XVIII. COGNITIVE INFORMATION PROCESSING

Academic and Research Staff

Prof. Murray Eden
Prof. Jonathan Allen
Prof. Barry A. Blesser
Prof. Chi-Hau Chen*
Prof. Goesta H. Granlund†
Prof. Francis F. Lee
Prof. William F. Schreiber
Prof. William Simon‡
Prof. Ching Y. Suen**
Prof. Hoo-min D. Toong
Prof. Donald E. Troxel
Prof. Ian T. Young

Dr. Jack E. Bowie
Dr. Brian E. Boyle
Dr. Ralph W. Gerchberg
Dr. Uri F. Gronemann
Dr. Martin J. Hewlett
Dr. Marvin E. Jernigan
Dr. Dennis J. Kfoury
Dr. Murat Kunt††
Dr. Perry L. Miller
Dr. David M. Ozonoff
Dr. Stephanie E. Sher
Dr. Robert J. Shillman
Dr. Robert D. Solomon

Dr. Armando Roy Yarza
Dr. Makoto Yasuhara‡‡
Marc Berthod
Francis X. Carroll
Caleb J. Drake
Allison G. Hagerich
M. Sharon Hunnicutt
Eric R. Jensen
Kenneth A. Okin
Robert A. Piankian
William E. Rodgers
Kenneth P. Wacks

Graduate Students

Jack A. Allweis
Robert P. Bishop
Ronald E. Boucher
Charles H. Cox
David R. Cuddy
James R. Ellis, Jr.
Irvin S. Englander
Richard S. Goldhor
Samuel M. Goldwasser
Monson H. Hayes
Eli Israeli
Malik Khan

Theodore T. Kuklinski
Jay S. Kunin
Sze-Ping Kuo
Donald S. Levinstone
Charles W. Lynn
Lee C. Makowski
José L. Marroquin
Dennis V. Marsicano
Joseph T. Olesik
Douglas O'Shaughnessy
Douglas B. Paul

Roger S. Putnam
John N. Ratzel
Robert E. Rouquette
Mark A. Shalloway
Chyi Shiau
Brian R. Stanley
Robert J. Steingart
Christopher E. Strangio
Joseph E. Wall, Jr.
Allen W. Wiegner
Tonny Y. Wong
Gregory W. Zack

1. TEXT-TO-SPEECH AND AUDIO ANNOUNCEMENT SYSTEMS   JSEP

Jonathan Allen, M. Sharon Hunnicutt, Francis X. Carroll, Douglas O'Shaughnessy, Richard C. Goldhor, Robert J. Steingart, Robert S. Rosenschein

Our current emphasis on text-to-speech conversion for unrestricted text and for audio announcement systems centers on improvements to the quality of phonemic synthesis and on the development of prosodic algorithms for the control of pitch and timing.

In the area of phonemic synthesis, an improved algorithm is being implemented that not only provides for better speech quality but also allows relatively easy editing of the   JSEP

---

*Visiting Associate Professor, Southeastern Massachusetts University.

†Visiting Associate Professor, Linkoeping University, Linkoeping, Sweden.

‡Visiting Associate Professor, University of Rochester.

**Visiting Scientist, Concordia University, Montreal.

††Visiting Scientist, Laboratoire de Traitement de Signaux, Lausanne, Switzerland.

‡‡Visiting Scientist, The University of Electro-Communications, Tokyo, Japan.

rules. With this capability, we expect to achieve considerably improved speech quality, to the point of acceptability in a wide range of applications. The design of two special-purpose hardware modules to support the processing load of the phonemic synthesis is nearly complete. One device is a microprocessor-based system for converting a narrow phonetic transcription to control parameters needed by a digital vocal-tract model, and the second module is an all-digital vocal-tract model capable of exercising 32 second-order sections (resonances) in real time at a 10-kHz sample rate. These digital systems are small, and after debugging we expect to consider possible custom LSI implementation.

While improved phonemic synthesis is necessary, it must be augmented by appropriate timing and pitch control. We are implementing a two-level parser that can detect phrases and clauses for use by the prosodic algorithms. This parser is designed for use on unrestricted text, and is very efficient both in space and time. Two major theses have been completed on prosodic parameters. In one case, it has been shown that the syntactic role of lexical items can be used to predict word durations, and in the second study a complete model for pitch control has been developed. We plan to integrate these two results with the phonemic synthesis algorithm in order to yield a complete sentence-level synthesis system capable of delivering highly intelligible and natural-sounding speech.

2. APPROXIMATE PLANAR DECOMPOSITION AND ENCODING OF TWO-DIMENSIONAL PICTORIAL INFORMATION

James R. Ellis, Jr., Murray Eden

Our research has two primary goals: to obtain more efficient coding methods for highly structured two-dimensional pictorial information, and to investigate novel methods, based on objective measurements and criteria, of defining the complexity of criteria. These complexity definitions could be used in lieu of subjective a posteriori judgments to determine adequacy of digital representation of pictures for storage, transmission or processing purposes.

Fourier analysis has been employed to develop sufficient criteria for the encoding, transmission, and reproduction of images. It has been found sufficient to use the linear-system and time-series concepts of frequency content for image description. Furthermore, the use of information theory for images has resulted in redundancy reduction. In applications of both approaches it has generally been assumed that light intensities sampled over a rectangular mosaic are the appropriate descriptive variables.

There have been several attempts, primarily by psychologists, to measure the information content of pictures and binary-valued figures. No clear picture seems to have emerged from these efforts.

We have found that most meaningful pictures have relatively few regions of interest. Often these regions are fairly simple in character. The characteristics are generally carried in informationless, objective, accessible form by intensity planes. Bit planes, which are binary-valued maps having the value of a given bit of a code, carry much more information for representation and reconstruction purposes than intensity planes. Precisely, we go from M intensity planes to $\log_2(M)$ bit planes.

Thus we may conjecture that procedures that encode fairly simple regions in real pictures as fairly simple regions in bit planes, followed by procedures that encode the regions in the bit planes efficiently, may well encode pictures in an efficient, information lossless (up to sampling and quantization accuracy) form. Such procedures may also lead to methods of specification of image degradation in terms of contour distortion.

