In this study, 44 patterns were used, and these are shown in Fig. 5. The associations between the patterns and the references for these 44 patterns are listed in Table I. The successive formation of the individual memories and the total memory are depicted graphically in Fig. 6.

It is useful to emphasize that when a large number of patterns are stored, subsequent presentation of any one pattern for recognition does not regenerate the associated reference by itself, but does in fact yield a complex pattern which might be predominantly the associated reference. This situation is illustrated in Fig. 7. For any specific memory, these characteristics can be followed as the number of patterns stored is increased. The results for the 64 storage site memory in question are shown in matrix form in Fig. 8. The results are presented in a nonnormalized form. When five patterns are stored, and when pattern #6 (associated reference #32) is presented for recognition, the response is primarily the associated reference #2. As seen from the appropriate row entries, the response for the correct reference is 68 versus -12.0, -12.0, 12.0 and -4 for the other references. However, by the time nine patterns have been stored, presentation of pattern #8 (associated reference #51) results in significant values of the coefficients for reference #19 and for reference #0.

Examination of these characteristics indicates that if recognition depended on the associated pattern being regenerated predominantly, then this scheme would be of limited use indeed. We show however, that this is not the case. In fact, presentation of a pattern for recognition generates a "star" of reference vectors. It is not necessary that the star have the associated pattern as the predominant feature. It is only necessary that the stars regenerated in this manner are different from each other. The final act of recognition consists of taking the scalar product of a discriminant vector f and the star of reference vectors to yield a single number. To each pattern there is, then, a specific number, and this list of patterns and corresponding identifying number is the simple dictionary which is used in the recognition and retrieval of patterns.

Using the pattern-reference associations listed in Table I, an associative memory was prepared, and a discriminant vector consisting of a linear combination of the reference vectors was used for determining which pattern had been presented for recognition. This is to say that when x_q was presented for recognition,

$$f = 0.1y_1 + 0.2y_2 + 0.3y_3 \cdots 0.44y_{44}$$

was used to yield

$$\langle x_q m_f \rangle = D_q$$

a number. The 44 patterns and the numbers associated with them are listed in Table II.

IV. Significance of Results and Direction of Future Investigations

The results of the previous section indicate that an 8×8 memory can be used to store many (at least 44) patterns in parallel and that when any of these patterns is presented for recognition, only 3×44 processing steps are required to yield a number which then readily identifies which pattern had been presented for recognition. This is in contrast to the demands of the template-matching technique where 64×44 storage sites and about $3 \times 64 \times 44$ processing steps are required for recognition. These savings become very significant once memories of the order of 10^6 storage sites are contemplated.

This scheme has many interesting properties, some of which are now known and others need to be explored further. Of the patterns depicted in Fig. 5, we note that patterns #2 and #25, and #3 and #26 form identical pairs. Referring to Table II, we see that the identifying numbers also form pairs, even though different references were used. Thus this memory is effective in sorting out which patterns look alike, and this result is achieved regardless of the references used. Similarly, patterns #7 and #19, #18 and #27, and #28 and #29 form pairs which are "gray scale" inverted images of each other. It is interesting that the corresponding identifying numbers form pairs identical in absolute magnitude bit with opposite signs.

This proposed scheme either in software or hardware implementation may be used to serve as content addressable memories for patterns in major automated industrial systems. Presentation of part of a pattern would lead to recognition and subsequent automatic retrieval from storage all that is associated with the pattern and that needs to be recalled. The addressing is based on the content matter of the pattern so presented and not on any additional address. It would seem that this scheme has potential as an important element in all major information processing systems.

The discriminant vector f may be used to accentuate the importance of certain patterns or to group certain others together. The way to accomplish this is quite obvious when f is expressed as a linear expansion of the linearly independent reference vectors. Analogous techniques for circumstances when f is not a linear sum of the reference vectors and/or when the reference vectors do not form an orthonormal set of basis vectors need to be investigated further. There are indications that the characteristics of those memories are of substantial practical interest.

In addition to studies of software implementation of this scheme, a hardware implementation has also been initiated, and corresponding characteristics for memories of much larger storage capacity will become available in the future.

