Scaling Binary Images with the Telescoping Template

ROBERT A. ULICHNEY AND DONALD E. TROXEL

Abstract—The importance of enlarging and reducing two-level images such as graphical and documentary matter by digital means continues to grow as more such images are digitally represented.

A nonlinear scaling scheme is devised which exploits the simplicity of this binary nature, treating images logically instead of arithmetically; a convolution-like effect is achieved without a single addition or multiplication! This method yields high-fidelity digital scaling and meets the objectives of being fast, conducive to hardware realization, and void of special pre-encoding requirements.

Index Terms—Computer picture processing, picture processing, printing–computer controlled typesetting, scaling binary images—fast nonlinear method, signal processing.

I. Introduction

Our investigation deals with the problem of taking a given set of image samples and producing another set of samples which represents the original image at a desired scale within the domain of binary images. For this effort a binary image is defined as a two-dimensional signal whose amplitude is precisely either black (numerically and logically represented as 1) or white (numerically and logically represented as 0). Furthermore, it is assumed that such an image is perceived visually as binary. This excludes high-frequency distributions of black and white which produce the illusion of a gray area as in halftone pictures.

Manuscript received June 23, 1980; revised November 16, 1981. This paper is based on a thesis submitted in partial fulfillment of the M.S. degree in the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, on January 19, 1979.

R. A. Ulichney is with the Corporate Research Group, Digital Equipment Corporation, Maynard, MA 01754.

D. E. Troxel is with the Research Laboratory of Electronics and the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

Digital scaling is the process performed on the digital input image, resulting in a digital output image of a different size at a *fixed resolution*. We devised such a process, which preserves the integrity of the original image while meeting the objectives of

1) high speed (ergo, conducive to hardware realization) and
2) no special pre-encoding of source images or a full two-dimensional intermediate memory space required.

The input parameters are s_x, s_y, and $G[m, n]$ which are, respectively, the horizontal and vertical scale factors, and the given $M \times N$ binary image. The output is the $K \times L$ scaled binary image $S[k, l]$.

To date, attempts at high fidelity digital scaling of binary images required that the source images be subjected to an extensive preprocessing stage where approximating curves are fitted to the contours [1]-[4]. Scaling, then, is a trivial matter, since all the real labor occurs in the pretranslation to a scale-insensitive code. Output quality is controlled only by the quality of the encoding processes.

This pre-encoding is suitable for some phototypesetting applications. However, many typesetters use facsimile coding techniques [5], [6] and decode the images to a bit-map of the video information prior to exposing the output medium. They then use relatively crude methods of scaling such as omitting the appropriate number of columns or rows.

II. The Linear Approach

All linear methods of digital scaling are comprised of two phases (at least internally) (see Fig. 1). The first is an attempt to reconstruct or "repaint" the original continuous image from the given samples. The resulting scaled image is then produced by sampling this reconstruction. We refer to this procedure as retrospective resampling, that is, looking back to the continuous image as it existed before the original sampling, then resampling it.

The first phase involves conversion of the given sampled digital image $G[m, n]$ to a sequence of weighted impulses (Dirac delta functions) which is then convolved with the interpolating filter $h(x, y)$. The resulting continuous amplitudes are then quantized to one of two levels by the binary amplitude quantizer yielding the reconstructed continuous image

$$Rl(x, y) = Q_2 \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G[m, n]\ h(x - m, y - n) \right\}.$$

The second phase of retrospective resampling involves the sampling of the reconstructed binary image with a finer (for enlargement) or coarser (for reduction) grid with sampling frequency

$$(X, Y) = \left(\frac{M-1}{K-1}, \frac{N-1}{L-1} \right)$$

yielding

$$S[k, l] = Q_2 \left\{ \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G[m, n]\ h(kX - m, lY - n) \right\}.$$

In practical systems, a convolution kernel $h(x, y)$, with as small an area as possible, is desirable. A tradeoff exists between reduction of computation time and reconstruction fidelity. For ease in implementation it is also desirable to have a kernel which is separable, i.e., $h(x, y) = h(x)\ h(y)$. A list of such functions, along with a discussion of interpolation error, is presented by Pratt [7].

