25. Cognitive Information Processing

Advanced Television Research Program

Academic and Research Staff

Prof. W.F. Schreiber[29], Prof. D.H. Staelin, Prof. J. Lim, Prof. S. Benton, Prof. J. Lang, J. Ward

Graduate Students

J. Bernstein, P. Cassereau, T.M. Chen, K. Christian, B. Hinman, M. Holzbach, S. Hsu, M. Isnardi, A. Kirkland, E. Krause, C. Lee, S. Marshall, H. MacDonald, Henrique S. Malvar, D. Martinez, D. Pan, D. Pian, G.N.S. Prasanna, G. Saussy, J.-P. Schott, B. Szabo, A. Tom, J. Wang

Sponsorship

Members of the Center for Advanced Television Studies, an industry group consisting of American Broadcasting Company, Ampex Corporation, Columbia Broadcasting Systems, Harris Corporation, Home Box Office, Public Broadcasting Service, National Broadcasting Company, RCA Corporation, Tektronix, and 3M Company

25.1 Goals

The purpose of this program is to conduct research relevant to the improvement of broadcast television systems. Some of the work is carried out in the Research Laboratory of Electronics and some in other parts of M.I.T., including the Media Laboratory. Audience research is carried out by Prof. W.R. Neuman in the Political Science Department. Reporting is by means of theses and published papers.

25.2 Background

The Japan Broadcasting Company (NHK) has demonstrated a high definition television system of 1125 lines, 30 frames, 60 fields, 25 MHz bandwidth, with image quality comparable to 35 mm motion pictures. Substantial improvements in image quality over that of the existing NTSC and PAL systems have been demonstrated by laboratories in Europe, Japan, and the United States, which require only signal processing and the use of special electronic components, such as frame stores, at the receiver. These systems do not require increasing the present 6 MHz bandwidth. Still other systems have been demonstrated that achieve nearly NHK quality by

---

[29] Director

adding a second channel to the present broadcast signal. In view of all these developments and of the economic importance of the U.S. television industry, it was deemed appropriate by the sponsors to fund a research program at an American university.

25.3 Research Activities in RLE

So far, research has concentrated on fundamentals of signal processing with emphasis on the special case where the signals are functions of two space variables and time. Motion-adaptive temporal interpolation has been discussed in two theses by Krause[1] and Hinman,[2] adaptive sharpening of color pictures by Marshall,[3] a new transform for signal processing by Bernstein,[4] quantifying the relationship between TV bandwidth and area of motion picture film by MacDonald,[5] and a study of motion perception was carried out by Hsu.[6]

A paper published by Schreiber[7] discusses some psychophysical and signal processing aspects of TV design and another was presented by Hinman, Bernstein, and Staelin[8] at the International Conference on Acoustics, Speech and Signal Processing on the short-space Fourier Transform. Further studies are underway on color and motion perception, and on certain aspects of three-dimensional signal processing. A programmable three-dimensional interpolator is under development for the up-conversion of digital TV signals to higher line and frame rates. A new computer system, located in the Media Laboratory, is being assembled for the simulation of moving high definition TV images. A system capable of performing digital experiments for NTSC video image sequences is also being developed.

References

1. E. Krause, "Motion Estimation and Interpolation in Time-Varying Imagery." M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., September 1984.
2. B. Hinman, "Theory and Application of Image Motion Estimation," M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., May 1984.
3. S. Marshall, "Color Image Enhancement," M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., August 1984.
4. J. Bernstein, "Properties and Applications of Short-Space Fourier Transform," M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., May 1984.
5. H. MacDonald, "The Relative Quality of Television and Motion Pictures," M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., January 1985.
6. S.C. Hsu, "Motion-Induced Degradations of Temporally Sampled Images," M.S. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., June 1985
7. W.F. Schreiber, "Psychophysics and the Improvement of Television Image Quality," JSMPTE 93, 8, 717-725, August 1984.
8. B. Hinman, J. Bernstein, and D. Staelin, "Short-Space Fourier Transform Image Processing," Proceedings of the International Conference on Acoustics, Speech and Signal Processing, March 1984.

