An Inertial Pen With Dynamic Time Warping Recognizer for Handwriting and Gesture Recognition

Ms.A.Elakya,ME.,
Assistant Professor,
Department of Electrical and Electronic Engineering,
Sri Krishna College of Technology,
Coimbatore,India.

UG scholar,
Department of Electrical and Electronic Engineering,
Sri Krishna College of technology
Coimbatore,India.

Abstract—This paper presents an inertial pen which is inertial-sensor based digital pen and associated with Dynamic Time Warping algorithm for handwriting and gesture recognition. With preferred handheld style and speed user holds inertial pen to write numerals or English lowercase letters and make hand gestures. The signals from the accelerometer and the gyro sensor are integrated and collected into a quaternion based complementary filter to reduce the integral errors from the gyro sensor which causes the intrinsic noises and signal drift are sometimes reduces the accuracy of the orientation estimation. The DTW-recognition algorithm procedure is composed of inertial signal acquisition, signal pre-processing, motion detection, template selection, and recognition. To obtain a superior class separation for improved recognition we have developed to minimal inter-class to maximal inter-class based template selection method. In the Experimental results, the effectiveness of the DTW-based recognition algorithm for online handwriting and gesture recognition using the inertial pen is successfully validated.

Keywords— Human Computer Interface (HCI), Inertial pen, dynamic time warping, quaternion-based complementary filter, handwriting recognition, gesture recognition.

I. ACCELEROMETER WITH GYROSENSOR

The sensors used in the inertial pen are the accelerometer embedded with a gyro sensor. Triaxial accelerometer is chosen as it measures the vibration in three axes X, Y and Z. They have three crystals positioned so that each one reacts to vibration in a different axis. The output has three signals, each representing the vibration for one of the three axes. The ACC301 has lightweight titanium construction and 10 mV/g output with a dynamic range of +/-500 g/s over a range of 3 to 10 kHz. Gyro sensors, are also know as angular rate sensor or angular velocity sensors, are devices that sense angular velocity. It can sense the rotational motion and also sense the changes in orientation, since the humans have the difficulty in sensing. In similar terms, angular velocity is the change in rotational angle per unit of time. It is generally expressed in Deg/s. The main three main application for gyro sensors which includes sensing the amount of angular velocity produced used in measuring the amount of motion itself, sensing angular velocity produced by the sensor’s own movement.
sensing vibrations produced by external factors, and transmits vibration data as electrical signals to a CPU.

**III. PROPOSED MODEL**

This proposed system, designs a wireless transmission using zigbee technology which controls PC with hand gestures (movement of hand). This is a low cost wireless technology which is very easy to operate or control PC. **CC2500** is wireless transmitter receiver. In this project, the input present at PORTD of transmitter atmega8 is transmitted wirelessly to the PORTD of receiver atmega8. The CC2500 RF module is a low-cost 2.4 GHz transceiver used in very low power wireless applications.

**V FUNCTIONAL DESCRIPTION OF INERTIAL PEN**

Our inertial pen consists of a triaxial accelerometer with gyro sensor, a triaxial magnetometer, a microcontroller, and an RF wireless transceiver (nRF24L01, Nordic). The accelerometer, gyro- scope, and magnetometer are used to detect accelerations, angular velocities, and magnetic signals generated by hand movements. The Accelerometer possesses a linear acceleration full scale of ±2g, ±4g, and ±8g, with data output rate from 0.75 Hz to 75 Hz. The accelerometer’s sensitivity is set from −4g to +4g in this study. The L3G4200D gyroscope simultaneously detects the X-, Y-, and Z-axis angular rates of the inertial pen, possesses a full scale of ±250, ±500, and ±2000 degree per second (dps) with data output rates from 100 Hz to 800 Hz, and is capable of measuring rates with a user-selectable bandwidth.

The microcontroller collects the digital accelerations, angular velocities, and magnetic signals, and transmits wirelessly the abovementioned inertial signals to a PC main processor via the RF wireless transceiver for further signal processing and analysis. The sampling rate of the abovementioned measurement signals is set at 75 Hz. Note that all signal processing procedure is performed on a PC. The overall power consumption of the hardware device is 30mA at 3.7V. The battery of the inertial pen is replaceable and rechargeable.

