

## XXI. COGNITIVE INFORMATION PROCESSING\*

### Academic and Research Staff

|                       |                    |                   |
|-----------------------|--------------------|-------------------|
| Prof. M. Eden         | Prof. D. E. Troxel | Dr. T. Yamamoto   |
| Prof. J. Allen        | Prof. I. T. Young  | Dr. M. Yoshida    |
| Prof. B. A. Blesser   | Dr. R. R. Archer   | F. X. Carroll     |
| Prof. T. S. Huang     | Dr. K. R. Ingham   | Deborah A. Finkel |
| Prof. F. F. Lee       | Dr. J. I. Makhoul  | E. R. Jensen      |
| Prof. S. J. Mason     | Dr. D. M. Ozonoff  | R. A. Piankian    |
| Prof. W. F. Schreiber | Dr. O. J. Tretiak  | J. S. Ventura     |

### Graduate Students

|                  |                  |                     |
|------------------|------------------|---------------------|
| W. R. Bogan      | M. Go, Jr.       | R. J. Shillman      |
| W. T. Borroz     | D. W. Hartman    | D. G. Sitler        |
| J. C. Borum      | M. E. Jernigan   | J. R. Sloan         |
| J. E. Bowie      | J. W. Klovstad   | A. A. Smith         |
| B. E. Boyle      | T. T. Kuklinski  | R. D. Solomon       |
| P. Coueignoux    | C. W. Lynn, Jr.  | J. H. Stahler       |
| C. H. Cox        | H. S. Magnuski   | W. D. Stratton      |
| J. R. Ellis, Jr. | L. C. Makowski   | H-m. D. Toong       |
| I. S. Englander  | P. L. Miller     | J. M. Tribolet      |
| A. M. Fakhr      | B. Minkow        | K. P. Wacks         |
| A. E. Filip      | J. A. Myers      | J. A. Walker        |
| S. G. Finn       | D. O'Shaughnessy | J. A. Whealler, Jr. |
| T. Fukuda        | I. Paskowitz     | H. M. Wolfson       |
|                  | R. S. Putnam     |                     |

### RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The Cognitive Information Processing Group works on problems related to the analysis, manipulation, and simulation of the sense data in certain human perceptual and cognitive processes. Our major interest is increased understanding of visual processes; in particular, the ways by which humans can recognize and interpret pictorial information. Thus we do research on electronic methods for simplifying pictures without altering significantly human subjective judgments of picture quality, on methods for altering pictures so as to enhance particular features and remove others, on the way humans perform sophisticated visual tasks, such as reading and writing, and on the way humans learn to recognize complicated patterns in microscopic images or photographs and categorize them without conscious effort.

These studies have a variety of applications, including computer input-output devices, communication systems, sensory replacement, and many others. The studies of language and picture processing have suggested ways to substitute other senses for sight, so that a blind person can "read" printed material in ways that are not too different from the ways sighted persons use such information sources. Image-processing and pattern-recognition studies are being applied to the classification of white cells in blood smears, to the detection of malignant cells in Papanicolau smears, to the diagnosis of blood dyscrasias by measurements on erythrocytes in smears, to the enhancement of x-ray images before radiological diagnosis, to chromosome classification, and to the location of marks in bacteriophage DNA.

---

\*This work is supported by the National Institutes of Health (Grants 5 PO1 GM14940-06 and 1 PO1 GM19428-01), the National Science Foundation (Grant GK-33736X), the Joint Services Electronics Programs (U. S. Army, U. S. Navy, U. S. Air Force) under Contract DAAB07-71-C-0300, a Grant from the Associated Press, and Peter Bent Brigham Hospital Purchase Orders F64088 and F65806.

## (XXI. COGNITIVE INFORMATION PROCESSING)

The mathematical methods, as well as the computer algorithms and technical processes developed within communication science, have proved relevant to a wide range of sensory and cognitive processes. Our research efforts have contributed toward the development of coherent theories of image processing and pattern recognition. We have found, for example, that research results, and even computer programs, relevant to speech analysis have a straightforward connection with research on the visualization of biological tissues. Thus our many discrete research activities share two unifying attributes: a common set of mathematical analytic techniques, and a core data-processing facility specifically created for research on two-dimensional signal processing and pattern recognition.

B. A. Blesser and D. M. Ozonoff have developed a procedure for discovering perceptual errors in the reading of medical x-rays when these have been made with and without techniques such as image enhancement. An experimental program is being carried out in collaboration with radiologists at the Peter Bent Brigham Hospital.

