Proc. 14th Int. Conf. on Global Research and Education, Inter-Academia 2015
JJAP Conf. Proc. **4** (2016) 011601

A Human Computer Interactions Framework for Biometric User Identification

Kamen Kanev[1,2], Maria De Marsico[2], and Paolo Bottoni[2,1]

[1]Graduate School of Science and Technology, Shizuoka University, Japan
[2]Department of Computer Science, Sapienza University of Rome, Italy

E-mail: kanev@rie.shizuoka.ac.jp

(Received September 24, 2015)

Computer assisted functionalities and services have saturated our world becoming such an integral part of our daily activities that we hardly notice them. In this study we are focusing on enhancements in Human-Computer Interaction (HCI) that can be achieved by natural user recognition embedded in the employed interaction models. Natural identification among humans is mostly based on biometric characteristics representing what-we-are (face, body outlook, voice, etc.) and how-we-behave (gait, gestures, posture, etc.) Following this observation, we investigate different approaches and methods for adapting existing biometric identification methods and technologies to the needs of evolving natural human computer interfaces.

1. Introduction

A number of core Human-Computer Interaction (HCI) methods and technologies lay out the foundation for the computer assisted functionalities and services that are penetrating our world up to the point that we deem them indispensable. However, like with familiar objects we use in customary tasks, we hardly notice them in our daily activities. In the world of changing HCI, the recognition and identification of humans engaged in the entailed interactions becomes a crucial issue both from the security point of view and as a vehicle for increasing computer awareness and adaptability for tailoring user-oriented services.

In this study we focus on exploring possible enhancements of natural HCI by introducing unobtrusive user identification and recognition capabilities directly into the employed interaction models. These are weaved into the fabric of the interface, deeply integrated into the interaction model, and highly transparent for the user. Note that when natural human interfaces are used, there might be no keyboard, no mouse, and even no direct contact with any physical interface component. Consider for example accessing a Smartphone through its touch screen and a keyboard image on it vs. a physical keyboard, or consider playing a game using Microsoft's Kinect where there is no physical interface contact whatsoever. Yet, even in such situations, users are often identified by usernames and authenticated by passwords (or some "visual" versions of them). Obviously, it would be much better if user identification and consequent authentication were integrated into the HCI process in a natural an unobtrusive way so that they remain mostly transparent to the user.

Natural identification among humans is most often based on biometric characteristics representing what-we-are (face, body outlook, voice, etc.) and how-we-behave (gait, gestures, posture, etc.). Following this observation, in our work we investigate different approaches and methods for adapting existing biometric identification methods and technologies to the needs of the novel HCI and evolving natural human computer interfaces. In particular we discuss the design and development of a specialized framework and a corresponding supportive environment for research and experimental work with a wide range of biometric identification methods in the context of natural HCI and related interfaces.

2. Natural User Identification Scenarios

In our earlier work [1] we have designed and implemented a user identification system employing a mouse with an embedded fingerprint scanner. The system provided for continuous identity tracking of the mouse operator that was completely transparent with respect to the employed interaction model. With the launch of iPhone5s we have a good example of a widely available button interface incorporating an embedded fingerprint scanner. Following this, one can envisage a keyboard with a fingerprint scanner gathering information from all its keys or even a general touch screen that will be able to scan fingertips when operated. All these examples illustrate the idea of natural user identification that occurs in the course of the normal use of an interface. Such identification and continuous re-identification requires no conscious effort by the user and is thus completely transparent to him.

Although biometric identification employing fingerprints and other human characteristics based on what-we-are is highly reliable when properly implemented, it is still prone to various security threats. Some biometric footprints, for example, can be easily recovered from the environment (e.g. fingerprints) while others can be obtained by simple surveillance techniques (face photos, etc.) This increases the risk of biometric identity theft which is a serious problem since user biometrics cannot be changed at will.

In our current work we aim to address some of these issues by giving precedence to biometric identification based on how-we-behave (gestures, posture, gait, etc.) rather than what-we-are. Using static handwriting and traditional signatures for human identification in graphology has a long history and is deemed to provide for reliable human identification. Modern multidimensional input devices add additional layers of security to this by collecting dynamic gesture data in the process of writing and/or signing, which makes forgeries extremely difficult. And, if really necessary, one can always change his/her signature (but not his/her fingerprints). That is why in this work we consider different methods and approaches for the collection of dynamic gesture data that can be integrated in various natural identification scenarios.