When a binary planar figure has a contour that is short compared with the square root of its area, chain encoding of the boundary is efficient. When the boundary is fairly simple, it can be encoded even more efficiently by polygonal approximation. We have shown[1] that the number of extreme points in such an approximating polygon is proportional to the fourth root of the area of the figure.

One obvious way to simplify the bit-plane images is to minimize the number of different regions that the bit planes represent. Clearly, unit-distance codes have the minimum possible total number of transitions for any n-bit code. The Gray code is the best known unit-distance code. It is simply generated, and is the only unit-distance code for 1, 2, or 3 bits. There are 9 'different' 4-bit codes, and many codes for 5 bits or more.

We have found a method of generating desired unit-distance codes, especially those with equalized bit significance. All nine 4-bit codes can be generated in this way, and the collection has a simple but unobvious structure. Equalized significance codes have also been generated for 5, 6, 7, and 8 bits.

These codes are now being used with the standard set of pictures to generate a collection of code-word length estimates for coding scheme comparison.

Our primary theoretical considerations at this time involve the relationships among statistical specification of pictures and the rate of convergence of their threshold plane, bit plane, and gray-scale image representations. It is not clear that the commonly employed measures of 'entropy' are adequate for this purpose.

References

1. J. R. Ellis and M. Eden, "On the Number of Sides Necessary for Polygonal Approximation of Black-and-White Figures in a Plane," Inform. Contr. 30, 169-186 (1976).

3. FONT-INDEPENDENT CHARACTER RECOGNITION

Barry A. Blesser, Robert J. Shillman, Murray Eden

The objective of this research is to design a computer algorithm capable of recognizing characters printed in a wide variety of type fonts.

During the past 10 years in the United States and abroad, considerable engineering work has been done on the design of a multifont machine.[1,2] Although there are many machines on the market capable of reading characters printed in a limited number of type styles (typically less than 10 different fonts) there has been little success in achieving an omnifont device (OCR). We attribute this failure to the lack of a general theory of characters upon which to base recognition techniques. Toward this end, in the past two years, we have developed a theoretical model for character recognition.[3]

We have approached the omnifont recognition problem from two directions. We have performed a wide variety of psychological experiments in order to discover a set of invariants for describing all 26 uppercase letters of the roman alphabet,[4] and we have been analyzing a wide variety of type fonts in an attempt to describe the consistency within a particular font and among different fonts.[5,6] In the past six months we have devised an algorithm for the PDP-9 computer to generate character trajectories, and we have utilized this computerized system to generate typical U-V trajectories for

JSEP

subsequent use in psychological experiments.[7] We have been working simultaneously on a model for describing roman type fonts. The present model is valid over a large class of textual fonts and is applicable to the multifont problem. This model accounts for consistency within a font, and hence has predictive power. It deals with simple parts of characters, and may help the OCR device to be less sensitive to certain kinds of noise such as kerning effects and missing parts.

References

1. L. D. Harmon, "Automatic Recognition of Print and Script," Proc. IEEE 60, 1165-1176 (1972).

2. R. J. Shillman, C. Cox, T. T. Kuklinski, J. Ventura, M. Eden, and B. A. Blesser, "A Bibliography in Character Recognition: Techniques for Describing Characters," Visible Lang., Vol. 8, pp. 151-166, Spring 1974.

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

4. R. J. Shillman, "Character Recognition Based on Phenomenological Attributes: Theory and Methods," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., 1974 (unpublished).

5. P. Coueignoux, "Generation of Roman Printed Fonts," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., 1975 (unpublished).

6. R. J. Shillman, T. T. Kuklinski, and B. A. Blesser, "Experimental Methodologies for Character Recognition Based on Phenomenological Attributes," Proc. Second International Joint Conference on Pattern Recognition, Copenhagen, Denmark, August 13-15, 1974, pp. 195-201.

7. C. Y. Suen and R. J. Shillman, "The Discrimination of Us from Vs" (in preparation for publication).

JSEP

4. PATTERN RECOGNITION OF CONVENTIONAL SYMBOL SYSTEMS

Murray Eden, Barry A. Blesser, Robert J. Shillman, Theodore T. Kuklinski

In our effort to design a machine that will read unconstrained hand-printed characters at an error rate that is equal to, or lower than, human performance we have developed a theoretical approach to character recognition[1,2] based on ambiguously shaped characters, that is, characters that can be perceived as either of two letters. These characters are important because they define the boundaries between letter subspaces in the character space; an analytical description of these boundaries could be incorporated into a recognition algorithm. Through extensive psychophysical experimentation we have found and examined more than 190 ambiguous characters and have developed a grammar for describing letters.[3] The grammar comprises a compact set of attributes and relational rules. The relation between the physical attributes of a character, derived from physical measurements, and the functional attributes, which specify the character's identity, are termed PFRs (Physical-to-Functional Rules). We have shown recently that a character's letter label need not be fixed but may vary, depending upon the graphical environment in which the character is embedded.[2,4] This indicates that although a given character has a unique physical representation, it can

have more than one functional representation, and hence more than one letter label. This effect is due, at least in part, to the modulation of the PFRs by the graphical context. Graphical context, a form of context not previously investigated in the character-recognition literature, takes into account the effect of stylistic consistency between characters and within a character itself.

Recent results of the past year's work are summarized as follows.

1. A test has been developed[2,4] for determining the relevance of an attribute for discrimination in any particular letter pair. This test involves shifting one interletter boundary and observing whether a different interletter boundary also shifts.

2. The PFR for LEG formed by a line addition has been determined.[5] This rule is important for discriminating E from F and C from L.

3. A two-parameter function has been found to describe the boundary between block-printed U and V.[6]

4. A transformation has been obtained for mapping characters with thick strokes (e.g., made by bold printing or blunt pencil) to phenomenologically similar skeleton characters.[7] This transformation is important for optical character recognition because the location of the intersection of two thick strokes is not defined.

In summary, during the past year we have studied the relation between physical and perceptual properties of conventional symbols in order to perfect procedures for recognizing alphabetic characters and other common symbols. We are attempting to draw together results and methods from both engineering and psychological disciplines so that eventually we can build a machine capable of reading unconstrained input.

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 Studies 6, 701-714 (1974).