Computer Graphics and Image Processing

Academic and Research Staff

Prof. W.F. Schreiber, Prof. D.E. Troxel, Dr. C.W. Lynn, Dr. L. Picard, Dr. K.P. Wacks, Dr. Yao-Ming Chao, Dr. M. Khan, Qiding Du, S. Corteselli, J. Ward

Graduate Students

J. Allen, D. Chan, D. Clemens, R. Fortino, A. Garcia, G. Hamilton, S. Hsu, Shiao-Li Huang, M. Isnardi, E. Krause, C. Lee, H. MacDonald, S. Marshall, D. Pan, D. Pian, G.N.S. Prasanna, J. Sara, G. Saussy, J.P. Schott, R. Ulichney, G. Vachon, R. Vinciguerra, J. Wang

25.4 Picture Coding

A. Sampling and Reconstruction

Most present-day picture coding is digital; virtually all input and output images are analog. Thus the conversion from one form to the other is everpresent and important to the quality and efficiency of the overall process. To elucidate the phenomena involved, a systematic study was carried out on a wide variety of pre-sampling and interpolation filters. It was found that, as suspected, the "ideal" low-pass filter was far from best, subjectively, since the least rms error criterion is not valid for human observers. Aliasing resulting from "violation" of the sampling theorem is visually traded off against sharpness and visibility of the sampling structure which occurs with other filters. Several filters studied performed substantially better than the ILPF. The best combination of performance and computational simplicity was found in the "sharpened Gaussian" filter, which has an impulse response not unlike that of human vision.[1]

B. Differential Pulse Code Modulation

This most common of image data compression systems has been widely studied. Its performance depends in marked degree on the characteristics of the signal source, since the presence of very sharp transitions requires an unfavorable trade-off between slope overload and the granular noise which occurs in blank areas. Nevertheless, when this factor is carefully taken into account, we have gotten good results with a form of DPCM in which the companding characteristic is adaptively adjusted to local image properties. In speech coding effective adaptation is possible based only on the coded signal. Because of the isotropic character of vision and images, this is not possible — extra data must be transmitted. Even including this extra data, we have obtained nearly "original" quality at 2.2 bits/pel, non-statistical, and 1.5 bits/pel statistical.[2]

C. Two-channel Coding System

For a number of years we have studied a picture transmission system in which the signal is divided into two channels, a two-dimensional low-pass signal ("lows") which is coarsely sampled and finely quantized, plus the remainder ("highs") which is finely sampled and coarsely quantized with the aid of a tapered randomized quantizer.[3] While this can be thought of as a crude form of transform coding, the two- or three-channel approach permits tailoring the separate channel coding parameters to well-known human visual properties. While the performance of the non-adaptive version of this system is not as good as the adaptive DPCM system mentioned above, it does feature rather simple implementation and excellent (PCM-like) performance in the presence of channel errors. Under a grant from the Sony Corporation, a real-time hardware system was implemented.[4]

D. Adaptive Two-channel Color Coding System

The above-mentioned system can be substantially improved by adapting the highs quantization to the local signal properties. Blank-area signal-to-noise ratio of 50 db at about 3 bits/pel is possible in most cases. The system is extended to color coding by transmitting a three-color lows signal plus an adaptively-companded achromatic highs signal. The color lows signal is further compressed by mapping to a perceptually uniform color space, akin to the Munsell system of subjective color notation. Excellent quality full color images are possible at four bits/pel, compared with 24 for the uncoded signal. Further compression is possible by statistical (entropy) coding of the two-channel coder output.[5]

25.5 Data Processing for the Graphic Arts

Providence Gravure Co. (Grant)

An eight year program of research and development directed towards prepress data processing was completed in August 1984. This project had as its main goal the application of modern methods to the solution of an important problem in the printing industry — namely the production of final printing surfaces (printing plates) given the mass of text, graphics, pictures, and instructions which are supplied to the printer by the customer. The hoped-for advantages of the use of modern methods include higher speed, lower cost, and better and more consistent quality. An important aspect is the elimination of all photographic steps in between the original copy and the final plates. In this case, printing is by the rotogravure process, with printing cylinders mechanically engraved on the Hell Klischograph, a kind of enormous rotating drum facsimile machine. The copy to be printed is normally scanned on a companion scanning machine, the diamond-tipped engraving heads being operated by analog video signals derived from the scanner. The monochrome aspects of the system which was developed have been described in a published paper.[6] The system automates page composition, eliminating all of the laborious cut-and-paste operations of conventional systems. This report concerns the color aspects of the process.[7]

Conventional prepress color processing

Color pictures to be printed by gravure are normally supplied as transparencies or prints, typically of good quality and often of rather large size. In the conventional process, these originals are scanned at high resolution on a rotating drum color scanner, producing four continuous tone color separations. The separations eventually control the amount of the four printing inks — yellow, magenta, cyan, and black — delivered to the final page. For use in gravure, the separations are made in the final size on the scanner and are then used to make black and white films. These films are assembled along with the other copy intended for each page and rephotographed to make the full-page pictures to be mounted on the scanning machine. Often, the color pictures do not turn out completely as desired and must be made over — in some cases rescanned by the color scanner. In virtually all cases, a great deal of painstaking hand work must be done at the page composition stage as well as in the processing of individual pictures to get all the colors right. The operation of the color scanner requires a great deal of skill and experience, since the operator must correct for the action of the particular inks and papers being used and also must make any desired editorial changes in the color and/or tone rendition of the pictures.