3. Design and Framework Development

3.1 Framework Building Blocks and Components

The main target of the experimental framework that we build is to integrate various input methods and corresponding interactive interfaces into a modular system for gathering, processing, and analyzing motion and gesture data. Such a modular approach allows for coverage of a wide span of input methods and technologies starting with traditional mouse-based input and its enhancements with gesture dynamics [2], going through classic 2-DOF digitizers, tablets, and sign pads for 2D graphical input and their enhanced multiple-DOF versions capable to track the stylus azimuth, elevation, and applied forces [3]. We also cover current touch screen based input both for single- and multi-touch surfaces [4].

While there are many applications that employ tablets and electronic sign pads to authorize access to bank accounts or credit card purchases, they are all limited to obtaining a user signature on a specialized flat surface [5]. The truly natural gesture-based user identification that we envisage, however, should be more flexible in respect to the user and thus allow for spatial 3D gestures. Our gesture identification framework should, therefore, facilitate the dynamic gathering of such 3D gesture data on multiple hardware platforms. In this respect we build upon our spatial motion tracking experience with Nintendo Wii remote controllers, Android Smartphones, and others.

We have used Nintendo Wii remote controllers connected to a Windows PC by Bluetooth as spatial input devices in a Kanji writing game for elementary school pupil called Kanji Sports [6]. A new, more advanced version of the Kanji Sports game for mobile devices is under development while a pilot implementation for Android Smartphones is already available for experiments [7].

3.2 Spatial Motion Tracking with Android Smartphones

A recent development of a general gesture tracking systems for authentication with Android Smartphones has been reported in [8]. Such Smartphone based general gesture tracking systems employ the motion and orientation sensors, embedded in the device, to gather data in local device coordinates. Based on the obtained data a device-to-world transformation matrix is dynamically computed and applied for coordinate transformation to global (world) coordinates. After that, numerical integration is carried out to calculate the 3D orientations, corresponding velocities, and the time-dependent 3D movement path of the device. Each obtained paths is finally mapped to the set of sample gestures and the closest match is determined. Following this procedure, we have successfully conducted experiments with simple Smartphone movements in horizontal and vertical directions and with drawing of circles, triangles, and rectangles, etc.
In respect to spatial gesture input, however, there is an inherent problem related to the fact that humans are simply not capable of writing a straight line in the air if no reference edge or a surface is provided in the vicinity. The reason is mainly related to the way our joints work e.g. an intentional straight vertical line drawn on an imaginary large screen in front of us appears as an arc in 3D due to our limited depth control capabilities. This implies that extracted spatial paths need some preprocessing before actual analysis could be done. We implement such preprocessing by projecting each path on a carefully selected surface e.g. imaginary screen on which the writing is mentally conducted. Projected data is then used for matching to the available gesture samples.

3.3 Kanji writing specificities

To avail of Kanji writing specificities we note that every Kanji employed in the Japanese writing system can be decomposed into a predetermined sequence of individual strokes. In fact Kanji handwriting rules prescribe lifting of the stylus at the end of every stroke which makes character decomposition and further analysis much more straightforward. In contrast, note that individual characters and even sequences of characters from the English alphabet are often written without lifting the stylus which obviously complicates handwriting analysis.

