Two-dimensional Haar thinning for data base compaction in Fourier proximity correction for electron beam lithography

M. E. Haslam, J. F. McDonald, D. C. King, M. Bourgeois, D. G. L. Chow, and A. J. Steckl
Rensselaer Polytechnic Institute, Center for Integrated Electronics, Troy, New York 12180-3590

(Received 28 June 1984; accepted 13 September 1984)

Fourier transforms have been successfully employed to produce fast, accurate calculations for dose precompensation in proximity effect correction for electron beam lithography. The principle advantage of this method is that it lends itself to implementation on special purpose computing hardware which is optimized for the generation of one- and two-dimensional Fourier transforms. However, the extremely accurate data base produced by the direct employment of the method is excessively large and, therefore, not practical. A controlled method for approximating this accurate solution is required which gives rise to the least mean-square error in the net exposure. This process of approximation can reduce the data base by "thinning" or selectively removing larger and larger features. In this paper we explore the use of two-dimensional Haar transforms for thinning the data base. An additional benefit of using Haar transforms is that they may also be computed with the same hardware accelerators which speed up the calculation of Fourier transforms.

I. INTRODUCTION

A previous paper[1] described a technique [which we term the Fourier precompensation (FP) method] that can be used to compute the adjusted doses for reducing the proximity effect in e-beam lithography. Kern[2] seems to have been the first to employ Fourier transforms in proximity correction. While the treatment in his paper is quite general, it appears he used the method in practice as an alternate means for repairing *regions* of artwork where the more traditional methods of pattern dimension adjustment (windage)[3] and pattern dose adjustment[4] failed to work properly. The accuracy of the Fourier precompensation approach suggests that a more widespread use of this technique might be feasible in which the FP method *completely* replaces the traditional calculations. Certainly a greater reliance on automated use of the FP method seems worth considering on the merits of its accuracy. Nevertheless, such calculations on ordinary Von Neumann computers are too slow to be attractive. On the other hand, the Fourier transform lends itself to implementation on fast, special purpose computing equipment termed attached array processors which accelerate this kind of calculation greatly. For example, the medium priced Floating Point Systems model 5430 can execute a two-dimensional Fourier transform for a 512×512 pixel pattern (262 144 points total) in only 0.5 s (ignoring disk transfer times). On the basis of this machine's performance the full processing time for proximity correction of patterns of typical dimensions would be on the order of 1 or 2 h depending, of course, on the actual size of the pattern and the minimum feature size.

II. THE NEGATIVE DOSE PROBLEM

One problem confronting the FP method is the presence of extremely large oscillatory positive *and* negative applied dose values which are to a large extent associated with the abrupt transitions between "on" and "off" regions of the desired pattern. The step discontinuities at the transitions become rounded by the proximity effect unless corrections are applied. The corrections attempt to force the original sharp edges to reappear in the net exposure. It is easy to see that some negative excursions in the applied dose are required to accomplish this. However, abrupt transitions also tend to excite an oscillatory Gibbs phenomenon in the corrected dose. Kern[2] suggested softening the discontinuities in the ideal dose by introducing a slope or bevel at the transition boundaries. The Gibbs phenomena, however, are extremely sensitive to the shape of this bevel. A somewhat more damped Gibbs oscillation is observed in the precompensated applied dose if the bevel is replaced by a smoother taper.[1] Ultimately an optimum taper can be developed for a given degree of proximity effect. To concentrate on minimizing the negative excursions an antisymmetric taper can be selected which emphasizes the smoothness of the lower portion of the transitions in the pattern.

In any case the negative excursions produced by the FP method never seem to completely disappear. Once the negative excursions are minimized by use of an optimum bevel or taper the question remains how best to handle the residual negative doses. Kern suggested clipping all negative doses to zero. This works in some cases but not always. Clipping introduces an uncontrolled nonlinear distortion whose effects are difficult to predict. The authors have suggested adding an offset sufficient to guarantee a positive dose everywhere. This latter method tends to adversely affect the beam on-time and to increase the complexity of the data base somewhat. Probably a combination of these two methods is best, but a formal method of deciding what clipping level or additive offset to use is not yet known.

III. EXTENSIONS TO TWO DIMENSIONS FOR THE FP CORRECTION METHOD

Previous papers by the authors on this subject have been limited to one-dimensional bar patterns for simplicity. Extensions to two-dimensional patterns are straightforward, but there are significant programming problems scaling up to the raw data base which arises in two-dimensional pat-

 0734-211X/85/010165-09$01.00 © 1985 American Vacuum Society

terns of practical size, and a different approach is required. Ignoring the discrete nature of the computer calculation for the moment, we may write the effective exposure $E(x, y)$ in the resist resulting from the applied dose $G(x, y)$ using the two-dimensional convolution:

$$E(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x',y')F(x-x',y-y')\,dx'\,dy', \quad (1)$$

where $F(x, y)$ is the two-dimensional point spread function resulting from an infinitesimally small beam spot [which can be represented by $G(x, y) = \delta(x, y)$]. This convolution is a linear transformation which has an inverse. The Fourier precompensation method amounts to convolving the entire desired pattern $E(x, y)$ with a precompensating filter $H(x, y)$:

$$G(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(x',y')H(x-x',y-y')\,dx'dy', \quad (2)$$

where $H(x, y)$ is the inverse Fourier transform of $h(j\omega_1, j\omega_2)$ which is the reciprocal of the Fourier transform $f(j\omega_1, j\omega_2)$ of the point spread function $F(x, y)$. We will find this form of the FP method advantageous for handling large patterns.

