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

L.M.MerlinLivingston[#1], P.Deepika[#2], M.Benisha[#3]

[#1]Professor,[#2]Assistant Professor, [#3]Assistant Professor, Department of ECE, Jeppiaar institute of technology
Jeppiaar Nagar, Kunnam, Sriperumbudur, Chennai, India

Abstract—*This paper presents an inertial-sensor-based digital pen (inertial pen) and its associated dynamic time warping (DTW)-based recognition algorithm for handwriting and gesturer recognition. Users hold the inertial pen to write numerals or English lowercase letters and make hand gestures with their preferred handheld style and speed. The inertial signals generated by hand motions are wirelessly transmitted to a computer for online recognition. The proposed DTW-based recognition algorithm includes the procedures of inertial signal acquisition; signal preprocessing, motion detection, template selection, and recognition. We integrate signals collected from an accelerometer, a gyroscope, and a magnetometer into a quaternion-based complementary filter for reducing the integral errors caused by the signal drift or intrinsic noise of the gyroscope, which might reduce the accuracy of the orientation estimation. Furthermore, we have developed minimal intra-class to maximal inter-class based template selection method (min-max template selection method) for a DTW recognizer to obtain a superior class separation for improved recognition. Experimental results have successfully validated theeffectiveness of the DTW-based recognition algorithm for online handwriting and gesture recognition using the inertial pen.*

KeyWords—Inertial pen, dynamic time warping, quaternion-based complementary filter, handwriting recognition,Gesture recognition.

I. INTRODUCTION

With the rapid development of computer technology, contemporary human-computer interaction (HCI) devices/techniques have become indispensable in individuals'daily lives. HCI devices/techniques have also dramatically altered our living habits with computers, consumer electronics, and mobile devices. The ease with which an HCI device or technique can be understood and operated by users has become one of the major considerations when selecting such a device. Therefore, it is necessary for researchers to develop advanced and user-friendly HCI technologies which are able to effortlessly translate users' intentions into corresponding commands without requiring users to learn or accommodate to the device. Technologies are being developed which are able to intuitively express users' intentions, such as handwriting, gestures, and human body language, to naturally control HCI devices. These technologies have manpplications in the fields of remote control, virtual reality, sign language, signature authentication, sport science, health care, and medical rehabilitation. Recently, a number of researchers have developed diverse technologies for inertial-sensing-based HCI methods such as activity recognition, gesture recognition, handwriting recognition, and motion tracking. Among inertial-sensing-based HCI methods, pen-based input devices embedded with accelerometers and/or gyroscopes can most easily provide intuitive expressions through capturing translational accelerations and/or angular velocities generated by hand movements. Most importantly, inertial-sensing-based pen-based input devices for recognizing handwritten characters and hand gestures can be operated without ambit limitations such as writing ranges, directions, or dimensions, while other pen-based devices such as electromagnetic and pressure types must limit the writing space. The major challenge of inertial-sensing-based handwriting and gesture recognition using acceleration or angular velocity signals is misrecognition, since different users have different preferred speeds and styles. Recent studies have shown that hidden Markov model (HMM) and neural network approaches are effective at increasing the recognition rate of the inertial sensing- based handwriting and gesture recognition. However, the computational complexity of HMMs and neural classifiers are directly proportional to the dimension of the feature vectors, and both require more than one training sample to obtain acceptable recognition rates. While some researchers have demonstrated the effectiveness of the DTW algorithm, which selects the best match from many samples for each class for recognition, most of these studies was based on accelerometer-based gesture recognition alone. For example Liu *et al*, described an accelerometer-based gesture recognitionsystem for categorizing 3700 samples collected from seven subjects. The system employed the DTW with affinity propagation methods to obtain class templates for each gesture during the training phase. The accuracy for recognizing eight gestures reached 96.84% and 100% for user-independent and user-dependent recognition, respectively.