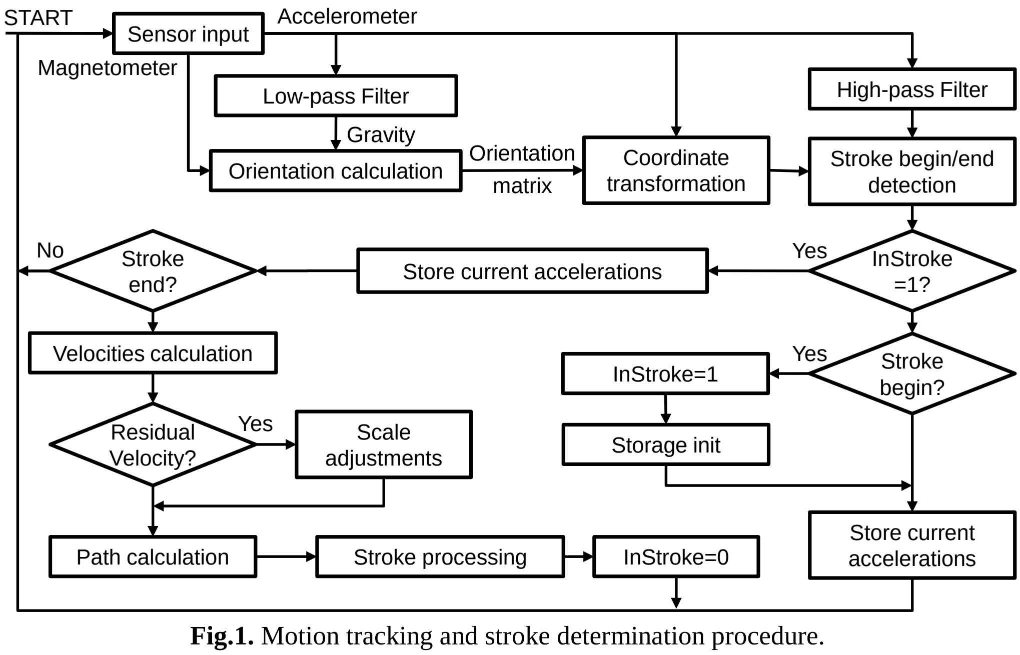
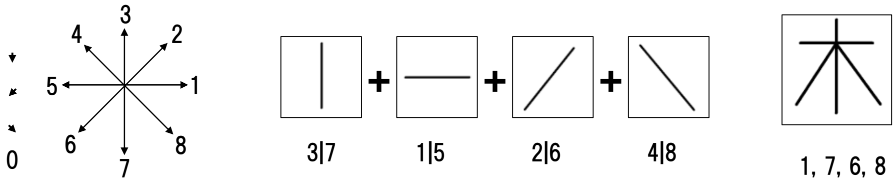


Fig.1. Motion tracking and stroke determination procedure.

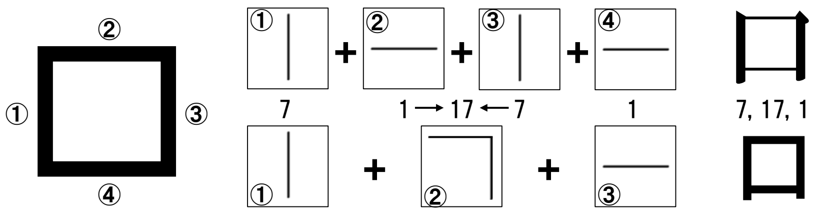
The idea is therefore to decompose the Kanji handwriting into individual strokes which are later combined into Kanji characters, etc. This is actually the standard way surface-based mouse, tablet, and touch screen Kanji input is implemented. When it comes to a truly 3-dimensional input, however, it turns out that in the absence of a tangible reference surface, the sense of lifting the stylus is almost entirely lost. To address this issue we have conducted experiments with buttons and other dedicated sensors to detect beginnings and ends of strokes but this puts additional burden on the user and interferes with the natural interface usage patterns. A deeper look into this problem reveals that stroke beginnings and ends could actually be quite reliably detected based on the observed motion acceleration patterns. In fact, Kanji writing mental models well learned by all Japanese children at school appear to enforce different acceleration patterns for writing of strokes and for movements between strokes. Based on such observations we have implemented the Kanji-oriented motion tracking and stroke determination algorithm shown in Fig.1 which encompasses natural stroke writing patterns and is thus posit to deliver superior extraction results.



(a) (b) (c)

Fig.2. Simple strokes and their codes.

The outcome of the algorithm is a sequence of time-dependent stroke descriptions that need to be classified following the Japanese Kanji writing rules. We will illustrate the stroke description and classification process by referring to the Japanese Kanji character for the English word “tree” shown in Fig.2(c). The character is comprised of four strokes namely a “horizontal”, a “vertical”, a “right-inclined”, and a “left-inclined” as shown in Fig.2(b). This textual description, however, does not indicate the directions in which the strokes should be drawn although such directions are an important part of the Kanji writing rules. To fully disambiguate the four stroke types and their directions we employ the encoding scheme shown in Fig.2(a). Following this scheme the stroke sequence of the Kanji in Fig.2(c) is expressed as 1, 7, 6, 8. Such unambiguously encoded stroke sequences and their extension, as discussed further on in this work, make the foundation of the handwritten Kanji analysis and recognition for biometric human identification and authentication that we develop.