I. T. Young has supervised the construction of a microscope with attached vidicon scanner and program-controlled stage so that large populations of blood cells may be studied in order to classify anatomical features and to study cell properties such as contact adhesiveness. Similarly, G. H. Granlund developed a novel technique for automatically classifying chromosomes which makes it possible to trace distinct contour patterns of chromosomes in differentially stained cells. The group has extended J. E. Green's work on red cells so as to compute the shape of normal red blood cells. This forms a basis for research on the mechanical properties of red cell membranes. Automatic methods have also been developed for the detection of malarial parasites in blood smears.

Several projects designed to develop aids for the handicapped have been completed this year. F. F. Lee has successfully applied tactile feedback of pitch information to allow deaf subjects to control their voices in an acceptable way, while D. E. Troxel has continued to develop techniques for TV scan acquisition of characters from printed text, as well as a model for the computer generation of handprinted characters.

J. Allen has continued the development of a system for conversion of unrestricted English text to speech. A complete algorithm for translating typed English sentences to synthetic speech has been completed.

B. A. Blesser and M. Eden are studying the relation between the physical and perceptual properties of conventional symbols with the view of perfecting efficient procedures for recognizing alphabetic characters and other common symbol types.

Brief descriptions of the principal projects now under way, and some thoughts on plans for future projects follow.

### 1. Image Compression Techniques

The reduction of the number of bits required to transmit a visual image is a continuing goal. Coupled with such reduction is the attendant increase in subjective degradation of these images. We have studied visual noise perception in such pictures with controlled spatio-temporal spectra. At present we are adapting these techniques to newspaper facsimile production.

We are now turning our attention to much higher quality images such as are encountered in x-rays and facsimile transmission. Results are expected to be somewhat different from those previously obtained, since it is well known that many aspects of image processing are strongly affected by resolution. We have interfaced a high-resolution facsimile transmitter and receiver with the PDP-9 computer, and are studying various aspects of image quality and coding efficiency, such as optimum linear and nonlinear filtering, coding of both color and monochrome images by digital and analog means, and ways to transmit such images with a high degree of control over their properties. A new PDP-11 system for specific use in this study is to be installed this year. A

high-speed digital transmission system for x-ray images is also under development.

## 2. Image Enhancement and Transformation

Most of our work in this area has been devoted to analyzing the mathematical and technical problems encountered in dealing with visual images degraded by a variety of physical processes. Potential applications for the techniques and processes under study include many situations in which visual images must be analyzed and interpreted. Initial applications have been to medical problems.

Although we have done some exploratory work on radiological images, relatively little biomedical work could be justified because some technical problems had to be solved first.

Based on work begun during the past year, increasing attention will be directed toward specific biomedical image studies in the future. Two such projects have been initiated.

a. The work on the computation of x-ray sections is being continued to include an error analysis. We are using more elaborate models of the image formation mechanism, and evaluating the effect of model parameters on the reconstructed image.

b. Studies of the characteristic image degradations of x-ray camera systems and development of two-dimensional recursive digital filters so as to compensate for certain aberrations by computation.

c. We are using sequential estimation theory to refine models for images and to develop new image-enhancement techniques.

There are some interesting possibilities that we expect to explore in the future.

a. Use of computer processing to correct degradation of electron micrographs resulting from lens defects that induce linear, shift-variant aberrations.

b. Construction of an x-ray transmission system for remote diagnostic examination, where a very high-quality replica is required. (The project will include the design of a laser scanner, a receiver, and display unit with hard-copy output and a coding scheme dependent on techniques of bandwidth compression.)

## 3. Pattern Perception and Recognition

Because our work deals with visual or auditory signals that have a meaningful interpretation to human observers, we have always had certain relevant psychophysical or psychological projects. We have also carried on studies in the theory of pattern recognition. The applications of pattern-recognition techniques include new health-care methods, oceanographic studies, and automated industrial quality control. At present, the following projects are being pursued.

a. The use of filtered random-noise visual patterns for the measurement of human spatial frequency response. Both psychophysical and objective measurements are made.

b. A study of the utility of a perceptual model for diagnostic radiology (in collaboration with the Department of Radiology, Peter Bent Brigham Hospital). This study is essential to any future work on x-ray image enhancement. Our effort to characterize the nature of radiologic errors made in the normal hospital context has continued. By using a pair wise disagreement algorithm, we have determined the average error rate for reading conventional chest films. We expect to expand our methodology, using a five-way comparison. Also, a concurrent effort is being carried out to characterize the errors in terms of their perceptual origin.

c. A study of the possibility of speech recognition based on the detection of higher order linguistic constructs (syllables and words rather than phonemes). Thus far, this scheme has worked successfully for the recognition of the spoken integers 1 to 999.

