XXI. COGNITIVE INFORMATION PROCESSING

Academic and Research Staff

Prof. Jonathan Allen
Prof. Barry A. Blesser
Prof. Murray Eden
Prof. Goesta H. Granlund
Prof. Francis F. Lee
Prof. William F. Schreiber
Prof. Donald E. Troxel
Prof. Ian T. Young

Dr. Brian E. Boyle
Dr. Martin J. Hewlett
Dr. Denis J. Kfoury
Dr. Tadashi Mimaki
Dr. Douglas O'Shaughnessy
Dr. Louis François Pau
Dr. Donald G. Ross

Dr. Stephanie E. Sher
Dr. Robert J. Shillman
Dr. Makoto Yasuhara
Caleb J. Drake
M. Sharon Hunnicutt
Robert A. Piankian
Kenneth P. Wacks
Allison G. Zack

Graduate Students

Robert T. Babcock
Gerald R. Barber
Robert P. Bishop
Ronald E. Boucher
Robert E. Bruccoleri
John C. Carter
Yao-Ming Chao
Charles H. Cox
Paul J. Curlander
James R. Ellis, Jr.

Irvin S. Englander
Dennis M. Freeman
Richard S. Goldhor
Samuel M. Goldwasser
Monson H. Hayes
Su Ling Hsu
Malik M. A. Khan
Theodore T. Kuklinski
Robert C. Laurenson

Donald S. Levinstone
Charles W. Lynn
Ulick O. Malone
Joseph T. Olesik
Leonard Picard
John N. Ratzel
James R. Seeley
Mark A. Shalloway
Donald A. Wood, Jr.
Gregory W. Zack

1. NATURAL LANGUAGE PROCESSING

National Science Foundation (Grant SED74-12653-A01)

Jonathan Allen

Our overall objective is to develop a comprehensive model for converting English text to speech. This requires development of a modular research system, and construction of special hardware.

The backbone of the project is the modular research system. We intend to construct a flexible configuration of algorithms, coded in a higher level language (BCPL), which uses well-defined file interfaces for module interconnection. These modules include a morphological analyzer, parts-of-speech determiner, parser, letter-to-sound converter, lexical stress analyzer, prosodic correlate determiner (pitch and timing), and phonemic synthesizer. Each is coded as an independent module with well-defined input and output data structure files. This enables independent development of the individual modules, and permits us to perform experiments using many different configurations of the total system. In this research system, emphasis is placed on flexibility of design, ease of understanding, ability to modify quickly and easily, and relative machine independence. In the coming year we shall continue to refine the system, with emphasis on prosodic algorithms, parsing, and phonemic synthesis. Although we do not intend to optimize the system for minimum memory space or execution time, some attention will be devoted to reduction of the morph lexicon.

Once the modular research system is developed, we shall seek to implement these algorithms in special-purpose hardware, which is intended to provide application-oriented performance that is efficient in space and time. Thus far, a microprocessor-based design has been constructed for computation of speech parameter values from an input phonemic segment specification, and an all-digital vocal-tract model has been built that produces a speech waveform from an input set of speech parameters. These are prototype designs that must be tested and evaluated extensively. In addition, we are designing a new special processor to perform the morphological analysis and parsing tasks, both of which are expressed in augmented transition network formalism. The eventual goal is to perform all of the tasks involved in text-to-speech conversion utilizing specially constructed processors suited to the structures involved. This should lead to practical implementations of this speech output capability, suitable for a wide range of applications.

2. PATTERN RECOGNITION OF CONVENTIONAL SYMBOL SYSTEMS

Joint Services Electronics Program (Contract DAAB07-76-C-1400)
National Science Foundation (Grant ENG74-24344)

Barry A. Blesser, Robert J. Shillman, Theodore T. Kuklinski, Makoto Yasuhara

The objective of this research is to design a computer algorithm capable of recognizing handprinted or machine-printed characters at error rates comparable with human performance. Our approach has been to investigate human cognition in an attempt to discover what features are crucial to letter identity and then to incorporate these features in an OCR algorithm.[1] Results of the effort of the past six months toward this goal are as follows.

We have devised an algorithm for low error rate recognition of handprinted U and V.[2] These characters, which are among the most difficult for both humans and computers to recognize, were investigated in detail by using various psychophysical techniques. The experiments indicated that the slope of the legs and the curvature of the base are important features in U-V discrimination. A computer algorithm incorporating these features was designed and resulted in a recognition rate of more than 94% on a standard data set of unconstrained handprinted U and V.

Studies based on an ABX paradigm show that letter discrimination peaks at inter-letter labeling boundaries.[3] This lends additional support to the distinctive-feature theory of letter recognition and provides another technique for investigating the perceptual letter space.

