REAL-TIME ACCELEROMETER SIGNAL PROCESSING OF END POINT DETECTION AND FEATURE EXTRACTION FOR MOTION DETECTION

Jong Gwan Lim, Young Il Sohn, Dong Soo Kwon

Department of Mechanical Engineering,
Korea Advanced Institute of Science and Technology

Abstract: A simple signal processing algorithm is proposed for a novel input device developed for the elderly. The device is conceptually based on a pen, since it is one of the most familiar communication tools for the elderly and is closely tied to Hand Writing Recognition. We propose a new algorithm to execute End Point Detection and feature extraction for pattern recognition simultaneously. This algorithm is based on the fact that the drift error during acceleration can be corrected using the Zero Velocity Compensation, which is sensitive to the performance of EPD. The performance of the proposed algorithm is compared to that of the conventional approaches and its recognition rate is found to be superior to that of a previous research.

Keywords: accelerometers, pattern recognition, computer interfaces, signal-processing algorithms, motion estimation

1. COMPUTER INTERACTION FOR THE ELDERLY

With the advent of the 21th century, information technology has rapidly evolved. Unlike TVs, which offer one-directional information to the audience, internet service removes the gap between the information provider and consumer and makes bi-directional communication possible. The benefits of bi-directional interactions made possible through the internet drive the conventional media such as broadcasting, communication, journals, etc, unified from a macroscopic view and have stimulated convergence of digital devices from a microscopic view simultaneously. The home has also borne witness to the technological revolution. In this sense, the home can be thought of as a terminal of a broad internet network, and is evolving into a network computer to process or create information. This situation was predicted at an early point by Gates, *et al*. (1996), who predicted that all information would be eventually delivered to a home server and shown on a large-scale screen at the center of an automated home. Users collect and produce information via the interaction with this large-scale screen. The commercialization of home theatre personal computers and IPTV constitute a gradual realization of this vision. This trend accelerates the convergence between the conventional home appliances and input devices for human computer interaction. Interaction technology with large screens is at the same time attracting considerable interests. Since the environmental changes at the home influence daily life, all family members will be expected to be exposed to them directly. In particular, they will have a ripple effect on the elderly, who may be reluctant to handle new devices. For this reason, TV remote controllers and computer mouses have been integrated into novel input devices in order to ease such reluctance and maximize its convenience.

2. THE PROPOSED DEVICE

Lim, *et al*. (2007) derived the design requirements of digital devices for the elderly from an analysis of their physical/ emotional characteristics and the functional features of commercialized home appliances specialized for such devices. Utilizing these design requirements, we establish the following four design guidelines: (i) the function and structure should be simplified to ease their reluctance to complexity; (ii) the delivered information modality should be diversified to remind them of their choice; (iii) intuitive interactions should be developed to enable the elderly user to use the device without training; and (iv) the application of stuffs should be taken into

Fig. 1. The proposed device is composed of an accelerometer, a micro controller and a few buttons and is connected in RS232C into a PC.

account to induce empathy.

In order to induce the elderly to project their emotions on the device, we borrow the metaphor of a pen, a common communication tool that is readily accepted by virtually anyone. Also, user acceptability is expected to be high given that the illiteracy rate is extremely low in Korea. For the functional convergence between the mouse and TV remote controller, the functions of each device are analyzed and the functional structure of the new device is assumed to be constructed in the form of a 3D-mouse with many buttons. Cursor positioning is a unique functional feature of a mouse whereas command execution by pressing buttons is not. Therefore, the new device is required to undertake cursor positioning and the execution of more commands than would be capable with a conventional mouse. We have considered the basic functions of a pen, including 'drawing', and verified that the two tasks noted above can be implemented by assigning two buttons to each function and various drawing patterns to various commands for individual execution. In addition, by replacing the 'selection' function in a mouse with a 'picking' motion based on behavioural similarity to press something, we reduce the number of buttons and simplify the design structure. The device has been implemented with an accelerometer (3axis, MMA7260Q / Freescale, 2g), ADC in ATMEGA8 (3ch, 100Hz, 0~3V, 8bits), and is connected to a personal computer (Celeron 2.66 Ghz, Ram 512B) by RS232C for the signal processing. Fig. 1 shows the implemented device.

