On-line Modelling of Drawing/Writing Behaviour for Verification of User Identity

Xiaojun Zhang and Iain Macleod

Computer Sciences Laboratory, RSISE
& TRUST Project[1]
Australian National University
Canberra, ACT 0200, Australia

E-mail: (xiaojun, iain)@cslab.anu.edu.au

Abstract As part of a larger project, we are assessing the ability of pen gestures (as used in pen-based computing) to provide user-specific information. Verifying identity by means of pen gestures is much more challenging than via a signature because of two intrinsic limitations: there is less information in a pen gesture than in a signature, and the derived spatial and temporal features tend to be unstable. This paper describes a gesture-based user verification system in which parameterised models (embodying both physiological and behavioural aspects of the user's drawing actions) are used to predict moment-to-moment pen body inclinations during drawing/writing actions. The parameters in these models are estimated during execution of specialised training tasks. Analysis of system performance has shown that for typical drawing of a single preferred gesture (ie. from among the set which best discriminates the given user from other users) the error rate for equalised Type I and Type II errors is no more than 15%. For multiple (say 3 or 4) non-preferred gestures, the true user gives a computed probability of better than 0.84 while "impostors" give a probability of less than 0.45. The analysis here was based on a total of 1200 examples of 8 standard gestures (as used in Microsoft Pen Computing) obtained from 9 subjects over a two week period. Particular advantages of our system are that the verification performance is relatively accurate and stable, the methodology used does not require specific training with each of the set of gestures employed, being applicable to general drawing/writing actions, and little storage is required for the user models.

1. INTRODUCTION

On-line verification of handwritten signatures is regarded as one of the best means of automated personal identification because a signature is a "piece of identification" that can be produced nearly anywhere and at any time, unlike passwords or identity cards that only need to be known or possessed (and can also be discovered, stolen or lost) [Raphael & Young, 1974]. Gesture-based interaction has recently become popular with pen-based systems; in this mode of usage, the user specifies commands by simple stylised drawing movements [Rubine, 1991]. In the application we are studying, various modes of input such as typing, speaking and pen or pointer movement are being assessed for their ability to provide user-specific information during normal task related activity (ie. transparent verification), with the overall goal of improved security of access to sensitive data. Verification based on gestures not only inherits the advantages of on-line signature verification, but can also be implemented in a way which is less intrusive.

Traditional signature verification methods have the same basic processing pattern: generate user models in a training phase and then, in the operational phase, filter input data to reduce extraneous variations, compare the filtered data to the model for the claimed user (and possibly other users) and make an accept/reject decision [Plamondon & Lorette, 1989]. In trying to adapt this paradigm for gesture verification, two main problems are encountered: the individual gestures are much simpler than a typical signature, giving less information on which to base the verification decision; and dynamic features of the type successfully employed in signature verification tend to be either unstable or to have low discriminatory power. As far as we aware, the system described below represents the first successful use of pen gestures to assert user identity.

During writing and drawing activity, the trajectory followed by a pen tip can be related to muscle activity

[1] Technology for Robust User-Conscious Secure Transactions.

in the arm and to the angles of the joints in shoulder, elbow, wrist and fingers [Thomassen & van Galen, 1992]. A significant component of pen body inclination to the writing surface can also be described in musculo-skeletal terms. It thus follows that given a model which captures important aspects of each user's physiology and writing behaviour, the pen body inclination can be predicted with reasonable accuracy from the coordinates of the writing tip. If the writer and the person modelled are different, then the accuracy of prediction can be expected to decrease. This expectation forms the basis of our gesture verification system.

Most people write with the wrist joint for movements in the horizontal direction and the thumb and forefinger joints for movements in the vertical direction. Left-handers sometime reverse the roles of the wrist and fingers so as not to occlude newly-written material. An accepted working hypothesis is that the hand can be regarded as having two orthogonal degrees of freedom, with associated directions of movement. In each of these directions, muscles can be modelled as a pair of opposing springs representing agonist and antagonist muscles [Hollerbach, 1981]. This model has been modified and successfully applied at the "ink" level to improve character recognition accuracy [Simard *et al.*, 1993; Singer & Tishby, 1994].

Unfortunately, Hollerbach's model simplifies drawing/writing behaviour too drastically to capture some of the fine detail necessary for effective gesture verification. First, there is no explicit representation of hand physiology and musculature and few relevant kinematic features are represented (ie. those which are both stable and differ between users). Second, oblique elastic movements for the wrist along the finger axis and for fingers along the wrist axis. which provide useful information for discriminating between users' drawing/writing habits [Zhang & Macleod, 1995], cannot be represented. Third, the relation between the initial state of wrist and fingers and the hand movements which follow is potentially user-dependent, but is difficult to represent in Hollerbach's model.