References

[1] L. D. Harmon, "Automatic recognition of print and script," *Proc. IEEE*, vol. 60, pp. 1165–1176, Oct. 1972.
[2] D. G. Willis, "Plastic neurons as memory elements," presented at 1959 Int. Conf. Information Processing, Paris, France.
[3] D. J. Willshaw, "Models of distributed associative memory," Ph.D. dissertation, University of Edinburgh, Edinburgh, Scotland, 1971.
[4] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins, "Non-holographic associative memory," *Nature*, vol. 222, p. 960, 1969.
[5] H. C. Longuet-Higgins, D. J. Willshaw, and O. P. Buneman, "Theories of associative recall," *Quart. Rev., Biophys.*, vol. 3, p. 223, 1970.
[6] H. F. Harmuth, *Transmission of Information by Orthozonal Functions*. New York: Springer, 1969. (A background discussion of Walsh functions may be obtained from this reference.)
[7] C. R. Rosen, "An approach to a distributed memory," in *Principles of Self Organization*, Von Foerster and Zopf, Eds. New York: Pergamon, 1962.

Computer Editing of News Photos

DONALD E. TROXEL

***Abstract*—A computer-based system for the editing of news photographs is described. Capabilities now implemented include cropping, enlarging, and the insertion of captions. A key element of this system is a display peripheral capable of posting a flicker-free display on a standard monochrome television monitor. A relatively small semiconductor memory of 2^{18} bits is used to store the picture information. The Roberts**

Manuscript received February 6, 1975; revised May 23, 1975. This work was supported by a Grant from the Associated Press.
The author is with the Research Laboratory of Electronics and the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

noise technique and linear interpolation are used to produce a video signal corresponding to 512 × 512 picture elements with each picture element represented by 8 bits or 256 shades of gray.

I. Introduction

Since August, 1970, the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, has been developing for the Associated Press (AP) a radically different news picture (Wirephoto) distribution system [1]. The primary motivation for this development is the desire to utilize direct digital transmission services now being installed instead of the voice telephone channels currently used for picture transmission. The use of digital transmission services provides the opportunity to achieve substantial improvements in image quality, speed, reliability, flexibility, and cost.

The system is to be introduced in stages, in such a way that at least the present standard of quality and service is 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, enhancement (contrast control, etc.), combining, and the addition of captions. No additional chemical photographic work will be required for the network operation.

Transmission over the "backbone" system linking the AP bureaus and the large metropolitan newspapers that have substantial computer facilities will be high speed and digital and generally will originate and terminate at computer-controlled digital storage devices. The combined storage capacity will accommodate approximately 100–200 pictures with film used for archival storage. Transmissions to subscribers will be analog or digital and at speeds and scanning standards appropriate to the existing transmission facilities. Control and scheduling of all transmissions will normally be accomplished automatically by the computer facilities located in New York City. Manual control of the system may be selected by AP personnel at New York.

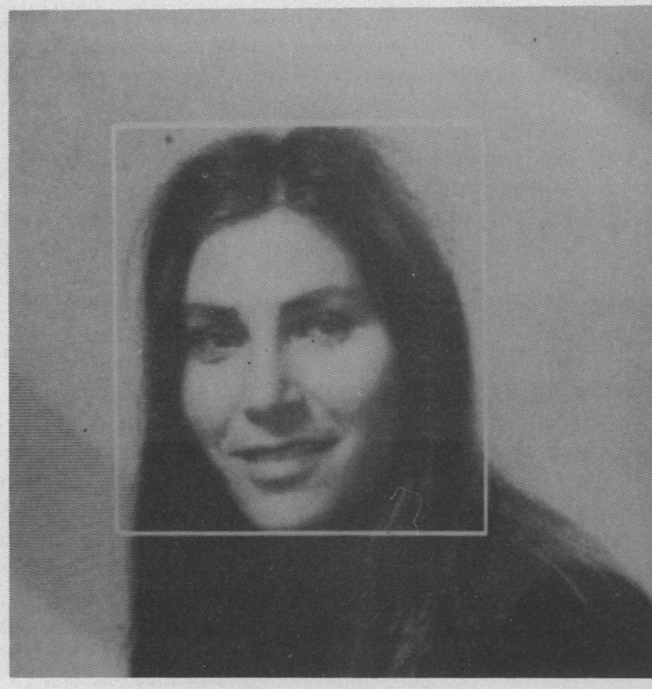
In the laboratory we have implemented some of these procedures. Pictures may be introduced from the AP network, a local analog transmitter (a 12 000 picture element/second CRT scanner, magnetic tape, or Dectape) and automatically stored on the disk. Pictures may be transmitted from the disk to comparable receiving points. Pictures stored on the disk may be viewed on a TV display utilizing a full frame semiconductor storage system. Editing facilities now in operation include cropping, enlarging, reducing, combining several pictures into one, and the addition of captions.