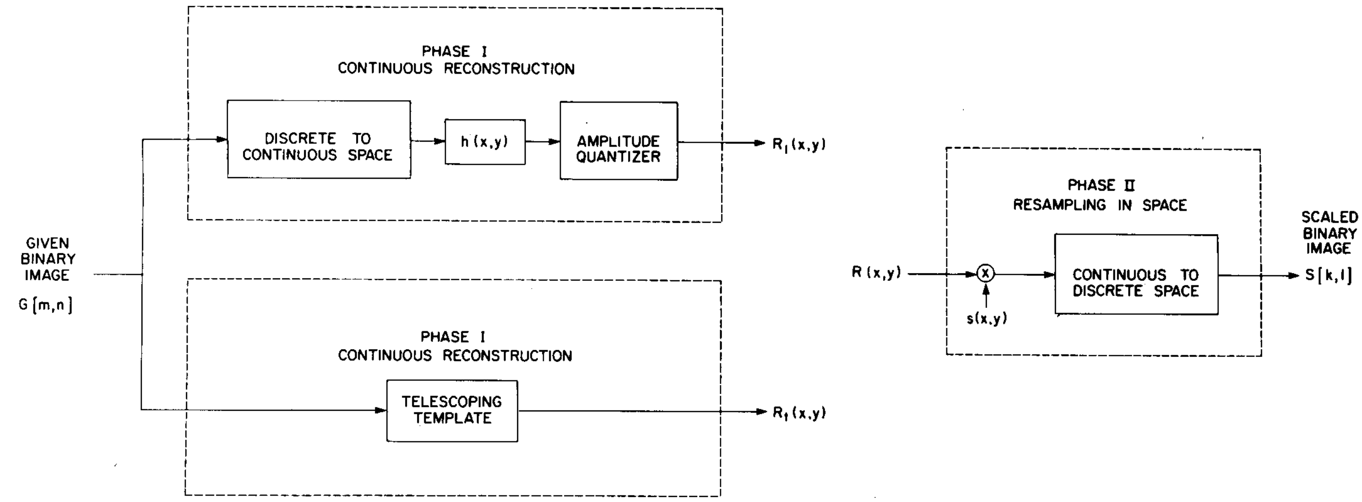


Fig. 1. Linear and nonlinear models of retrospective resampling.

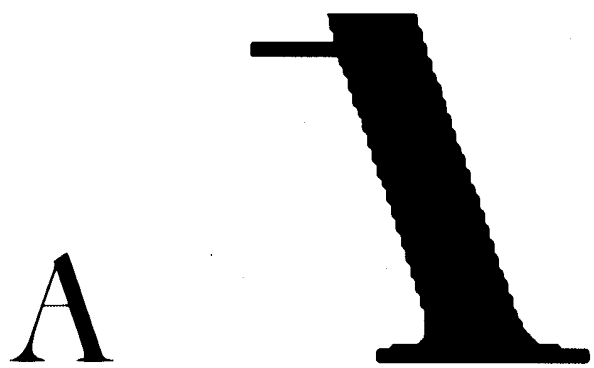


Fig. 2. Bilinear enlargement, $s_x = s_y = 5.5$

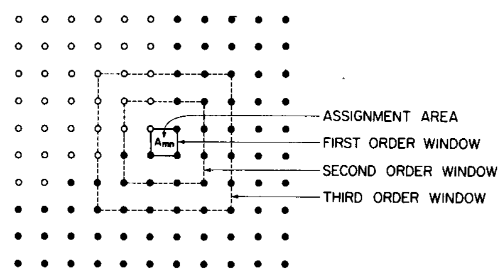


Fig. 3. Assignment area and some possible windows.