2. B. A. Blesser, T. T. Kuklinski, and R. J. Shillman, "Empirical Tests for Feature Selection Based on a Theory of Human Character Recognition" (to appear in Pattern Recognition).

3. R. J. Shillman, "Character Recognition Based on Phenomenological Attributes: Theory and Methods," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., 1974 (unpublished).

4. T. T. Kuklinski, "Plasticity Effects in the Perception of Handprinted Characters," S.M. Thesis, Department of Electrical Engineering, M.I.T., 1975 (unpublished).

5. T. Lee, "The Phenomenon of Line Addition in Character Recognition," S.B. Thesis, Department of Electrical Engineering, M.I.T., 1975 (unpublished).

6. C. L. Jans, "An Investigation of U-V Discrimination," S.B. Thesis, Department of Electrical Engineering, M.I.T., 1975 (unpublished).

7. C. Chow, "Effects of Line Width on the Recognition of Vs and Ys," S.B. Thesis, Department of Electrical Engineering, M.I.T., 1975 (unpublished).

5. USE OF COMPUTERS IN RECOGNITION OF BIOMEDICAL
    PATTERNS: CHROMOSOME STUDIES

Goesta H. Granlund, Gregory W. Zack, Jan S. Spriet,

Robert P. Bishop, Monson H. Hayes, Ian T. Young

The principal objective of our studies in automatic cytogenetics is to apply the high-speed and large data-handling capacity of computers to the problem of human chromosome analysis. Techniques of image processing, scene analysis, and statistical pattern classification are used to aid in chromosome research and to develop improved clinical procedures. The work falls into four categories: metaphase spread detection and classification, chromosome localization, sister chromatid exchange detection, and automatic karyotyping.

a. Metaphase Spread Detection

A metaphase chromosome spread, comprising all chromosomes from a single cell, is approximately 50 μm in diameter. Since only 40 such spreads can normally be prepared on a typical microscope slide, it is desirable to find a means of locating them automatically. To achieve this, a metaphase spread detector has been incorporated into the TEMPO automatic microscope facility of the Cognitive Information Processing Group. Run-length analysis is used to detect the presence of a chromosome spread in the television picture obtained from the microscope. For our purposes, a run is defined as a sequence of consecutive picture points having brightness below a given threshold. The distribution of lengths of all runs in a picture is characteristic of the spatial frequency content or texture of the picture and is used to determine whether a spread is present and, if so, whether it is a prophase, early metaphase, or late metaphase spread. Tests performed thus far have shown a high degree of reliability. In one test, 12 G spreads were classified correctly and only 4 were incorrectly identified. We have also tested 100 fields containing artifacts. All of these fields were correctly identified as artifactual.

b. Chromosome Localization

Isolation of the individual chromosomes of a metaphase spread is performed on the CIPG PDP-9 computer. A grammatical analysis of each of the objects found in the picture is performed to distinguish single chromosomes from touching or overlapping chromosomes, dirt or stain particles, or chromosome fragments. A second-order coordinate transformation is used to orient and straighten the chromosomes in preparation for further analysis. To make these localization procedures more accurate and reliable, we are developing programs to improve the original pictures, which have been degraded by bandlimiting and distortion in the electronics apparatus and by diffraction in the microscope. Techniques of background subtraction and both iterative and non-iterative inverse filtering are being explored.

c. Sister Chromatid Exchange Detection

Sister chromatid exchanges are sensitive indicators of chromosome damage caused by certain drugs.[1] Through use of certain cell culture and staining procedures the two chromatids can be made to exhibit different brightnesses, with reversal occurring at exchanges. Preliminary work has shown the feasibility of the automatic detection of such exchanges[2] and work is now under way to improve the methods and extend their application.

d. Automatic Karyotyping

The cost of clinically preparing a karyotype, or identifying each of the chromosomes contained in a cell, is approximately $75. Our studies have shown that this identification can be done automatically, and, in principle, less expensively, by using decomposition into Gaussian distributions to describe the profiles of optical density that are characteristic of chromosomes treated with a banding stain.[3] This procedure is now being adapted for application to chromosomes from many cells to obtain a statistical description of the chromosome complement, rather than a complete description of only one cell. An advantage of this method is that every chromosome from each cell need not be used; touching and overlapping chromosomes can be discarded; hence, the procedure is more compatible with the metaphase spread detection and chromosome localization segments of the system.

References

1. S. A. Latt, "Sister Chromatid Exchanges, Indices of Human Chromosome Damage and Repair: Detection by Fluorescence and Induction by Mitomycin C," Proc. Natl. Acad. Sci. USA 71, 3162 (1974).

2. G. W. Zack, J. A. Spriet, S. A. Latt, G. H. Granlund, and I. T. Young, "Automatic Detection and Localization of Sister Chromatid Exchanges" (to appear in J. Histochem. Cytochem.).

3. G. H. Granlund, "The Use of Distribution Functions to Describe Integrated Density Profiles of Human Chromosomes," J. Theor. Biol. 40, 573 (1973).

6. DIGITAL FACSIMILE EQUIPMENT FOR RADIOGRAPHS

William F. Schreiber, Uri F. Gronemann

a. Version I

Our first version of radiograph facsimile equipment, described in Quarterly Progress Report No. 112 (pp. 95-96), had one scanner unit and one reproducer unit. The scanner served as a transmitter, and the reproducer as a receiver. Transmission was in serial digital form, at 56 kbits/s over a short coaxial cable. Sampling density was 250/in in a square grid. Scan raster format was variable in width and height, in 1/2 inch increments, up to 14 inches wide and any height. Sampling rate was variable between 7000 and 14,000 per second, corresponding to variation in sample digitization between 8 bits and 4 bits. These parameters were controlled from switches in the front panel of the scanner, with the reproducer following by slave drive.

The optical systems of both units were identical, based on the principle of laser beam line scanning by means of pre-objective lens deflection, with objective lens auto-collimation ("double pass"). The objective lens was a symmetrical half of a commercial reprographic lens, and deflection was effected by a mirror attached to a galvanometer. Vertical scanning was effected by moving the film a linewidth (i.e., 4 mils) at a time by means of rubber rollers driven by a stepping motor. In the scanner, the laser beam was internally chopped at 56 kHz, and light intensity was stabilized by a feedback arrangement. Light transmitted through the film, or selectably reflected from an opaque picture, was collected by a strip of solar-cell photodiodes. The resulting signal was amplified, rectified ("dechopped"), converted to logarithmic (density) scale, and then digitized.

