A New Technique in Automatic Character Recognition

By M. B. Clowes and J. R. Parks

This paper describes a new method for identifying a character, which overcomes certain fundamental problems arising with existing techniques. The method involves the evaluation of an autocorrelation function of the character to be recognized. This function contains maxima corresponding to straight lines in the character, and is essentially unchanged by changes in the position and orientation of the character. Variations in the size and aspect ratio of the character also leave the function sensibly constant.

Introduction

This paper describes some of the experiments and results obtained during the last year from research on a fundamentally new approach to Character Recognition. The long-term aim is to produce a reading machine capable of recognizing hand-printed numerals. At the same time the Russian-to-English Machine Translation project at the NPL creates a specific requirement for a reader of texts printed in the Cyrillic alphabet. Both applications will require complex devices which would probably be uneconomic if we were to confine ourselves to existing techniques. We expect that the attainment of these objectives will proceed through intermediate devices, applicable to progressively larger alphabets. It is important, however, that any simplifications introduced in intermediate devices should be such as to retain the possibility of subsequent extension. Accordingly, the problem to which we are now addressing our efforts is that of reading numerals printed in a single style (or small range of styles). Printing control will be sufficient to maintain adequate *separation* of the characters on the paper. The precise position and orientation of the character, in the zones thus defined, is assumed to be subject to random variations. We expect also that the quality of the printing will be such as to produce low and variable contrast over the character, together with voids, ink splash and random paper noise. These are the conditions which we believe to obtain principally on tally rolls printed by cash registers, but, in general, on any cheap hand-operated printing device in regular commercial use.

Existing Methods

The most general, and in principle the most flexible, reading system starts from a break-down of the character into elemental areas whose reflectance is measured by a photocell. This analysis may be achieved directly by projecting an image of the pattern on to a square array of photomultipliers (Taylor, 1959) or by applying a raster scan* (see Wada, Takahashi, Iijima, Okumura, and Jonoto, 1960; and Grimsdale, Sumner, Tunis, and Kilburn, 1959) to the pattern.

* See the paper by Merry and Norrie on p. 137 of this issue.

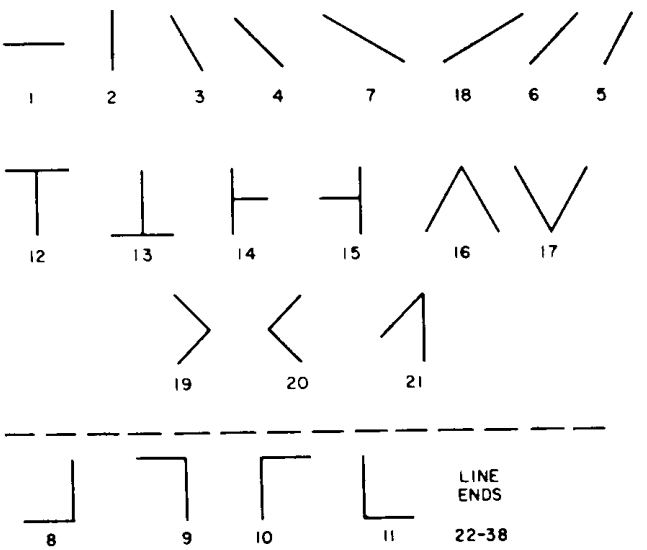

Fig. 1.—Template features used by Bomba (1959) to recognize alpha-numeric characters

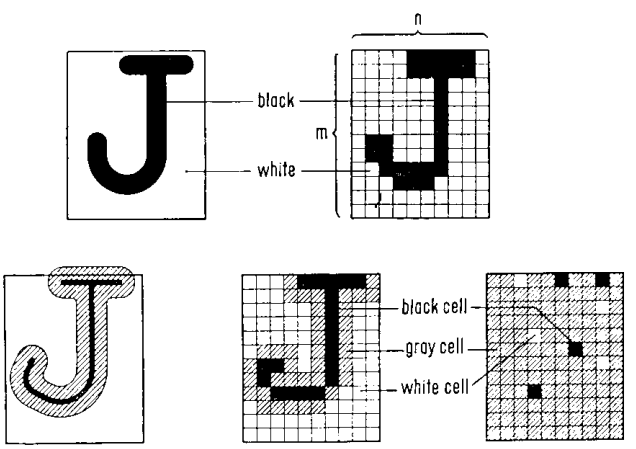

Fig. 2.—Spatial analysis employed by Wada *et al.* (1960). Black and white elements in lower right picture specify this character

The recognition of the unknown may then proceed in either of two different ways. We can *sum* the intensities recorded at selected points in this scan, these points being arranged to correspond to features of various kinds, as shown in Fig. 1. This approach essentially forms templates which are matched with selected areas of the pattern (Bomba, 1959). The alternative method of recognition involves the examination of a small number of key elements to see whether or not they agree with one of a set of logical statements defining the characters which can be recognized: see Fig. 2. In some cases a large number of elements—again organized into features—are examined, as shown in Fig. 3. While this technique is capable of much higher levels of discrimination than is the template system, it suffers the disadvantage that the input signals have to be regarded as binary if the amount of computing involved in recognition is not to become prohibitive. This quantization of intensity is a disadvantage, because in deciding whether a point in the pattern is black or white, it is easy to introduce artificial voids and splash which lead to a distortion of the actual pattern. For example, part of an unevenly-inked character could be clipped off.