Computer-based prepress processing

Several commercial computer-based systems are available to automate the process. The three most popular, Hell (West Germany), Crosfield (United Kingdom), and Scitex (Israel), all use the output of a conventional color scanner to make the major ink and paper corrections. A system being developed by Eikonix (Bedford, Massachusetts) but not yet in widespread use, does color correction in the computer and is intended to be used with simpler and cheaper scanners. As useful as these systems are, a missing link is a sufficiently accurate TV display ("soft proof") which predicts the appearance of the printed page. The combination of a soft proof accurate enough to be relied on, together with easily understood color editing capabilities, are the keys to getting the full advantage of computerization, and therefore were the main topics of investigation.

Color metrics

Perceived color has three degrees of freedom, in that a very wide gamut of colors can be produced by a combination of three (additive) primary lights or three (subtractive) colored inks or dyes. In addition, the <u>appearance</u> of a color, in the sense that it matches some other color under a fixed set of viewing conditions, is describable by a set of three "tristimulus values" which are proportional to the intensities of three particular primary lights. Color descriptions in terms of additive lights are linearly transformable from one set of primaries to another. The science of colorimetry is concerned with such transformations, which can be performed accurately by totally objective methods. Color descriptions in terms of inks, where four or more are often used, are much more difficult to manipulate and are best dealt with empirically by means of lookup tables (LUT's). Conventional color prepress systems invariably operate in terms of ink densities. To be sure of the appearance of a color to be printed, the operator looks up the ink density combination in a book of samples.

Interactive color editing

A key element in the system is the TV display, which serves as a "soft proof", and its companion "color translation module" (CTM) with which the editor makes desired changes in image appearance. This display is a colorimetrically accurate representation of the RGB signal values held in memory, and is independent of the ink, paper, or printing process to be used. An important factor in using the TV is controlling the state of adaptation of the viewer's visual system. This was accomplished by surrounding the TV and its companion comparison display of the original image, by a generous surround of reference white. To make the display even less affected by the room illumination, it would have been useful for the tube to be several times brighter.

The CTM performs digital processing on 512x512 digital video images interactively. The controls available to the operator are exclusively in terms of image appearance, and as a consequence, their operation can be learned very easily and quickly by persons without previous graphic arts experience. They permit changes in tone scale (contrast, brightness, etc), overall color saturation, and independent control of luminance, hue, and saturation of the separate colors. Local color control (i.e., of one object), if not possible on a color-selective basis, can be done by first drawing a contour which separates that object from other objects of similar color, but which generally does not have to be an accurate outline. When the TV resolution image has been adjusted to satisfaction, the high resolution disk image (typically up to 2000x3000 pels) is run through the same hardware and returned to disk storage.

Ink correction

The TV, as mentioned, is simply a colorimetrically accurate display of the stored image. In order that the display accurately predict the appearance of the page, the ink densities are chosen so that the page colors have the same tristimulus values as the TV. This is the opposite of the usual arrangement in which the TV is made to look like the page by a (complicated) transformation. That arrangement means that the TV calibration is different for each ink and paper and must be done by the user. In our arrangement, the TV calibration (it must, of course, be calibrated for tone scale linearity and white point) is fixed and done by the manufacturer. Correction for the inks requires a LUT whose addresses are RGB values (i.e., the displayed intensities) and whose contents are ink densities. The LUT contents are calculated by a computer program, using a data base obtained by printing a set of color patches and measuring their RGB values with a colorimeter. The color patches comprise all combinations of a number of density steps (typically 4 to 10) of each ink. The entire process of making and using the LUT is objective, except for nonprintable colors.