Liu *et a[14]l.* used a triaxial accelerometer with a DTW algorithm for personalized gesture recognition. Over 4000 samples with eight gestures collected from eight users were utilized for user-dependent recognition with 98.6% accuracy. In this paper, an inertial-sensor-based digital pen (inertial pen) and a dynamic time warping (DTW)-based recognition algorithm is presented for both handwriting and gesture recognition tasks. The portable inertial pen is composed of a triaxial accelerometer, a triaxial gyroscope, a triaxial magnetometer, a microcontroller, and an RF wireless transmission module. Users can utilize this inertial pen to write numerals or English lowercase letters, and make hand gestures at their preferred speed without any space limitations. Measuredaccelerations, angular velocities, and magnetic signals are transmitted to a personal computer (PC) via the RF wireless module.

The proposed DTW-based recognition algorithm is composed of the procedures of inertial signal acquisition,signal preprocessing, motion detection, template selection, and recognition. In the proposed recognition algorithm, we utilize the zero velocity compensation (ZVC) method and a quaternion-based complementary filter to reduce the integral errors caused by the intrinsic noise/drift of the accelerometer and gyroscope, which worsen the accuracy of the velocity, position, and orientation estimations. Furthermore, we have developed a minimal intra-class to maximal inter-class based template selection method (Min-Max template selection method) for a DTW recognizer to obtain a superior class separation for improved recognition. The advantages of this approach include the following: 1) with the inertial pen, users can deliver diverse commands through hand motions to control electronic devices anywhere without space limitations; 2) the DTW-based recognition algorithm only requires one training sample or class template for each class for highly accurate motion recognition; and 3) the DTW-based recognition algorithm can effectively reduce the integral errors of inertial Signals.

II. INERTIAL PEN

Our inertial pen consists of a triaxialaccelerometer (LSM303DLH, STMicroelectronics), a triaxial gyroscope (L3G4200D, STMicroelectronics), a triaxial magnetometer (LSM303DLH, STMicroelectronics), a microcontroller (STM32F103T8, STMicroelectronics), and an RF wireless transceiver (nRF24L01, Nordic). The accelerometer, gyroscope, and magnetometer are used to detect accelerations, angular velocities, and magnetic signals generated by hand movements.

The LSM303DLH possesses a linear acceleration full scale of •}2*g*, •}4*g*, and •}8*g*, with data output rates from 0.5 Hz to 1 kHz for all axes, and a magnetic field full scale of •}1.3, •}1.9, •}2.5, •}4.0, •}4.7, •}5.6, and •}8.1 gauss, 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 above mentioned 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. The size of the pen-type board is 130 mm × 15 mm × 8 mm (Fig. 1). Note that all signal processing procedures are performed on a PC. The overall power consumption of the hardware device is 30 mA at 3.7 V. The battery of the inertial pen is replaceable and rechargeable. The schematic diagram of the inertial pen hardware system is shown

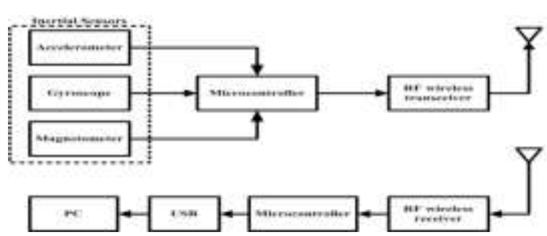

Fig. 1. Schematic diagram of the inertial pen

III. DYNAMIC TIME WARPINGALGORITHM

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. The DTW algorithm inthe 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 thehandwriting 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** = [$m1$, $m2$, . . . ,mp] and **N** = [$n1$, $n2$, . . . , nq]. First, a cumulative distance matrix **D** $\in Rp\times q$, 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**. Subsequently, a warping path $W = \{w1,w2, . . . ,wK \}$ can be calculated from the cumulative distance matrix (**D**), which is composed of the local cumulative distances $D(i , j)$.