3. MOTION ESTIMATION USING INERTIA MEASUREMENT UNIT

The application of accelerometer measurement in human computer interaction was initially found mainly in the area of signature verification (Herbst and Morrissey, 1976; Epperson, 1993). It has taken the form of a pointing device (Nonaka and Da-Te, 1990), a head tracker system (Strickland, *et al.*, 1994), a data glove system (Perng, *et al.*, 1999), Hand Writing Recognition (Ishikawa, *et al.*, 1993; Miyagawa, *et al.*, 2000; Milner, 1999), and information input devices for wearable computers (Cheok, *et al.*, 2002; Bang, *et al.*, 2003). As

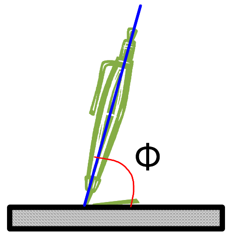


Fig. 2. The tilt change of the device

applications become more diverse, the used IMUs have evolved from 3 axis accelerometers to 6 DOF IMUs fabricated by MEMS. However, motion trajectory estimation is still available only with 3 axis accelerometers, because the cost of gyrosensors and IMUs is prohibitive with respect to commercialization (Choi, *et al.*, 2004; Oh, *et al.*, 2004).

3.1 Issues in algorithm implementation

The problem arising in trajectory estimation with accelerometers has been explained in Milner (1999) and Huddle (1998). The vulnerability in an inertial system is "drift" error, which is caused by a change in the deflection of the gravity vertical in the acceleration measurement. In the situation depicted in Fig.2, for the measured acceleration at given discrete time t = k, $A(k)$ is described as the follows

$$A(k) = A_{motion}(k) + \{g \sin\phi + g \sin(\phi - \theta(k))\} + A_{bias} \quad (1)$$

where $A_{motion}(k)$ is dynamic acceleration due to motion, A_{bias} is measurement error, the second term in Equation (1) is a static component by gravity, and $\theta(k)$ is the time-varying deviation about the mean angle. When the inclination angle of the pen at the motion starting point k_1 is not identical to the angle at the motion ending point k_2, the double integration of the acceleration from k_1 to k_2 goes to infinity, resulting in position drift (Milner, 1999). To correct the drift error without the help of a secondary device, filters are generally used such as reported in Milner (1999) due to the simplicity of this approach. However the application of a filter may damage the information. Otherwise, an external measurement

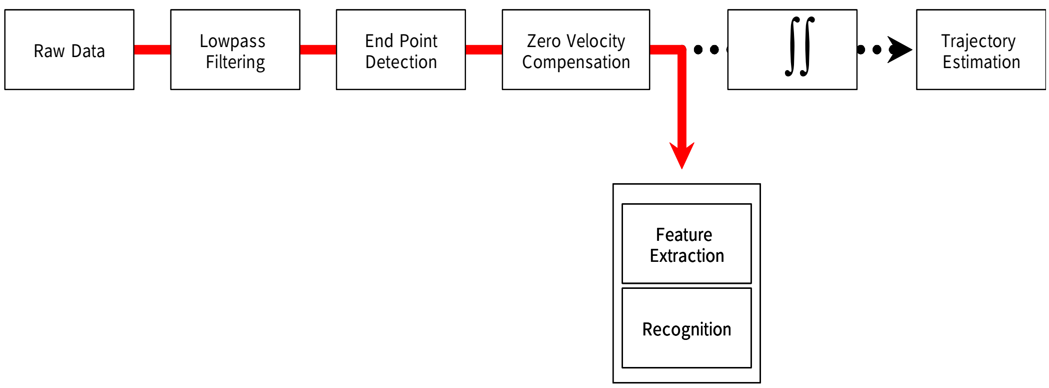


Fig. 3. Procedure for hand writing recognition. A general approach is to classify the writings with the estimated trajectory. The conventional approach is critically affected by drift error and numerical error during double integration.