II. Editing

Pictures to be edited are put into the computer via a standard facsimile transmitter, and the digitized picture is stored at full resolution on the bulk storage device (disk) of the computer. A photo editor can name the picture he wishes to edit or simply request the next picture in an editing queue. A lower resolution version of the picture is posted on the television picture display.

Cropping

To facilitate the cropping of news photographs, a white rectangle is superimposed on the picture display. The photo editor adjusts the position, size, and shape of this rectangle to indicate the desired cropping. These adjustments can be made by a series of keyboard commands or with the aid of an electronic tablet and pen. The rectangle on the picture display continuously follows these commands, allowing the photo editor to arrive quickly at the desired cropping specification. An example of the final position of the rectangle is shown in Fig. 1(a).



(a)

(b)

Fig. 1. Cropping.

When the photo editor is satisfied with the position of the displayed rectangle, he enters a keyboard command to create a cropped version of the original full resolution picture. This picture can then be enlarged to any desired size by invoking a scaling routine which uses a linear interpolation process to create the magnified picture. This final result may be inspected by posting it on the picture display (Fig. 1(b)).

Another editing mode enables a photo editor to specify a combination of up to four different croppings to be combined into a single composite picture. An example of this is shown in Fig. 2.

Captions

The addition of an appropriate caption is the final step before transmission of the edited news photograph. The captions are stored in the computer as ASCII characters and are only converted to a facsimile video signal at the time an analog transmission is initiated.

III. Picture Display

The picture display is a computer peripheral which can be connected to any computer capable of supporting a serial interface of length 16 bits or more. It is capable of posting a flicker-free display on a monochrome television monitor. The outline of a rectangle of arbitrary dimensions may be superimposed on the picture. The picture display includes sufficient memory to support a gray-scale picture on a 256 × 256 matrix with each picture element (PEL) consisting of 4 bits, thus allowing 16 shades of gray. The capacity of this memory is equivalent to 16384 16-bit words. The cost of this memory dominates the cost of the picture display. Accordingly, some picture processing operations have been implemented to permit a reasonably high quality image to be supported by a relatively small memory. It is primarily these operations which distinguish this picture display from ones previously reported [2].

Resolution

One obvious way to increase the quality is to double the resolution or matrix size to 512 × 512 PEL's. This would quadruple the required memory, substantially increasing the cost. As an alternative to storing a higher resolution picture, linear interpolation is used to synthesize a 512 × 512 matrix of PEL's from the stored matrix of 256 × 256 PEL's. This, of course, does not actually increase the resolution of the displayed picture. It does, however, increase the apparent image quality. Each PEL is now small enough so that it is difficult to discern the sampling structure of the picture at normal viewing distances. Compatibility with a standard television monitor is maintained by using normal interlace to display the stored and interpolated lines for alternate fields. It is interesting to note that the use of interlace without interpolation results in the rather annoying effect of a picture that jumps up and down.

Number of Gray Levels

Another way to increase the displayed picture quality by increasing the memory capacity is to use more bits per PEL, allowing for a greater number of shades of gray. The dominant bad effect of a small number of gray levels is the perception of contours in low frequency areas of the picture. The technique of adding pseudorandom noise described by Roberts [3] completely eliminates the contours. Implementation of this scheme allows the synthesis of an arbitrary number of gray levels in the displayed image for essentially any number of bits per PEL actually stored. The signal to noise ratio of the resultant picture is a function of the number of bits per PEL. Our choice of 4 bits per PEL is a compromise between memory cost and picture quality.