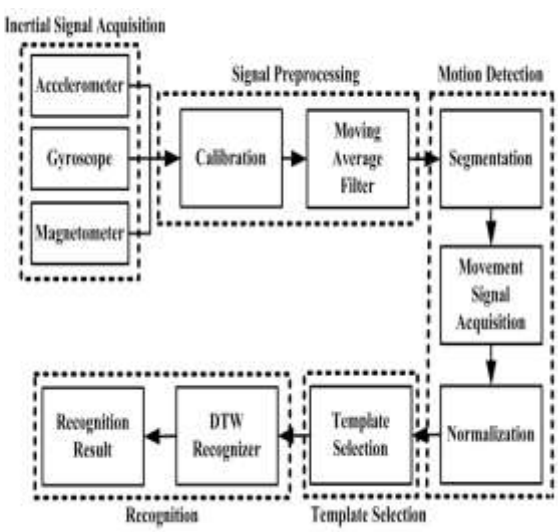

Fig. 2. Block diagram of the DTW-based recognition algorithm.

3.1. Signal Preprocessing

Since the measured signals are always contaminated not only by the sensors' error sources but also with users' unconscious trembles, signal preprocessing composed of calibration and a moving average filter is an essential procedure after inertial signal acquisition.

a) 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 acceleration 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. Then, we can obtain the scale factor (*SFacc)* and bias (*Bacc)* for each axis of the Accelerometer from these measurements as follows.

b)***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.

3.2. Motion Detection

The motion detection procedure involves the following steps: 1) segmentation, 2) movement signal acquisition, and 3) Recognition by the DTW Recognizer. We describe each step in detail as follows.

a) 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 movement 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. The magnitude of the filtered acceleration signal, $MAb(k)$, where k denotes the time steps. Then, we set the multiple of the mean valueof $MAb(k)$ in the interval of kts as the threshold.$T H = K\text{mean}k\ (MAb\ (k))$,where k denotes the time steps in the interval of kts, and 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.

b) Movement Signal Acquisition:The movement signal acquisition step includes following sub steps: 1) orientation estimation, 2) coordination transformation and gravity compensation, and 3) velocity and position estimation, which are elaborated as follows. *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. Then, the magnetic signals of the X-, Y -, and Z-axis can be transformed back to a horizontal plane (hx ,hy)

c) 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 i ×3 and the testing pattern Tj of size j × 3, is computed.

V. 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. Extending the results of our previous work with a minimum selection method with a DTW recognizer, 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.

A) Handwritten digit recognition 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 • } 2.01 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.

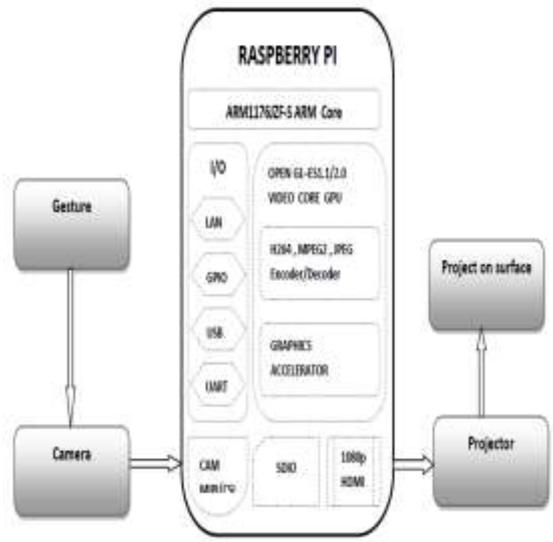

Fig 3: Raspberry PI

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. 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. Table II shows that the best recognition rate was achieved through the use of velocity signals with 92.0% accuracy by leave-one-out cross-validation. The results shown in Table III demonstrate that the proposed Min-Max template selection method with the DTW recognizer can obtain better recognition results compared to the random and minimum template selection methods.

