XXII. COGNITIVE INFORMATION PROCESSING*

Academic and Research Staff

Prof. M. Eden
Prof. J. Allen
Prof. B. A. Blesser
Prof. T. S. Huang
Prof. F. F. Lee
Prof. S. J. Mason
Prof. W. F. Schreiber
Prof. D. E. Troxel
Prof. I. T. Young
Dr. R. R. Archer
Dr. G. H. Granlund
Dr. J. E. Green
Dr. E. G. Guttmann
Dr. K. R. Ingham
Dr. J. I. Makhoul
Dr. D. M. Ozonoff
Dr. D. Sankowsky
Dr. O. J. Tretiak
F. X. Carroll
D. A. Fay
Deborah A. Finkel
C. L. Fontaine
L. Hatfield
C. C. Hsieh
E. R. Jensen
C. Nehring
Christine C. Nora
D. E. Robinson
J. M. Sachs
Sandra A. Sommers
J. S. Ventura

Graduate Students

A. K. Bhushan
R. R. Bochner
J. E. Bowie
B. E. Boyle
R. L. Brenner
J. R. Ellis, Jr.
A. E. Filip
A. C. Goldstein
W. B. Grossman
D. W. Hartman
L. P. A. Henckels
P. D. Henshaw
T. Kitamura
J. W. Klovstad
G. P. Marston
G. G. Matison
G. Poonen
N. Rashid
R. Singer
D. G. Sitler
A. A. Smith
R. D. Solomon
J. H. Stahler
W. W. Stallings
W. D. Stratton
K. P. Wacks
D. L. Whelan
T. R. Willemain
Y. D. Willems
K. H. Yung

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The primary interests of the Cognitive Information Processing Group are related to gaining an understanding of visual processes; in particular, to the ways by which humans can process pictorial information. Thus we do research on electronic methods for simplifying pictures (still or motion), 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 led to 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.

During the past year substantial progress has continued on the development of a reading machine for the blind. This system will be able to provide a blind user with

*This work was supported principally by the National Institutes of Health (Grant 5 PO1 GM14940-04), and in part by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 28-043-AMC-02536(E), the National Institutes of Health (Grants 5 TO1 GM01555-04 and 5 SO5-RRO7047-05), and by a grant from International Business Machines, Inc.

a "translation" from a printed page of text into synthetic English speech. The integrated system was first demonstrated in November 1968. Since then, there have been improvements both in the accuracy of the recognition of printed material and in the quality of the speech produced by the machine.

A stand-alone Braille translation facility is under development. It is anticipated that the system would be used to produce only a single Braille copy of the typed input. Thus far, the necessary hardware and software have been completed and English text typed on a teletype keyboard has been successfully converted into Braille. It now appears that the minicomputer which is central to this facility will support multiple translation stations.

The sensory aids program, during the last year, has completed the four-phone audio response system for the blind and the physically handicapped. A talking desk calculator and Fortran programs were added to the system's application software. A new keyboard incorporating an integrated circuit interface that transforms ASCII code to the 3 out of 14 touch-tone code was added to the system input devices. The training of blind persons in the use of the system has begun on a preliminary basis in order to assist in the specification of a stand alone service bureau system to be established with local agencies for the blind.

Development of a three-station time-shared Braille Translation System has reached the end of its initial implementation phase. A Grade 2 Braille and a limited mathematics Braille translator has been implemented with the input to the translators taken from keyboard and paper tapes. This system is designed to require no previous training by the sighted typists who would make the initial input of text to be translated. The ultimate goal is a machine translation system sufficiently low in cost to be within the means of the smaller agencies serving the blind.

The addition of a new drum memory to the reading machine system has permitted a marked expansion of the dictionary. We are now developing this new dictionary in a systematic way. Prefixes, suffixes, function words, high-frequency roots, and exceptions to letter-sound rules form the base lexicon. A large corpus of 40,000 words is then decomposed into this nascent lexicon by first sorting the corpus according to length, shortest word first. If a word from this list can be decomposed into morphs in the lexicon, then there is no need to add the word, but if not, the word is either added to the lexicon or it will be converted by letter-sound rules. An explicit algorithm for letter-to-sound conversion is under development, so that high-frequency exceptional words are converted by dictionary look-up, whereas low-frequency regular words are directly converted by letter-sound rules.

Two projects that deal with the syntactic analysis of sentences have been completed. Parsing is needed to provide the basis for proper stress, intonation, and pauses. The parser used in the system has been extended to yield a complete analysis of the sentence, including complicated effects such as ellipsis and embedding. Furthermore, a transformational rule tester has been developed to provide insight into the operation of grammatical transformations. It is expected that experience with the tester will lead to further improvements in the parser.