related to the trajectory parameters such as zero velocity updates is used (Huddle, 1998). Frank (2001) applied zero velocity updates developed for distance estimation in UAV navigation to hand writing recognition, an approach dubbed Zero Velocity Compensation. This method was adopted by Bang, *et al*. (2003) again and the subsequent researches (Choi, *et al*., 2004; Oh, *et al*., 2004) have applied it as an alternative to drift error in trajectory estimation for HWR. Among many solutions, given that the trajectory is estimated for HWR, Milner (1999) is worthy of note. After removing the offset with a bandpass filter, he applied a speech recognition technique by transforming the 2 axis acceleration signal into 2-dimensional vectors and recognizing the signal pattern with a Hidden Markov Model without exact trajectory estimation. He reported the recognition rates of 96.2% out of 70 data for 7 classes. In this approach, a pen with an accelerometer loses its function of free drawing or writing and thus becomes a device to input predefined symbols. In engineering, restricting the degrees of freedom of mathematical modelling in return for the estimation improvement is not uncommon. In similar sense, Oh, *et al*. (2004) attempted to find the optimized features among the various combinations that could be composed by 3 axis accelerometers and 3 gyro sensors. They showed that patterns derived from the acceleration itself could be sufficient features for HWR. Therefore, we also adopt the same procedure, as outlined in Fig. 3

3.2 Zero Velocity Compensation

Even if the trajectory is not estimated, it is still important to correct the error resulting from the orientation change of the pen with ZVC, because the orientation varies in accordance with the users and may deteriorate the recognition rates. ZVC is conceptually simple. It is clear that a motion period is located between non-motion periods and acceleration and velocity are zero in a non-motion period. Based on this, we can distinguish the motion period from the non-motion periods. The parts where acceleration is stationary are regarded as the non-motion period and after the acceleration is integrated into the velocity, a slope connected with two successive stationary points (the motion starting point and the motion ending point) is assumed as the velocity offset error due to the inclination angle change of a pen. Velocity compensation is achieved by adding or subtracting this slope from the integrated velocity to make the velocity at two stationary points the same as zero, because the velocity obtained by integrating acceleration must be identical to zero when the motion stops. The differentiation of the compensated velocity gives the compensated acceleration, and the drift error is finally corrected without the help of extra devices using only a computational trick (Bang, *et al*., 2003).

The ZVC algorithm itself has several restrictions owing to a few assumptions. Initially, the orientation change of the pen that occurs throughout the motion, $\phi - \theta(k)$, is assumed to be monotonically increased or decreased. Therefore, the longer the motion period takes, the larger the difference shows in the orientation estimation. This implies that the time difference between the starting point and the ending point of the motion, $\Delta t = k_1 - k_2$, should be short. Furthermore, since the orientation change is estimated as a first order linear function to connect two points, failure in finding them, degrades the reliability of ZVC. That is, it is essential to find the exact motion starting and ending point in ZVC. It is very common in time series pattern recognition to extract the period of interest, which is called End Point Detection in automatic speech recognition (Li, *et al*., 2001). While many EPD algorithms have been introduced for speech, those developed for speech are not applicable to the acceleration signal due to their spectral and temporal differences.

4. EXTREME POINT SAMPLING: A NEW ALGORITHM FOR END POINT DETECTOIN

In the previous chapter, it was explained that EPD is closely related to the performance of ZVC and the reliability of the EPD algorithm may critically affect the pattern recognition rates. In this chapter with a simple introduction of EPD algorithms developed in the previous researches (Bang, *et al*., 2003; Yang, *et al*., 2004), we present a new algorithm, Extreme Point Sampling for EPD.