Our Joint State Model [Zhang, Macleod & Glenn, 1994] is based on the finding that handwriting/drawing movements can usefully be described in terms of the state of wrist and finger joints. The state vector can be extended to describe elastic movements (*FEW, finger elastic equivalent to wrist direction movement*, and *WEF, wrist elastic equivalent to finger direction movement*). What we propose here is to employ enhanced user models to estimate the initial state vector and its sequence associated with a global task description and movements of the pen tip during gesture drawing, and to compare the estimates with the input signals at each sampling interval. We assume that a separate gesture recognition process supplies the verification module with the recognised gesture identity prior to verification analysis. Our approach can be contrasted with the conventional paradigm, in which the input data sequence would be analysed to recognise either a gesture or user.

Before we commence verification, we need to estimate the initial state. We do this on the basis of data acquired during a training stage which relates (in a user-dependent way) the starting posture to the range of wrist and finger movements required to draw the just-completed gesture. In this way, both global features (dependent on the user's plan of motor movement made before the gesture is drawn) and complete local information (relating to moment to moment joint movement) are combined to assert user identity. Note that the operation described here is relatively independent of the detailed form of the drawn gesture (apart from its spatial extent).

2. GVM: GESTURE VERIFICATION MODEL

2.1 Gesture Verification Model

The form of model is as follows:

$$\mathbf{JointS_{i+1}} = \mathbf{H}(\mathbf{M}(\mathbf{ts_{i+1}}), \mathbf{JointS_i}); i = 0, 1..., n$$

$JointS_i$ is the joint state composed of wrist angle, finger angle, FEW and WEF on the $\mathbf{i}^{\text{th}}$ node of the state chain; $\mathbf{ts_i}$ is the detected task (described in terms of its direction and amplitude) at the $\mathbf{i}^{\text{th}}$ step; and M is a state transition probability map (set up during the model training stage). Thus, assume that the $\mathbf{i}^{\text{th}}$ joint state is determined with the on-line detected task; the model would infer the $\mathbf{i{+}1}^{\text{th}}$ joint state. If the initial state can be determined the whole state chain is created.

Our drawing experiments revealed a strong correlation between task size and initial joint state. Regression analysis showed that finger angle was linearly correlated with finger task size, confirming the relationship apparent in Fig. 1.

$$\mathbf{Initial} = \mathbf{Na} - \frac{(\mathbf{Na} - \mathbf{FRI})(\mathbf{Final} - \mathbf{Na})}{\mathbf{FRA} - \mathbf{Na}} + \varepsilon \qquad (1)$$

$$\varepsilon \sim N(0, \sigma^2)$$

where the three parameters *Na*, *FRI* and *FRA* are each subject's natural finger state, minimum finger angle and maximum finger angle (determined in a training stage as described below); *Final* was the final finger angle required by the task. Projecting the state chain into pen angle space (pen body tilt to x and y axis) we get a concluded chain of $T_{x'}$ and $T_{y'}$ which we can then directly compare with the input signal T_x and T_y. Fig. 2 shows our experimental results for gesture "Cut" with GVM.

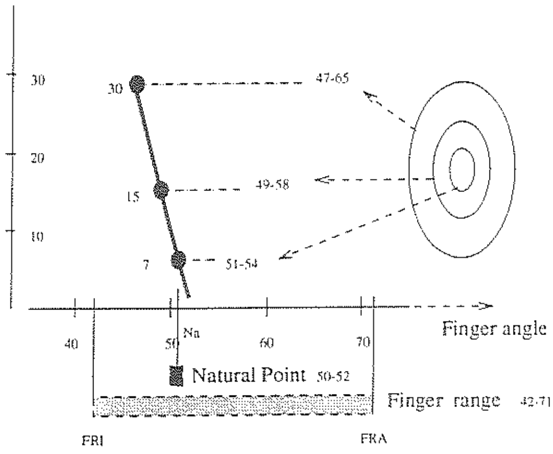

Fig.1 Initial Finger State for Circle Drawing

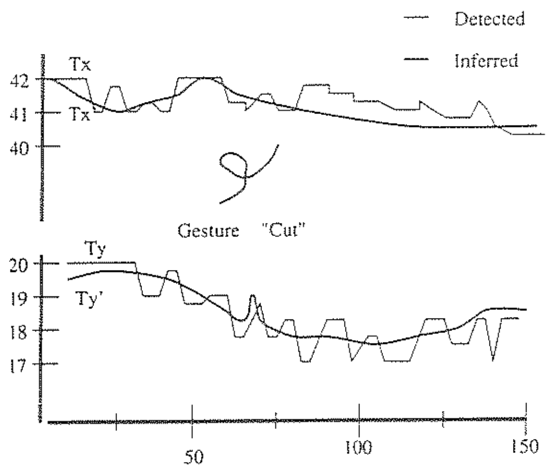

Fig.2 Detected and Inferred Chain

2.2 Model Training

Three special training tasks were designed: (i) finger flexion and extension on max/min *FEW*; (ii) wrist abduction and adduction on max/min *WEF*; and (iii) free-hand circle drawing. Important parameters for each user (such as finger and wrist angle range, natural point, palm height and length, pen body rest point, state possibility map, and coefficient slope of *FEW* and *WEF*) were obtained via regression analysis from the user's pen movements during execution of the training tasks.