Bilinear interpolation is a typical scheme for the scaling of multilevel monochromatic images [8]. The results are good for small changes in size. For gross enlargements, however, the quality of the scaled image suffers from an apparent "blockiness." This undesirable artifact follows from the fact that the span of the unit impulse response of the effective two-dimensional filter is only 2 × 2 pels and, thus, cannot have circular symmetry. The results are much more severe when this method is applied to binary images as illustrated in Fig. 2. Other interpolation schemes for multilevel images have been reported, e.g., [9], which more nearly achieve circular symmetry but require significantly greater computation time.

III. The Telescoping Template Method

The process of binary thresholding constitutes an appreciable irreversable nonlinear contribution. The resulting abrupt edges, which are the essence of binary images, make scaling by linear means impractical. We introduce the following nonlinear method which exploits the simplicity of the binary nature.

A priori knowledge of general contour characteristics that can occur within a given image "window" is stored in a "telescoping template," which aids in the internal continuous reconstruction of the original image to be resampled at a different spatial frequency. This nonlinear method differs from the linear approach only in phase *I*, the continuous reconstruction, as shown in Fig. 1. The intermediate signal $R_t(x, y)$ is also a spatially continuous binary image representing the original image. However, the entire process producing $R_t(x, y)$ is nonlinear, and all internal amplitude values are restricted to either 1 or 0.

The continuous reconstruction $R_t(x, y)$ can be thought of as a concatenation of many unit squares called "assignment areas." Fig. 3 illustrates an assignment area with its associated first-, second-, and third-order window. The reconstruction is complete once we select an "assignment rule" for each assignment area describing how each assignment area is painted. This selection is based on the neighborhood of samples defined by a window.

Because our input samples are binary, they can be treated as bits of a code which is used to map a given window arrangement to one of a finite set of assignment rules. If the window order p is relatively small, direct enumeration of all $2^{(2p)^2}$ window arrangements would be tedious but nonetheless manageable. Fortunately, the particular assignment rules can be determined by simply interpreting how continuous contours and cusps are represented by discrete binary samples. The result is a convolution-like process without a single multiplication or addition.

The second phase of either model consists of resampling in space. Clearly, we need not reconstruct the entire image before resampling. Assignment areas are considered only when a new sample falls within that area.

A. Enlargement

The general flavor of the "telescoping template" nature of this scheme is illustrated by Fig. 4. The example shows one particular assignment area and the appropriate assignment rule for windows of increasing order. Note that, as the window size grows, the increased information is used to determine more intelligently the assignment rule. In order to enlarge an image, we must consider where a reconstructed edge should lie in relation to original picture element (pel) centers. Assuming that we are dealing with black objects on a white background, Fig. 5(a) illustrates two choices for a thin, 3 pel wide vertical

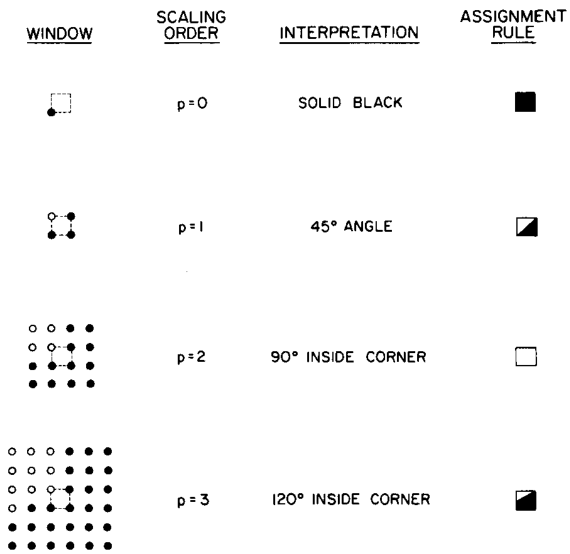


Fig. 4. Telescoping templates.