4.1 EPD by Bang, et al. (2003)

Bang, *et al*. (2003) developed an EPD algorithm for acceleration and named it “Stroke Segmentation”. In their algorithm, they attempt to find the motion starting point and the ending point by comparing the standard deviation of $|A|$ computed on the basis of S samples, $\sigma^s_{|A|}(K)$, where $K_2 = K_1 + S \cdot k$ with a threshold, σ_{th}. H is the minimum number of samples to keep $\sigma^s_{|A|}$ greater or less than σ_{th}, and W is the minimum number of samples for making a motion. In short, estimation of the starting point and the ending point depends on whether $\sigma^s_{|A|}$ maintains a value

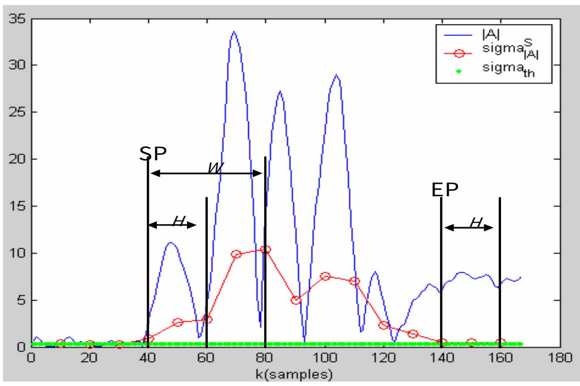


Fig. 4. Typical profile of the standard deviation and the magnitude of acceleration at S=10, H=20, W=50, σ_{th} =0.28

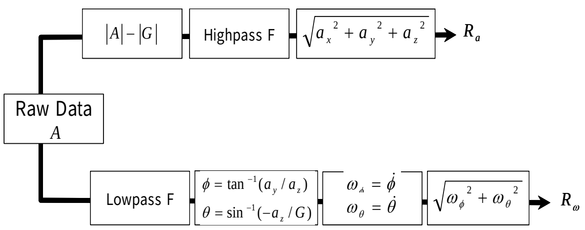


Fig. 5. Transformation from raw acceleration into the norm of the acceleration and the angular rate

greater or less than σ_{th} during H period. Fig. 4 illustrates how to find the starting and ending point. Using this algorithm, it is difficult to find the optimized combination that it restricts the user's motions since there are 4 parameters to be tuned. Furthermore, because the standard deviation is calculated on the basis of S samples, the starting point and the ending point are located at multiples of S samples and two points detected by this algorithm indicate only approximated starting/ ending points.

4.2 EPD by Yang, et al. (2004)

Yang, *et al*. (2003) also proposed an EPD algorithm for acceleration and they transformed raw acceleration into two components: the norm of the acceleration, R_a, and the norm of the angular rates, R_w. The process to obtain these components is depicted in Fig. 5. After $R_a(k)$ and $R_w(k)$, the detection region of the starting point [1, k_{start}] and ending point [k_{end}, end of data] are defined by choosing two constants, k_{start} and k_{end}. Then, in the range of $R_a^2(k)$ and $R_w^2(k)$, the starting point and the ending point are detected by applying the thresholds of TD_a and TD_w, respectively. While they reduced the computing load and the number of parameters to two without using a statistical approach,

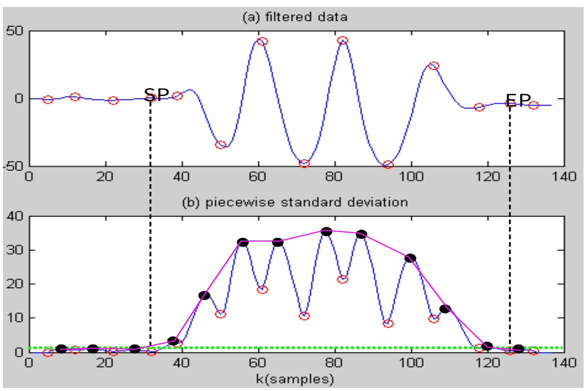


Fig. 6. Extreme Point Sampling. Circles indicate the extreme value in the above and $\sigma_{LV}^p(n)$ in the below. Filled circles indicate $\sigma_{\max}^p(n)$ and the starting/ ending point are detected by comparison of $\sigma_{\max}^p(n)$ and σ_{th}^p

this algorithm is still cumbersome to optimize. Furthermore, experiments reveal that the performance is very sensitive to high frequency noise.