**VI. DYNAMIC TIME WARPING ALGORITHM**

Dynamic time warping (DTW) algorithm is developed to ensure a minimal cumulative distance between the aligned sequences and to find the similarity for the optimal alignment between two temporal sequences [25]. The DTW algorithm in the current paper is used to classify time sequences (movement signals) of different digits, letters, or gestures based on the nature of the movement signals generated from the handwriting and gesture trajectories. Most importantly, the training procedure of the DTW recognizer only needs one class template for each class. Thus using the DTW recognizer to handle recognition problems is potentially much simpler and faster, providing significant advantages for HCI devices.

The DTW process is described as follows: Let M and N be two similar temporal sequences with the same sampling rate and different lengths, where $M = [m_1, m_2, \ldots, m_p]$ and $N = [n_1, n_2, \ldots, n_q]$. First, a cumulative distance matrix $D \in R(p \times q)$, in which each element represents a mapping and alignment between $M(i)$ and $N(j)$, is constructed for measuring the similarity between the sequences M and N. The local distance

---

**Simplified Accelerometer Functional Block Diagram**

[Diagram of Accelerometer Functional Block]
d(mi, nj) is used to measure the local distance between the two temporal sequences and varies with different applications. In this paper, the local distance d(mi, nj) is defined as the Euclidean distance. The distance DTW(M, N) between the two temporal sequences is then calculated by summing the local dis-tances over the optimal warping. The computational complexity of this dynamic programming algorithm is O(pq).

VII. DYNAMIC TIME WARPING-BASED RECOGNITION ALGORITHM

A DTW-based recognition algorithm has been developed in this study to deal with inertial-sensing-based handwriting and gesture recognition, and is composed of the following procedures: 1) inertial signal acquisition, 2) signal pre-processing, 3) motion detection, 4) template selection, and 5) recognition. First, signals generated by hand movements are measured by the accelerometer, gyroscope, and magnetometer embedded in the inertial pen and then transmitted to a PC via the RF wire- less transceiver. Second, the effects of sensor uncertainty and the influence of users unconscious trembles are eliminated through the signal pre-processing procedure.

Third, the motion detection procedure is used to segment each inertial signal properly to extract an exact motion interval during which handwriting and gestures are performed. Subsequently, the inertial signals, including the filtered accelerations, compen- sated accelerations, velocities, and positions in the motion interval, are acquired for the recognition tasks. The Z-score method is utilized to reduce the signal amplitude biases of each waveform for each movement signal due to individual differences of writing speeds or styles.

Fourth, a Min-Max template selection method is used to sieve out reliable class templates for each class of sequences (i.e., movement signals) during the training phase of the DTW recognizer. Finally, the recognition results are processed through the DTW recognizer by measuring the similarity between the testing data and the selected class templates. The block diagram of the proposed DTW-based recognition algorithm is shown in Fig. 3. We now introduce the detailed procedures of the proposed DTW-based recognition algorithm.

A. Signal Preprocessing Since the measured signals are always contaminated not only by the sensors’ error sources but also with users’ uncon- scious trembles, signal preprocessing composed of calibration and a moving average filter is an essential procedure after inertial signal acquisition [22], [23].

1) Calibration: The accelerations, angular velocities, and magnetic signals are calibrated to reduce sensitivity and offset errors from the raw signals. When the inertial pen is stationary, the triaxial accelerometer measures the gravitational acceler- ation only. On the basis of this fact, we align each axis of the triaxial accelerometer with the Earth’s gravity to calibrate the accelerometer. To execute the calibration, we first place the triaxial accelerometer on a level surface and then point each axis alternately upward and downward.

2) Moving Average Filter: The second step of the signal preprocessing procedure is to reduce the high-frequency noise from the calibrated signals by using a moving average filter angular is the number of points in the average filter. In this study, we set N = 5 based on our empirical tests.

B. Motion Detection The motion detection procedure involves the following steps: 1) segmentation, 2) movement signal acquisition, and 3) normalization. We describe each step in detail as follows. 1) Segmentation: After filtering the measurements, we first segment each inertial signal properly to extract a precise motion interval, since the size of measurements of each move- ment frequently differs between fast and slow writers. For this study, we segment the motion intervals of all inertial signals to obtain the accurate locations of the start and end points of each measurement based on an adaptive magnitude threshold generated from the filtered acceleration signal. We determine the value of the threshold by using the filtered acceleration during the time steps at the beginning of a motion, denoted as kts. Note that the filtered acceleration includes the gravitational acceleration that should be subtracted from the filtered acceleration for avoiding the effect of gravity. Then, we set the multiple of the mean value of MAb(k) in the interval of kts as the threshold. K is an empirical value (K = 2 is used in this study). Once we obtain the threshold, the motion interval can be determined by selecting the start and end points (or time steps) whose magnitudes are higher and lower than the threshold, respectively. Finally, the total time steps when performing handwriting and hand gestures can be partitioned into two time intervals: a non-motion interval and a motion interval.