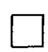
bar. Since each pel is considered to be one unit square, "ideal" boundaries are so named because they are separated by three units. The alternative boundary placement through the center of edge pels is designated "practical" and is the one used. This choice has the effect of reducing the weight of black strokes, which is consistent with the aesthetic considerations of an artist's design of a larger version of a small type face [10], [11]. A consequence of this is that thin lines of one pel thickness will not be enlarged. A technique remedying this is presented later.

We consider assignment rules composed of straight line segments in order to minimize computational complexity. We avoid unwanted discontinuities in the continuous reconstruction by ensuring that the endpoints of assignment rules will be equal at the boundaries of adjacent assignment areas.

It should be noted that cusps will always occur on original pel centers, eliminating the need for a second equation in the assignment rule. Fig. 5(b) again supports our boundary selection convention for this reason.

Window decoding is simple and has two steps. In the first step, the first-order window is considered. Only 16 possible arrangements exist and fall into two groups, enumerated here with their associated assignment rules.

1) *Solid Area Group:*

$$\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \left(\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \quad \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right)^{1}$$ 

$$\begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix}$$ 

2) *Edge Area Group:*

$$\begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} \quad \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix} \quad \begin{matrix} 1 & 0 \\ 0 & 0 \end{matrix} \quad \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}$$ 

$$\begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \quad \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix} \quad \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix} \quad \begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}$$ 

$$\begin{matrix} 0 & 1 \\ 1 & 1 \end{matrix} \quad \begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix} \quad \begin{matrix} 1 & 1 \\ 1 & 0 \end{matrix} \quad \begin{matrix} 1 & 0 \\ 1 & 1 \end{matrix}$$ 

[1] Thin lines of one pel thickness are ignored.

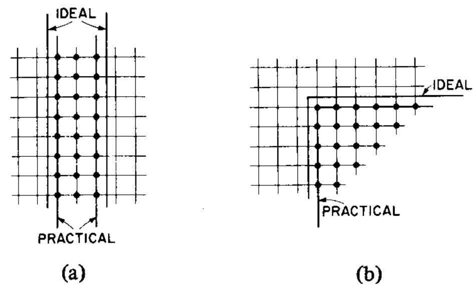


Fig. 5. Selection of image boundary. (a) Vertical bar. (b) Corner.

Fig. 6. Telescoping template enlargement, $s_x = s_y = 5.5$. (a) Result of 0th-order scaling. (b) Result of first-order scaling. (c) Result of second-order scaling. (d) Result of third-order scaling.

An important shortcut follows from the fact that most of a given binary image is solid black or white and that edges seldom occur within an assignment area. The second step of finding a template match is taken only when an edge area is detected and $p > 1$. Fig. 4 illustrates how the assignment rule is modified as the order of the window is increased. For a second-order system 13 templates and eight assignment rules were found to handle all such exceptions. A third-order system requires 45 templates and 22 assignment rules. A complete list of templates and assignment rules is given in [12].

Software Simulations: The simple special case of $p = 0$, or 0th-order enlarging, corresponds to a sample-and-hold process comparable to scaling by change of resolution. There are only two templates and two obvious assignment rules (all black or all white). Fig. 6(a) shows the quite unacceptable result for a symmetric magnification of 5.5 times. Results for $p = 1$, 2, and 3 are also shown in Fig. 6 along with a graphical representation of pel centers and assigned edges.

Another example is given in Fig. 7. Notice how as the order increases the gentle curves smooth out, while corners and cusps remain crisp.

B. Reduction

Reduction is fundamentally "easier" than enlargement because information is being removed rather than inserted. The principal difference in the approach to the problem of reduction stems from the fact that at most only one new sample will fall within a given assignment area. The management of boundaries within each assignment area is not necessary; it is sufficient to simply decide whether an A_{mn} is "on" or "off."