(a)

(b)

Fig. 2. Composite resulting from four croppings.

Companding

Companding, or tone scale compression and expansion, was used in addition to pseudorandom noise by Roberts [3]. The basis for use of this procedure is the variation of apparent noise visibility as a function of the gray level. Noise is more visible in the blacks or shadows than in the highlights or white portions of a picture. Roberts used a square law compandor. Hashizume [4] concluded that the "optimum" compandor characteristic is proportional to log $(1 + aB)$, where B is the brightness and a is approximately 0.02. The expandor characteristic of the CRT in the television monitor was measured and found to be quite

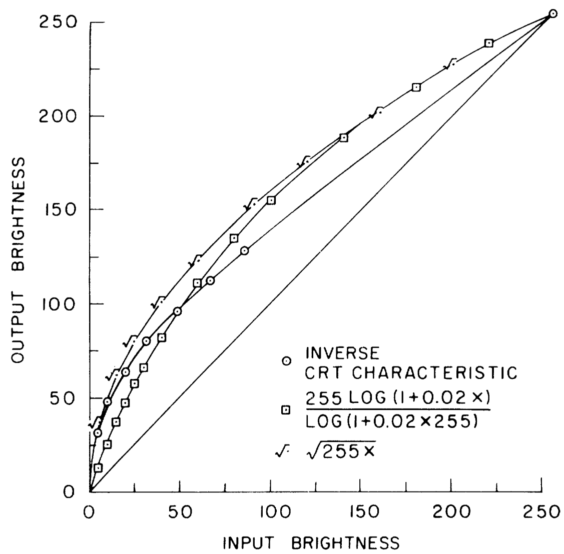


Fig. 3. Potential compressor characteristics.

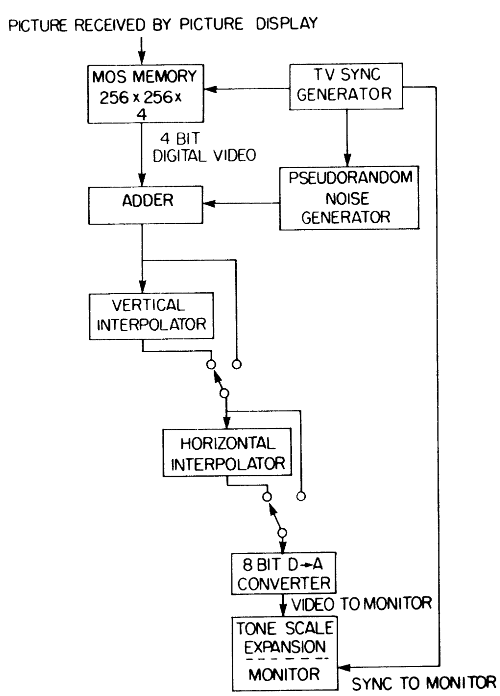


Fig. 4. Picture processing by picture display.

stable and repeatable. Our original aim was to perform compression in the computer and to incorporate expansion along with correction for the CRT characteristic in hardware. Inspection of three potential compressor characteristics shown in Fig. 3 resulted in the judgment that there was little to be gained by using other than the inverse CRT characteristic for the compressor, which choice happily minimizes the complexity of the hardware. Pseudorandom noise is subtracted after tone scale compression, and each picture element is quantized to 4 bits for transmission to the picture display.

Picture Processing

The picture processing performed by the picture display is illustrated in Fig. 4. The picture is stored in a memory which receives timing signals from a television sync generator. As the 4-bit per PEL digital video is retrieved from the memory, the pseudorandom noise is added. The picture now consists of 256 × 256 8-bit PEL's. Via a panel switch, the user may enable the image to be linearly interpolated in the vertical direction to yield 512 lines, each consisting of 256 8-bit PEL's. The additional horizontal lines are made possible by using standard TV interlace, resulting in the reduction of the picture repetition rate from 60 to 30 Hz. Optionally, the user may select an additional (horizontal) interpolation step to increase the number of PEL's for each line to 512.