Improvements in the quality of synthetic speech have centered in two areas. Previously, the control parameters for synthesizing vowels and fricatives have been computed according to acoustic theory. Recently, a thorough analysis of the stop consonants has been completed. The sound quality has improved considerably, but certain weaknesses in the synthesis-by-rule algorithm have become apparent. As a result, we have started to revise the basic algorithm to provide more flexible control of the acoustic parameters. The difficult problem of determining the correlates of stress has also been investigated. A doctoral thesis has been completed on the relation of vowel duration to syntactic context for multisyllabic words, so that this parameter can be accurately controlled by the output of the parser.

While not related to the reading machine project, we have also investigated several problems in speech recognition. A new technique for optimally normalizing vowel

formant frequencies to account for different sized vocal tracts has been developed. This procedure can be used as an input pre-processor to a speech recognition system. Another project is concerned with the recognition of the spoken integers 1 to 999. By means of a linguistic description of the utterance, a segmentation is possible into syllabic units that can then be more readily recognized.

Research in the fundamental issues related to image processing and enhancement continues to be an important aspect of our work. During the past year, we extended the work on coding of continuous tone images by contour transmission, the contours being approximated by quadratic curves. While some improvement in coding efficiency (compared with previous curve-following techniques) was found, the much greater complexity of the process appears to rule out this technique. It would appear that efficient means of adding texture to contour-coded pictures is a much more promising avenue of investigation. We have found efficient means of representing texture by approximating the Fourier transform, but have not yet added this to coded images.

We have also studied the optimization of nonstatistical PCM picture transmission. For the first time, this was done in a manner that permitted accurate comparison between two systems over a wide range of channel capacities. The idea is to generate, for each system studied, a set of pictures, ranging from low to high quality, and, of course, from low to high channel capacity. At each capacity/quality level, the system parameters have been optimized to give the best subjective quality for each of a variety of subjects. When we have such a set of pictures of the same subject for each of two or more processing techniques, they can readily be intercompared and matched by observers. If two pictures have the same subjective quality, then the relative coding efficiency for that particular subject is simply the ratio of their channel capacities, without regard to such matters as scanning density or bits per sample. This study has shown not only that this is an effective way to compare systems but also that optimization of parameters (associated with linear and nonlinear filtering and quantizing noise randomization) results in very substantial improvements over "ordinary" PCM. In the future this work will be carried further, including higher quality pictures.

Several techniques for restoring images degraded by linear shift-variant systems (such as the coma aberration of a lens, and atmospheric turbulence) have been developed. They have been applied successfully to the restoration of images corrupted on a computer by linear shift-variant systems with additive noise.

Two techniques have been developed for estimating the impulse response of a linear shift-invariant degrading system by looking at the degraded image. The first technique consists in measuring edge response and then calculating therefrom the impulse response. In the second technique, the degrading image is divided into small regions, and the degrading impulse response is estimated by taking advantage of the fact that this impulse response is the same for all regions while the intensity of the original image varies from region to region.

A scheme of optimally approximating a two-dimensional digital filter by a sum of separable filters has been devised. In some cases, this can save a considerable amount of computer time in doing two-dimensional filtering.

As part of our long-range goal of understanding image transmission, including the role of the observer of the output image, we have carried out a study of threshold visibility of noise of known two-dimensional spectrum, both alone and superimposed on images. The objectives have been to find a description of visual spatial frequency response that is alternative to that obtained with highly structural test patterns (sine-wave charts, etc.). At the same time it was to be more closely related to ordinary viewing conditions. The "noise MTF's" obtained thus far are similar to, but not exactly like, "sine-wave MTF's" previously obtained. We have also tried to ascertain whether a linear weighting function existed which could operate on a broad spectrum of individually subthreshold noise components to predict their joint visibility, but the results are inconclusive.

While the understanding of the problems of picture compression and image

enhancement is important for the broad problems related to extracting useful information in pictures, it has particular relevance to medicine and biology. This is because the processes of diagnosis and analysis in the health sciences depend heavily on the qualitative description of visual objects, such as microscope preparations, x-ray images, and so forth. Thus image-processing basic research is very closely tied in with the Biological Image Processing projects of our group.

Work has continued on automatic methods for analyzing the morphology of red blood cells in blood films. It has become obvious that manually prepared blood films are too poor in quality for precise automatic analysis. Accordingly, we have developed methods for automatic preparation of blood specimens which produce films exhibiting very uniform cell distributions.