2) Movement Signal Acquisition: The movement signal acquisition step includes following substeps: 1) orientation estimation, 2) coordination transformation and gravity compensation, and 3) velocity and position estimation, which are elaborated as follows. a) Orientation estimation: Once the non-motion and motion intervals are separated during the segmentation step, we can calculate the orientation angles within those two intervals. The purpose of the orientation estimation for the non-motion interval is to obtain the initial orientation angles for the motion interval. The signals measured from the accelerometer and magnetometer are utilized to estimate the orientation angles during the non-motion interval since the initial orientation angles cannot be directly obtained from the signals of the gyroscope. The roll angle (ϕ) and pitch angle (θ) of the inertial pen can be derived directly from the filtered accelerometer signals [18]. After obtaining the initial Euler angles (ϕ, θ, Ψ), we can compute the parameters of the initial quaternion representation. Generally, the orientation angles in the motion interval can be obtained through the single integral of the filtered angular velocities measured by the gyroscope. Once the initial quaternion of the pen’s orientation is obtained in the non-motion interval, we can obtain the quaternion representing the orientation angles at each time step within the motion interval by using the filtered angular velocities.

In this paper, the quaternion-based complementary filter is employed to integrate the filtered accelerations, angular veloc- ities, and magnetic signals for correcting the orientation of the inertial pen in the motion interval. The flowchart of the
quaternion-based complementary filter is shown in Fig. 4. However, the filtered accelerations may always contain acceleration signals generated from hand trembles. The filtered magnetic signals may also be influenced by ferrous interference in the environment. It is necessary, therefore, to decide whether or not the filtered accelerations and magnetic signals will be used in the quaternion-based complementary filter through setting empirical magnitude thresholds where k denotes the time steps. Note that the filtered magnetic signal in includes the Earth’s magnetic field which is normalized to 1 in this paper and should be subtracted from the filtered magnetic signals for avoiding the effect of Earth’s magnetic field. The empirical magnitude thresholds of the filtered acceleration and magnetic signals established in this paper are THAb =0.01g and THmb =0.3 gauss, respectively. When the magnitude of the filtered acceleration (MAb(k)) is lower than the magnitude threshold of the filtered acceleration (THAb), this indicates that the accelerometer detects only the gravitational acceleration. The accelerations will be used to calculate the roll angle (ϕ) and pitch angle (θ), which are then transformed by the quaternion representation (qA (k)). When the magnitude of the filtered magnetic signal (Mmb(k)) is lower than the magnitude threshold of the filtered magnetic signal (THmb), this indicates that the outputs of the magnetometer are not influenced by ferrous materials and are reliable for estimating the yaw angle (ψ). The obtained yaw angle will then be transformed by the quaternion representation. The orientation of the inertial pen estimated by the gyroscope through the quaternion-based complementary filter. Once the estimated orientation of the inertial pen is obtained by the quaternion-based complementary filter, we must first derive a transformation matrix based on the quaternion via equation to transfer all the filtered acceleration signals from the body coordinate to the reference coordinate. Due to the fact that the transformed acceleration in the reference coordinate is composed of both the gravitational acceleration and motion acceleration, the gravitational acceleration must then be subtracted from the transformed acceleration to obtain the compensated acceleration generated by the movements alone by movements in the reference coordinate. Velocity and position estimation: The estimated velocity of the inertial pen during a motion can be obtained through the single integral of the compensated acceleration in the motion interval as follows. However, the integration of the drift errors generated from the accelerometer causes a cumulative error in the velocity, which becomes extremely large after a time interval. Particularly, the velocity must be zero during the non-motion intervals, which can be used to compensate the error of the velocity within the motion interval. In this paper, we utilize the zero velocity compensation (ZVC) method to compensate the error of the velocity, which models the velocity error accumulation through a linear function [3]. At first, the estimated velocity in the non-motion interval must be set to zero. Then by finding the difference between the first and last velocities of the motion interval, the slope of the linear error model can be obtained by equation. The offset term of the linear error model is set to be equal to zero. Once the linear error model is derived, the following equation can compensate the error caused by the drift of the acceleration. Note that, in this paper, the movement signals used for the recognition tasks include the filtered accelerations (Ab), compensated accelerations (A), velocities (v), and positions (p). 3) Normalization: The signal amplitude biases of the filtered accelerations, compensated accelerations, velocities, and positions are generally inconsistent due to individual differences of writing speeds or styles. In order to avoid extreme amplitude scaling, normalization of the amplitude of the abovementioned movement signals is required to ensure that the DTW distance calculated by utilizing the local distance measurement is representative. This paper utilizes the Z-score method to normalize the movement signals by first subtracting the mean value of each movement signal from each movement signal to eliminate the offset effect, which is then divided by its standard deviation.