The resulting 8-bit PEL's are then converted to an analog video signal and fed to the television monitor. The tone scale expansion indicated in Fig. 4 is accomplished by the physical nature of the CRT in the television monitor. The reproducibility of this characteristic is assured by a switch on the monitor which selects precalibrated brightness and contrast settings for the display picture. Optionally, the user may manually adjust these controls.

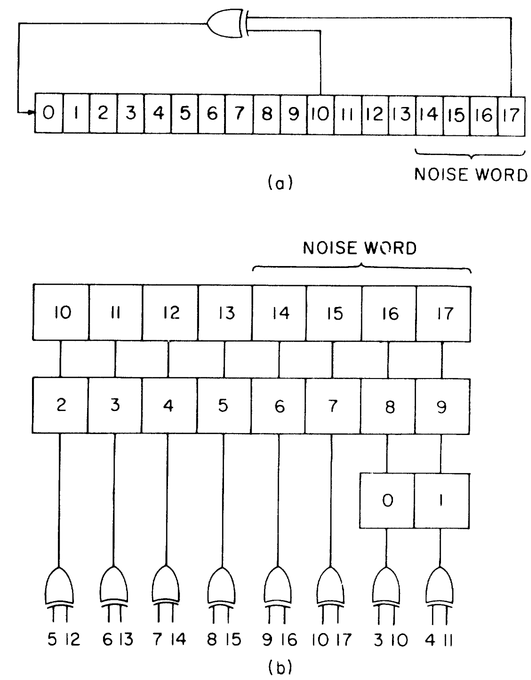


Fig. 5. Pseudorandom noise generator.

Pseudorandom Noise Generator

The pseudorandom noise generator implementation is based on a feedback shift register with the generator polynomial $1 + D^7 + D^{18} = \phi$. This corresponds to the logic diagram shown in Fig. 5(a).

This register is initialized to $760544_{(8)}$ at the beginning of each frame. Successive noise words are obtained by shifting eight times. The actual implementation is equivalent to this scheme, but by using more combinational logic and feedback, it generates successive noise words with only one clock pulse per PEL (see Fig. 5(b)).

Horizontal and Vertical Interpolator

Horizontal and vertical interpolation are accomplished with arithmetic functional units. The units alternately perform the functions A and $(A + B)/2$, where A and B are the inputs to the unit. These functional units are also used to blank the video during retrace and to force the video signal to maximum amplitude for displaying the cursor. One-PEL and one-line storage are included to provide the appropriate operands to the functional units.

IV. Summary

A laboratory version of a system for computer editing of news photos has been implemented. Simple commands put in by keyboard and tablet allow an editor to crop, enlarge, and combine pictures. Appropriate captions can be associated with each picture before transmission. We anticipate that additional capabilities such as tone-scale enhancement will be easily added at a later time. The primary advantages of this system are the reduction of time required to crop and enlarge a picture and the elimination of chemical photographic work normally required for these operations. Additionally, if the cropping is not severe, the computer processing will result in less degradation than the usual photographic processes.

References

[1] W. F. Schreiber and D. E. Troxel, "Digital wirephoto system," M.I.T. Research Laboratory of Electronics, Quart. Progress Rep., No. 112, pp. 96–98, Jan. 1974.
[2] N. H. Kreitzer and W. J. Fitzgerald, "A video display system for image processing by computer," *IEEE Trans. Comput.*, vol. C-22, pp. 128–134, Feb. 1973.
[3] L. G. Roberts, "Picture coding using pseudo-random noise," *IEEE Trans. Inform. Theory*, vol. IT-8, pp. 145–154, Feb. 1962.
[4] B. Hashizume, "Companding in image processing," S.B. Thesis, Mass. Inst. Tech., Cambridge, June 1973.