Work has also progressed in the collection of data from large numbers of cells from individual subjects. Tests on the accuracy and reproducibility of these results indicate that cell features are measured within mean errors of ±1%. This work has been impeded by the necessity of photographing every microscopic field that we investigated, but this impediment will be removed soon upon the completion of a microscope-scanner facility. This facility will enable us to collect cell data much more rapidly. Substantial progress has been made in streamlining the execution of the computer algorithms that process the cell images, to the point where the photography and scanning procedures present a bottleneck.

In the neural anatomy projects, we have developed computer algorithms that display data from serial sections as collated three-dimensional stereographic images.

We have found that in a Papanicolau preparation of exfoliated vaginal cells, integrated chromatin density is a remarkably effective single parameter for distinguishing between normal and suspicious smears.

In the analysis of stained histological sections, we have developed algorithms for locating nuclei, and for tracing cell boundaries. These procedures are now being integrated into a system that will parse images of a section.

We are continuing our efforts to reconstruct three-dimensional images from collections of x-ray photographs. New algorithms for this process have been developed, and we have a better theoretical understanding of the errors implicit in this procedure.

A simple procedure for distinguishing between mature and immature polymorphonuclear leukocytes has been developed.

We have investigated the feasibility of recognizing leukocytes from diffraction and spectral absorption measurements. It seems that such measurements are adequate to classify the five normal categories.

An investigation into the perceptual dynamics of reading diagnostic x-ray image has been initiated in conjunction with the Radiology Department of Peter Bent Brigham Hospital. It is felt that an understanding of the perceptual errors is necessary before any attempt to use modern technology for image enhancement can yield significant improvements. In particular, a methodology is needed to evaluate a system, since the image "quality" is not necessarily related to diagnotic performance. These studies will concentrate on generating a standard calibrated reference library of chest x-rays for use in evaluation.

The audio test facility now connected to the PDP-9 (FRM) computer is essentially completed. Experiments involved with on-line computer-controlled listening of speech and music have begun.

Our research objectives for the coming year are extensions of work that has been in progress for many years.

1. The development of a Reading Machine for the Blind is one of our major projects.

(a) Work continues on the new Vidicon scanner, which will be able to scan intact pages of a book or magazine. This project is also aimed at developing more efficient means of inputting data to the computer, so that the cost of the scanner part of the system can be reduced.

Experience during the past year has demonstrated effective methods of employing Vidicon scanners for reading-machine input. We have substantially revised our plans in this connection and are now putting together a system that will handle intact books and magazines with quality adequate for accurate recognition.

(b) We are continuing our study of character-recognition algorithms. We are trying to develop inexpensive techniques to recognize several fonts, and to deal with the problems of poor-quality printing, such as missing segments and touching letters.

(c) We are attacking the problem of speech quality from a fundamental point of view. The classes of speech sounds (such as stop consonants) are being described from an acoustic-phonetic point of view, and the synthesis algorithms are being designed in accordance with these data. We expect to be able to generate phonemes accurately, and also to provide a better approximation to the suprasequential cues, including stress, intonation, and juncture.

(d) To be useful, a reading machine for the blind must have convenient controls to allow for page positioning, backup, and other functions that would otherwise require a sighted operator. Experiments are under way to determine what controls are needed and how they should be implemented.

(e) Tests will be conducted with the Braille translation facility to determine the number of keyboard input and Braille stations that can be supported. The use of keyboard inputs that are more suitable than a teletype will be investigated.

(f) We are expanding our dictionary with new entries in order to be able to handle a wide variety of texts. Present efforts are centered on selection of new words, provision for correct sound representations, and examination of all new words for possible decomposition into prefixes, roots, and suffixes.

2. Image-processing Techniques continue to be studied in several areas.

(a) 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. Study of visual noise perception in such pictures is being done with controlled spatio-temporal spectra.

(b) In the field of coherent optics, we have generated holograms of two-dimensional objects by computer. We are now turning to computer-generated holograms of three-dimensional objects.

(c) In the field of image enhancement, we have been studying the effect of linear systems that are not shift-invariant. Initial work has shown that the distortion induced by such shifts can be reduced by use of a form of the Mellin transform on the data.

(d) Previous work on x-rays has shown that simply improving the appearance of radiographs is of doubtful value. To improve diagnosis, it appears that either the processing must reveal previously hidden features, or else the radiologist must be trained long enough to get added benefit from an altered presentation. With respect to the first possibility, we shall do some experiments with a much higher quality picture input/output system to see if this is practical. With respect to the second, we plan to carry out some psychophysical experiments to try to understand what radiologists do when they examine x-rays.

