SOME EXPERIENCE IN THE REAL-TIME PROCESSING OF HANDWRITING

B.R.Gaines, I.D.McKellar, W.P.Dinger, S.R.Fast,
B.J.Fowles, M.A.Fraccaro, G.C.Jolivet & A.B.Maludzinski

CADRE Information Transfer Systems,
339, Canarctic Drive, Downsview, Ont. M3J 2P9.

In recent years the costs of digitizing and recognizing handwriting have become low enough to make these technologies available in office workstations. As digital networks become more common in the office there is growing interest in using them for a range of different communications purposes. This has been particularly noteworthy with speech where what was once seen as an informal, interactive and transient medium has been developed by analogy with textual messages into a very much more flexible medium. This paper describes a similar development for handwriting and shows how the application of digital techniques to this medium provides a new office technology that integrates well with existing media and provides new resources of its own. Speech and handwriting together provide the appropriate natural user interface for fifth generation systems that can be used by everyone with very few concessions to the technology.

INTRODUCTION

There has been considerable research on handwriting in the past but it has concentrated on character recognition[1] and signature verification[2]. The role of digitized handwriting as a medium for carrying information in computer networks has been neglected, as was the comparable role of digitized speech until commercial systems were introduced in recent years. This paper is based on system research that commenced with the development of a low-cost, real-time hand character recognizer as part of an office image communications network[3], but has extended the technology to encompass all aspects of handwriting in communication systems. The system developed can handle forms containing a mixture of fields for: drawing and handwriting that need not be recognized; signatures for verification; mark sense fields; and alpha-numeric character recognition for communication to the computer.

The digitization of handwriting itself leads to substantial benefits:-

(a) Digitized handwriting (produced by writing on a piece of paper on a digitizing clip-board) can be transmitted over normal communication media, i.e. police radio, telephone lines, local area networks. In particular, speech and handwriting may be freely intermixed.

Hence, documents and diagrams may be generated naturally by people without computer or keyboard skills and sent rapidly to one or more remote locations.

(b) Digitized handwriting may be stored electronically and retrieved at computer terminals or reproduced on paper.

Hence, paper filing systems involving manual access to large numbers of documents may be replaced by electronic systems requiring no physical access.

(c) Multiple copies of any required document may be produced as required in geographically separated locations.

Hence, the users of the documents do not have to be close to the filing system and may be located where most convenient for their other activities.

(d) All or part of any document can be changed in size and reproduced as part of another document.

Hence, condensations of multiple documents may be prepared simply and easily.

(e) Where part of the document has to be encoded for computer processing (as contrasted with storage and communication only) this can be done in a number of simple and natural ways. Firstly, **positional encoding** may be used for much information, i.e. the user ticks a box and the computer notes that a mark has been made at a defined location. Secondly, **numeric character recognition** may be used, i.e. the user writes numbers in the box; since there are only ten digits the recognition can be done with high accuracy. Thirdly, **alphanumeric character recognition** may be used; the range of possible characters makes the task more difficult and the user has to be careful in writing but this works well with trained users.

Hence, it is also possible to use the system for data entry and access.

(f) Computer checking of a handwritten signature may be used when verification of the authority of a user to enter or access a document or computer data is required.

Hence, it is possible to use the document entry system to check the authority of users to perform requested actions.

These findings are consistent with similar experience by IBM and others of the benefits of digitizing speech[4,5]. The ability to digitize spoken messages, store them, communicate them, edit them, run through those received rapidly, add comments to them, forward them to others, and so so, has proven to be an important new office technology. The same is true for handwriting. Digitized speech and digitized handwriting are more flexible forms of conventional speech and handwriting. They are true fifth generation computer technologies in that they give users a simple and natural interface to the computer using their existing skills.

DEVELOPMENT OF A TELIDON INPUT PAD (TIP)

The objective of our research has been to develop a low-cost tablet for handwriting digitizing and processing similar in shape and operation to the "clip-boards" used when writing on a pad of paper that has to be carried around. The tablet has on-board real-time processing for handwriting filtering, smoothing and encoding, field detection, block character recognition, editing graphics and text, and signature verification. It also has local display and storage facilities together with communications so that it can be used either on-line to a computer or to batch and store a reasonable amount of data (e.g. one day's forms, some 20K characters say) for later block-transmission.