A mathematical model for incorporating contextual information has been devised and successfully tested. The model, based on the range-frequency theory of Parducci and Perrett,[4] should provide a basis for predicting changes attributable to graphical

context of the Physical-to-Functional Rules (PFRs) that operate in letter recognition.

References

1. B. A. Blesser, R. J. Shillman, T. T. Kuklinski, C. Cox, M. Eden, and J. Ventura, "A Theoretical Approach for Character Recognition Based on Phenomenological Attributes," Int. J. Man-Machine Stud. 6, 701-714 (1974).

2. C. Y. Suen and R. J. Shillman, "Low Error Rate Optical Character Recognition of Unconstrained Handprinted Letters Based on a Model of Human Perception" (to appear in IEEE Trans. on Systems, Man, and Cybernetics).

3. M. Yasuhara and T. Kuklinski, "An Investigation of Discrimination Enhancement at Interletter Boundaries" (in preparation for publication).

4. A. Parducci and L. Perrett, "Category Rating Scales: Effects of Spacing and Frequency of Stimulus Values," J. Exp. Psychol. 89, 427-452 (1971).

3. DIGITAL WIREPHOTO SYSTEM

Associated Press (Grant)

Donald E. Troxel

Since August 1970, we have been developing a news picture (Wirephoto) distribution system that is entirely new for the Associated Press. It is to be introduced in stages, in such a way that at least the present standard of quality and service will be maintained everywhere, with improvements spreading gradually to all locations.

The ultimate system as now envisioned will operate as follows. Pictures will be stored under computer control. An editor can then view any picture on a TV display in order to select, discard, edit, transmit, or store that image for later automatic dispatch. Editing may include cropping, enlarging, reducing, tone-scale enhancement, sharpening, combining, and addition of captions. No additional chemical photographic work will be required for any of these picture-processing operations.

Transmission over the "backbone" system linking AP bureaus and large metropolitan newspapers that have substantial computer facilities will be via high-speed digital links and will originate and terminate generally at computer-controlled digital storage devices. Transmission to subscribers will be analog or digital and at speeds and scanning standards appropriate to the existing transmission facilities. Complete control will be exercised by the New York network monitor. In the absence of manual interventions, transmission to all points among the bureaus, from point to point, and to regional networks, will be accomplished automatically.

We have implemented some of these procedures in the laboratory, using a PDP-11 computer (80k core, 38 megabit disk). The input may be a picture from the AP network, from a local analog transmitter, magnetic tape or Dectape, and is stored on a disk. Pictures may be transmitted from the disk to comparable receiving points. Pictures

stored on the disk may be viewed on a TV display utilizing a full-frame semiconductor storage system. Editing facilities already in operation include cropping, enlarging or reducing, combining several pictures into one, addition of captions, and sharpening.

The multitask software operating system permits new picture-processing operations to be integrated easily, and we plan to keep incorporating additional picture-processing routines into the system.

We are particularly interested in picture-processing operations in which the processing depends on the local content of the picture. That is, the detailed parameters of a coding or enhancement scheme vary for different local areas. In this type of processing it is of prime importance to avoid artifacts such as contours outlining these local areas. We are also accelerating our interest in color picture processing, both from the viewpoint of coding for bandwidth compression and enhancement or manipulation.

4. RECOGNITION OF PARTIAL DENATURATION MAPS (PDM) OF BACTERIAL PHAGE

National Institutes of Health (Grant 1 RO1 GM22547-01)

Ian T. Young, Donald S. Levinstone

In our effort to determine the PDM of long DNA molecules we have attempted automatic determination of the location of the PDM of short DNA molecules from a virus such as P22. Thus if the P22 or λ phage were inserted in the DNA of E. coli in vitro, then by discerning the location of the inserted molecule we ought to be able to determine the PDM of the E. coli DNA molecule in the neighborhood of the insertion site.

A key requirement of an automatic procedure such as this is to be able to identify the PDM histogram of the inserted molecule against the background PDM histogram of the longer molecule. To determine whether the procedures that we have developed previously[1] have this sensitivity, we performed the following experiment.