There are many variations on these methods, and more sophisticated approaches have been reported in the U.S.A. All of these methods, however, have a basic inability to recognize displaced or mis-orientated characters. Generally speaking, they need to regard each different position of a character as a different character. This intolerance of positional changes, and, in the high-discrimination systems, the setting of arbitrary intensity levels, are the principal factors which militate against the successful application of these methods to the general problems of character recognition.

Pattern Recognition by Auto-Correlation

The method which we are studying offers in principle a solution to most of these difficulties. It starts from an assumption common to many approaches (especially Grimsdale *et al.*, 1959) that a character may be described in terms of the *relative* location of a small number of *features*. For example, a "2" is: "a downward-facing hook joined at its right-hand edge by a relatively-straight diagonal, which touches the left-hand end of a horizontal feature at the bottom of the pattern." In an early phase of our work, we examined the use of stylized features of this kind in a simple template system of the kind described earlier. The results of this small investigation were sufficiently encouraging for us to proceed with this feature description. However, the use of such templates is completely inadequate so far as positional variations are concerned, and we, therefore, introduced the notion of self-matching or *autocorrelation*. If we consider an idealized numeral (see Fig. 4) then we can make the straight lines in that character *self-detecting* by comparison with a displaced copy of the original character. If we imagine the character and its copy to be transparent and placed one above the other, then there will be certain

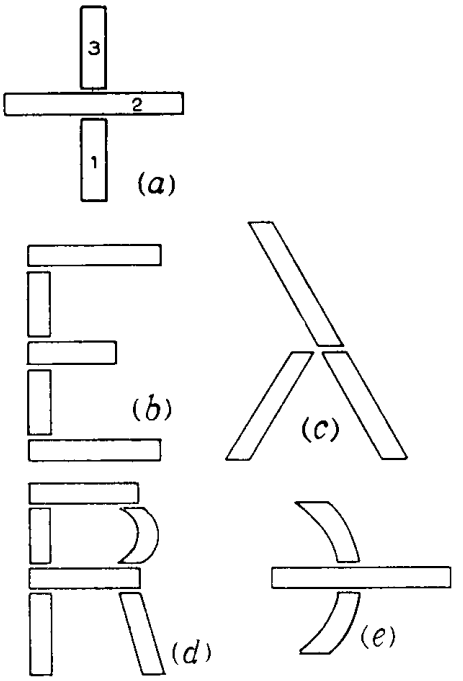

Fig. 3.—Features employed by Grimsdale *et al.* (1959)

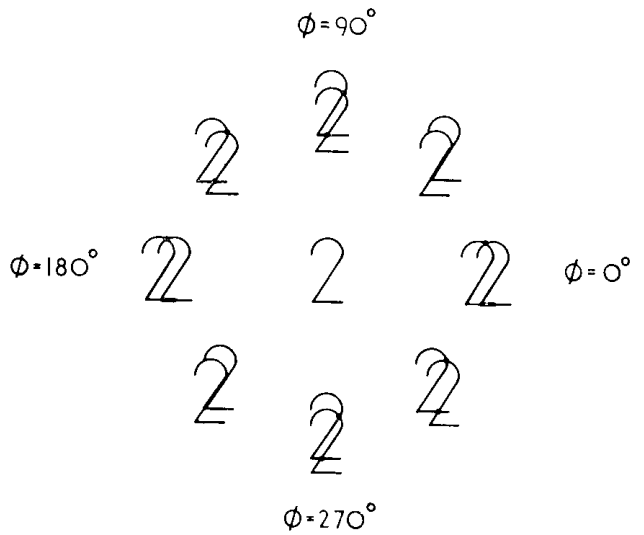

Fig. 4.—Auto-correlograms formed by comparison of input character (centre) with an identical copy displaced in different directions $\phi = 0°$ to $\phi = 180°$

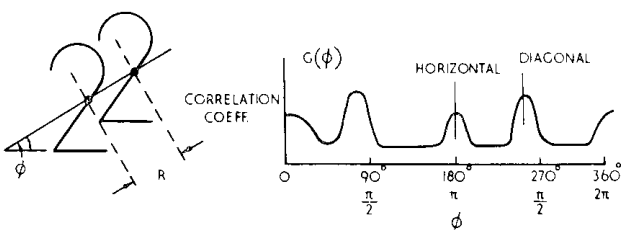

Fig. 5.—Autocorrelation coefficient (schematic) of an ideal "2" as a function of direction of displacement (ϕ), displacement magnitude (R) being constant

positions of the copy where a lot of light will be transmitted. This occurs when corresponding straight lines in the two copies lie over one another. In practice, the copy is displaced through some preset distance in a direction which is slowly varied. Fig. 5 illustrates the resultant variation of illumination. The maxima in this function indicate the presence of straight lines, and the separation of these maxima their relative orientation.