(a) (b) (c)

Fig.3. Combined strokes and their encoding

While a large number of Kanji can be written by using the simple strokes shown in Fig.2 there are also other more complex combined strokes. For example, as shown in Fig.3(a), the Kanji character for the English word of “mouth” looks like a simple square. One may think that it is constructed from four segments as shown in the top sequence of Fig.3(b) but the Japanese writing rules prescribe that it should be written by three strokes, the second of which is a combination of a horizontal stroke followed by a vertical stroke without lifting the stylus essentially making a corner or an angle as shown in the bottom sequence. We will encode such combined strokes by simply putting together the codes of their comprising strokes in the right order e.g. a stroke encoded by 1 followed by a stroke encoded by 7 without lifting the stylus results to a combined stroke encoded by 17. In some cases either a simple stroke or a combined stroke could be acceptable. In Fig.4, for example, the third stroke of the Kanji for a “person” could be a simple right-inclined 6 as shown in the top sequence. But it can also be a vertical 7 followed by a right-inclined 6 without lifting the stylus which makes to a combined stroke with code 76 as shown in the bottom sequence. This is reflected in the different possible and acceptable appearances of the corresponding Kanji character depending on the employed font. Note that a large variety of combined strokes built of two and more comprising strokes can be encoded following the above presented rules. In the discussion that follows, however, we will limit the scope to the strokes that are found in the Kanji set employed in our experiments.

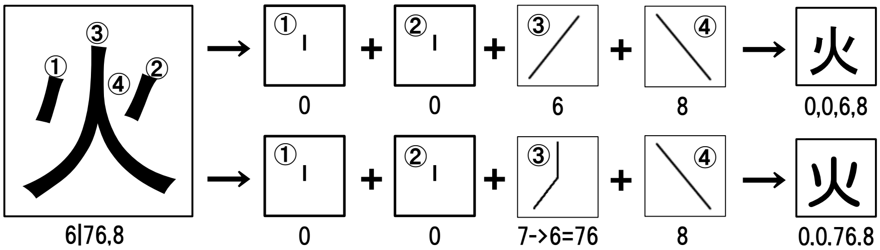


Fig.4. Simple and combined stroke alternatives

4. Pilot Implementation and Experiments

The pilot software implementation based on the procedure outlined in Fig.1 enables us to obtain extensive data fully describing the dynamics of the complex motions the mobile device is subjected to. While our ultimate objective is to use such data for natural user identification and authentication, the first set of experiments that we have conducted focuses on spatial Kanji input and its use as a complement or a substitute of the standard touch screen based user authentication essentially allowing users to authenticate by spatial gestures rather than via a touch screen. Our next step will be to employ spatial signatures for direct user identification and consequent authentication without the need of user names and passwords.

Kanji			Stroke No	Stroke Sequence	Rcognition Rate	
Font1	Font2	Font3			1st time	2nd time
十	十	十	2	1, 7	88.2%	100.0%
口	口	口	3	7, 17, 1	64.7%	64.7%
山	山	山	3	7, 71, 7	64.7%	52.9%
小	小	小	3	74, 6|76, 8	11.8%	35.3%
王	王	王	4	1, 7, 1, 1	64.7%	100.0%
木	木	木	4	1, 7, 6|76, 8	76.5%	76.5%
火	火	火	4	0, 0, 6|76, 8	5.9%	17.6%
右	右	右	5	6|76, 1, 7, 17, 1	11.8%	47.1%
言	言	言	7	0, 1, 1, 1, 7, 17, 1	17.6%	29.4%
町	町	町	7	7, 17, 7, 1, 1, 1, 74	5.9%	29.4%

Fig.5. The Kanji set employed in the experimental testing.