A set of 60 curves of P22 phage waveforms was divided into a training group and a test group, each with a set of 30 curves. A set of 55 curves of BP5, a deletion phage of P22, was divided into a training set of 25 curves and a test set of 30 curves. It can be seen from Figs. XXI-1 and XXI-2 that the training patterns are quite similar to the histograms of the complete sets of curves. The two test groups were combined into a single shuffled test set of 60 molecules. Then each curve was compared with the training set histograms of P22 and BP5 to obtain a correlation coefficient as a

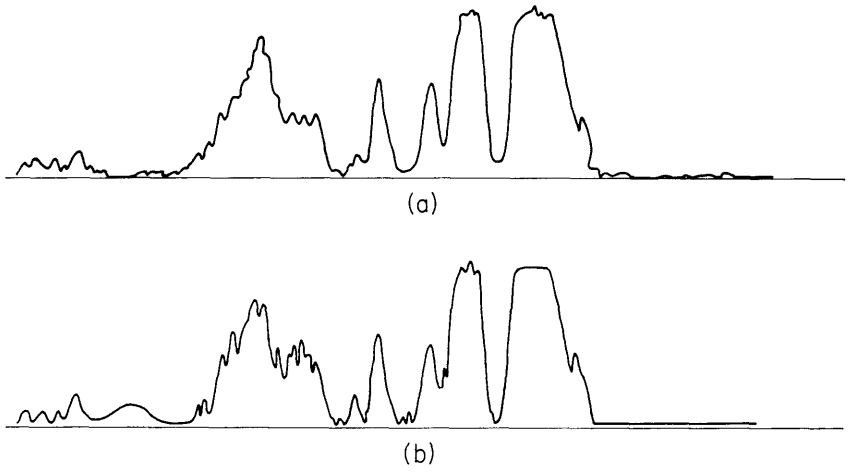

Fig. XXI-1. PDM histograms for P22 molecules. (a) Complete set of 60 molecules. (b) Training set of 30 molecules.

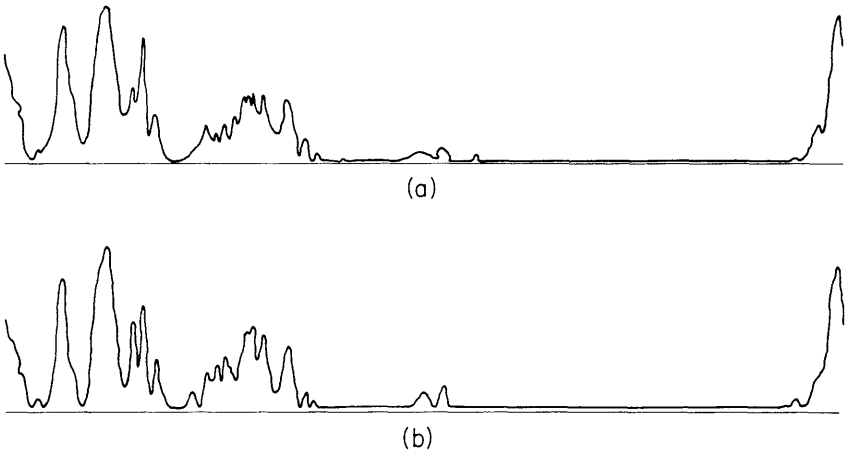

Fig. XXI-2. PDM histograms for BP5 molecules. (a) Complete set of 55 molecules. (b) Training set of 25 molecules.

measure of similarity. The i^{th} curve (molecule) from the test set was classified according to the following decision rule:

$$\rho_{i,\,P22} \underset{i \in BP5}{\overset{i \in P22}{\gtrless}} \rho_{i,\,BP5}$$

An analysis of the results shows that 29 of the thirty P22 curves and 26 of the thirty BP5 curves were correctly assigned. For the set of P22 curves, the quantity $(\rho_{i,\,P22} - \rho_{i,\,BP5})$ had average .19 with variance .09; for the set of BP5 curves, the average was

-.12, with variance .11. This is quite significant because biologically the BP5 curves would be expected to have some portions of their patterns in common with those of P22, since BP5 is a deletion phage of P22.

References

1. I. T. Young, D. Levinstone, M. Eden, B-K Tye, and D. Botstein, "Alignment of Partial Denaturation Maps of Circularly Permuted DNA by Computer," J. Mol. Biol. 85, 528-532 (1974).

5. ANALYSIS OF CHROMOSOME IMAGES: MULTIPLE-CELL KARYOTYPING

National Institutes of Health (Grant 2 PO1 GM19428-04)

Goesta H. Granlund, Ian T. Young, Gregory W. Zack, Murray Eden

In our studies of the automated karyotyping of human chromosomes we have formed a new potentially powerful concept of statistical karyotyping, which is that a karyotype can be developed from the chromosome complement of several cells. An obvious advantage is that we can take for analysis from each cell only chromosomes that are not touching, overlapping, or otherwise distorted. An even more important advantage is that information about the average behavior of descriptive parameters of each chromosome class can be assembled.