4.3 Extreme Point Sampling

EPS based on a statistical approach uses the standard deviation, as in Bang, *et al*., (2003). The significant difference from their approach is that, while the standard deviation on every S sample with respect to $|A|$ is computed in their approach, the piecewise standard deviation is calculated on the basis of S samples with respect to acceleration A for EPS. While we analyze the piecewise standard deviation of A, it is observed that local minima in the piecewise standard deviation correspond to the extreme values in acceleration. The detailed algorithm is described as the follows.

1) For a given acceleration, A, calculate the piecewise standard deviation, $\sigma^p(k)$

$$\sigma^p(k)=std\left(\left[A(k-S+1),A(k-S+2),\cdots,A(k)\right]\right) \quad (2)$$

2) Find every local minimum in $\sigma^p(k)$, $\sigma_{LV}^p(n)$
3) Calculate the local maximum between two consecutive local minima in $\sigma^p(k)$, $\sigma_{\max}^p(n)$

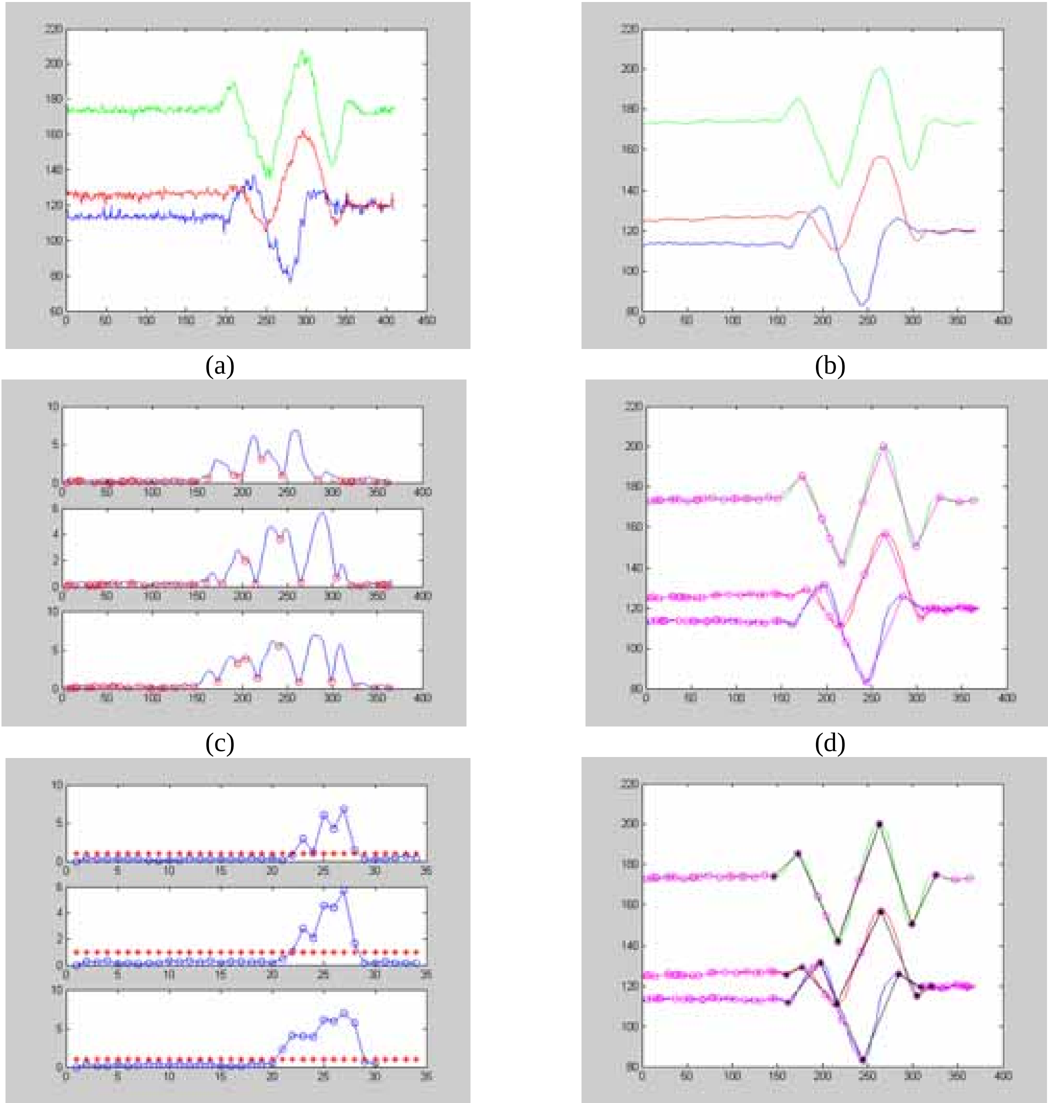
$$\sigma_{\max}^p(n)=\max[\sigma_{LV}^p(n-1),\sigma_{LV}^p(n)] \quad (3)$$