There are two important aspects of this function: changes in the position of the original character (which are necessarily accompanied by corresponding changes in the position of its copy) do not affect the function at all. Changes in the orientation of the original displace the function along the x-axis without altering its shape. In addition, simple variations in the size or the aspect ratio of the character will leave the function essentially unchanged. The technique also has the advantage that it does not involve a preliminary quantization of intensity levels in the character. The analysis is essentially analogue.

Fig. 6 shows the optical system used to generate these auto-correlation functions experimentally. The character, in the form of a transparent negative, is illuminated by a parallel beam of light. The "light image" of the character is reflected by M_2 back through the character. The light which emerges, therefore, represents the comparison of two copies of the character. The displacement of the copy relative to the original is controlled by the orientation and position of M_2. By rotating the character, or M_2, we can scan through the different directions of displacement successively. The resulting variation of intensity is recorded by the photomultiplier P_2, and is plotted in Fig. 7 for a standard set of numerals (dotted line). It will be clear that for some styles, these waveforms alone will suffice to identify the character. However, confusions are likely to arise, especially with 6 and 9 and perhaps 2 and 7.

An improvement in the resolution of these waveforms results from the introduction of a further copy into the comparison process. This second copy is displaced in the same direction as the first, but by a smaller amount. Correlation now occurs only between the same straight line in the three patterns, thus eliminating the "cross-talk" between different adjacent features of the pattern (see Fig. 8). This scan procedure may be carried out in the optical system described earlier (Fig. 6) by measuring the result of two successive reflections at M_2 and M_1. P_3 records the corresponding waveform. In Fig. 7, autocorrelation with one and with two copies is compared.

The waveforms generated in this process are not sufficient in general to discriminate between all characters. Ambiguities arise because no positional information is present in these waveforms, only information about orientation. One attempt to restore some positional information involved the division of the character space into zones, corresponding roughly to Top, Middle, Bottom; Left, Middle, Right, etc. The details of this division are illustrated in Fig. 9. The three zones are rotated in phase with the scan to give three waveforms.

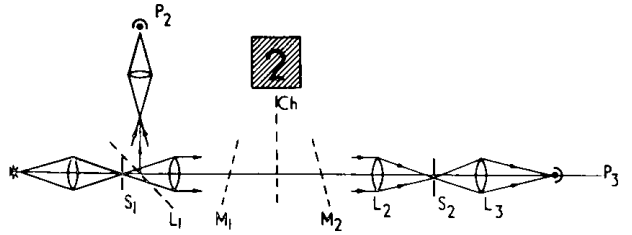

Fig. 6.—Optical system used to generate autocorrelation functions. Transparent negative of character (Ch) is transilluminated by collimated light. M_1 and M_2 are half-silvered mirrors whose separation and position control spacing of character and its copies. P_2 and P_3 measure autocorrelation for one and two copies respectively

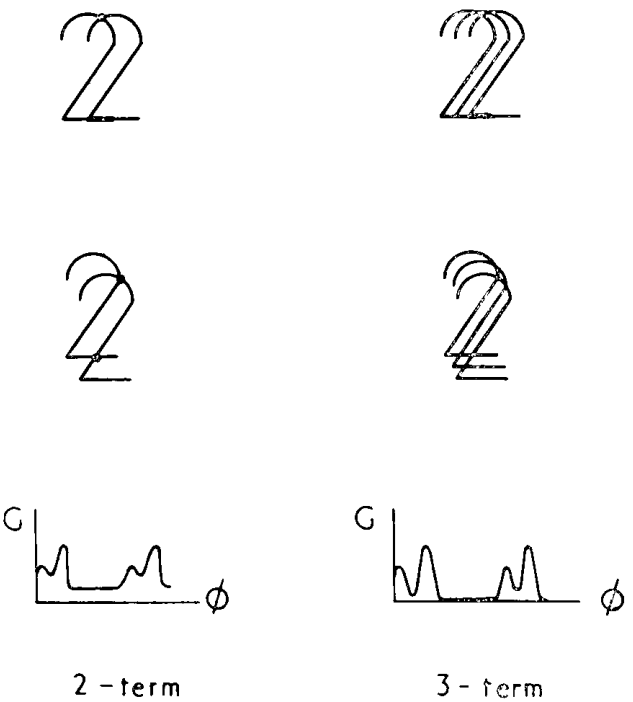

Fig. 8.—One-copy (2-term) autocorrelation contains spurious correlations (circular dots on figures in left-hand column), which mask the peaks due to straight lines. Two-copy (3-term) autocorrelation eliminates these correlations (right-hand column), and improves resolution of function

The waveforms generated from the outer zones are 180° out of phase, and one of these channels may therefore be discarded. The results of applying this scan, with two copies only, to the standard set of numerals is shown in Fig. 10. This zoning suffers from the fact that these waveforms now depend upon the position and aspect ratio of the character. This arises because we are measuring the *absolute position* of the features, whereas we really only need their positions *relative to one another*. We are currently developing techniques to obtain this "relative" information by the use of three or more copies in the original scanning process. This should restore the original invariance of the waveforms to changes in character position.