The conversion from RGB, or appearance representation, of colored images to ink densities for printing, has several major complications. A minor problem is that certain colors, particularly some brilliant yellows, cannot be produced on a TV display. In this case some operator judgment is called for. A more serious problem is that some of the colors of original transparencies cannot be reproduced by even the best inks and in virtually all cases, the dynamic range of the printed page is less than desirable. We dealt with this problem by causing the display screen to flicker whenever a nonreproducible color is displayed. This function is performed by the excess gamut alarm (EGA), a truncated version of the RGB-to-ink density LUT. The operator then attempts to use the controls of the CTM to reduce or eliminate the nonreproducible colors while at the same time maintaining good picture quality. The remaining out-of-range colors are dealt with in the LUT by any one of a number of fairly simple algorithms.

A final complication concerns the use of black ink. Black is used primarily to save money, in that one part of black replaces three parts of colored ink, and the black ink is much cheaper. The replacement of colored ink by black, termed "undercolor removal" (UCR), is difficult to do accurately because of the nonideal nature of printing inks. In this project a process was successfully developed for 100% UCR with virtually no change in appearance from the printed version which used only colored inks.

System description

The process in divided into two major portions. In the scanning process, color pictures are scanned, digitized, processed as necessary, and stored in a large disk memory in a form which accurately represents the size and colors of each of the complete final pages. This requires, of course, combining all of the material to appear on each of the pages, an operation called page composition. In the engraving process, this data is retrieved from the disk and is used to drive the engraving machine. Because the engraved cylinder contains many pages, the engraving process must combine the pages in the desired manner (this process is called imposition) and this is done on-the-fly while engraving. Composition and imposition are discussed in reference.[6] The output process must also convert the stored data, which is in color appearance form, into actual ink densities. This is done with the aid of the LUT whose contents have previously been calculated from the results of a large color patch printing test.

The layout of the system is shown in Fig. 25-1. A color scanner measures the RGB values of the original, but does not do any color correction. These signals are digitized in the analog-to-digital converter and processed as necessary by Table 1 to convert to accurate tristimulus values. (This step may be omitted. but more operator intervention is then required.) The full resolution image goes to the disk and a 512x512 version to TV memory and is displayed. The operator then uses the CTM, conditioned by the EGA, to edit the image. When satisfied, the full resolution image is retrieved from the disk, processed by the CTM hardware, passed to the page composition and coding module, and returned to disk storage. Color information at this point is preferably in

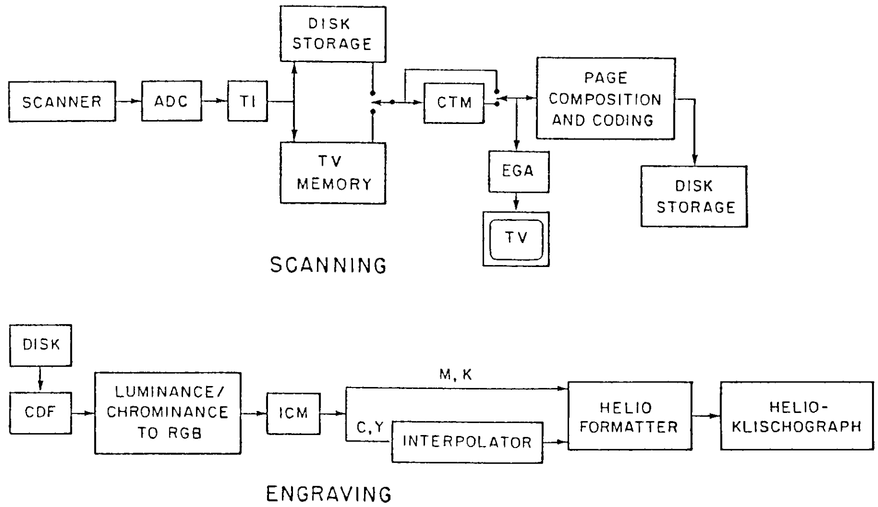

Figure 25-1:

luminance-chrominance form (as in the NTSC color TV system) where the chrominance components are at lower resolution than luminance.

After all the pages for a cylinder have been processed, engraving can be performed. The page information is retrieved from the disk in the order in which it is needed by the cylinder, and organized into appropriate format by the color data formatter (CDF). It is then converted back to full resolution RGB form and passed to the ink correction module (ICM) which includes the LUT and other processing circuitry required to obtain the ink densities. In this particular engraving process, the "raster" (corresponding to the halftone screen of offset printing) for cyan and yellow is different from that for magenta and black, necessitating an interpolation process when the yellow and cyan cylinders are engraved. Finally, the data goes to the Helio formatter where it is converted into analog form for driving the engraving machine.