The recognition rates obtained by 2-fold cross-validation, 5-foldcross-validation, 10-fold cross-validation, and leave-one-out cross-validation strategies were 70.3%, 72.5%, 81.4%, and 92.0%, as shown in Table IV. In addition, the user-dependent recognition rate for 2D handwritten English lowercase letter recognition evaluated by leave-one-out cross-validation was 94.3%, as shown in Table V. In addition, recognition performance was evaluated when 2D digits and 2D English characters were written simultaneously. Ten participants were asked to hold the inertial penand 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(=36×105) data were collected for this experiment. From The recognition rate of the proposed Min-Max template selection method with the DTW recognizer

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 motion 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%.

VI. CONCLUSION

This paper has presented an inertial pen with a systematic time alignment algorithm framework for inertial-sensing based handwriting and gesture recognition. The proposed DTW-based recognition algorithm consists of inertial signal acquisition, signal preprocessing, motion detection, templateselection, and recognition. To obtain better movement signals, we have utilized a quaternion-based complementary filter to reduce orientation errors and the ZVC method so as to minimize the undesirable error accumulation of velocity signals. Subsequently, to improve the performance of the DTW recognizer, all movement signals are normalized via the Z-score method and the class template is selected via the proposed Min-Max template selection method. During experimental validation, 2D handwritten digits, 3D handwritten Digits, 2D handwritten English lowercase letters, 2D handwritten digits and English letters, and 3D hand gestures were collected to evaluate the effectiveness of the proposed inertial pen and algorithm. The user-independent recognition rates for the abovementioned experiments were 97.9%, 87.3%, 92.0%, 92.1%, and 98.1%, respectively. In addition, the user dependent Recognition rates of the experiments were 99.4%, 94.6%, 94.3%, 93.0%, and 99.8%, respectively. Based on the above experimental results, we believe that the inertial pen and its associated DTW-based recognition algorithm can be considered an innovative and effective HCI device.

REFERENCES

[1] K. Altun, B. Barshan, and O. Tunçel, “Comparative study on classifying human activities with miniature inertial and magnetic sensors,” *PatternRecognit.*, vol. 43, no. 10, pp. 3605–3620, 2010.

[2] A. Akl, C. Feng, and S. Valaee, “A novel accelerometer-based gesture recognition system,” *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6197–6205, Dec. 2011.

[3] W. C. Bang, W. Chang, K. H. Kang, E. S. Choi, A. Potanin, and D. Y. Kim, “Self-contained spatial input device for wearable computers,” in *Proc. IEEE Int. Conf. Wearable Comput.*, Oct. 2003, pp. 26–34.

[4] G. Bailador, C. Sanchez-Avila, J. Guerra-Casanova, and A. de Santos Sierra, “Analysis of pattern recognition techniques for in-air signature biometrics,” *Pattern Recognit.*, vol. 44, nos. 10–11, pp. 2468–2478, 2011.

[5] C. M. N. Brigante, N. Abbate, A. Basile, A. C. Faulisi, and S. Sessa, “Towards miniaturization of a MEMS-based wearable motion capture system,” *IEEE Trans. Ind. Electron.*, vol. 58, no. 8, pp. 3234–3241, Aug. 2011.

[6] M. J. Caruso, “Application of magnetoresistive sensors in navigation systems,” in *Proc. SAE*, 1997, pp. 15–21.

[7] R. Xu, S. Zhou, and W. J. Li, “MEMS accelerometer based nonspecificuser hand gesture recognition,” *IEEE Sensors J.*, vol. 12, no. 5, pp. 1166–1173, May 2012.

[8] Z. Dong, C. Wejinya, and W. J. Li, “An optical-tracking calibration method for MEMS-based digital writing instrument,” *IEEE Sensors J.*, vol. 10, no. 10, pp. 1543–1551, Oct. 2010.