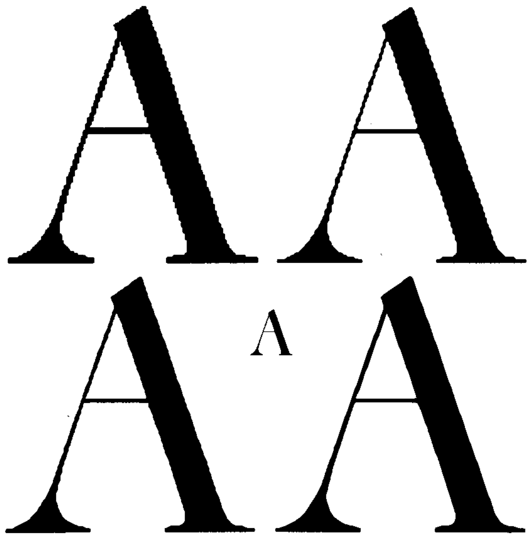


Fig. 7. Telescoping template examples, $s_x = s_y = 5.5$. Given image in center; $p = 0$ in upper left; $p = 1$ in upper right; $p = 2$ in lower right; $p = 3$ in lower left.

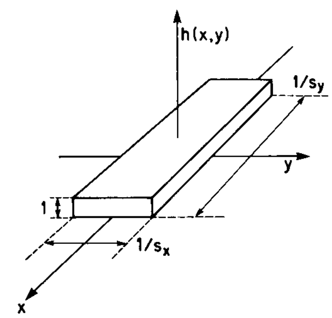


Fig. 8. Reduction convolution kernel.

The 0th-order scaling presented in the last section is one possible solution. This may be adequate for some applications, but in general a more acceptable solution is sought for two reasons.

1) The weight of black strokes should be increased or thickened proportionately for both aesthetic and practical considerations [11], [12]. This does not occur with 0th-order scaling.

2) The chance of a thin line being lost or unrecognized increases as the scale factors decrease.

Both of the above shortcomings are alleviated by employing the linear reconstruction process with the convolution kernel shown in Fig. 8. This process can be easily achieved nonlinearly in a process we call "binary convolution" which does not require a single multiplication or addition. The process uses the same tools developed previously with a few modifications.

1) The decoding of a given window is skipped if the associated A_{mn} is not occupied by a new sample.

2) Instead of decoding the window and searching the repertoire of templates for a match, the state of the new sample in A_{mn} is simply equal to the logical OR of the bit-by-bit ANDing of the associated window with a master reduction template.

This reduction template is determined by the scale factors and can be thought of as samples of the convolution kernel in Fig. 8. It is a d_H by d_V rectangle of black samples where

$$d_H = \text{integer}\,(1/s_x + 0.5)$$

$$d_V = \text{integer}\,(1/s_y + 0.5).$$

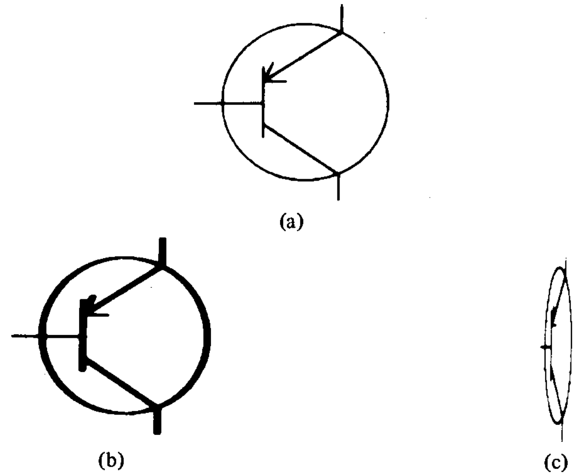


Fig. 9. Reducing a thin line drawing. (a) Given image. (b) $s_x = s_y = 1.0, d_H = 6, d_V = 1$. (c) $s_x = 1.0, s_y = 0.167, d_H = 6, d_V = 1$.