We can decompose elemental movements in the x-y plane into wrist and finger requirements in a hand-based coordinate system. The wrist and fingers do not move in exact accordance with these calculated requirements because of elastic action. Correlations between (i) wrist and *WEF* and (ii) finger and *FEW* can be derived for each subject. Based on these correlations, a state transition probability map can be estimated (for detailed definitions and algorithms here see Zhang, Macleod & Glenn [1994]). Standard deviations of the various parameters and state transition probability maps are also estimated during this stage.

A model set up in this way will show variable sensitivity with different subjects/gestures. To improve the model's discrimination power, a distance distribution of each gesture for each subject was also estimated during the training stage. We can thus calculate moment to moment probability values during gesture drawing.

2.3 Decision Mechanism

Measurement of the fit between the directly detected signal and the model output estimate for the purported user is performed using a weighted Euclidean distance metric, similar to that used in signature verification [Crane & Ostrem, 1983]. Let $\vec{s}$ be the observed sequence of pen tilt, $\vec{t}$ be the model's output sequence (determined by the method described above) and $\vec{\sigma}$ the associated standard deviation.

$$d(\vec{s},\vec{t}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{s_i - t_i}{\sigma_i}\right)^2} \qquad (2)$$

where n is the length of the chain.

Each writer has certain preferred (idiosyncratic) gestures, determined during training, which distinguish him/her more reliably than others. Our system has two modes of working, according to the determined identity of the just-completed gesture and whether or not it falls within the set of preferred gestures for the claimed user. In the first mode, a single (preferable) gesture is identified according to the following decision rule:

If $P > d^{thres}$, the user is verified.
If $P <= d^{thres}$, the user is judged to be an impostor.

The quantity d^{thres} is preselected according to the measured dependencies of Type I/Type II errors on the threshold value. These error dependencies allow us to estimate the discrimination power of the model. Our "impostor" population was chosen at random from among the experimental subjects available to us and did not (to our knowledge) include any practised forgers. Very little is known about the characteristics of deliberate impostors/forgers; the generality of the data we used to represent potential forgers thus remains to be established. To circumvent this issue in practical applications, we compute a probability that the writer was the claimed user instead of making an accept/reject decision.

The second working mode operates with multiple gestures. As a user interacts with the computer, the series of gestures employed may not include any of the preferred gestures (ie. the gestures which for a given user have high discriminatory power). In this case, a series of gestures with lower individual discriminatory power may well give better overall discrimination than a single preferred gesture. Probability values P that the gestures were drawn by the true user are estimated as follows:

$$P_k = \frac{\sum_{i=1}^{k} W_i * P_i(d(\vec{s}, \vec{t}))}{\sum_{i=1}^{k} W_i} \qquad (3)$$

where k is the number of detected gestures. W_i is the i^{th} gesture weight as calculated in the training stage.

3. PERFORMANCE EVALUATION

3.1 Experimental Details

Nine right-handed subjects took part in evaluation experiments. First, they were told the purpose of the experiments. They then attempted to forge the true user's gesture after being shown its graphic form and practising up to 20 times.

The handwriting/drawing movements were captured by sampling x and y coordinates, axial pressure, and two-axis tilts of the pen tip (represented as $x(i)$, $y(i)$, $p(i)$, $tilt.x(i)$, and $tilt.y(i)$) simultaneously at a frequency of 100 Hz, using a CalComp "DrawingBoard III" digitiser. The specifications of this device are: worst-case position signal accuracy, 0.2 mm; tilt, 3°; and pressure, 1 force unit on a scale of 0 to 30.

Three training tasks and drawings of a standard gesture set were sampled in two sessions over two weeks. Each session comprised a training stage, gesture drawing and forgery attempts. There was a 5 minute rest break between stages.

3.2 Results

The discrimination power of the model was evaluated by joint consideration of Type I and Type II error rates defined as follows:

$$\text{Type I error} = \frac{R^t}{T^t} \text{ and Type II error} = \frac{R^f}{T^f}$$

T^t represents the total number of trials in the true-drawer data base and R^t the number of trials for which a true drawer was falsely rejected. Similarly, T^f is the total number of forger trials and R^f is the number of trials for which a forged gesture passed the verification criteria. The result of typical circle drawing for a subject is shown in Fig. 3. By varying the acceptance criterion, the equalised error rate (Type I/Type II) was found to be no more than 15%. For such a simple gesture, this is a most encouraging result!