C. Template Selection for the DTW Recognizer Recognition performance greatly depends on the quality of the selected class templates. During the training phase of the DTW recognizer, it is most important to select reliable training class templates for recognition from all the templates within the same class. In this paper, we have developed a minimal intra-class to maximal inter-class based template selection method (Min-Max template selection method) to perform the template selection task. This approach utilizes both intra-class and inter-class DTW distances to select the reliable patterns for the class templates. The intra-class DTW distance is calculated as the sum of the DTW distance between the template and all other patterns within the same class, while the inter-class DTW distance is calculated as the sum of the DTW distance between the template and all other patterns from the different classes. The class template is chosen according to how well it represents its own class by minimizing the intra-class DTW distance and maximizing the inter-class DTW distance. First, we calculate the mean values and standard deviations of the intra-class and inter-class DTW distances. C template is the optimal class template for each class. In other words, by using the DTW recognizer there is only one training pattern or class template for each digit, English lowercase letter, or gesture.

D. Recognition by the DTW Recognizer Once each class template of each digit, English lowercase letter, or gesture is selected, the similarity between each class template and the movement patterns will be measured through the DTW recognizer. Since each movement pattern, including the filtered accelerations, compensated accelerations, velocities, and positions, is composed of three signal sequences (X-, Y-, and Z-axis), the distance DTW, which denotes the similarity between the class template Ci of size ix3 and the testing pattern Tj of size j ×3, is computed. Finally, the minimal DTW(C,T) represents that the testing pattern Tj and the class template Ci are within the same class.

VIII. EXPERIMENTAL RESULTS
In this section, the effectiveness of the inertial pen and its associated DTW-based recognition algorithm is validated by...
three experiments: 1) handwritten digit recognition, 2) handwritten English character recognition, and 3) gesture recognition a minimum selection method with a DTW recognizer [24], this paper adopts a Min-Max template selection method for obtaining superior class separation and improved recognition performance of 2D and 3D handwritten digits, 2D handwritten English lowercase letters, and 3D hand gestures. The proposed DTW-based recognition algorithm consists of the following procedures: inertial signal acquisition, signal preprocessing, motion detection, template selection, and recognition. We collected movement signals for the three experiments from ten subjects (3 females, 7 males; aged 23.5±2.0 years old) in a laboratory environment. Due to the size of the inertial pen, participants were asked to practice writing with the pen before the experiment. Once they felt comfortable with the pen, data for each type of movement were collected. In addition, we compared the recognition results of the filtered accelerations, compensated accelerations, velocities, and positions (trajectories) separately to identify reliable movement signals for the online handwriting and gesture recognition tasks. The digital output signals of the accelerometer, gyroscope, and magnetometer are all sampled at 75 Hz. Our experiments were performed on a PC running Microsoft Windows 7 operating system with an Intel® Core Processor i5-2400 and 8-GB RAM.

A. 2D/3D Handwritten Digit Recognition For handwritten digit recognition, two experiments were designed to demonstrate the effectiveness of the proposed Inertial Pen and its associated DTW-based recognition algorithm. In these experiments, each participant was invited to hold the inertial pen and draw Arabic numerals (shown in Fig. 7) in a 2-dimensional (2D) space and a 3D space. Each participant was asked to write 10 digits (from 0 to 9) with each digit written 10 times in each space condition. Therefore, a total of 1000 (10×10×10) data were generated for each experiment. 1) 2D Handwritten Digit Recognition: This experiment required that the pen-tip touch a table in the writing of Arabic numerals. The performance comparison of our proposed Min-Max template selection method and two existing methods, the random selection method and the minimum selection method [9], is summarized in Table V.