4) If $\sigma_{\max}^p(n)>\sigma_{th}^p$
   Starting point = $k^{(n-1)}$,
   If $\sigma_{\max}^p(n)<\sigma_{th}^p$
   Ending point = $k^{(n)}$
5) Although the extreme values in acceleration are determined by the local minima of $\sigma^p(k)$, the signal contaminated by noise indicates the wrong positions for extreme values. As such, these must often be removed by checking whether a point $\left(k^{(n)},A\left(k^{(n)}\right)\right)$ satisfies the relation, as in Equation (4).

$$\Delta A(n-1)\cdot\Delta A(n)<0, \quad (4)$$

where $\Delta A(n)=A(k^{(n)})-A(k^{(n-1)})$, whenever a new $\sigma_{LV}^p(n)$ is found between the staring point and the ending point.

Fig. 6 describes the physical meaning of σ^p, σ_{LV}^p, and $\sigma_{\max}^p$. σ_{LV}^p is depicted by circles and it is shown that the locations of circles in the bottom part of Fig. 6 correspond to the extreme values of the acceleration above. Filled circles denote $\sigma_{\max}^p$ and we can find the starting/ ending points by taking the threshold in $\sigma_{\max}^p$. It is noted that while the EPS algorithm performs EPD, it produces the location of the extreme points in acceleration at the same time and these extreme points can be applicable to pattern recognition as features.



(a) (b)

(c) (d)

(e) (f)

Fig. 7. Prodedure of Extreme Point Sampling. (a) Raw acceleration (b) Lowpass-filtered acceleration (c) Local minima found in σ^p. Circles correspond to σ^p_{LV} (d) Extreme points indicated by σ^p_{LV} in acceleration (e) Thesholding in $\sigma^p_{\max}$ (f) Successful EPD and feature extraction(pointed with *)

Therefore, additional feature extraction does not need to be followed by EPD.

5. EXPERIMENT FOR PERFORMANCE COMPARISON

In order to verify the performance of EPS, data have been collected from 12 adults (male 6, female 6, age varying from 20 to 40) 11 times with respect to 14 hand writing gestures (Fig.7), for a total of 1848 data. Among 1848 data, 168 data are used as a reference template and the remaining, 1680 data are used for a recognition test. Dynamic Time Warp is chosen as a recognizer since it is appropriate to compare time series signals with different lengths and shows critical sensitivity to the performance of EPD. Also, This approach is recommended by Milner (1999), who noted that the number of states should be increased for non-stationary signal modelling when we use HMM and HMM, and it becomes similar to DTW in this case. Therefore, it is plausible that the recognition rates by DTW are in proportion to EPD performance.