For our experiments we have prepared a set of 10 Kanji characters as shown in Fig.5. The set includes some very simple characters comprised of 2 or 3 strokes as well as some fairly complex ones with up to 7 strokes. The stroke types used in the writing of the selected Kanji subset are shown in Fig.6. Employed strokes include the five simple stroke types with codes 0, 1, 6, 7, and 8 two of which are used most often i.e. 15 code 1 (horizontal) strokes and 10 code 7 (vertical) strokes. From the combined stroke types which are used less often, the employed ones are with codes 76, 17, 71, and 74. As indicated in Fig.6, different recognition rates have been observed for different stroke types, those with codes 0 and 74 appearing to be the most difficult to draw.

We have conducted extensive experiments with discriminating the 9 different stroke types from Fig.6. A group of 17 test subjects was formed and each participant was asked to write "in the air" the selection of 10 different Kanji characters in Fig.5 while holding an Android Smartphone in his hand as if it were a brush. While recognition failures were common in the initial writing attempts, repeated writing lead to significant improvements thus demonstrating a steep learning curve for this novel input interface.

Strokes			Recognition Rate		Strokes			Recognition Rate	
Visual	Code	No	1st time	2nd time	Visual	Code	No	1st time	2nd time
	0	3	29.4%	41.2%		8	3	94.1%	92.2%
	1	15	96.1%	96.1%		17	4	67.6%	83.8%
	6|76	4	83.8%	86.8%		71	1	64.7%	58.8%
	7	10	89.4%	91.8%		74	2	26.5%	47.1%

Fig.6. Stroke patterns included in the testing Kanji set.

5. Conclusion

Our framework implementation has been used so far for conducting experiments with 3D Kanji writing and various stroke-based Kanji games. We are continuing our research with more detailed analysis of the extensive 3D data that we collect, aiming to better understand the true dynamics of the 3D gestures that might lead to more advanced transparent user identification.

Acknowledgment

This work was supported in part by funding for exchange of researchers provided by the Ministry of Foreign Affairs of Italy, funding for a Cooperative Research Project at Research Institute of Electronics, Shizuoka University, and KAKENHI Grant 25560109.

References

[1] K. Kanev, N. Kamiya, N. Mirenkov, G. Kanda, A. Takagi, Adaptive E-learning Applications with Transparent User Identification, *Proceedings of the Tenth Int. Conf. on Humans and Computers HC2007*, Dusseldorf, Germany, December 7,13-15, 2007, pp. 25-30.

[2] B. Sayed, I. Traore, I. Woungang, M. Obaidat, Biometric Authentication Using Mouse gesture Dynamics, *IEEE Systems Journal*, Vol.7, No.2, 2013, pp.262-274. DOI:10.1109/JSYST.2012.2221932.

[3] D. Sakamoto, H. Morita, T. Oishi, Y. Komiya, T. Matsumoto, On-line Signature Verification Algorithm Incorporating Pen Position, Pen Pressure and Pen Inclination Trajectories, *Proceedings of the 2001 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, USA, May 7, 2001, pp. 993-996. DOI:10.1109/ICASSP.2001.941084.

[4] N. Sae-Bae, K. Ahmed, K. Isbister, N. Memon, Biometric-rich Gestures: A Novel Approach to Authentication on Multi-touch Devices, *CHI 2012 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin, Texas, USA, May 5-10, 2012, pp. 977-986. DOI:10.1145/2207676.2208543

[5] A. Jain, F. Griess, S. Connell, On-line Signature Verification, *Pattern Recognition*, Vol.35, No.12, 2002, 2963-2972. DOI:10.1016/S0031-3203(01)00240-0.

[6] H. Takaki, T. Yanagida, K. Kanev, Kanji Sports: Acquisition of Kanji Strokes in Three Dimensions, Human Interface Society Proceedings, Vol. 12, No. 1, Hamamatsu, Japan, March 4-5, 2010, pp. 21-24. (In Japanese)

[7] K. Kanev, K. Oyaizu, M. De Marsico, P. Bottoni, Stroke Analysis for Kanji Learning with Mobile Devices, *2015 International Workshop on Serious Gaming = Serious Business*, Hamamatsu, Japan, March 5, 2015.

[8] O. Medina Reyes, Desarrollo de un Sistema de Autenticacion para Dispositivos Moviles con Sistema Operativo Android por Medio del Accellerometro, Seminario de Investigacion, Departamento de Ciencias de la Computacion, Universidad Autonoma de Aguascalientes, Mexico, Junio, 2014.