Steps toward Handwriting Analysis and Recognition

Makoto Yasuhara

In order to introduce to the field of human handwriting recognition formal concepts which correspond to those of the analysis by synthesis method in speech analysis, a dynamic model of handwriting process is proposed. Discussion to support the proposed scheme is presented both from the theoretical and from the experimental points of view.

1. *Introduction*

Undoubtedly, writing and speech play the most important roles in human communications. These two forms of verbal behavior have appreciable similarities as well as differences. Although their physical appearances are different—and although different sets of muscles are used in the execution or articulation of the appropriate gestures, the problems they present in the higher levels of language processing seem to be the same. That is, at some level of the language processing in the brain, both must relate to the same grammatical structures of the language in question. Thus, processes corresponding to the production or perception of handwriting and to those of speech may be in a one-to-one relation at that level. So, many of the problems concerning speech processing would probably be related closely to the corresponding problems concerning visual language processing.

In this paper, an active method for handwriting analysis and recognition is proposed, which corresponds to the so-called analysis by synthesis method (henceforce abbreviated to A-b-S) in speech research.

K. N. Stevens (1960) first proposed A-b-S method as a guiding principle of his speech analysis and recognition. This is one of the typical and practical techniques that execute the idea of active analysis. The first attempt to apply A-b-S scheme to handwriting and pattern recognition was made by M. Eden (1962) at MIT. He considered two alternative models of the description of cursive script. One was based on a set of defined primitive stroke segments analogous to the distinctive feature in speech analysis; the other was based on a sinusoidal model of practiced writing. However, neither of these models can be considered to simulate the actual human handwriting.

2. *Synthetic Model of Handwriting*

2.1 Van der Gon's Model

Cursive handwriting may be considered as a highly skilled process which is executed by means of a rapid sequence of motion. As early as 1917 Lashley studied such processes, including the principle of position feedback. Although this principle doubtlessly works in guiding some human movements to a certain extent, it is well-known that it does not apply to quick and well-practiced movements. Thus Lashley (1961) concluded that in these cases, an effector mechanism can be primed to discharge at a given intensity or for a given duration, independent of any sensory control. According to the above background, Denier van der Gon postulated the following assumptions:

(1) The effects of position feedback can be neglected for cursive handwriting.

(2) The timing of muscle contraction determines the shape of the pattern to be generated; the magnitude of the applied muscular force does not play an important role.

Van der Gon proposed a simple model of the human handwriting process, in which the hand-pen couple was assumed as a mass point that moved according to the following simple equations,

$$\left.\begin{aligned} \ddot{x} + R_x\dot{x} &= F_x(t) \\ \ddot{y} + R_y\dot{y} &= F_y(t) \end{aligned}\right\} \quad (1)$$

where R_x and R_y represent time invariant model parameters. The handwriting patterns are considered as the resultant loci of the movement of a equivalent mass point.

2.2 Refined Model of Handwriting

In order to have more refined and generalized discussions on the dynamics of human handwriting movement, consider the process

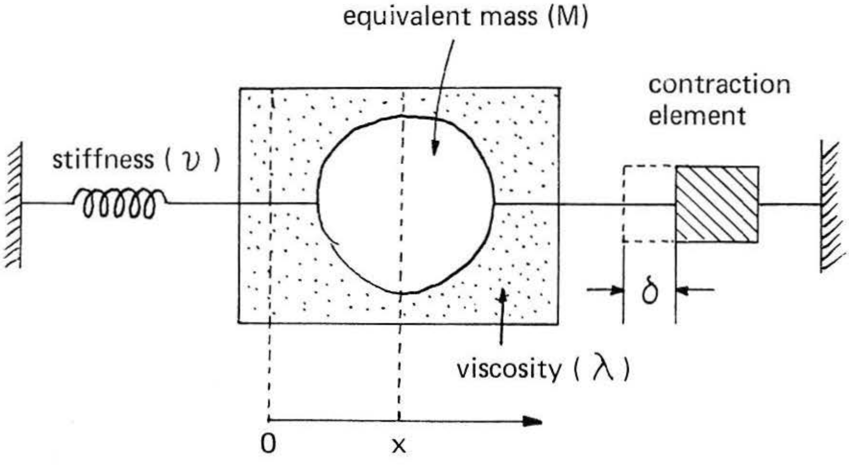


Figure 1. Schematic illustration of the dynamic model of the human handwriting process.

schematized in Figure 1, in which the effects of stiffness, internal friction of muscle, and friction force between the surface of paper and a pencil point are taken into account. The motion of a mass point, when the muscle contracts by δ, is expressed by the following dynamic equations,

$$\left.\begin{aligned} \ddot{x} + r_x\dot{x} + n_x x = h_x\delta_x = F_x(t) \\ \ddot{y} + r_y\dot{y} + n_y y = h_y\delta_y = F_y(t) \end{aligned}\right\} \quad (2)$$

where,

$$\left.\begin{aligned} & F_x(t) = f_x(t)/M, F_y(t) = f_y(t)/M \\ & r_x = 1_x + m_x p(t)/v \\ & r_y = 1_y + m_y p(t)/v \end{aligned}\right\} \quad (3)$$

$1_x = \lambda_x/M,\ 1_y = \lambda_y/M$	viscosity coeff. of muscle
$m_x = \mu_x/M,\ m_y = \mu_y/M$	friction coeff.
$n_x = \gamma_x/M,\ n_y = \gamma_y/M$	stiffness coeff. of muscle
$v = (\dot{x}^2 + \dot{y}^2)^{1/2}$	writing speed
$p(t)$	writing pressure
M	equivalent mass

If the writing pressure $p(t)$ and the writing speed v are time-invariant, eqs. (3) become

$$\left.\begin{aligned} r_x &= l_x + m_x P/V = R_x \\ r_y &= l_y + m_y P/V = R_y \end{aligned}\right\} \quad (4)$$

Consequently, the normalized effective viscosities R_x, R_y are proved to be time-invariant. And further, by neglecting the stiffness term of the muscle which is actually considered to be of less importance compared with the effective viscosity term, eqs. (2) become

$$\begin{aligned} \ddot{x} + R_x\dot{x} &= F_x(t) \\ \dot{y} + R_x\ddot{y} &= F_y(t) \end{aligned}$$

These are exactly identical to the model proposed by Van der Gon. It is clear, therefore, that the tacit assumptions have been introduced in Van der Gon's model, that the speed and pressure of a pencil point during handwriting movements are time-independent and the stiffness term of muscle is small enough to be neglected compared with the effective viscosity term.