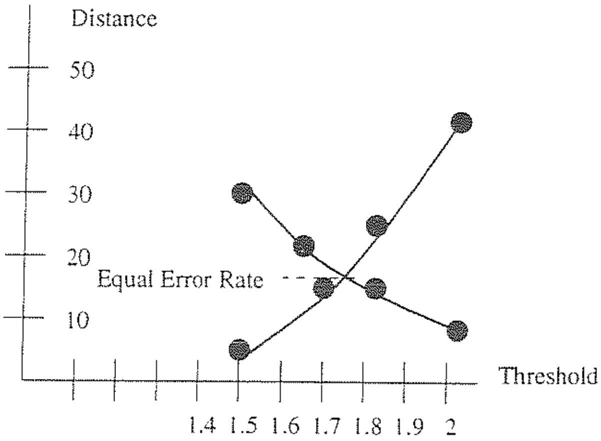

Fig.3 Type I/Type II error curves

Fig. 4 shows a multiple gesture mode result with a series of four gestures: 7–10–11–8 (InsertNL–Copy–Paste–Cut). As the number of gestures in the series increases, the P value of the true user remains approximately constant at about 0.87, but that of an impostor generally decreases.

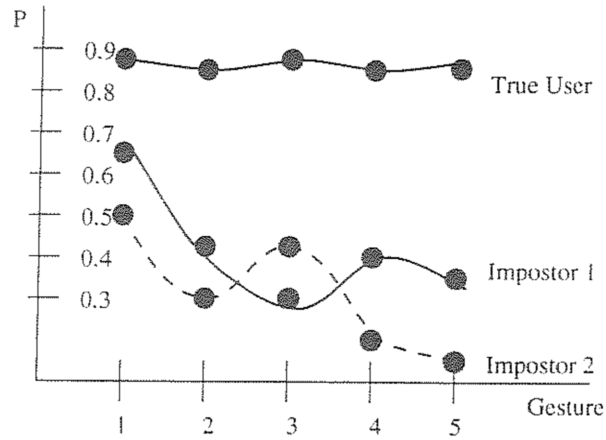

Fig.4 Multiple Gesture Verification

4. SUMMARY AND DISCUSSION

Our proposed GV model is based on a chain of joint state vectors. Unlike usual verification systems, the true user's reference data are not resident in memory but are generated on-line according to the detected task. The main advantage of this strategy is that new gestures can be included which were not part of the training procedure (performance will generally be improved with comprehensive training, however). Relatively invariant drawing/writing motor features were employed as model parameters to improve model stability. Such features include: (i) grasp posture, handedness, wrist and finger movement and their elastic contributions; and (ii) hand physiology features such as palm height and length, wrist and finger range and so on. Distance measurements are taken at each sampling interval by comparing the detected state chain with the estimated state chain. Without having to use additional feature extraction or time warping, global and local information with practical user-specificity is derived from the limited strokes of gesture drawing.

The techniques described are applicable to general drawing/writing actions, meaning that user identity can still be assessed (but with lower confidence) across the range of pen actions employed with a pen-based computer.

5. REFERENCES

Crane, H. D. and Ostrem, J. S., Automatic signature verification using a three-axis force-sensitive pen, *IEEE Trans. Sys. Man & Cybern.*, *13(3)*, 329–337, 1983.

Hollerbach, J. M., An oscillation theory of handwriting, *Biol. Cybern.*, *39*, 139-156, 1981.

Plamondon, R. and Lorette, G., Automatic signature verification and writer identification — the state of the art, *Pattern Recognition*, *22(2)*, 107-131, 1989.

Raphael, D. E. and Young, J. R., Automated Personal Identification, *Long Range Planning Service*, Stanford Research Institute, Menlo Park, CA, 1974.

Rubine, D., Specifying gestures by example, *Computer Graphics*, *25(4)*, 329–337, 1991.

Simard, B., Prasada, B. and Mahesh R., On-line character recognition using handwriting modelling, *Pattern Recognition*, *26(7)*, 993–1007, 1994.

Singer, Y. and Tishby, N., Dynamical encoding of cursive handwriting, *Biol. Cybern.*, *71(3)*, 227–237, 1994.

Thomassen, A. J. W. M. and van Galen, G. P., Handwriting as a motor task: experimentation, modelling, and simulation, in *Approaches to the Study of Motor Control and Learning*, 113–144, Elsevier Science Publishers B.V., 1992.

Zhang, X. J., Macleod, I. and Glenn, B., Transparent verification of user identity with pen based systems, *Proceedings of ICCT'94, Shanghai*, *2*, 1081–1085, 1994.

Zhang, X. J. and Macleod, I., Oblique elastic movement for wrist and finger while writing/drawing, *in submission*, 1995.