In the reproducer the received signal was converted back to analog video (with additional nonlinear processing to compensate for film characteristics), which in turn modulated the laser. The reproduction medium was 3M Company dry silver film. Following exposure, this was automatically processed in continuous fashion by moving through a slotted oven whose temperature was accurately controlled.

The equipment was tested early in 1974 with both facsimile test patterns and actual original x-radiographs. The results were quite satisfactory and very encouraging. The test patterns showed the expected limits of resolution, with aliasing occurring at the fine line patterns. The reproduced chest radiographs, in the judgment of several radiologists, were of diagnostic quality. A very high resolution original of a hand (done by direct exposure of industrial x-ray film) reproduced fairly well, with a surprising amount of bone detail. Some artifacts were visible in the reproduction, but they could all be traced to minor faults and deficiencies in various components of the equipment. An interference fringe pattern was also visible; we think it resulted from lack of antihalation backing on the dry silver film. The reproducible density was limited to approximately 2.5 optical density units, because of inherent noise in the photodiode-preamplifier combination in the scanner.

b. Current Versions

During 1975, the first version of the equipment has been partly redesigned and rebuilt with the following modifications.

1. Electronics

a. Sample density has been made variable by factors of 2, that is, between 1000 and 31 per inch (horizontally and vertically).

b. Sampling rate has been made selectable, that is, 7000, 14,000, and 28,000 per second.

c. Additional controls have been provided to position the raster automatically horizontally and vertically with relation to the film. Pushbutton manual control of vertical film position has been provided.

d. Some of the video processing circuits have been modified and/or improved.

e. Interface circuits to a PDP-11 computer (via DMA channel) have been built for parallel format data input and output. The necessary handler for software has been partially written.

f. Interface circuits to a Bell DDS System have been built for long-distance transmission.

2. Optical-Mechanical

a. The optical system has been entirely redesigned, to accommodate the variable sample density. To achieve this, the scanning spot size has been made to vary by factors of 2, that is, between (nominally) 1 mil and 32 mils. This is effected by switching into the laser beam one of several beam-expansion lenses of various focal lengths, thereby controlling, in effect, the f number of the system. Two new versions have been designed and built. Version II is essentially similar to Version I and uses the same objective lens; it is expected to produce the 1-mil spot over the central 4 inches. The optical system is shown schematically in Fig. XVIII-1. Version III is based on a specially designed objective lens, which will achieve a 1-mil spot over the entire 14-inch width; it uses a separate collimating lens (i.e., single pass through the objective).

b. A different make and type of internally modulated laser is being used, which includes feedback control of light intensity as part of the package.

c. An entirely new light-collecting system has been designed, by using a specially made cylindrical reflector that images the objective aperture onto a photodiode, which
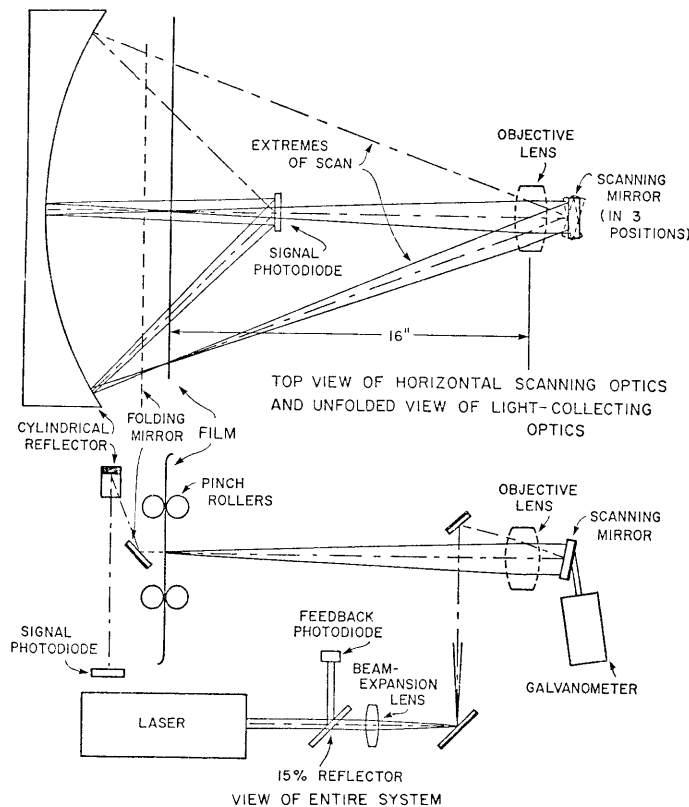
Fig. XVIII-1. Redesigned optical system.

allows the diode to be much smaller and decreases overall light loss. The photodiode itself is PIN-type, 1 inch in diameter, with low capacitance and low dark current. The overall effect is a much higher signal-to-noise ratio, thereby enabling scanning greater densities; we expect good 3.5 o.d.u. reproduction.

d. The stepping motor drive has been changed to provide variable film motion steps down to 1 mil.

e. A special attachment to the scanner has been designed to handle cineradiographic 35 mm films; it allows accurate frame-to-frame registration.

c. Present Status

A scanner with Version II optical system and a cine attachment has been built and is to be loaned to the Radiology Department of the Peter Bent Brigham Hospital, in Boston. It will be connected there to a PDP-11 computer and will serve as the input device in various projects of research into computer-aided diagnosis. A reproducer with Version III optical system has been built and tested. Its computer interface is being built. The two units have been tested together on direct transmission, as well as on transmission via a DDS system simulator.

Work is under way to build another scanner, this one with Version III optical system. This scanner and the reproducer will stay at the Research Laboratory of Electronics as computer facsimile input and output devices, and will be useful for critical evaluation of transmitted radiographic quality.

7. SOURCE CODING FOR X-RAY PICTURES

Murat Kunt

New technological developments in electronic devices make the idea of storing visual images in electronic form reasonable. The laser-scanner-receiver unit described in Section XVIII-6 enables us to transform a high-resolution, large dynamic range picture such as an x-ray into an electric signal. A typical 14" × 17" x-ray picture scanned at a rate of 250 points per inch and at 256 gray levels requires ~10^8 bits, a rather large number. Therefore, it is useful to inquire whether some data-reduction techniques can be used. The purpose of this study is to investigate the feasibility of various coding schemes for compact storage or rapid transmission of radiographic images.