The hand writing gesture in Fig. 8 was proposed by Bang, *et al*. (2003) for the first time and later adopted by Yang, *et al*. (2004). Since the issues and approaches are different from one another, it's impossible to compare their experimental results directly with the present results. Accordingly, the test

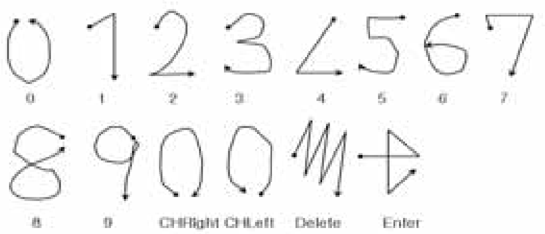


Fig. 8. Hand writing gesture for the experiment.

Table 1 Recognition rates comparison

Gestures	Bang's	Yang 's	EPS
0	41.67	65.00	64.17
1	88.33	91.67	95
2	87.5	76.67	93.33
3	99.17	94.17	95.83
4	86.67	80.83	97.5
5	88.33	90.83	85.83
6	49.17	30.00	69.17
7	95.83	95.00	97.5
8	95	95.83	92.5
9	94.17	97.50	99.17
CHRight	90.83	97.50	96.67
CHLeft	97.5	94.17	96.67
Delete	95	97.50	95
Enter	81.67	80.83	100
Total	85.06	84.82	91.31

is conducted by restricting our interest to EPD performance with identical hand writing gestures. The results are given in Table 1. It is noted that 0 and 6 show much lower rates due to their natural ambiguity. In some cases, Bang and Yang record better rates, but it is clear that EPS is superior to the others on the whole. Rotational translation is shown to be less recognized and this is not due to constraints of the algorithm but due to a constraint of the linear accelerometer. More rigid scrutiny and analysis are left for further work.

6. Discussion and Further Work

The strengths of the EPS algorithm are not limited to EPD performance itself. It should be emphasized that extreme points acquired during EPD computation are suitable features to describe the acceleration signal in reduced dimensions. Furthermore, EPS is also applicable to the implementation of cursor positioning with an accelerometer.

The functional requirements of the proposed device include cursor positioning such as in application of a mouse, as stated in chapter 2. Since we do not estimate the trajectory against drift error, quantitative computation of double integration cannot be made for this aspect in our approach. In order to address this matter, we turn an EPD tool of EPS into an activated acceleration detector on each axis based on the findings that cursor positioning is specified by the translating direction, the force with respect to the direction, and the time that elapses during the translation in the case of a linear translation.

Given that EPS is computed in 3 axes, extreme points detected by EPS in acceleration are given as $A_x^{t_1}$, $A_x^{t_2}$, $A_x^{t_3}$, and $A_x^{t_4}$ on the x axis; direction D_x, force F_x and elapsed time T_x are given as follows.

$$D_x = \begin{cases} right & for \quad A_x^{t_2} - A_x^{t_1} > 0 \\ left & for \quad A_x^{t_2} - A_x^{t_1} < 0 \end{cases} \tag{5}$$

$$F_x = \left| A_x^{t_2} - A_x^{t_1} \right| \tag{6}$$

$$T_x = t_4 - t_2 \tag{7}$$

If the direction, force, and time are accurately computed with respect to the 3 axes, their combination can be utilized to fully estimate the linear translation. Since a rotational translation can theoretically be decomposed to several linear translations, the former can be approximately estimated with this method if sensitivity of the accelerometers is allowed.

Although quantitative recognition rates are given in Table 1, the causes of these results and solutions to improve the rates have not yet been examined in detail. Additionally, a performance comparison with a conventional 3D mouse regarding cursor positioning should be accomplished. Moreover, the ultimate goal of our research, providing enhanced accessibility for the elderly, has not been yet verified. These tasks will be undertaken in the subsequent studies.

Acknowledgment

This research was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Commerce, Industry and Energy of Korea. The authors would also like to thank Dr. Won-Chul Bang, Dr. Sung-Jung Cho, and Dr. Jing Yang, authors who contributed to Bang, *et al*. (2003) and Yang, *et al*. (2004), and Dr. Geehyuk Lee at the Information and Communication University for advice on this work.

