Oh, W. C. Bang, W. Chang, E. S. Choi, Y. Jing, J. K. Cho, and D. Y. Kim, "Magic wand: A hand-drawn gesture input device in 3-d space with inertial sensors," in Proc. Samsung Tech. Conference 2004. Giheung, Kyonggi, Korea: SAIT, Nov. 2004, p. 35.
[60] F. Jensen, An Introduction to Bayesian Networks. New York: Springer, 1996.
[61] K. H. Kang, D. Y. Kim, W.-C. Bang, E.-S. Choi, and W. Chang, "Control device of electronic devices based on motion," Korea A 2004-0 081 270, Sept. 21, 2004.
[62] S. Park, "Apparatus and method for move recognizing of mobile phone," Korea Patent 2001-0 017 203, Mar. 5, 2001.
[63] D. H. Kang, J. H. Baek, W. B. Park, and K. H. Lee, "Method and device for executing pedometer in wireless terminal," Korea A 2004-0 093 811, Nov. 19, 2004.

Wook Chang
Wook Chang received the B.S., M.S., and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 1994, 1996, and 2001, respectively. Since September 2001, he has been with Samsung Advanced Institute of Technology, Suwon, Korea, where he is currently a Research Staff Member. He has worked in several areas, including capacitive touch sensing, inertial navigation, fuzzy modeling, genetic algorithms, neural networks, and fuzzy logic control. His current research interests are in the design and synthesis of sensory mobile devices especially equipped with touch sensors.

Phone: +82-31-280-6965
Fax: +82-31-280-9257
E-mail: wook.chang@samsung.com

Won-Chul Bang

Won-Chul Bang received the B.S. degree in electrical engineering from Hanyang University, Seoul, Korea, in 1994, and the M.S. and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1996 and 2001, respectively. In 2001, he was a Research Associate with the Human-friendly Welfare Robot System Engineering Research Center, KAIST. Then, he joined Samsung Advanced Institute of Technology, Yongin, Korea, in 2002. He has published 10 journal papers and 45 conference papers, and has 38 patents. His current research interests include human-computer interaction, soft computing techniques, pattern recognition, and inertial sensor applications. Dr. Bang is a member of Korean Institute of Electrical Engineers, Korea Fuzzy Logic and Intelligent Society, and the Institute of Control, Automation and Systems Engineers.

Phone: +82-31-280-6506
Fax: +82-31-280-9257
E-mail: wc.bang@samsung.com

Eun-Seok Choi

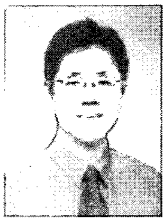
Eun-Seok Choi received the B.S. degree in control and instrumentation engineering from Kwangwoon University, Seoul, Korea, in 1996, and the M.S. and the Ph.D. degrees in mechatronics from Gwangju Institute of Science and Technology, Gwangju, Korea, in 1998, and 2002, respectively. Since 2002, he has been with Samsung Advanced Institute of Technology, Suwon, Korea, where he is currently a Senior Research Engineer. His research interests include robust control, nonlinear system control, robotics, inertial navigation system, and real-time control applications. His current research interests center on gesture interaction design and implementation using inertial sensors in mobile devices.

Phone: +82-31-280-1793
Fax: +82-31-280-9257
E-mail: eunseok.choi@samsung.com

Jing Yang

Jing Yang received the B.S. degree in Electrical Engineering and the Ph.D. degree in Precision Instruments and Machinery from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1996 and 2002, respectively. Since January 2003, she has worked as a Research Staff Member at Samsung Advanced Institute of Technology, Suwon, Korea. Her research area includes guidance, navigation and control, neural networks, signal processing and fault diagnosis. Her current research interests are in the motion-based interaction, especially using inertial sensor.

Phone: +82-31-280-1795
Fax: +82-31-280-9257
E-mail: jing.yang@samsung.com



Sung-Jung Cho

Sung-Jung Cho received the B.S., M.S., and Ph.D. degrees in Computer Science from Korea Advanced Institute of Science and Technology (KAIST), in 1996, 1998 and 2003 respectively. He was a research summer student at IBM Watson Research Center in 2002. Since on March 2003, he has been a research staff member at Samsung Advanced Institute of Technology (SAIT) and studying intelligent interaction on mobile phones with sensors and context information. His research interests include artificial intelligence and machine learning, especially handwritten character and gesture recognition based on Bayesian networks, hidden Markov models and support vector machines.

Phone: +82-31-280-9275
Fax: +82-31-280-9257
E-mail: sung-jung.cho@samsung.com

Jong-Koo Oh

Joon-Kee Cho received the M.S. degrees in sensor engineering from Kyungpook National University, Taegu, Korea, in 2001. Since October 2001, he has been with Samsung Advanced Institute of Technology, Suwon, Korea, where he is currently a Research Staff Member. He has worked in several areas, including intelligent sensor, active vision, neural networks, and mobile robot.

Phone: +82-31-280-1794
Fax: +82-31-280-9257
E-mail: handle.cho@samsung.com

Jong Koo Oh

Jong Koo Oh had worked as a Research Staff Member from May 2003 to September 2005.

Dong-Yoon Kim

Dong-Yoon Kim received Ph.D. degrees in mechanical engineering from Graduate School of Yonsei University, Seoul, Korea, in 1994. He is taking care of sensing strategic technology unit these days. He has been researching the gesture interactive technology since 2000 using the inertial sensor.