We have investigated three specific coding schemes. The first is the so-called synthetic high system developed previously by William F. Schreiber. This elegant technique takes advantage of the properties of the human visual system and the two-dimensional nature of images. The synthetic high system is one of the most powerful information-lossy techniques for redundancy reduction in gray-level pictures. An average compression ratio of 7.5:1 was obtained in 4 test pictures. But the large amount of computation required in this scheme makes it impractical for real-time applications with present technology.

The second technique, called block coding, was used by the author for two-level facsimile pictures. In this study, it is generalized for encoding bit planes in gray-level pictures. This technique is not as efficient as the synthetic high system, but its simplicity makes it suitable for real-time applications and for economical hardware implementations. According to the value of a preset parameter, this method can be either information-lossless or information-lossy. In the latter case, distortions are like impulse noise affecting small isolated areas of the picture and their effect can be reduced easily by appropriate filtering. The average compression ratio is approximately 2.5:1 in the information lossless case and can be increased up to 5.5:1 by information-lossy coding. The dependance of the compression ratio on image resolution was investigated. If the resolution is doubled, an average relative increase of 50% of the compression ratio can be obtained. Furthermore, with a slight increase in complexity, block coding can be made adaptive according to the local "activity" in bit planes.

The third method is the well-known run-length coding, which can be applied also to bit planes. This technique is information-lossless. Based on the statistical analysis of run lengths in the bit planes, the exponential run-length code is used. This code is a fixed-word-length code and leads to an average compression ratio of 2.5:1. The dependence of the compression ratio on image resolution was also investigated. If the resolution is doubled, the average relative increase of the compression ratio is approximately 70%. The exponential run-length code can also be made adaptive according to the statistics of each bit plane, and leads to a relative increase of 10% of the compression ratio.

8. STUDY OF THE RADIOLOGICAL DIAGNOSTIC PROCESS

Barry A. Blesser, David M. Ozonoff, Murray Eden

The aim of this research project is to study the process by which radiologists make their findings from radiographic images. In this way we hope to discover those aspects of the process for which performance can be improved by appropriate image-enhancement techniques.

Work in collaboration with the Radiology Department of the Peter Bent Brigham Hospital, Boston, continues. Results of a five-way comparison of independent readings of the same set of 100 chest x-rays were reduced to a form compatible with computer manipulation of the data base, and preliminary analysis is now being carried out. Data from an investigation of the biasing effects of clinical history are now being collected.

Baseline information on the radiologic diagnostic process gained in this work will be useful in evaluating technological intervention in a clinical setting.

9. STOCHASTIC MODELING OF PARTIALLY DENATURED DNA MOLECULES

Ian T. Young, Donald Levinstone, Murray Eden

Since the Watson-Crick model for the structure of the DNA molecule and the associated concept of nucleotide base-pairing were discovered, one of the fundamental problems in molecular biology has been the decoding of the DNA nucleotide sequence. A powerful procedure applied to solve that problem has been analysis of the products from partial denaturation of DNA molecules. We have recently developed a technique for automated computer analysis of such partially denatured molecules. We use classical signal-processing techniques to provide a low-resolution picture of the DNA double helix reflecting the local base concentrations (A:T/G:C). Alkali- or formamide-induced denaturation of a double-stranded molecule proceeds most rapidly in the regions of high double-hydrogen-bond concentrations (A:T rich regions). The distribution of denatured regions is observed and recorded from electron-photomicrographs of individual molecules. The partial denaturation map of a molecule thus obtained is a fingerprint of genetic information encoded in DNA.

We have applied our signal-processing methodology to study the mechanism of DNA encapsulation by the temperature-sensitive Salmonella phage P22, and have results that support a sequential encapsulation model. Our main objective is to develop a method for taking random fragments of a large (e.g., yeast chromosome) DNA molecule and aligning the partial denaturation maps of these fragments by computer to reconstruct a partial denaturation map of the entire structure. The high correlation found between aligned partial denaturation maps of P22 viral DNA lends support to this approach.

The biological aspects of this research were performed in the laboratory of Professor David Botstein, Department of Biology, M.I.T.

10. APPLICATION OF PATTERN RECOGNITION TECHNIQUES TO
    MEASUREMENT OF HISTOCOMPATIBILITY

Stephanie E. Sher, Allison G. Hagerich, Murray Eden

A cooperative research program with the tissue typing laboratory at Peter Bent Brigham Hospital, Boston, was begun in July 1975 to devise a new measurement of histocompatibility that requires less time than the standard mixed lymphocyte culture (MLC) assay. The standard MLC assay is performed by measuring the uptake of tritiated thymidine by proliferating lymphocytes in tissue culture. This assay requires from five to seven days to perform. Cadaver kidneys can be preserved in vitro for a maximum of three days. Therefore, in order to provide histocompatibility testing for transplants involving cadaver kidneys, the assay must be modified to give results in less than three days.

Our approach involves applying computer techniques for picture processing to photomicrographs of mixed lymphocyte cultures. We are using two-dimensional Fourier transforms and run-length histograms to analyze textural differences observed in the MLC. In the past six months we have identified culture methods that allow discrimination between histocompatible and histo-incompatible cultures after 24 hours of incubation. These picture-processing techniques are being employed to correlate textural differences with histocompatibility differences in order to identify the best possible transplant match.

11. ELECTRON MICROGRAPH PHASE-PROCESSING

Ralph W. Gerchberg

In 1971, Gerchberg and Saxton showed that it was possible to achieve a solution to the phase problem of electron microscopy by using pictures taken in the Gaussian image plane and in the back focal (diffraction) plane of the objective lens. The processing scheme was embodied in a computer algorithm which proved capable of reducing a defined error whose value was zero when an acceptable solution was realized. Electron microscopy data on an organic crystal of catalase was successfully processed to a resolution of 30 Å and it was noted that the algorithm produced an acceptable solution only utilizing data from the back focal plane and no real image data. This observation bears not only on the phase problem of electron microscopy but also on that of x-ray crystallography. At present, this aspect of the problem is being pursued. Some positive results have been achieved and we hope that electron diffraction data now taken on unstained wet catalase crystals (supplied by Roswell Park Memorial Institute) to 2 Å resolution can be processed successfully by the end of this project.