Although our understanding of the physiology of the actual human handwriting process is not complete, we can understand some of operations required in human handwriting process, and a limited model of handwriting can be simulated—as demonstrated in the Appendix.

3. *Handwriting Analysis and Recognition*

In this section, the formal concepts of the active analysis of handwriting—analyzer-synthesizer model—are introduced to the field of human handwriting analysis and recognition with reference to the actual human handwriting process.

3.1 General Description of Handwriting Analyzer

The block diagram shown in Figure 2 illustrates the basic structure involved in the handwriting analyzer to be described. The first stage of the analyzer accepts the handwritten specimen to be analyzed as its input and provides the physiological descriptions (or equivalents) that are necessary to generate the input handwritten specimen. At the next stage, a description in terms of strokes is derived, into which

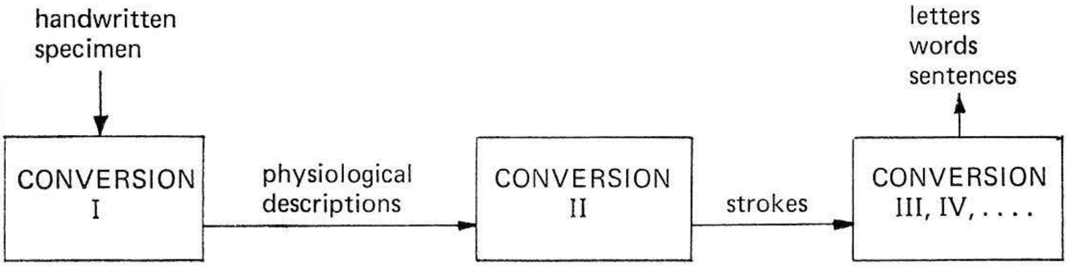


Figure 2. Block diagram illustrating several stages of analysis in the handwriting analyzer.

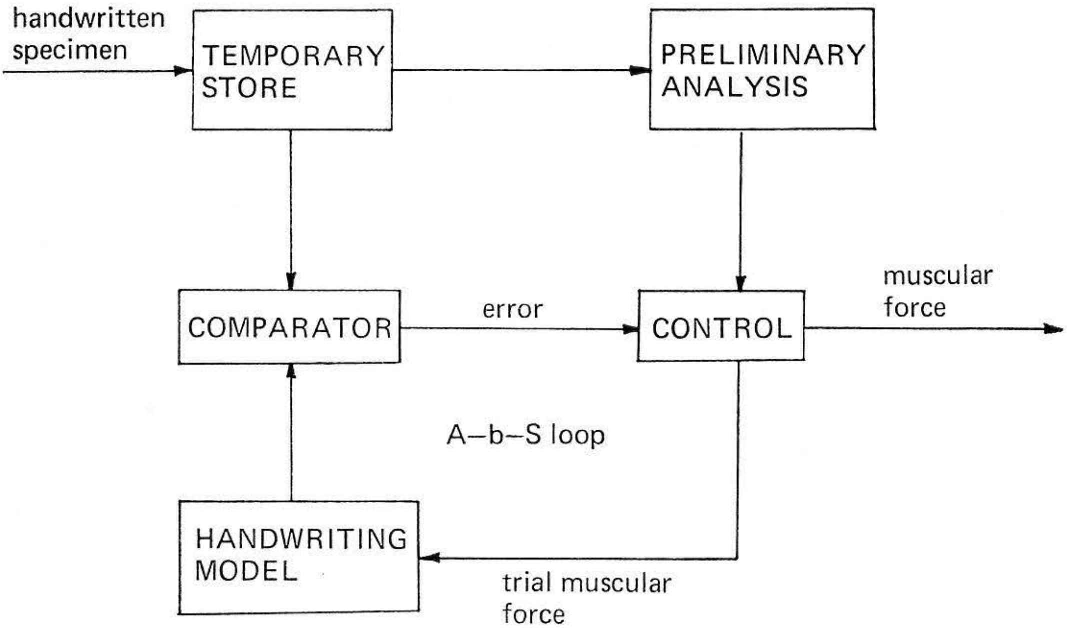


Figure 3. Block diagram of an analysis-by-synthesis procedure for extracting muscular force signals from an input handwritten specimen.

handwritten specimens are segmented (Appendix), and in the higher stages the sequence of strokes is converted into letters, words and sentences. In this paper, we shall engage principally in the conversion of handwritten specimen into a sequence of strokes. We shall note here that the method of analysis at each stage can be described as "active" rather than "passive"; that is, the analysis is performed by synthesis according to stored rules and by comparison of the synthesized signals with the input signals to be analyzed.

3.2 Conversion to Physiological Descriptions

Before describing the details of Conversion I in Figure 2, there must be some comments on the signal representation at the output of this stage.

We consider two levels, called roughly "muscular" and "neurophysiological." The representations at the muscular level are in terms of the actual muscular forces applied to hand-pen couple to generate a sequence of strokes. We expressed these as $F_x(t), F_y(t)$ in the previous section. When discussing effects of a writing pressure, we must add the muscular force $F_z(t)$ to cause a pen-point pressure $p(t)$. The second and perhaps more fundamental representation, termed neurophysiological, should specify in some sense the neural control signals that must be transmitted to the motor unit of muscles to cause them to generate the prescribed force.

In order to simplify the discussion, we shall refer to the representations at this stage simply as "physiological descriptions."