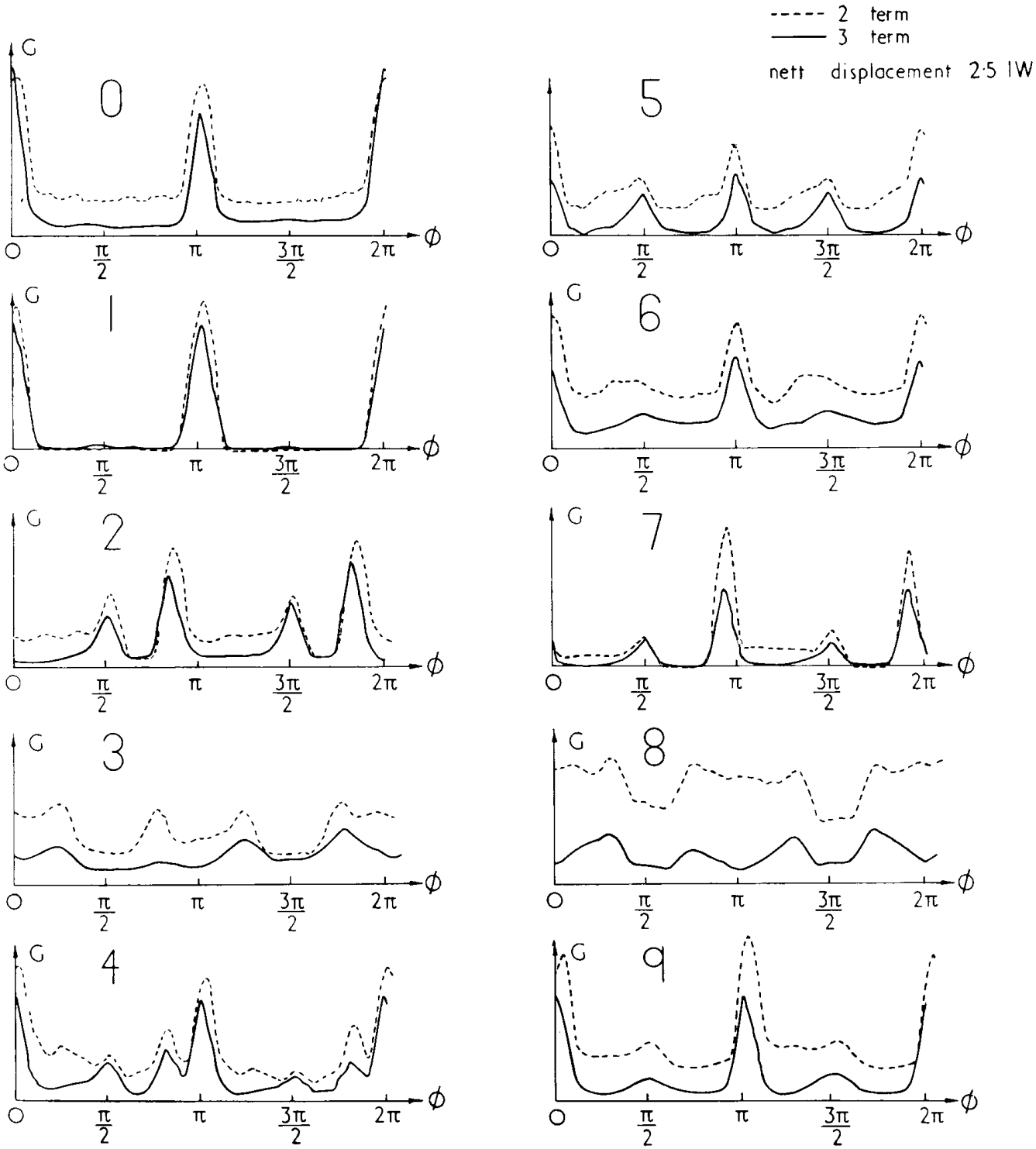

Fig. 7.—Auto-correlation functions obtained for standard numerals illustrated. Dotted line indicates function obtained with only one copy, full line generated by use of two copies