We have called the system a **Telidon Input Pad** (TIP) because the NAPLPS[6] encoding scheme is very well-suited to the mixture of handwriting, graphics, text, status and positional data, which the pad generates. It is particularly important that the presentation level protocol is standard so that the pad can operate as a component of open interconnection networks in conjunction with processes and output devices of unknown origin but operating to an industry standard. We see one product resulting from this research being a desktop pad for true electronic mail in which the user can mix text and graphics freely, possibly in conjunction with voice communication, either directly to another user or via a computer-based mail system.

The following sub-sections describe the architecture of our system, the problems encountered and the approaches developed.

Digitizer

Our requirements for a writing surface were that it should be: low in cost, below $20 say; durable, capable of some 10,000 or more writings; easily replaced if damaged; compact and light in weight, similar to a normal hardboard clipboard; give adequate resolution, 1000 by 1000 say; and have adequate linearity, 1% say. The resolution requirement is based on having at least 8 linearly resolvable elements (pels) on a 3mm high character on an 8" by 11" sheet. The linearity requirement is based on allowing the separation between boxes on business forms to be as low as 3mm.

We currently use a potentiometer technique in which pressure from the pen through the form forces contact between a conducting rubber layer acting as the wiper arm and a resistive layer across which a voltage is applied. An intervening mesh ensures that only point pressure leads to contact, not area pressure from the hand resting on the tablet. Voltages are switched across the X and Y axes of the resistive surface alternately at 2 millisecond intervals and the output voltage digitized through 10-bit dacs. This gives an effective sampling rate of 250 samples/second at the required resolution.

The output from the dacs is interfaced to a simple microcomputer system: a 4MHz Z80 with 64K ram, up to 56K rom and up to 48K battery-backed cmos. This, together with a real-time clock and an RS232 serial port, are contained on a a single printed circuit card at the back of the digitizing surface so that the complete unit can be packaged as "clipboard" about 1/2" thick. A liquid crystal display is used to give information and show characters recognized to the user. The Z80 operating system is specialized for handwriting processing, essentially interrupt-driven multi-tasking.

Filtering and Encoding

The use of an analog input surface with intermittent contact through a conducting layer results in low cost but some problems of noise. In particular, spurious readings can occur as the pen makes contact initially or loses contact in a low-pressure stroke. Fortunately the dynamics of handwriting make filtering this noise fairly simple. Like all human movement hand motion is produced by bursts of muscular contraction, effectively bounded acceleration acting on inertia[7,8]. Writing is characterized by the time sequence of quantized accelerations in two axes, and different sizes of writing are produced by the same time sequences with different levels of acceleration[9,10].

With the bang-bang movement pattern described in the references, the peak velocity is twice the average velocity. A typical stroke in block writing takes between 200 and 400 ms and the table below shows the peak velocities reached in pels/sample for our 1000 pel resolution, 4msec sampling tablet. It can be seen that for normal handwriting of 3 to 6 mm characters there will usually be 0 or 1 pel change between sample instants showing that our sampling rate is adequate. For large hand movements in drawing there may be up to about 20 pel changes. Thus a noise threshold of about 20 pels is adequate to reject spurious inputs without losing data. These design calculations have been confirmed in our empirical studies of handwriting.

Distance Moved			Peak Velocity in pels/sample	
			200ms Movement	400ms Movement
Pels	mm	in	50 samples	100 samples
10	3	1/8	0.4	0.2
25	6	1/4	1	0.5
50	12	1/2	2	1
100	25	1	4	2
250	63	2.5	10	5
500	125	5	20	10
1000	250	10	40	20

TABLE I Peak Velocity as Function of Distance Moved

This table is useful also in considering encoding schemes for handwriting. The preponderance of 0 and 1 pixel moves suggests a chain encoding scheme. The inertia of the hand limits the speed at which changes in direction can be made and suggests that source encoding in terms of the acceleration patterns would be optimal. However, these are difficult to recover and are more complex than the simple bang-bang orthogonal model suggests[7]. The unsigned increments of the Telidon incremental line take advantage of the velocity stability of handwriting and provide a fairly efficient encoding.

Field Recognition

The TIP is intended to be used with structured forms where some fields may be for character recognition, others for mark sensing and others for graphics. Hence the next stage processing after the writing has been captured depends on where it is being written on the form. **Form definitions** are held in rom, cmos or remotely for downline loading. They specify a number of non-overlapping rectangular areas together with associated characteristics such as numeric data, alphanumeric, graphic, or function call. An interactive form definition utility enables users to define the structure of their forms for use with the TIP.