The detailed structure of the first stage "Conversion I" in Figure 2 is shown in Figure 3. It is the block diagram of analysis-by-synthesis procedure for extracting muscular force signals from a handwritten specimen. The input specimen, which may be placed in temporary storage, is compared in the comparator with signals synthesized by the model. Instructions as to the muscular forces to be tried are transferred to the model by the control component, which bases its decisions on the results of a preliminary analysis of the input specimen and on the output of the comparator for previous trials. When the best match is obtained in the comparator, the control component reads out the muscular force which, through the model, produced that match.

As mentioned above, the function of the model is to synthesize handwriting specimens from the physiological descriptions. The model must, furthermore, be able to simulate the handwriting by not just one but by a number of writers. Thus part of physiological descriptions must include the characteristics that are unique to a particular writer and that are relatively invariant in the particular handwriting material.

To program completely a machine to perform all these functions is clearly beyond our capabilities at present. But the model proposed in the previous section has been confirmed to be effective as shown by experimental studies (Appendix).

3.3 Conversion to a Sequence of Strokes

Since the active analysis procedure seems to have certain attractive features when applied to extracting the so-called physiological description, it is reasonable to examine whether the same general approach can be applied to a higher analysis stage in the analyzer; that is, the conversion of physiological descriptions to representations in terms of a sequence of discrete symbols—strokes.

Figure 4 shows the internal operations that would be involved in the analysis stage labeled "Conversion II" in Figure 2, if the method shown in Figure 3 were extended to this level of analysis. The "model" stores various rules that relate a sequence of strokes to physiological descriptions. The box labeled "control" determines the order in which different sequences of strokes are selected and converted to physiological descriptions for comparison with input data. The output is the sequence of strokes that produces a minimum error at the comparator.

Once the handwriting specimen has been decoded into a sequence of strokes, these strokes must be converted to letters and grouped into words and sentences in the later stages. We shall not be concerned here with this transformation into letters, words, and sentences.

3.4 Two-stage Scheme for Handwriting Analysis

The components discussed in Figures 3 and 4 are shown connected together in Figure 5. To begin with, assume that we have a sequence of strokes. If the output of the handwriting synthesizer is restricted to a particular writer's handwriting, the rules stored in the box labeled "Model II" would provide the conversion from stroke

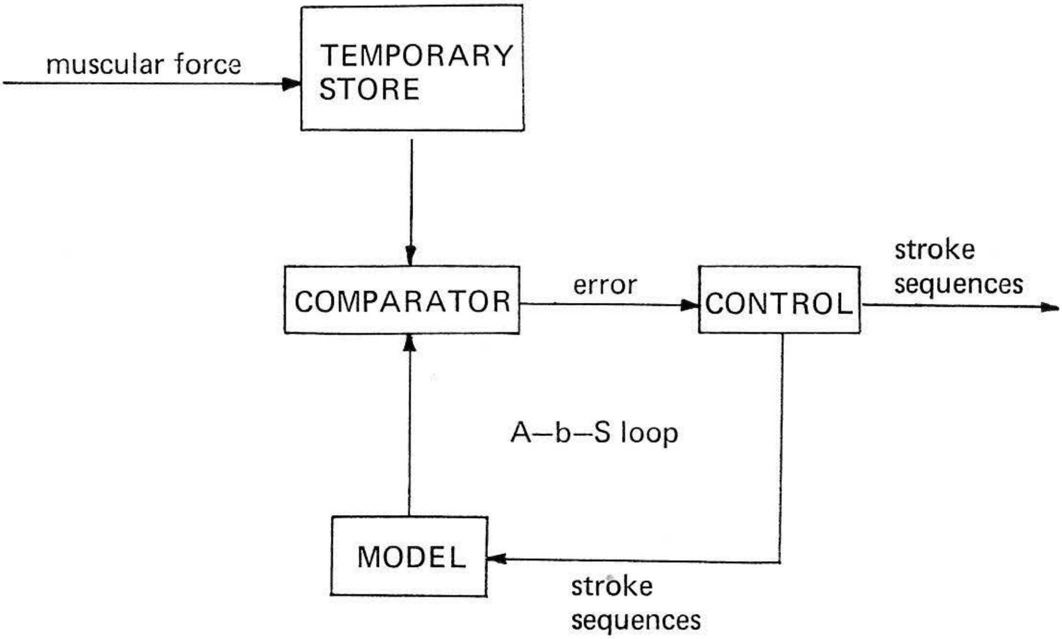


Figure 4. Block diagram of an analysis-by-synthesis procedure for extracting a stroke sequence from an input muscular force.

sequence to a muscular force description for that writer. Thus conversion to the muscular force description is accomplished by the Model II of Figure 5. This model is used both for analysis and for synthesis. If the handwriting output of the synthesizer is connected to the input of the analyzer, the over-all A-b-S loop may be completed. Such a feedback loop would permit comparison of derived muscular forces with those actually used. Thus adjustment of computations in Model II could minimize the difference between these two measures.

Figure 5 may be viewed as a proposed model of a complete analyzer-synthesizer model for human handwriting analysis and recognition.