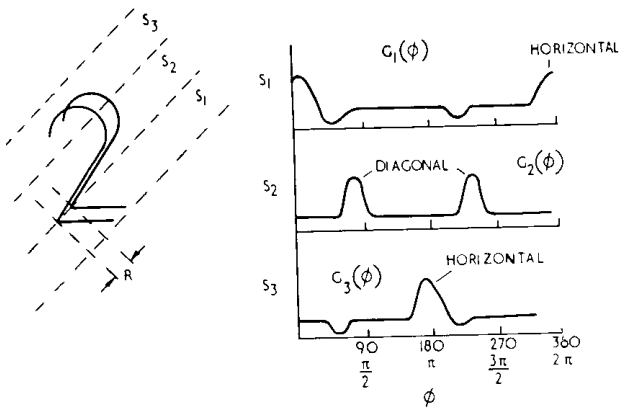

Fig. 9.—Details of sampling zones used to restore information about relative position of features within the character. Zones S_1, S_2, S_3 lie parallel to the instantaneous direction of ϕ. Functions generated in S_1 and S_3 are 180° out of phase, and S_3 may be discarded

Automatic Generation of Auto-Correlation Functions

We will now discuss methods of performing the scanning operations automatically. The optical system described above is essentially a research tool and is being used to explore the logical problems involved in the technique. We are at present using high-contrast symbols which are not typical of the characters one normally expects. Symbols of this quality are being used to demonstrate positional and orientation properties of the technique. Subsequently more realistic characters will be introduced.

As a first attempt to evaluate the auto-correlation function electronically, a TV link was used to generate a copy of the unknown character. The TV camera (see Fig. 11) views the character through a red filter. An image of the resultant CRT display is reflected back on to the original through a green filter. A photomultiplier, "viewing" the paper through another green filter, "sees" the copy superimposed upon the original. If the copy is now displaced in the usual way, the photomultiplier output generates the autocorrelation function. In practice, the output is highly irregular, as a result of random fluctuations in the CRT brightness (see Fig. 12). If we use the variations in CRT brightness to scale the photomultiplier output, these fluctuations can be eliminated and the resultant function averaged over a number of cycles, as shown in Fig. 13. The copy used was a negative of the original, so that the autocorrelation function is inverted. While this result demonstrates a possible scan process, it is difficult to extend it to a number of copies, and since a complete cycle of the scan requires some 20–30 TV frames, it is necessarily slow. The solution of these difficulties requires a radically different approach.

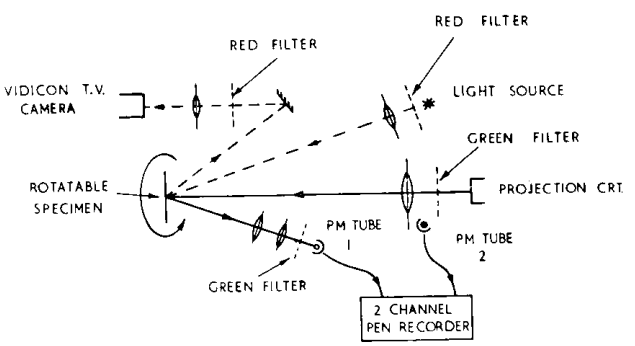

Fig. 11.—Automatic generation of autocorrelation functions using a TV link. Copy of "specimen" character is generated on CRT via the Vidicon camera. Photomultiplier tube 1 views this copy imaged on to the original. Use of red and green filters prevents saturation of PM tube by ambient light