12. OBJECTIVE VISUAL FIELD PLOTTING

National Institutes of Health (Biomedical Sciences Support Grant)

Marvin E. Jernigan

Information about a patient's visual field has long been acknowledged to be important for the diagnosis of visual and neurologic pathology, specifically of glaucoma and lesions affecting the visual pathways in the brain.

The visual field plotting techniques employed in most clinics suffer from problems of insensitivity and unreliability.[1] A major source of error is the necessity for manual or verbal subjective responses from the patient, to indicate perception of the test stimulus. A technique for plotting visual fields based on the analysis of eye movement response has been developed. This technique encourages normal eye movement responses and distinguishes between responses indicative of perceived stimuli and those indicative of no perception. Rather than relying on simple subjective response, an automatic decision is made based on the nature of the eye movement response. Results indicate a significant improvement in the detection of early visual-field defects. We have shown that aspects of the eye movement response signal, revealed by appropriate processing, provide a more reliable indication of stimulus visibility than the unqualified acceptance of the subject's verbalized response.[2]

The present research activity is the development and clinical evaluation of a practical visual field plotting instrument. Ultimately it will be available for commercial production and distribution. When used as part of a visual screening examination, the instrument will facilitate early diagnosis of glaucoma, allowing reversal of visual field damage and the prevention of further loss of visual function.

Clinical evaluation, in cooperation with opthalmologists in the Institute for Computer Sciences, Mt. Sinai School of Medicine, New York, is planned. Comparison will be made between visual field plots obtained with the eye movement response scheme and plots obtained with traditional clinical techniques. Recommendations based on the results will be made with respect to instrument design and clinical procedures to achieve sensitive and reliable performance in the clinical setting.

References

1. P. Nisel, "Streuungen perimetrischer untersuchungsergebnisse" (Scatter of results of perimetric studies), Opthalmologica 161, 180-186 (1970).

2. M. E. Jernigan, "Eye Movement Analysis in Plotting the Visual Field," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., June 1975.

13. DIGITAL WIREPHOTO SYSTEM

Associated Press (Grant)

Donald E. Troxel

Since August 1970 we have been developing an entirely new news picture (Wirephoto) distribution system 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 gradually spreading in time and location.

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 generally will originate and terminate 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. During the coming year we shall prepare for the first installation at a regional bureau, as well as continue research into methods of picture quality enhancement.

14. COLOR PICTURE CODING FOR FACSIMILE

Robert D. Solomon

A fundamental color coding system suitable for facsimile transmission has been developed, based upon practical engineering and psychophysical constraints. This system achieves information compression by first accurately transforming the measured red, green, and blue reflectances at each picture element (pel) into a luminance and two CIE UCS chromaticity components that are linear in the sensation domain. The chromaticity components are spatially filtered, coarse sampled, and coarse quantized. The compressed chromaticity components are then transmitted and linearly interpolated at the receiver and combined with the luminance to yield red, green, and blue signals, which are further processed to correct for the nonlinear photographic process.

The processing operations of color scanning, chromaticity quantization, spatial filtering and coarse sampling, and prereproduction color correction have been examined in detail to optimize their individual and interactive performance characteristics. Color pictures processed by the system proved in subjective tests to be of excellent quality compared with the original. They contain, however, only 1/16 the chromaticity information and 37% of the total information of the original. Thus the encoded picture can be transmitted in one-third the time by multiplexing the chromaticity information in the unused 11% scan-blanking region of a conventional facsimile signal.

XVIII. COGNITIVE INFORMATION PROCESSING

A. COLOR PICTURE CODING FOR FACSIMILE

Robert D. Solomon

1. Color Facsimile System Constraints

Since the infancy of facsimile there have been few innovations in the transmission of color pictures. Color photographs are still sent by successively scanning a color print three times with a conventional rotating drum transmitter. This scan requires separate transmission for the red, green, and blue components. In view of recent advances in color reproduction (e. g. , color office copying equipment), it is desirable to encode a color facsimile signal so that it can be sent in a single picture scan. Furthermore, the color facsimile signal should be compatible with the numerous existing black-and-white facsimile receivers, as is the case with color television systems.

The telephone bandwidth used for facsimile is very narrow relative to the wide television video bandwidth. Thus it is difficult in a facsimile system to use the frequency multiplexed coding techniques of color television, which encode the compressed chrominance components onto the standard luminance video signal. Only one frame of a picture is transmitted in facsimile, and hence the scanning is synchronized by a single pulse at the start of the picture. Most scanning systems have a blanking period in the direction of scan to enable the scanning beam to retrace or to allow a region of the rotating drum to contain a clip to secure the picture. Since facsimile does not require the elaborate television horizontal synchronization signal, the scan blanking region, which is 6-12% of the duration of the scan, can contain the compressed chrominance information.