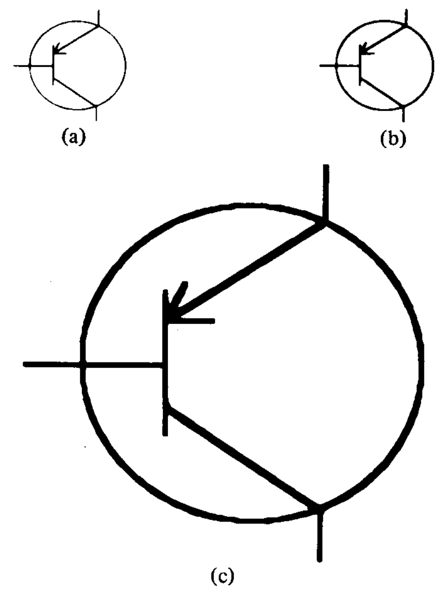


Fig. 10. Enlarging a thin line drawing. (a) Given image. (b) $s_x = s_y = 1.0$. $d_H = d_V = 3$. (c) $s_x = s_y = 3.5$.

For example, if $s_x = \frac{1}{2}$ and $s_y = \frac{1}{5}$ the master reduction template for a third-order system would be

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{matrix} \qquad (d_H = 2, d_V = 5).$$

For $s_x = 0.31$, $s_y = 0.95$, and $p = 3$, it would be

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{matrix} \qquad (d_H = 3, d_V = 1).$$

For a pth-order system, the reduction template will saturate when s_x or s_y is less than $1/(2p + 0.5)$. In most cases the ill

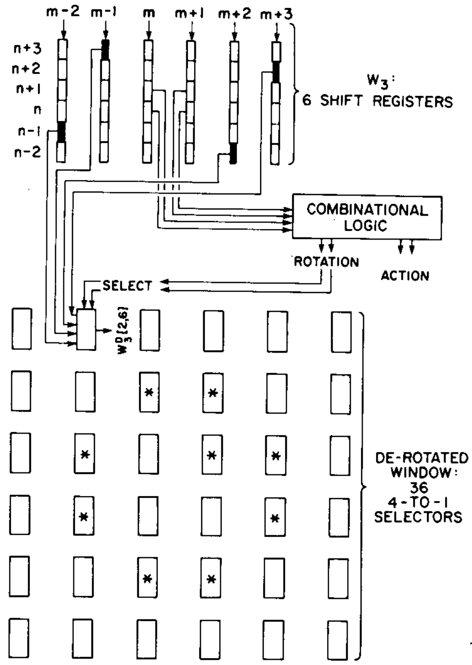


Fig. 11. Parallel hardware implementation.

effects of using a saturated template are not appreciable. Otherwise, a cascade approach or higher order scaling can be used.

Software Simulations: To reduce a binary image we first thicken the given image [Fig. 9(a)] by binary convolution to yield Fig. 9(b) and then resample to arrive at the desired reduction shown in Fig. 9(c). None of the thin lines were lost in the subsampling.

In the last section, we stated that the enlarging of thin-line drawings would have to be handled later because our templates do not recognize lines of one pel thickness. We solve this problem by using binary convolution to thicken the image first. Fig. 10 illustrates this technique.

IV. Implementation

A goal of this investigation was to devise a scaling scheme which is fast. In developing the telescoping template method, a premium was placed on ease of parallel hardware realization.

Fig. 11 illustrates the design of the decoding phase for a third-order system. The window W_3 is assumed to move from the image bottom to the top, six lines at a time. The required storage could consist of six 6 bit shift registers. For reduction, the 36 bit contents are simply logically ANDed bit for bit with the master convolution kernel, and ORed to yield a new sample if one happens to occupy the current assignment area. For enlargements the process is much more involved.