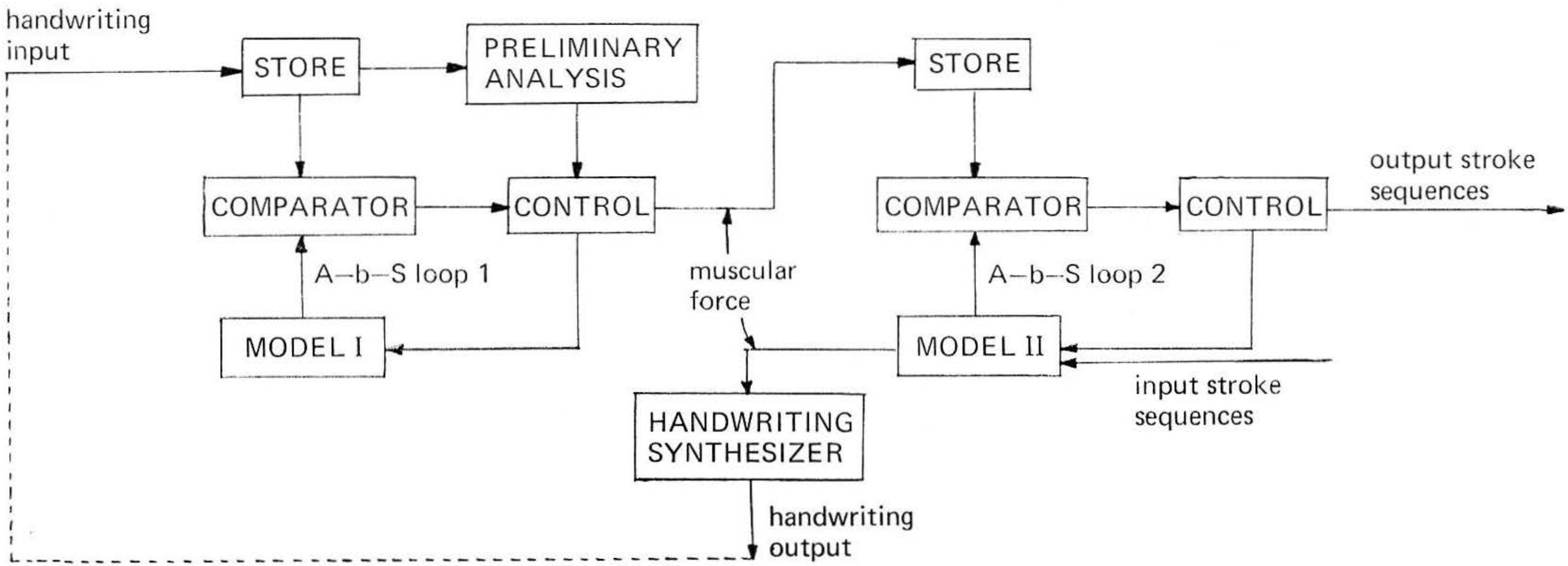


Figure 5. Block diagram of a two-stage scheme for handwriting analysis—overall system block diagram of the proposed handwriting analyzer.

Concluding Remarks

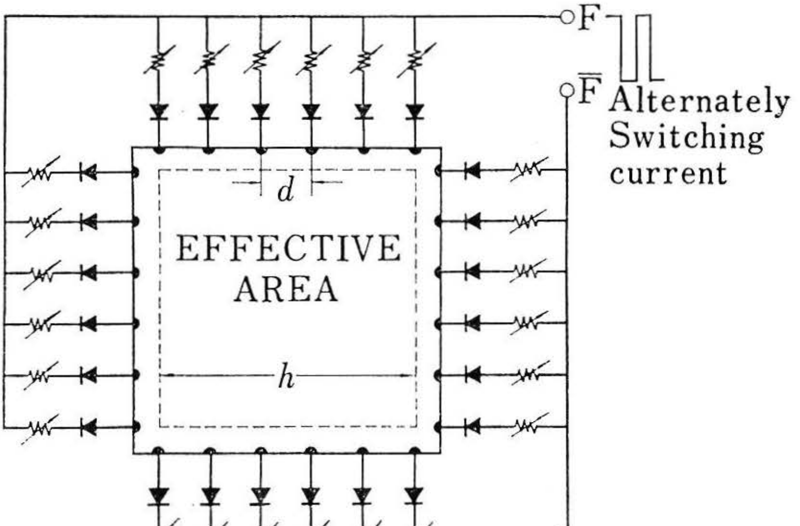
We have proposed herein an approach to the design of certain portions of a machine for the analysis and recognition of handwriting. The main features of the proposed system are the following:

(1) Between the handwritten specimen and the stroke output, we have two stages of conversion. The descriptions at the output of the first stage have been termed "physiological descriptions" which may take one of two forms termed "muscular" and "neurophysiological."

(2) At each stage of conversion, the analysis is performed by the active synthesis of signals which are compared with the input patterns that are under analysis.

(3) The analyzer-synthesizer system (Figure 5) has certain properties in common with the actual human handwriting process.

Figure 6. The principle of an X-Y position transducer.



APPENDIX

1. Handwriting Analyzer

We must, first of all, measure the displacement, velocity, acceleration of handwriting movements, and the writing pressure of the pencil point using a device for transforming X–Y coordinates and a pressure of a pencil point into electrical signals as functions of time. The handwriting analyzer generates such signals, and these signals can be connected to X, Y, and Z axes of a cathode ray tube display as well as to an input channel of a computer.

The principle for picking up this information with the stylus is illustrated in Figure 6. The writing plate consists of a square of uniform resistive material fed with equal magnitude current. The contacts are identically shaped and equally spaced along the four edges. A uniform horizontal and vertical current flow is established alternately by switching the operational mode of the respective contact sets. Because the contact separation (d) is small compared with the writing plate width (h), there will be a uniform potential gradient between mutually faced contact sets, except for local deformations in the vicinity of the edges. Thus, the time variant analogue samples of X–Y coordinates are generated by placing or moving the stylus over the resistive surface. A durable, highly transparent stannic oxide (SnO_2) coating on a carefully polished plate of glass shows no physical or chemical deterioration after several months of use. The stylus used is a soft lead pencil which can move freely over the coating. The writing plate is held up at the four corners by pressure sensitive gauges—strain gauges of non-contact type—and the signals from them are correlated. The multiplexer and analogue to the digital converter convert simultaneously three channels of analogue signals, corresponding to X–Y coordinates of the stylus and the writing pressure Z, respectively, into eleven bits of digital samples at a sampling rate of 400Hz.

Figure 7 shows the system block diagram. Figure 8 shows the electrical potential distribution generated over the resistive surface. Figure 9 shows an example transferred to the computer, intending to represent the Japanese letter "ru." The intensity of the pencil-point pressure is also shown in the figure by the printed numerics. Figure 10 represents some examples of the measured writing pressure wave forms.