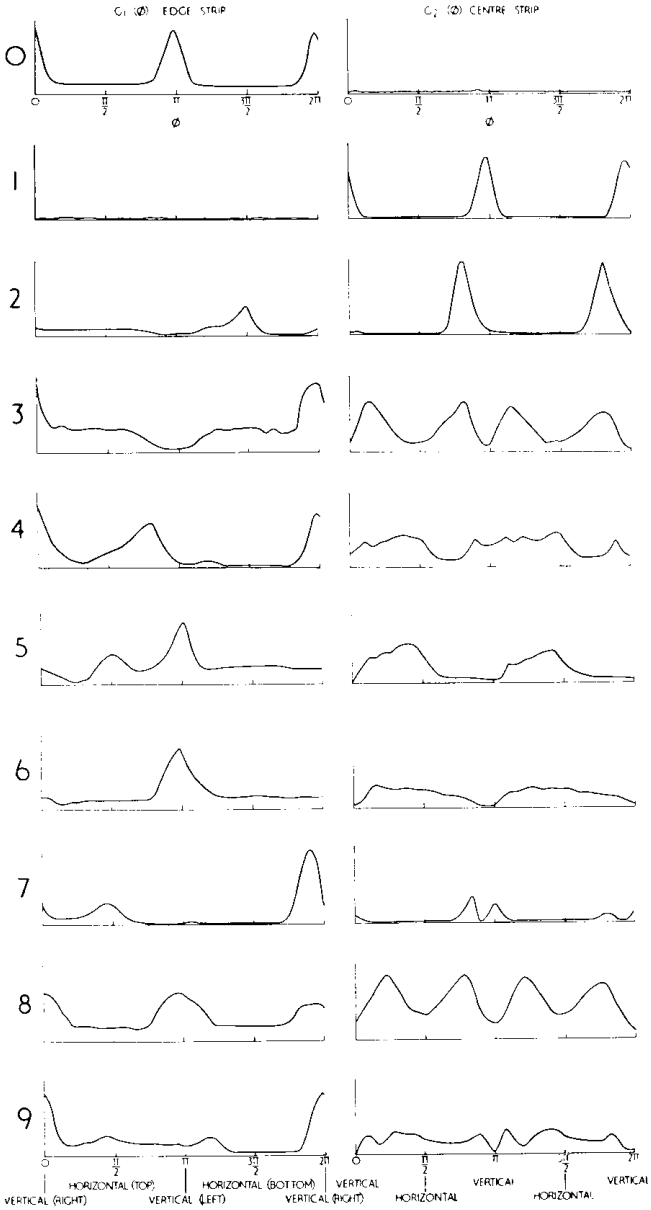

Fig. 10.—"Zoned" autocorrelation functions for standard numerals

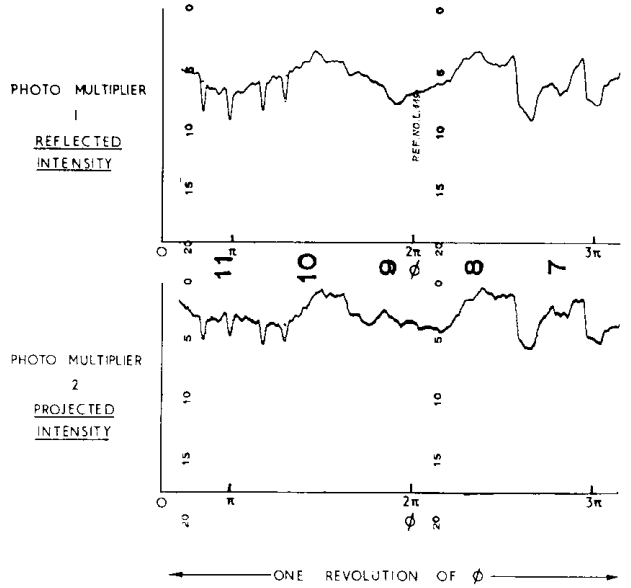

Fig. 12.—Waveforms recorded by PM1 and PM2

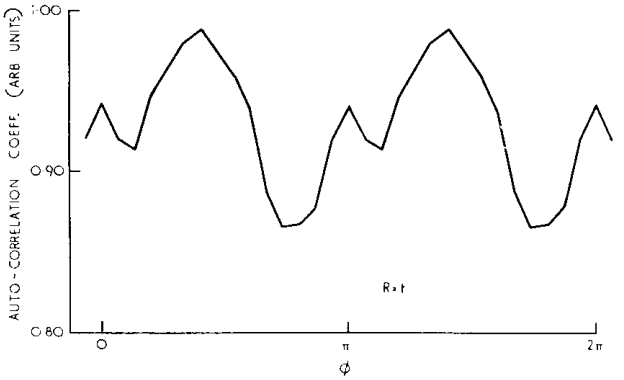

Fig. 13.—Autocorrelation function recorded by PM1 after scaling (by PM2) and averaging over several cycles. Double peaks arise from horizontal and diagonal in the specimen "2"

It seems likely that this will involve the application of image-intensifier systems which effectively handle pictures " in parallel."

Conclusion

To sum up, we believe that we have a recognition process which offers the possibility of overcoming the basic problems confronting existing devices. These problems concern the positioning, orientation, size, and print quality of numerals. The technique also offers the prospect of extension to the alphabet, and may prove tolerant of style variations. Its practicability physically —and ultimately economically—is the problem we are now investigating.

Acknowledgements

The authors wish to thank Mr. M. Wright and Mr. R. Riggs for technical assistance; and Dr. A. M. Uttley for many valuable discussions. The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

References

BOMBA, J. S. (1959). "Alpha-numeric Character Recognition using Local Operations," *Proc. Eastern Joint Computer Conf.*, p. 218.

GRIMSDALE, R. L., SUMNER, F. H., TUNIS, C. J., and KILBURN, T. (1959). "A System for the Automatic Recognition of Patterns," *Proc. I.E.E.*, Vol. 106, Pt. B, No. 26, p. 210.

TAYLOR, W. K. (1959). "Pattern Recognition by means of Automatic Analogue Apparatus," *Proc. I.E.E.*, Vol. 106, Pt. B, No. 26, p. 198.

WADA, H, TAKAHASHI, S., IIJIMA, T., OKUMURA, Y., and JONOTO, K. (1960). "An Electronic Reading Machine," *Proc. UNESCO Conf. on Information Processing*, Oldenbourg Verlag, Munich; Butterworth, London

Summary of Discussion

The Chairman, introducing the speaker, said Dr. Clowes had done his research work for his Ph.D. at Reading on studies of visual processes and eye movements. He had joined the National Physical Laboratory about two years previously and was leading a small group on investigations into character-recognition methods.