## (XXI. COGNITIVE INFORMATION PROCESSING)

We expect these three projects to continue for some time. In particular, the application of image-processing analysis to radiology is expected to require several years of research effort. It is also quite possible that the studies of visual psychophysics may be applicable to the rapid and objective diagnosis of certain visual defects, such as astigmatism, myopia, and hyperopia.

Consideration of problems of biological image processing has given us several interesting ideas about perceptual processes that we intend to study in the next five-year period. They include the following.

- a. The visual correlates of texture; that is, what physical signals are interpreted as textures by human observers. A related study concerns human judgment of shape factors.
- b. A comparison of the utility of color values (a subjective judgment) with chromaticity as parameters for making discriminations by machine; for example, distinguishing cells by color of nucleus or cytoplasm.
- c. The development of a theory of hand-printed characters. In our approach, we experimentally derive functions which map between the "goodness" of the letter and some physical attribute. Ultimately, we expect to be able to apply our results to an automatic character recognizer.

We are developing algorithms for the following of lines and contours in a noisy field. These algorithms are intended for chromosome analysis and for the detection of cell boundaries in tissue sections.

### 4. Quantitative Microscopy

Major projects in this area will be outlined separately.

- a. Quantitative study of cell morphology. We plan to extend our identification algorithms from the five principal types of leukocytes (neutrophil, lymphocyte, monocyte, eosinophil, and basophil) to the next five most common types: atypical lymphocyte, metamyelocyte, band cells, hypersegmented polyocytes, and nucleated red cells. As in previous work, the algorithms for identification of the cells will be developed to measure those properties of the cells observed in light microscopy, that is, nucleus and cytoplasm size, nucleus and cytoplasm color, nucleus shape, and possibly nucleus texture. These 10 cells account for the cells observed on 97% of all slides that we have examined.

We also plan an in-depth study of the five principal types of leukocytes. We propose to perform large population studies of these cell types to understand better their distributional properties and inherent biological variability.

- b. Study of leukocyte stickiness. Our recent studies have shown that the contact adhesiveness of leukocytes may be measured by preparing blood slides on a special centrifuge and then comparing the number of leukocyte clumps (two or more leukocytes touching) with a number projected from a stochastic model. The excess of clumps over the projected number is a measure of the cell "stickiness." During the next eighteen months we plan a thorough clinical study to calibrate our measure and determine the nature of the difference between leukemic and nonleukemic cell adhesiveness. Also included, of necessity, will be a study of the effects of anticoagulating agents and preparation times.

- c. Distributional property of cell hemoglobin. Using techniques that we have already developed for the quantitative analysis of red cell morphology, we plan to study the distribution of hemoglobin per cell per patient in a large cell population. At present, the only quantitative information at our disposal is the mean corpuscular hemoglobin (MCH) and the mean corpuscular hemoglobin concentration (MCHC), both measured as the ratio of extremely global parameters. Our techniques yield the actual distribution of corpuscular hemoglobin so that more than just the mean value might be available.

Such distributional information would be extremely valuable in evaluating the effects of therapy, for example, in anemias; or when the distribution itself is not unimodal, for example, in thalassemia.

d. Characterization of biological shape. One of the important morphological features of the human blood cell is the shape of its constituent parts. Hence a simple characterization of the shape of an object (by a scalar quantity) is desirable. Such a quantity for the assessment of shape complexity has been developed. This quantity, the bending energy, is a measure of how much energy must be expended to deform a circle into a more complex simply connected closed contour. The energy, derived from a Fourier expansion of a parametric representation of the contour, will be used as a robust descriptor of cell shape in a series of planned cell-classification experiments.

e. We have developed a procedure for analyzing band-stained human chromosomes into a terse data set that summarizes the intensity, position, and width of the bands. These features have been used for karyotyping human cells. Continuing work is aimed at refining these features so that they may be used to distinguish among individuals.

## 5. Text-to-Speech Conversion

In the area of text-to-speech conversion, we have now demonstrated a complete algorithm for converting an arbitrary English word in text form to speech. At the start, an attempt is made to look up the word in a dictionary. If this fails, a further attempt is made to split up the word into its constituent morphemes (enables  $\rightarrow$  en + able + s), which are found in the lexicon. This is an involved process, and many refinements have been added in order to secure the correct result. When a word can be represented by these morphemes, the dictionary provides the names of the sound elements involved, as well as the parts of speech. If the word cannot be mapped into dictionary entries, however, it is converted to sound elements by means of letter-to-sound rules. These are undergoing continuing revision, and an assessment of the utility of each of the 400 rules is currently being made.