In an effort to gain some knowledge of muscular force during writing movements, some electromyographic (EMG) records of muscle activities in the forearm are also taken. The equipment used can record two channels of EMG's simultaneously. A pair of surface electrodes is placed on the dorsal part of the forearm, that is, directly over the muscle, flexor

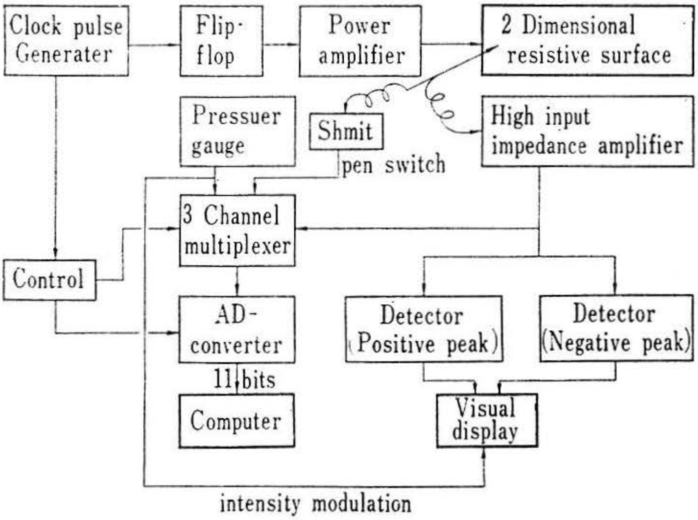
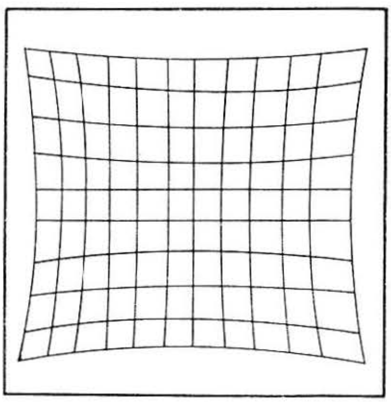


Figure 7. The system block diagram of handwriting analyzer.



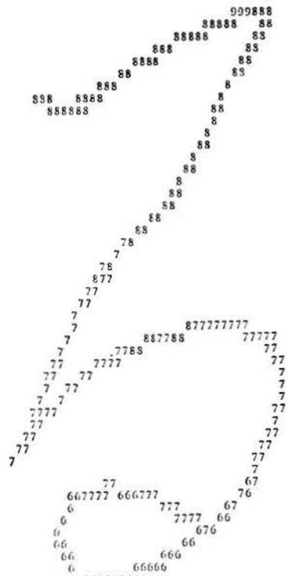


Figure 8. The electrical potential distribution generated over the resistive surface (120 mm $\times$ 120 mm). The central part is selected to use.

Figure 9. Japanese letter "ru" which is transferred to the computer, representing the writing pressure by the printed numerics.

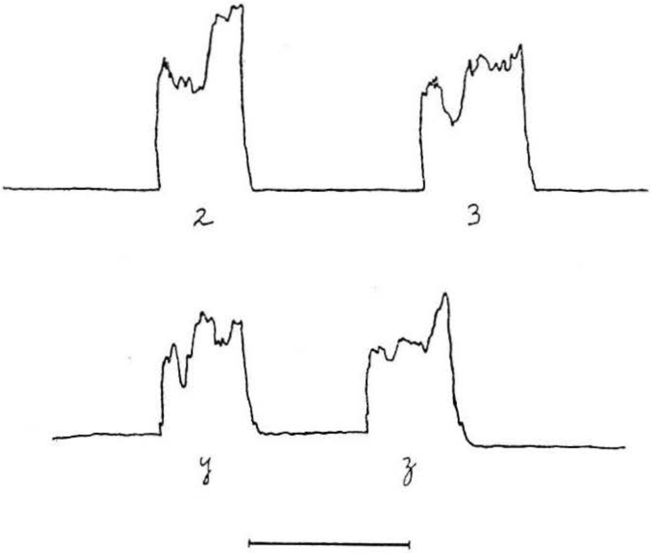


Figure 10. Some examples of characters and their measured waveforms of writing pressure.

carpi unlaris (EMG 1); the other is placed over the muscle, extensor carpi ulnaris (EMG 2). Figure 11 shows one of the records of the observed left-right displacement $x(t)$ and pressure waveform $P(t)$ of a pencil point together with two channels of the corresponding EMG's for subject Y (a 29-year-old male).

Case A. First of all, in order to estimate the optimum value of parameter R_x appearing in the Van der Gon's model eqs. (1), we define the approximation error $\epsilon_A(R_x)$ as follows,

$$\epsilon_A(R_x) = 1/T \int_T \{F_x{}^A(t, R_x) - M_x(t)\}^2 \, dt \qquad \text{(I–1)}$$

where,

$$F_x{}^A(t, R_x) = \ddot{x} + R_x \dot{x}$$

and $M_x(t)$ is derived from the corresponding EMG's. By minimizing ϵ_A with respect to R_x, we can estimate the optimum value R_x^* which is considered to be the intended parameter value. Thus $R_x^* = 18\ [T^{-1}]$, $\epsilon_A^* = 2.5 \times 10^{-2}$ were obtained for the data shown in Figure 11.

Figure 11. Simple repetitive right-left wrist motion.