The reason for analyzing the chromosome complement is that clinical judgment is necessary rather than just an assembly of the karyotype picture of a particular cell. If we have only the chromosome spread image of a particular cell, we get very little information about the chromosome complement. If something unexpected is observed, we must obtain one or more images to confirm or reject a suspected irregularity.

The system that we have built is shown in Fig. XXI-3. It has been described elsewhere[1-5] and tested on Giemsa-banded metaphase cells. Because of its organization

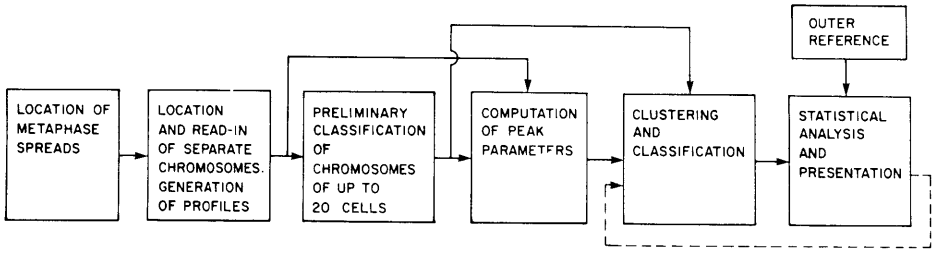

Fig. XXI-3. Automated karyotyping system.

and function, it is difficult to compare it with conventional systems or to specify its accuracy in well-known terms. An important factor in evaluating a conventional karyotyping system is the recognition rate of the system; that is, the percentage of correctly classified chromosomes. With our system, by observing the data output from the clustering and classification block, we have achieved a recognition rate of 90%, plus or minus a few percent depending upon the quality of the preparation and the care taken in the selection of spreads. This matter, however, is not pertinent for our system. Instead of error, we have an overlap of classes, as well as in homologues within classes. It can be shown that the error rate is directly related to the overlap of the distributions. The situation illustrated in Fig. XXI-4 shows two classes of chromosomes

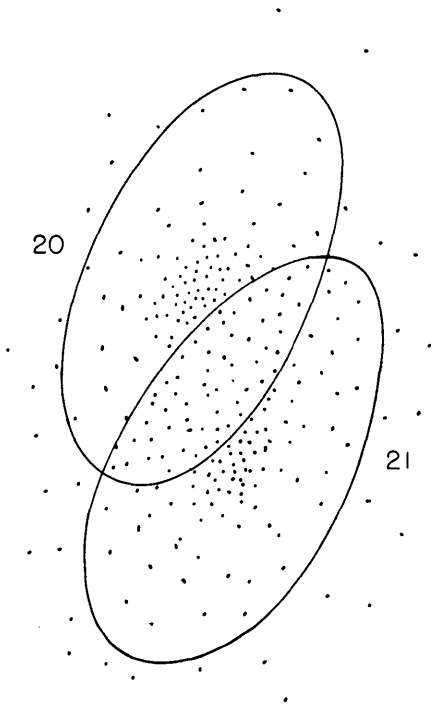

Fig. XXI-4.
Example of feature overlap of two classes: chromosomes No. 20 and No. 21.

with an overlap in the distributions of two features. We cannot tell to which distribution the points in the overlap region belong, but that is not important for our purpose. What is important is to find the means and variances of these two distributions, since they reflect the average behavior of these features, and we can now obtain this information.

These methods can be used to detect aberrations in the chromosome complement. Since the discovery of the banding stains, great variability in banding patterns has been observed, which is sometimes linked to clinical syndromes. Several hospitals now make investigations routinely in order to find chromosomal abnormalities that can be linked to observed syndromes. Some of these abnormalities are visible in the

karyotypes, while others can only be suspected. The methods that we have described may be and have been used to resolve such problems.

References

1. I. T. Young and I. L. Paskowitz, "Localization of Cellular Structures," IEEE Trans. on Biomedical Engineering, Vol. BME-22, No. 1, pp. 35-40, January 1975.
2. R. P. Bishop, "The Automated Detection and Classification of Human Chromosome Spreads," S.M. Thesis, Department of Electrical Engineering and Computer Science, M. I. T., April 1976.
3. G. H. Granlund, "The Use of Distribution Functions to Describe Integrated Density Profiles of Human Chromosomes," J. Theor. Biol. 40, 573-589 (1973).
4. G. H. Granlund, "Statistical Analysis of Chromosome Characteristics," Pattern Recognition 6, 115-126 (1974).
5. G. H. Granlund, G. W. Zack, I. T. Young, and M. Eden, "A Technique for Multiple-Cell Chromosome Karyotyping," J. Histochem. Cytochem. 24, 160-167 (1976).