Mr. J. R. Cartwright (*International Computers and Tabulators Limited*) said that Dr. Clowes seemed to suggest that one started by saying one could specify the characters one's self and then accepted the case where characters were specified by someone else. Then one went on to handwriting and texts. All that seemed to suggest that Dr. Clowes thought that it was a good thing to aim at a final system which would deal with almost anything that would take a dirty scrap of handwritten paper and translate it. On the other hand, using the standardized codes of existing data-handling systems—even simple things like punched tape—one finished up with a document that could be transformed readily from one code or system to another.

The lack of standardization had happened with automatic programming. A year or two earlier everyone had been running wild and had had an automatic programming system for their own machine which was quite incompatible with anyone else's. Now there was ALGOL which was a real attempt to produce a programming system which could be applied not only to all existing facilities but to future com-

puters which might be designed. He had no doubt the same kind of approach should apply to character-recognition generally.

In other words, as far as possible, one should produce a very efficient machine with a very low error rate on relatively few types of character. Obviously, one would have to deal with a limited number of different character sizes and faces, as between, say, Russian script and Arabic numerals, and so on. That should be the aim, rather than seeking to produce a machine capable of dealing with everything, including badly handwritten characters.

Dr. Clowes said he did not believe any rightminded professional worker in the character recognition field would ever try to read dirty pieces of paper if clear pieces were available. On the general point, the format of the document determines how the paper must be moved about, how it should be placed under the scanner, and can seriously affect any results obtained. It was a very relevant problem.

When he had started work at the National Physical Laboratory he had been told that the great problem lay in the categories proposed, i.e. numerals only. More and more people were going to say, however, that while the numerals were all right, there were only 10 of them. Indeed, that was already being said in connection with certain existing coding systems; and therefore one faced the fact that there was the need for a larger alphabet. If this implied a more complicated format, they must learn to live with that requirement.

Mr. Newman said that he believed that what had just been said showed a fallacious way of looking at things. He had attempted to put the point earlier. Character recognition was not a thing to do for its own sake, being in itself wasted work. It was forced on one by absence of a better method of procedure, and was best avoided if possible.

Where one's control of the media and system is so great that the use of stylized characters is possible, in general the need for character recognition is at most temporary. But where a source of input data is not under control, one might well get hand-written characters on dirty pieces of paper. This sort of thing therefore is the important data for automatic systems to read.

Mr. Cartwright said he certainly agreed that the approach advocated by Mr. Newman was the one for the future. The question of the necessity for reading handwritten characters involved a rather theoretical approach, from the commercial point of view. There had been a lot of publicity for a system which was to do character recognition on cash-register tally rolls. One could not argue that that was essential.

There was no great technical difficulty in producing a coded cash-register tally roll, in the form of a punched card or tape, or a strip of magnetic tape, which could be used directly with computers and calculators. Hence, a person wishing to install such a system would need to have a good business reason for wanting character recognition rather than some existing form of recording data. It would appear unlikely, for example, that a publisher of some particular journal would ever provide a punched-tape enabling the translation of that journal into some other language.

There was a great deal of interest in character recognition for economic reasons. Bank character-reading projects were very much under way. They were an example of a commercial application. Theoretically, a punched card could be produced for every cheque drawn; and yet character-reading systems were reported to be going into American banks on a large scale.

Mr. W. E. Norman (*IBM U.K. Limited*) said the American banks seemed to have the best of all three worlds. If one got a cheque from an American bank one might find in it one of three types of character. At the side there might be punched holes, because a great many of the American banks, especially those with nation-wide branches, punched the account number and the amount into the cheque making the cheque a punched card which could be processed by a punched-card or computer system.

The second kind of character would be in magnetic ink. His own company had a magnetic-character reader; and there was a sorter which could sort cheques magnetically at high speeds. Finally, some banks were working on an optical system.

Dr. Clowes said he would regard those characters as in the coded category; that certainly applied to magnetic ink and punched holes. The question was whether it cost more to convert all existing printing devices to a new standard and new style (bearing in mind that in printing with magnetic ink the tolerances in printing operations were a good deal stricter than what was involved in, say, a simple cash register), than to buy a package device which would read characters and would have a certain amount of flexibility in respect of the kind of printing tolerances with which one had to live.

The difficulty was that all the coded things were leading to more expensive processing, and there was a "break-even" point at which it would cost more to produce that kind of material than it would to get a character recognition machine. That "break-even" point might come for characters in category 2.

Mr. Nadler (*Cie des Machines Bull*) said that although Dr. Clowes had invited hard criticism, at this stage people could really only put forward their personal philosophies on the subject, which naturally differed. He was sure that in any case research on the auto-correlation techniques would continue and would prove fruitful. He had been particularly interested by the suggestion of using an image-converter.

He was struck by the fact that in the system one was transforming one analogue quantity into another analogue quantity with the dimensions amplitude and time (or angle, which is equivalent to time in the proposed system). It is certainly much easier to analyse analogue quantities, which are along a linear axis, for it is necessary only to detect the peaks and their relative amplitudes and the times at which they occur in the various channels; one is thereby using a great deal of the information available with the auto-correlation technique described, which seems more fruitful than other optical correlation systems one has seen, which produce new *optical* images, which appear to the eye to be pretty much alike.