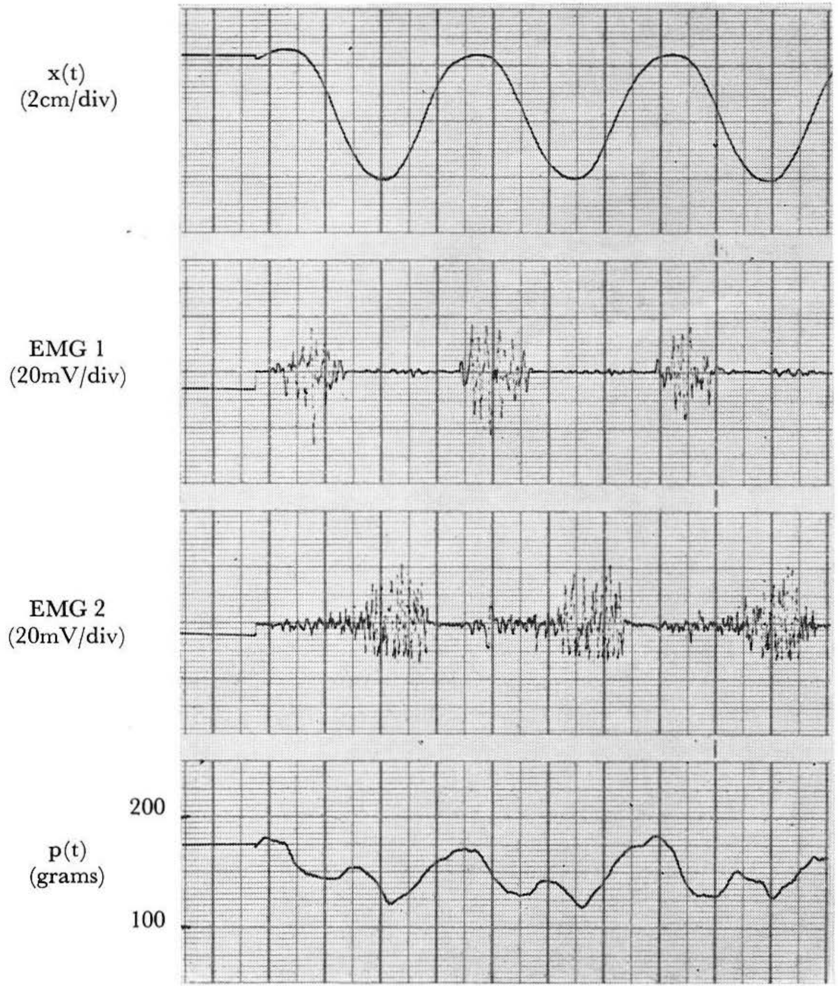
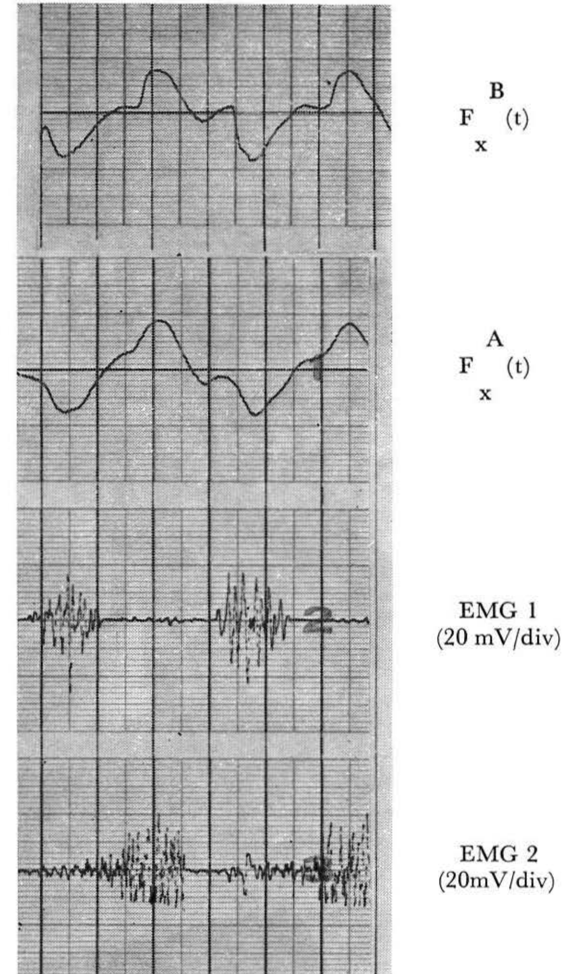


Figure 12. Estimated force waveforms for Case A and for Case B, together with the corresponding EMG signals.



Case B. In order to evaluate the effects on handwriting movements of a pencil-point pressure, consider the following model,

$$\left.\begin{aligned} \ddot{x} + m_x p(t)/v\, \dot{x} &= F_x(t) \\ \ddot{y} + m_y p(t)/v\, \dot{y} &= F_y(t) \\ v^2 &= \dot{x}^2 + \dot{y}^2 \end{aligned}\right\} \qquad \text{(I–2)}$$

which are derived from eqs. (2) by neglecting both the stiffness and viscosity terms of the muscle. In the case of simple left-right motions, eqs. (I–2) become

$$\ddot{x} \pm m_x p(t) = F_x{}^B(t, m_x) \qquad \dot{x} \gtrless o$$

since $\quad v = |\dot{x}|$.

By minimizing the expression

$$\epsilon_B(m_x) = 1/T \int_T \{F_x{}^B(t, m_x) - M_x(t)\}^2\, dt \qquad \text{(I–3)}$$

$m_x^* = 2 \times 10^{-3}\, [M^{-1}]$, $\epsilon_B^* = 1.3 \times 10^{-2}$ were obtained for the same data as in Case A.

Figure 12 shows the finally estimated waveforms of muscular force $F_x{}^A(t)$ and $F_x{}^B(t)$ together with the corresponding EMG signals. It may be recognized, as a whole, that although the estimated muscular force waveforms have rather good correspondence with EMG patterns for both cases, the fine structures differ between the two. The difference between ϵ_A^* and ϵ_B^* for Case A and for Case B tells fluently the fact that the effects on handwriting movements of the writing pressure are not so small to be ignored. This will be discussed again in the synthetic studies.

II. *Synthetic Studies of Handwriting*

It is easy to say that handwriting signals are continuous time functions themselves. But, handwriting signals are composed of a sequence of strokes that are digital time segments. The stroke is defined as the time eliminated locus of a pencil-point movement caused by a muscle contraction.