The 2 bit rotation code indicating the one of four possible rotations is decoded via combinational logic from the first-order window as shown in Fig. 12. This logic also produces one of three "action" codes: all white, all black, or edge (further decoding needed). If an edge action code is encountered, the window is derotated by 36 4-to-1 selectors. The explicit selection wiring is shown only for one location.

Once the derotated window is achieved, the repertoire of templates must be searched for a possible match. A very efficient way of doing this is to subdivide the set of templates into many very small linearly linked lists. Each list contains pointers to those templates which have the same bit values in the nine locations in the derotated window indicated by stars in Fig. 11. Thus, these nine bits can be used as an address to obtain the first pointer of such a list.

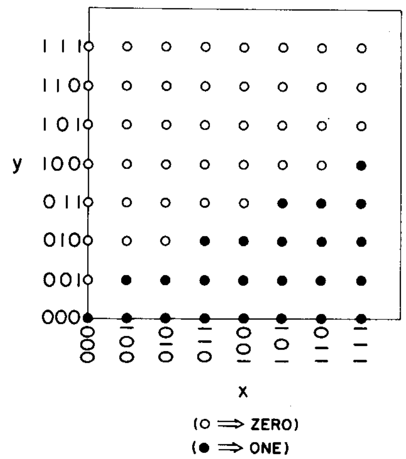


Fig. 12. Assignment area ROM.

The assignment area A_{mn}, originally defined to be spatially continuous, need not be, since the local coordinates of a new sample within it have a finite precision. If only three bits of precision are required, an assignment rule then would only need 2^3 by 2^3 or 64 discrete locations, as illustrated in Fig. 12. A 64 bit read-only memory could be used for this purpose.

In general, with b bits of precision, a maximum enlargement of 2^b could be accommodated. The memory needed to store assignment rules would be the number of assignment rules times $(2^b)^2$ bits. A total of 22 assignment rules were found necessary for a third-order system. Thus, only 2816 words of ROM need be used for enlarging up to 32 times.

In this study, the details of the telescoping template method were worked out only to a third-order system. While still preserving the corners and cusps, a fourth- or higher order system would yield even more satisfying smoothing of curves. Increased order of scaling may be complex in design, but when it is implemented with parallel hardware there is no loss of processing speed.

References

[1] G. W. Evans and R. L. Coswell (Rockwell Int. Corp.), U.S. Patent 4 029 947, June 14, 1977.
[2] "The Linotron 202: Better than anybody anticipated!" *Seybold Rep.*, vol. 7, July 17, 1978.
[3] P. J. M. Coueignoux, "Compression of type faces by contour coding," M.S. thesis, Massachusetts Inst. Technol., Cambridge, Jan. 1973.
[4] D. E. Knuth, "Mathematical typography," Dep. Comput. Sci., Stanford Univ., Stanford, CA, STAN-CS-78-648, Feb. 1978.
[5] T. S. Huang, "Coding of two-tone images," *IEEE Trans. Commun.*, vol. COM-25, pp. 1406–1424, Nov. 1977.
[6] Special Issue on Digital Encoding of Graphics, *Proc. IEEE*, vol. 68, pp. 757–924, July 1980; in particular, an extensive bibliography appears on pp. 923–924.
[7] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1978, ch. 4.
[8] D. E. Troxel and C. Lynn, "Enhancement of news photos," *Comput. Graphics Image Processing*, vol. 7, pp. 266–281, Apr. 1978.
[9] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 508–517, 1978.
[10] R. Pell, Type Design Expert, Rockwell MGD, Downers Grove, IL, private conversation, June 27, 1978.
[11] J. I. Beigeleisen, *Art Directors' Workbook of Type Faces.* New York: Arco, 1976.
[12] R. A. Ulichney, "Digital scaling of binary images," M.S. thesis, Massachusetts Inst. Technology, Cambridge, Jan. 1979.