No matter how the sound elements are computed, the stress contour must be derived by a further set of rules. Finally, a set of low-level morphophonemic adjustment rules adjusts the values of the sound elements at morpheme boundaries.

At the conclusion of this process, a complete phonetic specification of the word is obtained. This description must then be converted to speech. In order to accomplish this change, a synthesis-by-rule algorithm is used to transform the sound element labels into a set of control parameters for a vocal-tract model. This involves the specification of vocal-tract source configurations, resonances, timing, and pitch control, so that intelligible speech is produced.

We are now able to demonstrate the entire system described above. The largest remaining problem at the word level is vowel reduction, which involves careful control of vowel duration, as well as of vowel color. As experience accumulates, further editing of the dictionary will eliminate faulty transcriptions for individual words.

The problems in synthesizing phrases and sentences involve control of pitch, pauses, and durations so as to reflect the desired syntactic and semantic structure of the utterance. First, the sentence is parsed to reveal phrasal elements. Phrase-level stress is then assigned, followed by the computation of the appropriate pitch, durations, and pauses. At present, our knowledge of these effects is very incomplete, and a large portion of our research is devoted to more powerful techniques of syntactic analysis, and to the prediction of the prosodic contour of sentences. Increasingly, the need for semantic analysis is felt.

We have designed a very fast special-purpose computer for speech synthesis, and it is currently under construction. The heart of this machine is a very high-speed multiplier ( $16 \times 16$  in under 100 ns) which enables the device to perform real-time simulation of elaborate vocal-tract models. The computer provides both the necessary speed for

## (XXI. COGNITIVE INFORMATION PROCESSING)

real-time continuous speech synthesis, and also the flexibility for experimenting with different vocal-tract models.

In the related field of speech recognition, we have shown that fundamental frequency and energy envelope waveforms are powerful indicators of syntactic structure. Research continues on the use of this structure to facilitate the speech-recognition task. It seems clear now that the prosodic correlates must be used together with local syntactic analysis routines in order to overcome the high error rates produced by "front end" phonemic detectors.

### 6. Sensory Aids

Our activities here center around language-acquisition tools. These include a reading machine for the blind, and speech parameter feedback for the deaf. There are several independent projects.

a. Developing hardware and algorithms for more reliable and more flexible character recognition.

b. Devising a complete scheme for converting written text to speech, including a dictionary, letter-to-sound rules, syntactic analysis, and the mapping of linguistic structures into acoustic correlates.

c. Developing high-quality intelligible and natural-sounding synthetic speech from a phonemic description of the constituent morphemes of the utterance. (This work is being done in collaboration with the Speech Communication Group of the Research Laboratory of Electronics.)

d. Developing an inexpensive minicomputer-based Braille translation facility.

During the coming year, we anticipate that the direction of this work will change in the following ways:

a. More blind subjects will be involved in test situations designed to evaluate synthetic speech and text-scanning equipment, and in the evaluation of linguistically based Braille as well as computer-generated Braille pictures.

b. The linguistic aspects of the program will be extended by organizing the relation of linguistic structures (phrases, words, syllables) to timing, fundamental frequency, and pauses, as well as research in speech training for deaf subjects, and studies of the reading-writing process, e. g., design of systems of potential value in the training of dyslexics.

c. We expect to develop a high-speed Braille line printer as part of a Braille production facility. In addition to the embossing of Braille text, this machine would also be used for the creation of Braille pictures automatically from black and white photographs.

d. A more recent objective is to develop devices and training methods that will enable those who are born deaf or with severe hearing loss to learn to speak so that they can be more easily understood. It is our belief that a continuous, prolonged usage of a suitable tactile feedback apparatus should be introduced to a deaf child for use throughout the language-acquisition phase of his life. Preferably this apparatus could be used both for monitoring important parameters of his own speech, and extracting the same parameters of the speech of others.

A dynamic pitch-to-tactile display device has been built and performs well. Other schemes for the detection of speech sounds (e. g., fricatives) and for frequency compression will be studied. Those realizations which are successfully evaluated will be built into practical, portable devices.

M. Eden, J. Allen, B. A. Blesser, T. S. Huang, F. F. Lee,  
S. J. Mason, D. M. Ozonoff, W. F. Schreiber,  
O. J. Tretiak, D. E. Troxel, I. T. Young