Figure 13 shows the muscular force waveforms calculated from,

$$F_x(t) = \ddot{x} + m_x p(t)/v\, \dot{x}$$

where $m_x = 2 \times 10^{-3}$ for subject *Y* writing the Japanese character "か", together with the corresponding EMG's. It might safely be assumed from the inspections of these data that the muscular contraction causes the force with an exponential increase and decrease in amplitude. The introduction of this assumption makes it possible to determine uniquely

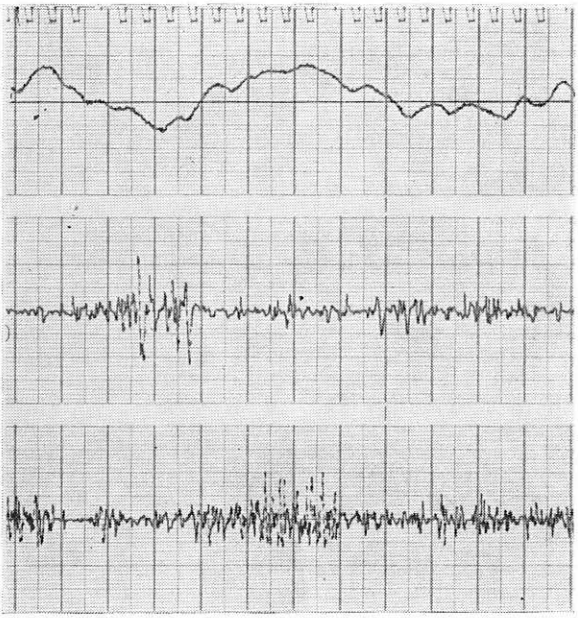


Figure 13. Muscular force waveforms calculated for writing the Japanese character "mi," together with the corresponding EMG's.

the force function by only giving the contraction timings τ_i's of muscle, at which handwriting signals are divided into a sequence of stroke segments.

Subjects are asked to write a certain Japanese letter "hiragana" without any constraints, and the samples of $(x\text{–}y)$ position and pressure of a pencil point during handwriting are transferred to the computer.

First of all, from the data obtained by above processes, a rough guess of timings

$$\boldsymbol{\tau}_x^0\ (\tau_{x0}^0,\ \tau_{x1}^0,\ \tau_{x2}^0,\ \ldots.),\ \boldsymbol{\tau}_y^0\ (\tau_{y0}^0,\ \tau_{y1}^0,\ \tau_{y2}^0,\ \ldots.),$$

which are considered to correspond to those of muscle contractions, is made by extracting zero crossing timings of $F_x(t), F_y(t)$ calculated from,

$$\left.\begin{aligned} F_x(t) &= \ddot{x} + m_x p(t)/_v\,\dot{x} \\ F_y(t) &= \ddot{y} + m_y p(t)/_v\,\dot{y} \end{aligned}\right\} \qquad \text{(II–1)}$$

which are the copies of eqs. (I–1) and the parameters m_x, m_y are assumed pre-known for each subject.

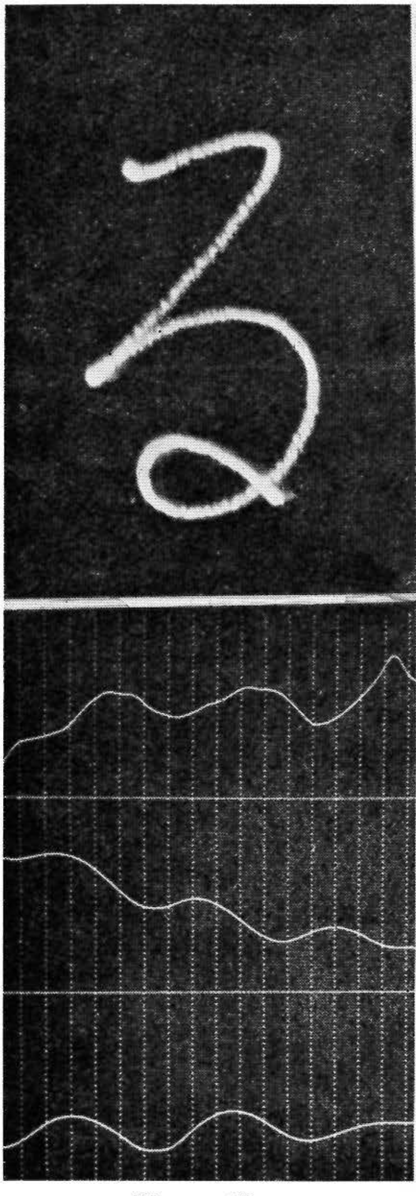
A.

B.

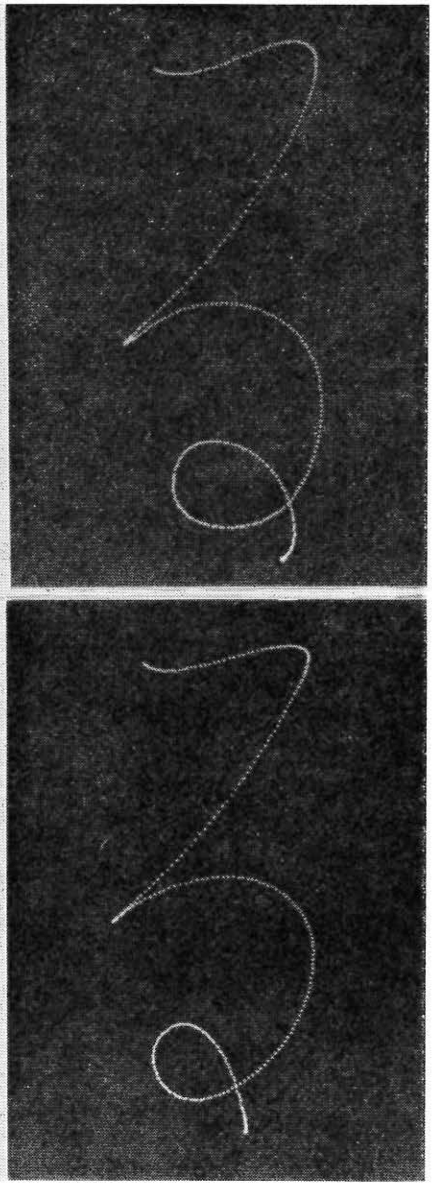
$p_o(t)$

$y_o(t)$

$x_o(t)$



80 m sec./div



A.