REFERENCES

Bang, W.C., W. Chang, K.H. Kang, E.S. Choi, A. Potanin and D.Y. Kim (2003), Self-contained spatial input device for wearable computers, In *Proceeding of 7th IEEE International Symposium on Wearable Computers*, pp.26-34, White Plains, NY

Cheok, A.D., K. G. Kumar and S. Prince(2002), Micro-accelerometer based hardware interfaces for wearable computer mixed reality applications, In *Proceeding of 7th IEEE International Symposium on Wearable Computers*, Seattle, WA, USA

Choi, E.S., W. Chang, W.C. Bang, J. Yang, S.J. Cho, J.K. Oh, J.K. Cho and D.Y. Kim (2004), Development of the gyro-free handwriting input device based on inertial navigation system(INS) theory, In *Proceeding of SICE Annual Conference 2004*, pp.1176-1181, Sapporo, Japan

Epperson, M. (1993), Autonomous computer input device and marking instrument, US patent 5 247 137

Frank, M. (2001), Position refinement algorithm, US patent 6 292 751

Gates, B., P.M. Rinearson and N. Myhrvold (1996), *The Road Ahead*, Penguin Books, USA.

Herbst, N.M. and J.H. Morrissey (1976), Signature verification method and apparatus, US patent 3 983 535

Huddle, J.R. (1998), Trends in inertial systems technology for high accuracy AUV navigation, In *Proceeding of Workshop on Autonomous Underwater Vehicles*, August. 20-21, pp.63-73

Ishikawa, H., Y. Yonezawa, K. Itoh and M. Hashimoto (1993), Extraction of character information from acceleration of handwriting pen, In *Transaction of Society of Instruement and Control Engineers*, vol. J76D-I, pp.541-543. In Japanese

Li, Q., J. Zheng, Q. Zhou and C.H. Lee (2001), A Robust, Real-time Endpoint Detector with Energy Normalization for ASR in Adverse Environments, In *Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, USA

Lim, J.G., Y.I. Sohn, Y.G. Kim and D.S. Kwon (2007). Motion Detection-based Intuitive Mediate Interface, In *HCI2007*, pp.924-930. Pyeongchang, Korea. In Korean

Milner, B. (1999), Handwriting recognition using acceleration-based motion detection, *IEE Colloquium on Document Image Processing and Multimedia*

Miyagawa, T., Y. Yonezawa, K. Itoh and M. Hashimoto (2000), Handwriting pattern reproduction using 3d pen acceleration and angular velocity, In *Transaction of Institute of Electronics, Information and Communication Engineers*, vol. J83D-I, pp. 1137 – 1140

Nonaka, H. and T. Da-Te (1990), The development of a pointing device with a cursor in inertial motion, In *Transaction of Information Processing Society of Japan*, vol. 31, no. 2, pp.268-274. In Japanese

Oh, J.K., S.J. Cho, W.C. Bang, W. Chang, E.S. Choi, J. Yang, J.K. Cho and D.S. Kim (2004), Inertial Sensor Based Recognition of 3-D Character Gestures with an Ensemble of Classifiers, In *Procceeding of 9th Int Workshop on Frontiers in Handwriting Recognition*

Perng, J.K., B. Fisher, S. Hollar and K. S. J. Pister (1999), Acceleration sensing glove(asg), in Proc. 3rd International Symposium on Wearable Computers, pp.178-180, Los Alamitos, CA, USA

Strickland, D., A. Patel, C. Stovall, J. Palmer and D. McAllister (1994), Self tracking of human motion for virtual reality systems, In *Proceeding of SPIE International Society for Optical Engineering*, vol. 2177, pp.278-287.

Yang, J., E.S. Choi, W. Chang, W.C. Bang, S.J. Cho, J.K. Oh, J.K. Cho and D.Y. Kim (2004), A Novel Hand Gesture Input Device Based on Inertial Sensing Technique, In *the 30 th annual conference of the IEEE Industrial Electronics Society*, Korea