(e) 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, as it is well known that many aspects of image processing are strongly affected by resolution. We are interfacing a high-resolution facsimile transmitter and receiver with the PDP-9 computer, and plan to

study 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 means by which such images may be transmitted with a high degree of control over their properties.

(f) We are interested in several problems related to the restoration of images degraded by atmospheric turbulence and image motion: (a) How to estimate the characteristics of the degrading system; (b) How to combat film noise, which imposes a fundamental limitation on all image enhancement techniques; and (c) Practical ways of handling linear shift-variant degrading systems.

(g) We are continuing to investigate the problems of efficient representation of color images. Initially, statistics are being collected for the color components in several images. From these data, it is expected that efficient coding schemes can be developed.

3. A major objective of our research is the Automatic Analysis of Biological Objects.

(a) One of the red cell indices now used in clinical analysis is the mean corpuscular hemoglobin (MCG). Current methods can only provide the mean value for many millions of cells. Based on automated visual examination of the peripheral blood film we have begun to investigate the quantitative assessment of hemoglobin on a per cell basis.

(b) Another of the clinical indices is the mean corpuscular volume (MCV) of the red blood cell. Since this calculation yields only the mean value of the corpuscular volume distribution (as opposed to the distribution itself), we propose to investigate the measurement of the volume of each cell. This requires that we establish a criterion that will specify shape parameters for the red blood cell.

(c) The investigation of leukocyte types will be extended to include primitive cells as well as abnormal cells. We shall investigate to see whether there is a quantitative difference in the cell morphology of a leukemic cell as contrasted with a normal cell.

(d) The completion of the microscope scanning facility will allow us to by-pass the photography process that is now necessary. This will facilitate more rapid collection of data in the erythrocyte morphology analysis project and eliminate errors inherent in a photographic intermediate. This facility will accelerate our effort to develop algorithms for diagnosis of abnormal types of cells, as well as our effort to develop a facility for quantitative individual cell measurements.

(e) Work will continue to standardize and characterize the automatic blood film preparation method, and will investigate instruments that are reported to produce much more consistent Wrights' staining of the blood films.

(f) A new project will be started to determine the feasibility of the automatic detection of malarial parasites in blood films. This task is of considerable importance to malaria control programs.

4. A variety of projects in Pattern Recognition continue.

(a) Several years ago, we were able to provide a formal description of the dynamics of hand motion that is used in producing English script. This was used as the basis for a highly successful series of recognition programs, but these programs required information about the time course of the writing. Since readers of cursive script do not have such information, it is highly desirable to do the recognition from spatial information alone. We are re-examining these algorithms to see if they can be extended to reading script based on spatial information alone.

(b) Work continues on a program to read musical scores. Some insight into the use of contextual information has been obtained, and the costly procedure of template matching has been avoided completely. Virtually all symbols on a musical score can now be identified. This work will now be extended to encompass dynamic marks

and chord structures.

(c) We have been trying to improve the performance of speech recognition systems by instructing the speaker to modify the manner in which he produces speech sounds. The system has been tested on a restricted set of vowels and consonants, and testing will continue for a larger class of English speech sounds.

(d) Research on the invariant properties of three-dimensional visual illusions is being carried out. Results, thus far, on the geometry of the trapezoidal window and on three-dimensional analogues of the Nekker cube have revealed interesting inconsistencies in interpretation. Attempts are being made to interpret these results within a more general theory of recognition of surfaces and textures in depth perception.

5. Our psychophysical and psychological studies of cognitive tasks continue.

(a) Our previous studies in noise visibility have shed some light on the role of the spectra of the noise and the image on which it is superimposed. If conditions permit, we shall attempt some further work along the lines of improving the quality of the experiment and extending the studies to include temporal variations.

(b) An important question in both picture coding and image enhancement is, How do we measure picture quality? Ideally, we would like to have a mathematical measure of picture quality. Although this ideal is hard to attain, we have gained much insight by studying many special cases. We plan to continue our effort in this area.

(c) Work proceeds on visual perception with the goal of applying this knowledge to improving character pattern recognition. In particular, we are trying to find those features that are invariant to type font.

(d) The role of context in perception is being studied in both auditory and visual modalities. We are most interested in understanding how a complex stimulus can be perceived without specific perception of its constituent individual elements. Such results will be of direct value in understanding reading and speech perception.

(e) Recent work in our group on the use of tactile pitch feedback for deaf speakers has given us the confidence to believe that deaf persons can acquire reasonably accurate pitch control through the other sensory modalities. The objective of this research project is to develop devices and training methods to enable those who are born deaf or have sustained severe hearing loss to learn to speak in a voice that can be more easily understood.

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