B.

Figure 14. A. Original handwriting sample that the synthesized output is supposed to match. B. Estimated force waveforms for the original sample.

Figure 15. Synthesized patterns: The writing speed and pressure are assumed constant in Case A, but not in Case B. The normalized approximation error $\mathcal{J}^*$'s $(\mathcal{J}_x^* + \mathcal{J}_y^*)$ were $\mathcal{J}_A^* = 1.21 \times 10^{-3}$ and $\mathcal{J}_B^* = 0.72 \times 10^{-3}$ respectively.

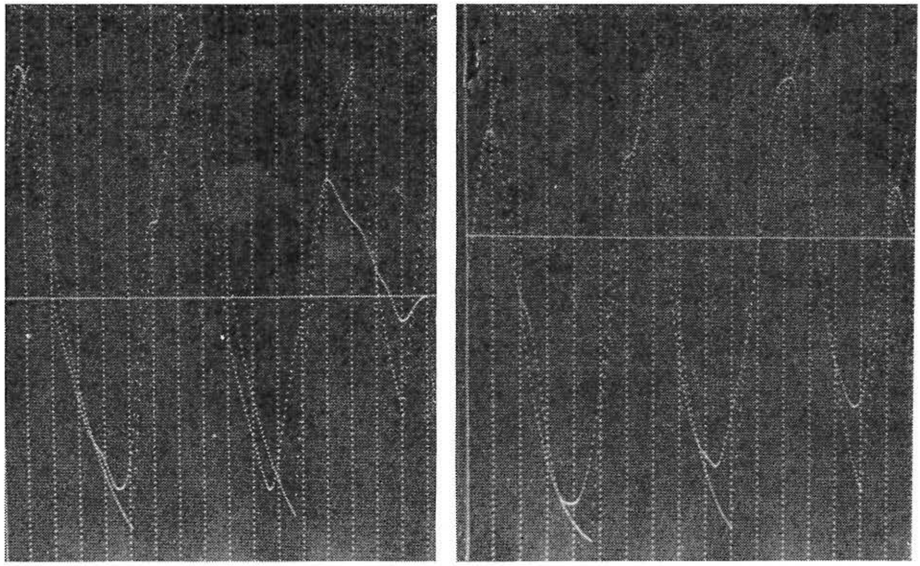
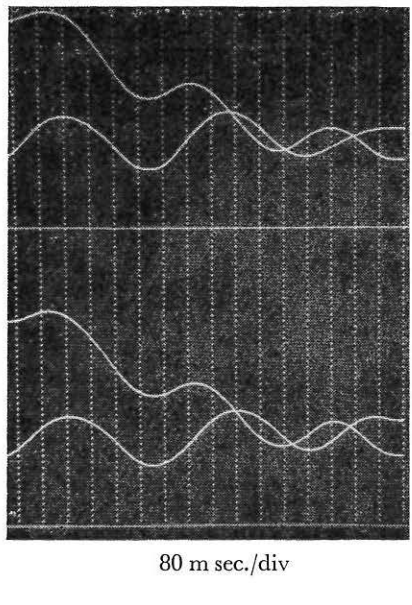


Figure 16. (A) Finally extracted force waveforms (cf. Figure 14B, estimated force waveforms for the original sample). (B) Original and synthesized *x–y* displacement waveforms, both of which correspond to the Case B in Figure 15.



Now, let $x_s(t), y_s(t)$ be the synthesized position functions by the proposed model of handwriting. The integral functional J_d ($d = x, y$) is introduced which evaluates the distance between the actual handwriting and the synthesized handwriting as follows;

$$J_d(\tau_d) = 1/T \int_T [\{d(t, \tau_d) - d_s(t)\}^2 + \{\dot{d}(t, \tau_d) - \dot{d}_s(t)\}^2] dt \qquad \text{(II–2)}$$

Our concern here is to find the optimum timing τ_d^* which minimizes the functional J_d. Iterations of A-b-S algorithms can be executed by setting previously guessed τ_d^0 as a initial value of τ_d. Thus, the optimally estimated value τ_d^* will be obtained and its corresponding force function $F_d^*(t)$ will also be calculated.

Figure 14A is the original handwriting that the synthesized output is supposed to match. Figure 15 shows the synthesized patterns in which the writing speed and pressure are assumed constant in Case A, but not in Case B. The synthesis in Case A is quite successful except for the tail of "ʒ" where the remarkably high writing pressure is observed (Fig. 14B). Further it was found that the last stroke on the tail of "ʒ" could not be made to have the same shape as the sample without changing other parts so that they failed to agree with the sample. On the other hand, the example shown in Figure 15B, for which the independent variation of the pressure are provided, is found to be matched quite well.

These reveal that the constant pressure model is not completely compatible with an arbitrary sample of cursive script; that is, the writing pressure plays an important role in the actual human handwriting.

Acknowledgements. The author wishes to record his deep gratitude for valuable suggestions to Dr. Hiroshi Sato, professor of the University of Electro-communications, and to Dr. Osamu Fujimura, professor of the University of Tokyo.

REFERENCES

Denier van der Gon, J. J., and Thuring, J. Ph. A handwriting simulator. *Phys. Med. Biol.*, 1962, *6*, 407–414.

Eden, M., and Halle, M. Handwriting and pattern recognition. *Trans. on IT.*, 1962, *8*, 160–166.

Lashley, K. S. The accuracy of movements in the absence of excitation from the moving organ. *Amer. J. Physiol.*, 1917, *43*, 169–194.

Macdonald, J. S. Experimental studies of handwriting signals. *Tech. Rep. MIT.*, 1966, *443*.

Stevens, K. N. Toward a model for speech recognition. *J. Accoust. Soc. Amer.*, 1960, *32*, 47–51.

Yasuhara, M. Handwriting analyzer and analyses of human handwriting movements. *Jap. Psychol. Res.*, 1969, *11*, 103–109.

Yasuhara, M. A synthetic model of handwriting letters. *J. Inf. Processing Soc. Jap.*, 1970, *11*, 584–591.