Conclusion

Virtually all of the hardware, software, and processing procedures described here were completed and operated successfully. Although the final system was not put into commercial production by the sponsor except for some trials, our belief is that the principles of operation, which are at great variance from general practice in the field, were demonstrated to be both feasible and economical, and are likely to be adopted in great part by the printing industry in the future.

25.6 Computer Graphics Architectures

International Business Machines, Inc.

We are investigating architectures suitable for high resolution, high speed interactive graphics displays. We consider cost-performance trade-offs both for today's costs and those projected for the future. Both memory and processing costs are expected to decrease dramatically.

However, reduced cost of special purpose computation suitable for high speed graphics displays by means of custom VLSI circuits may well be achievable only with volume production.

The primary focus of our research has been to investigate appropriate trade-offs for a hierarchically- structured graphics memory architecture consisting of a combination of disk, bulk random-access memory, and high speed image refresh memory. The overall aim is to provide a cost-effective architecture which will still permit rapid access to very large images. We are aiming to provide windowing, zooming, and roaming capabilities for images of arbitrary size, where these images are generated from a combination of data structures appropriate to scanned-in continuous tone images (contones), scanned-in line art, and algorithmically generated images, e.g., from a vector display list and/or a text file.

There has been considerable literature and increasing use of anti-aliasing in order to generate more pleasing displays for a given spatial sampling grid. This, of course, impacts architectural considerations as the size of the required image refresh memory can be reduced, thus altering the trade-offs. An interesting question concerns the trade-off between the spatial and gray scale resolution, i.e., how available memory should be allocated to optimize quality. Another area of research concerns the computation required to produce an anti-aliased display as pertaining to the benefits to be derived from anti-aliasing. When roaming or zooming quickly through an image, it is more important to give fast response than to optimize quality. When motion ceases, then the highest quality is desired.

We have worked on the development of polygon shading algorithms for raster displays. Of all the methods used to generate computer images, those rendered by ray-tracing are the most realistic and spectacular. This method combines the effects of hidden surfaces, shadows, reflection, and refraction in a way unmatched by other techniques. The intensity of each pixel in the image is calculated by extending a tree of rays from the viewer, through each pixel, to the first surface encountered. Then the ray is split to form new rays representing specular reflections and transmissions. This process is continued for each new ray until none of the new rays intersect any surface.

One advantage of the ray-tracing approach is that, in general, there is no need to transform the surface since the view is encoded in the geometry of the light rays. For example, a perspective projection is performed by passing all rays through a common vantage point, while in a parallel projection all rays travel in the same direction. Ray-tracing has been applied to images modeled by planar and quadric surfaces, bicubic surfaces, and more recently to algebraic and fractal surfaces.

A major problem with the ray-tracing method is its big appetite for floating-point arithmetic, which makes it unsuitable for any interactive computer graphics work. Our research examines various ray-tracing algorithms and investigates the use of hardware support to enable the ray-tracing methodology to be more suitable for interactive applications. One approach is to implement the line-surface intersection task in hardware. Another approach is to exploit the inherent independence that exists between picture elements and devise suitable parallel architectures for ray-tracing algorithms. Yet another approach is to develop algorithms that combine ray-tracing with other shading methods.

We are also interested in achieving faster computation by taking advantage of the frame-to-frame correlations that normally exist. One goal might be to develop "constant computation time" algorithms where the image display quality would vary with image complexity and/or motion rather than varying the frame rate.

References

1. J.N. Ratzel, "The Discrete Representation of Spatially Continuous Images," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., August 1980.
2. A. Zarembowitch, "Forward Estimation Adaptive DPCM for Image Data Compression," S.M. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., May 1981.
3. D.E. Troxel, et al., "Bandwidth Compression of High Quality Images," International Conference on Communication, June 1980, pp. 31.9.1-31.9.5.
4. D.E. Troxel, et al., "A Two-Channel Picture Coding System: I - Real Time Implementation," IEEE Trans. Comm. COM-29, 12, 1841-1848, December 1981.
5. W.F. Schreiber and R.R. Buckley, "A Two-Channel Picture Coding System: II - Adaptive Companding and Color Coding," IEEE Trans. Comm. COM-29, 12, 1848-1858, December 1981.
6. D.E. Troxel, et al., "Automated Engraving of Gravure Cylinders," IEEE Trans. Syst. Man Cybern. SMC-11, 9, 585-596, November 1981.
7. W.F. Schreiber, "Color Processing System," U.S. Patent Application Serial Number 06/374,804, May 5, 1982. This patent will be issued in March 1985.