IX. ALGORITHM USING VELOCITY SIGNALS
The proposed Min-Max template selection method with the DTW recognizer outperforms others for all movement signals. From Table II, the overall recognition rates obtained by leave-one-out cross-validation were about 79.1%, 93.7%, 97.9%, and 85.6% for 2D handwritten digit recognition using the filtered accelerations, compensated accelerations, velocities, and positions, respectively. Obviously, the DTW-based recognition algorithm using the velocity signals outperformed other movement signals. The worst recognition rate appears in the filtered accelerations since the filtered acceleration in the body coordinate is influenced by gravitational acceleration during the recognition tasks, which was not generated from the hand motions. In addition, the recognition rate using velocities is higher than that of the compensated accelerations and positions because the velocities are compensated by the ZVC method while the others cannot be compensated by any method or natural phenomenon. To further investigate the robustness of the proposed method, we evaluated the recognition performance by 2-fold cross-validation, 5-fold cross-validation, 10-fold cross-validation, and leave-one-out cross-validation strategies. The recognition rates are shown in Table IV. Moreover, the user-dependent 2D handwritten digit recognition using velocity signals achieved about 99.4% accuracy by leave-one-out cross-validation, as shown in Table V. 2) 3D Handwritten Digit Recognition: In this experiment, the ten participants were asked to hold the inertial pen to write 10 digits (shown in Fig. 7) without any ambit restriction in a 3D space. The same validation procedure as that of the first experiment was conducted for the 3D handwritten digit trajectories. The overall recognition rates evaluated by leave-one-out cross-validation were 79.5%, 83.1%, 87.3%, and 80.8% for 3D handwritten digit recognition using filtered accelerations, compensated accelerations, velocities, and positions, respectively, as shown in Table II. Similar to the 2D findings above, the recognition rate of the DTW-based recognition algorithm with velocity signals is better than other movement signals. Recognition of 87.3% was found for the proposed Min-Max template selection method, compared to a 74.8%, and 82.9% recognition rate. Fig. 8. Pictorial trajectories of English lowercase letters.

for the random and minimum template selection methods, respectively. The proposed Min-Max template selection method with the DTW recognizer using velocity signals outperformed other methods. The recognition rates obtained by 2-fold cross-validation, 5-fold cross-validation, 10-fold cross-validation, and leave-one-out cross-validation strategies are shown in Table IV. From Table V, the user-dependent recognition rate for 3D handwritten digit recognition evaluated by leave-one-out cross-validation was 94.6%. Thus, the results validate that the proposed Min-Max template selection method with the DTW recognizer can serve as an effective tool for 3D handwritten digit recognition.

B. 2D Handwritten English Character Recognition This experiment was designed to demonstrate the effectiveness of the proposed inertial pen and its associated DTW-based recognition algorithm for recognizing handwritten English lowercase letters. Ten participants were asked to hold the inertial pen and to draw English lowercase letters in a 2D space. Pictorial trajectories of English lowercase letters are shown in Fig. 8. Each participant was asked to write 26 letters (from a to z), and each letter was to be written 5 times for this experiment. Therefore, a total of 1300 (=26×10×5) data were collected for this experiment. In addition, recognition performance was evaluated when 2D digits and 2D English characters were written simultaneously. Ten participants were asked to hold the inertial pen and draw Arabic numerals and English lowercase letters in a 2D space. Each digit and letter was to be written 5 times. Therefore, a total of
1800(=36×10×5) data were collected for this experiment. From Table II, the best recognition rate obtained by leave-one-out cross-validation was about 92.1% for 2D handwritten digit and 2D English characters recognition using the velocities.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

| → | → | ↓ | ↓ | ← | ← | ← | ← |

**Trajectories of eight hand gestures.**

The user-dependent recognition rate for recognizing 2D digits and 2D English characters evaluated by leave-one-out cross-validation was 93.0%, as shown in Table V.

C. 3D Gesture Recognition

In the third experiment, the participants were invited to hold the inertial pen and perform eight hand gestures in a 3D space. The trajectories of the eight hand gestures. The participants were asked to repeat each of the hand gestures 10 times. Hence, a total of 800 (=8×10×10) hand gestures were generated. The same validation procedure as that of the first experiment was conducted for the gesture motions signals. Table II shows that the proposed DTW-based recognition algorithm using velocities can effectively recognize different hand trajectories that can be defined as various commands for HCIs. From Table III, the recognition rate of the proposed Min-Max template selection method with the DTW recognizer using velocity signals is superior to alternative methods. The overall user-independent and user-dependent recognition rates evaluated by leave-one-out cross-validation were 98.1% and 99.8%, as shown in Tables II and V. As shown in Table IV, the recognition rates obtained by multiple cross-validation strategies ranged from 82.3% to 98.1%.

**References**