Shape-Oriented Chromosome Classification

E. T. LEE

***Abstract*—An algorithm for classifying a chromosome image as an "approximate median chromosome," "approximate submedian chromosome," or "approximate acrocentric chromosome" is presented. Three distributive lattices and six most unsymmetrical chromosome images are found. Advantages of shape-oriented chromosome classification over chromosome classification done by human beings are stated. The results may have useful applications in pattern recognition, information retrieval, and artificial intelligence.**

I. Introduction

Chromosomes have been classified by centromeric index (ratio of arm lengths to total body length), ratios of body lengths from one chromosome to another, chromosome areas, etc. In particular, it has been done by applying the "rubber-mask" technique by Widrow [1], and the stochastic linguistic approach by H. Lee and Fu [2]. A rubber-mask classification has been made using various versions of the "Denver Standard" chromosomes [3] as stereotypes. The chromosome distortion parameters used in [1] are LENGTH, WIDTH, ANGLE, and CURVE, which are adjusted individually and independently in each of the four quadrants of a stereotype. Human chromosome stereotypes and skeletal illustration of chromosome distortion parameters were presented in [1]. Widrow obtained the best fit of a chromosome image by:

a) lengthening or shortening each arm;
b) thickening or thinning each arm;
c) offsetting at an angle along the median line of each arm;
d) curving with a second degree function along the median line of each arm;
e) a combination of these effects.

Manuscript received March 19, 1975; revised May 30, 1975.
The author is with the Department of Electrical Sciences, State University of New York at Stony Brook, Stony Brook, N.Y. 11790.

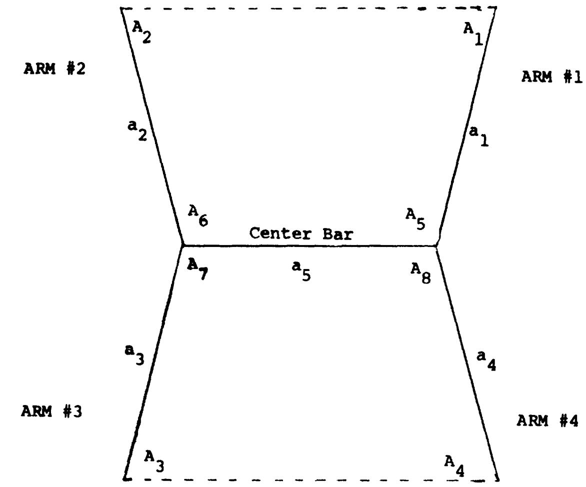


Fig. 1. Best fit skeletal "length and angle only" transformation.

As stated in [1], arm lengths, widths, areas, and centromeric index of human chromosome could be estimated in terms of the lengths, widths, and areas of the distorted stereotype by including the length and width factors derived from the fitting process. However, the parameters ANGLE and CURVE are merely used in the fitting and have no value in measuring and classifying chromosomes. In this correspondence, the angle and length factors derived from the fitting process are applied to the classification of chromosome images. The best fit skeletal "length and angle only" transformation with angles A_i and sides a_j is shown in Fig. 1. After connecting the tips of arms No. 1 and 2, and the tips of arms No. 3 and 4 as indicated by dotted lines, the best fit skeletal "length and angle only" transformation becomes a hexagon.

A stochastic context-free grammar was used to classify a string-encoded chromosome image into median, submedian, or acrocentric classes by H. Lee and Fu [2]. There is no sharp boundary between the class of median and submedian, nor between the class of submedian and acrocentric. Because the transition from membership to nonmembership in these three classes is gradual rather than abrupt, the concepts and techniques developed in fuzzy sets [4]–[8] and fuzzy languages [9] may be applied to classify chromosome images.

A preliminary study of applying shape-oriented similarity measures defined over a pair of chromosome images to the classification problem was presented in [10]. In this correspondence, the classification problem is studied through the use of shape-oriented similarity measures of a given chromosome image to symmetrical chromosomes, median chromosomes, submedian chromosomes, and acrocentric chromosomes. The test results of this theory on actual chromosome images will be presented in subsequent papers.

II. Symmetry of Chromosome Images

The preparation of chromosome images and the definition of metaphase chromosome images may be found in [1]. At metaphase each chromosome has a twin, normally identical counterpart.

Definition 1: A chromosome image with angles A_i and sides a_j is a *symmetrical* chromosome image if and only if $A_{2i-1} = A_{2i}$, for $1 \leq i \leq 4$, $a_1 = a_2$, and $a_3 = a_4$.

A shape-oriented quantitative measure of the similarity of a given chromosome image A to all symmetrical chromosome