2. Choice of Luminance and Chrominance Variables

Numerous operations are required to insure high-quality color reproduction and achieve the required chrominance compression. The red, green, and blue (R, G, B) sample values at each picture element (pel) of the scanned print must be converted to a luminance, L, and two independent chromaticity components, C1 and C2. (See Fig. XVIII-2.) The luminance is derived as a simple linear combination of R, G, B where the luminosity coefficients L_r, L_g, L_b are determined from the least-mean-square fit of the R, G, B detector spectral response functions to the photopic luminous efficiency function[1] $L = L_r R + L_g G + L_b B$. Quantization of the L, C1, C2 components will be most efficient if equal quantum steps along the coordinates correspond to nearly equal changes in visual sensation. In the case of luminance, a modified logarithmic function seems to distribute perceived picture noise and contour artifacts evenly over the entire gray
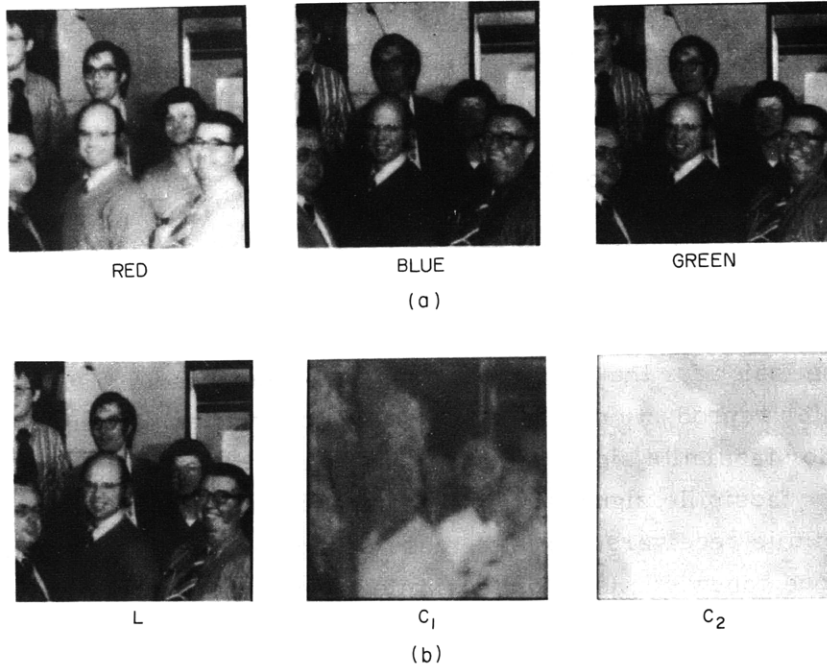
Fig. XVIII-2. (a) R, G, B separations of group picture.
(b) L, Cl, C2 components of group picture.

scale: $\hat{L} = [\log (1+aL)]/[\log (1+a)]$. The CIE (1960) Uniform Chromaticity Scale (UCS)
transform yields chromaticity components closely approximating a linear sensation
space. The transform is a bilinear function of R, G, B, yielding the chromaticity coor-
dinates u and v.

$$u = (u_r R + u_g G + u_b B)/S \qquad v = (v_r R + v_g G + v_b B)/S$$

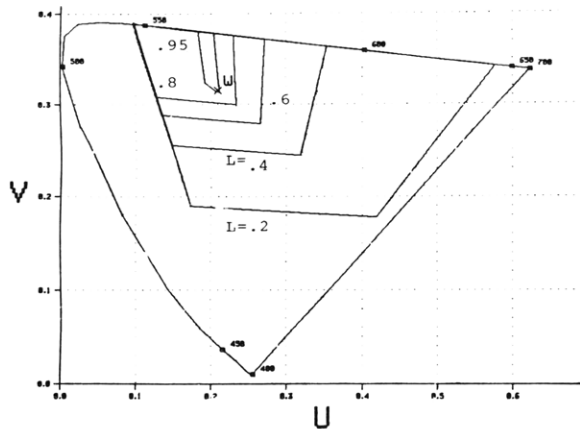where $S = S_r R + S_g G + S_b B$. In transforming the R, G, B components (Fig. XVIII-2a)

Fig. XVIII-3.

Loci of constant luminance on
UCS chromaticity diagram.

of a picture to the luminance and two chrominance components representation (Fig. XVIII-3), the correlation between pictures is greatly reduced.

3. Chrominance Spatial Compression

The ability to compress the color components of a picture comes primarily from the eye's lower spatial acuity for chrominance detail relative to luminance acuity.[2] In color television, because of this psychophysical property of the eye,[3] the chromaticity components are one-dimensionally lowpass filtered. Earlier research has shown that the chrominance information in a color picture can be greatly compressed without unacceptable subjective degradation, by two-dimensional spatial filtering and coarse sampling of the chromaticity components.[4] The author has succeeded in compressing the two chrominance components by a factor of 16:1 without compromising subjective picture quality.[5] Thus, based on the psychophysical properties of the human eye, it is possible to encode the compressed chrominance into an 11.1% blanking region, where the active scan region contains the full resolution luminance information that is compatible with monochrome facsimile receivers.

4. Chrominance Quantization

A method for greatly improving the efficiency of quantizing the chrominance components, based on the deterministic interdependence of luminance and chrominance information in any reproduction system, has been developed.[6] If the two chrominance components are digitized, then the chrominance plane is divided into a grid of boxes (Fig. XVIII-3). At luminance values below the primary blue luminance, all of these chromaticities within the triangle formed by the primaries can be reproduced. As the luminance increases, however, the gamut of synthesizable chromaticities decreases until the maximum luminance value is reached where only the achromatic point can be reproduced. Thus the luminance component supplies chromaticity information; for example, it is strongly (negatively) correlated with the relative blue component. The luminance scaled chromaticity (LSC) transform is derived by determining the minimum and maximum values of the two (u, v) chromaticity components as a function of luminance. The region of attainable chromaticities can thus be scaled and translated appropriately so that the number of reproducible chromaticities remains nearly constant at all luminance levels (Fig. XVIII-4). Therefore, it is not necessary to have an excessively fine chromaticity quantization to insure good color rendition of frequently encountered and critical pastels such as light flesh tones. By using the LSC transform, the chromaticity components can be digitized to 4 bits each with no objectionable chrominance-contouring or color errors.
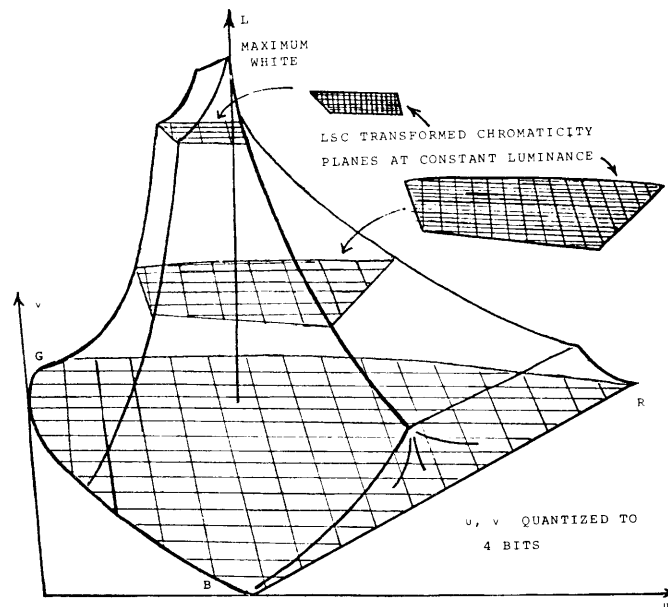
Fig. XVIII-4. Gamut of reproducible colors for the CIE UCS chro-
maticity coordinates as a function of luminance.