Since the user may write in any part of any field, transfer to any other, and then return to add to the first, and the interpretation of the writing depends on the field, it is important that field recognition is accurate and takes place early. It is generally not possible for it to be done sample by sample since the data rate is too high and hence field recognition has to be run as a background task to data collection and character recognition. This is a significant real-time load that is usually not taken into account in studies of character recognition alone.

Character Recognition

Field recognition may be used to collect mark sense data and to activate functions within the TIP such as calculation of extensions or totals on invoices from entered data. If the field is noted as **graphic** in type then the encoded data is buffered in chain encoded form ready for editing or transmission. If the field is noted as **numeric** or **alphanumeric** then recognition logic is used to process the data as character shapes.

The TIP recognition algorithms are simple but effective. Stroke detection is used to recognize the path of the pen from pen-down to pen-up. This is then boxed within a very coarse grid, only 4 by 4 giving 16 cells. The path is matched against sequential masks to give first level recognition. Specific second-level tests are then done for features dependent on the first level result, for example ratio of X and Y dimensions to discriminate a 1 from a /. Third-level tests are then done to see whether successive strokes may be grouped to form a multi-stroke character. Fourth-level tests are then done to see whether the character lies on top of one already written and hence over-writes it, whether there is a definite space between it and the preceding character, and so on.

Multi-stroke characters raise interesting problems in real-time feedback. The TIP display shows the user its intermediate results since they may well be the final ones. So the display will show 1, 1-, 1-1, H, in the final example. The recognition rate of these fairly simple real-time algorithms varies with the amount of practice of the user. A skilled user gets 100% in numeric fields and about 97% in alphanumeric. A first-time user with an ornate script gets 70% but rapidly adjusts to the TIP if motivated.

CONCLUSIONS

Digitized handwriting is an important fifth generation computer technology. It enables computer-based information systems to become **user natural** and to demand the minimum of change in their users' operations while providing new benefits. There are already developments of simple workstations combining the capabilities of voice and handwriting communication[11] for direct person-person interaction. The hardware technology is now sufficiently low in cost to bring such terminals into common use. The standardization of the open systems interconnection protocol, particularly through Telidon at the presentation level, makes a universal communication network possible using these terminals. The twenty five years of investment in research in handwritten pattern recognition can now become effective in adding additional features to these terminals for direct computer interaction.

ACKNOWLEDGEMENTS

The TIP project has been partially funded by an IRAP grant from the National Research Council, Ottawa.

REFERENCES

[1] Suen, C.Y., Berthod, M. and Mori, S., Automatic recognition of handprinted characters - the state of the art, **Proceedings of the IEEE, 68**(4) (1980) 469-487.

[2] Herbst, N.M. and Liu, C.N., Automatic signature verification based on accelerometry, **IBM Journal of Research and Development, 21**(3) (1977) 245-253.

[3] Gaines, B.R., From word processing to image processing in office systems, **Proceedings 1983 International Electrical, Electronics Conference, IEEE 83CH1955-4**(2) (1983) 622-625.

[4] Gould, J.D. and Boies, S.J., Speech filing - an office system for principals, **IBM Research Report, RC9769** (IBM Research Center, Yorktown Heights, NY, 1982).

[5] Allen, R.B., Composition and editing of spoken letters, **International Journal of Man-Machine Studies, 19**(2) (1983) 181-193.

[6] X3, **Draft Proposed American National Standard for Videotex/Telextext Presentation Level Protocol Syntax** (X3 Secretariat/CBEMA, Washington DC, 1983).

[7] Gaines, B.R., Linear and nonlinear models of the human controller, **International Journal of Man-Machine Studies, 1**(4) (1969) 333-360.

[8] Hollerbach, J., Understanding manipulator control by synthesizing human handwriting, in Winston, P.H. (ed.), **Artificial Intelligence: An MIT Perspective. Volume 2** (MIT Press, Massachusetts, 1979) 311-332.

[9] van der Gon, J.J.D. and Thuring, J.Ph., The guidance of human writing movements, **Kybernetik, 2**(4) (1965) 145-148.

[10] van der Gon, J.J.D., Thuring, H.Ph. and Strackee, J., A handwriting simulator, **Physics in Medicine and Biology, 6**(3) (1962) 407-414.

[11] Hasui, K., Katsuyama, T., Kato, M. and Hattori, S., Handwritten message switching via an integrated EPBX, **Proceedings Localnet'83** (New York, 1983) 19-29.