Would it not be more direct, however, to avoid going through a second analogue stage but to use as much as possible of the two-dimensional information obtained in the scanning process to generate a logical (discrete) function having some of the properties of the auto-correlation?

On the question of learning-machines, asked by the representative of IBM, we could say that Nature had taken a long time over developing the equipment which enabled human beings to read any language. When they had gone sufficiently far in their experiments at N.P.L. and IBM to be able to build a machine which, by virtue of its structure, would be able fairly simply to read one set of characters, they might then be in a position to consider the attachment of something for learning to read other characters.

He did not think that the complicated network systems

("neural networks"), which had been proposed, could be made to fit in with the requirements. Such a machine has been "taught" to distinguish between differently oriented straight lines, or between a few simple geometric forms, a few letters of the alphabet in a single fount or similar founts, but to construct such a machine to recognize up to 128 characters would require something almost the size of the hall in which they were sitting.

At Bull, the problem of learning-reading machines was considered in a rather different manner. There would probably be a special reader unit, which would carry out the optical analysis and the decoding into computer language as an attachment to a general-purpose computer. This attachment would not itself be capable of learning, but would have variable logic under the control of the computer it was feeding, so that the latter could contain the learning program proper (once we had learned how to code such a program). One could imagine the mode of operation would be to introduce a few pages of the text, containing the alphabet to be learned, into the computer simultaneously from keyboard and optically through the optical attachment. Then the machine could read along on its own, stopping for instruction at each further unlearned character. A multi-program machine of the type of the Gamma–60 seemed indicated for such work.

Dr. D. A. Bell (*University of Birmingham*) said it always appeared easier to put information into the machine deliberately rather than wait for the machine to learn. There were some very ingenious devices for semi-permanent storage. Recently a store designed to work on that basis had been installed in a Manchester machine. That was easier than having an elaborate process enabling a machine to read a trial piece to produce the information from statistical analysis.

Although it was convenient to enter material in binary form from character-reading, each time one made a binary decision one was throwing away information. One should try to preserve the pattern in analogue form until the latest possible stage before making any kind of decision.

The Chairman said the morning's discussions had been interesting and fruitful. Everyone present would wish to express their appreciation to Dr. Clowes and the other speakers for addresses which had given rise to such interesting questions and comments.

Correspondence

To the Editor,
The Computer Journal.

Sir,

"Predicting Distributions of Staff," by Andrew Young and Gwen Almond.

The Model given in the paper by Young and Almond (Vol. 3, p. 246) has been given previously by Prais (4, 5). The results on the latent roots of non-negative matrices have been given by Wielandt (7) and Debreu and Herstein (1), and are well known in the theories of Markov chains (2) and Leontief models (3). The setting up of such as difference equations and the properties of the solutions are well known to mathematical economists (6). These models have often been used for short-term planning or prediction.

As far as the Leontief models are concerned the closed model is a prediction model and the open model is primarily a planning model. The model discussed by Young and Almond is purely a prediction model. If the probabilities are made to reflect alternative policies of promotion then the model can be converted to a planning model in the same way as the Leontief models generalize to linear programming. This might save the authors from feeling the need to infer periodic components of 53 and 79 years—an inference that would have been considered rash even by the most audacious writers on closed Leontief models.

Yours faithfully,
Z. Herzenstein and F. E. A. Briggs.

International Computers and Tabulators Limited.
London, S.W.6.

21 February 1961

References

1. Debreu, G., and Herstein, I. N. (1953). "Non-negative Square Matrices," *Econometrica.*
2. Feller, W. (1950). *An Introduction to Probability Theory and its Application*, Chapter 15. Wiley.
3. Leontief, W. W. (1951). *The Structure of American Economy.* New York: Oxford U.P.
4. Prais, S. J. (1955). "The Formal Theory of Social Mobility," *Population Studies.*
5. Prais, S. J. (1955). "Measuring Social Mobility," *Journ. Roy. Stat. Soc.*, Series A.
6. Samuelson, P. A. (1947). *Foundations of Economic Analysis.* Harvard U.P.
7. Wielandt, H. (1950). "Unzerlegbare, nichtnegative Matrizen," *Math. Zeitschrift.*

The authors' reply

Our problem arose because the particular institution could see trouble ahead in a few years' time when many of the staff simultaneously reach the tops of their respective grades. We were not asked to gaze into a crystal ball for several decades ahead and, we think, it is clear from the paper that we do not expect accurate long-term forecasts. We therefore find it surprising that the letter-writers seem so concerned that we "rashly" drew attention to the periodic components of 53 and 79 years. Our point was that because the periods were long, no big oscillations in the numbers of staff in various statuses need be expected in the short run with the existing recruitment and promotion patterns.

It may be of interest to report that the predictions made for the institutions for this year have now been verified and have proved extremely accurate. This very fact will encourage the institution to amend its staffing policy and we cannot expect such accuracy in a few years' time when the probabilities are based on a changing staffing policy.