5. Color-Coding System

Chrominance compression not only necessitates the complex operations of two-
dimensional filtering and interpolation but also requires conversion to and from
photographic and psychophysical (e. g., chromaticity components) variables. At the
transmitter (Fig. XVIII-5) the red, green, and blue signals from a color scanner must
somehow approximate the human eye's visual responses. At the receiver (Fig. XVIII-6)
the luminance and chrominance components must be transformed and processed to yield
R, G, B signals that precorrect for the nonlinearities and chromatic distortions of the
photographic dye or printing ink reproduction process.

There are further system constraints in addition to the requirement of compressing
the chrominance information by a factor of 16:1 and packing it into the 11.1% scan
blanking region. Picture-coding operations may be classified according to the portion
of the picture that is operated on at one time. At one extreme are transform coding
processors (e. g., Fourier and Hadamard) that necessitate storing the entire picture in
a random access memory (RAM).[7] Digitizing each of the R, G, and B components of a
standard 8" X 10" color picture to 6 bits with a facsimile resolution of 100 pels/inch
results in more than 10^6 bits of storage. Such processors are not feasible from an
economic standpoint, and they cause severe delays between transmission and reception
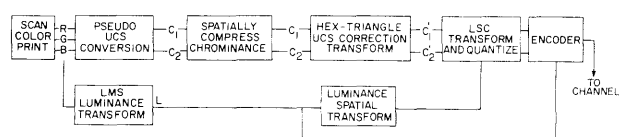
Fig. XVIII-5.

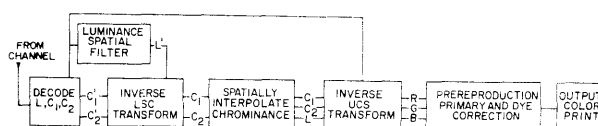Color facsimile transmitter encoder.

Fig. XVIII-6.

Color facsimile receiver decoder.

of the picture. A "hot news" picture must be scanned first for 8 minutes, then processed for at least another 8 minutes, and finally transmitted. Thus there is a threefold delay in effective transmission time that offsets the gains of color compression.

The simplest type of processor operates only once on a single pel, which is then transmitted immediately. This limited operation precludes linear filtering and inter-polation in two dimensions, which are crucial for effective color compression. An inter-mediate solution which effectively permits real-time transmission (i.e., the picture is reproduced at the receiver as it is scanned at the transmitter) is to store and process a small group of lines. The resulting transmission delay of several lines is negligible for a facsimile scan rate of 100 lines per minute.

The coding algorithms must be designed to process pels in the 500-μs intersample periods of the scanner. Since the entire picture is not stored, the window processor alone cannot use adaptive coding picture-processing techniques (e.g., Karhunen-Loève statistical vector basis) and so the coding algorithms must be designed to reproduce a wide variety of color pictures equally well.

The interpel sample period for the existing facsimile system is 500 μs (2 kHz tele-phone bandwidth) or a few thousand times the corresponding period for television. There-fore, in addition to the operations of sampling, filtering, quantizing, and packing there is sufficient time to perform complex chrominance transforms with a high-speed digi-tal processor. Color coders for television systems cannot afford this computational luxury, but color television is an additive primary system, and hence does not need many of the chrominance transforms that correct for the subtractive dye systems encountered in facsimile.

Real-time stream processing of the data restricts the possible ways in which the lowpass chrominance spatial-filtering operation can be implemented. The simplest way to compute such a function is with a superposition filter, since it results in the least number of computations. As each chromaticity component is calculated, it is weighted and accumulated. Coarse sampling is achieved by shifting the superposition filter every 4 pels and every 4 lines and storing the accumulated value, thereby compressing the data by a factor of 4 in each direction.

At the receiver, the chromaticity components are unpacked, inverse LSC trans-formed, two-dimensionally linearly interpolated, and inverse UCS transformed to yield

the CIE R, G, B values. These R, G, B coordinates are related to the color response functions of the eye and must be transformed to a new primary system where the R, G, B take on only positive values. Moreover, the new R, G, B values must be further processed in the logarithmic density domain to yield R, G, B exposure values that are precorrected for the nonideal dye or ink reproduction process.[8-10]

Color pictures processed by the facsimile encoder and decoder appear to be subjectively equivalent even though the processed picture contains only 1/16 of the chrominance information of the original and 37% of the total information.

References

1. R. D. Solomon, "A Color Facsimile Optics Head Using Solid-State Detectors," a paper presented at the Society of Photographic Societies and Engineers 27th Conference, Boston, Massachusetts, April 30, 1974.

2. W. E. Middleton and M. C. Holmes, "The Apparent Colors of Surfaces of Small Substance — A Preliminary Report," J. Opt. Soc. Am. 39, 582-592 (1949).

3. G. H. Brown, "The Choice of Axes and Bandwidths for the Chrominance Signals in NTSC Color Television," Proc. IRE 42, 58-59 (1954).

4. U. F. Gronemann, "Coding Color Pictures," Technical Report 422, Research Laboratory of Electronics, M. I. T., June 17, 1964.

5. R. D. Solomon, "Color Coding for a Facsimile System," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., August 1975.

6. R. D. Solomon, "The Luminance Scaled Chromaticity Transform," 1975 Annual Meeting, Optical Society of America, Boston, Massachusetts, October 21-24, 1975.

7. W. K. Pratt, "Spatial Transform Coding of Color Images," IEEE Trans. Vol. Com-19, No. 6, pp. 980-992, December 1971.

8. D. L. MacAdam, "Subtractive Color Mixture and Color Reproduction," J. Opt. Soc. Am. 28, 466-480 (1938).

9. J. A. C. Yule, Principles of Color Reproduction (John Wiley and Sons, Inc., New York, 1967).

10. R. D. Solomon, "Real-Time Color Picture Coding," a paper presented at the 1976 Picture Coding Symposium, Asilomar, California, January 28, 1976.