When programmed, these steps must be carefully discretized and specialized for patterns of finite extent. This leads to the use of the discrete Fourier transform (DFT) with a change of spatial dimension units to accommodate integer indexing:

$$f_{pq} = \sum_{x=0}^{L-1}\sum_{y=0}^{M-1} F(x,y)e^{-2\pi j(px/L + qy/M)}, \quad (3)$$

with similar expressions for g_{pq}, e_{pg}, and h_{pq}. For the reader unfamiliar with digital signal processing we note that the discretized version of the continuous Fourier transform of Eq. (1)

$$e_{pq} = g_{pq} f_{pq}, \quad (4)$$

is *not* the DFT for the discretized *linear* convolution corresponding to Eq. (1) which would be

$$E(x,y) = \sum_{x'=0}^{L-1}\sum_{y'=0}^{M-1} G(x',y')F(x-x',y-y'). \quad (5)$$

Instead, Eq. (4) is the DFT corresponding to a *circular* convolution[5,6]

$$E(x,y) = \sum_{x'=0}^{L-1}\sum_{y'=0}^{M-1} G(x',y)F(|x-x'|_L, |y-y'|_M), \quad (6)$$

where the vertical bars are the modulus function.

This latter expression includes a wrap-around effect at the ends of the convolution interval which is due to the modulus in Eq. (6). This wrap-around effect may be avoided for signals of finite duration by extending the signal data set size and padding with zeros. For example, if G is of finite extent in the interval $[0, L_1 - 1]\times[0, M_1 - 1]$ and F is of finite extent in the interval $[0, L_2 - 1]\times[0, M_2 - 1]$ then the wrap-around effect in Eq. (9) may be avoided by extending both data set sizes to

$$[0, L_1 + L_2 - 2]\times[0, M_1 + M_2 - 2],$$

and setting the extra trailing array spaces to zero in both cases. The resulting data set is big enough to prevent the wrap-around effect from altering the desired data and just large enough to contain all of the potentially nonzero results.

In executing the FP correction as in Ref. 1, it would be tempting to execute a single large transform for each of the functions of interest in order to arrive at the discretized implementation of Eq. (4). While simple in concept,[7] the execution of such a single large transform is impractical even on the most powerful of the available array processors. For example, the Floating Point Systems 164/MAX has only 16×10^6 (64 bit) words of fast directly accessible memory when maximally configured. Allowing for storage of real and imaginary parts of all four variables g_{pq}, e_{pq}, h_{pq}, and f_{pq} the typical image or pattern size would only be 1000×1000 pixels. Furthermore, the point spread function $F(x, y)$ and its transform f_{pq} really do not require such a large data set size. Therefore, the most straightforward approach is inefficient.

To cope with the large number of pixels required to specify the typical patterns of interest it becomes necessary to "stitch" together the results of the many small subregions. Note that these regions can be as small as a single memory cell pattern, or as large as a significant portion of the pattern for a large chip. In effect, Eq. (4) represents a convolution of a very large pattern with an inverse precompensating filter $H(x, y)$ which has a fairly small data set. The stitching is accomplished using either the overlap-and-add or overlap-and-save techniques of digital signal processing theory.[5] For the overlap-and-add method the desired pattern $E(x, y)$ is sectioned into a mosaic of functions $E_{rs}(x, y)$ each of which is zero outside of rectangular *nonoverlapping* regions (or "tiles") of dimension L_1 by M_1, and such that the union of these intervals covers the entire pattern. Assuming that there are RS of these subregions we can write

$$E(x,y) = \sum_{r=0}^{R-1}\sum_{s=0}^{S-1} E_{rs}(x,y), \quad (7)$$

where

$$E_{rs}(x,y) = \begin{cases} E(x,y) & rL_1 \leqslant x < (r+1)L_1, \text{ and } sM_1 \leqslant y < (s+1)M_1 \\ 0 & \text{elsewhere} \end{cases}. \quad (8)$$

The single large convolution to produce G is then written as follows:

$$\begin{aligned} G(x,y) &= \sum_{x'=0}^{RL_1-1}\sum_{y'=0}^{SM_1-1} E(x',y')\,H(x-x',y-y') \\ &= \sum_{x'=0}^{RL_1-1}\sum_{y'=0}^{SM_1-1} \left[\sum_{r=0}^{R-1}\sum_{s=0}^{S-1} E_{rs}(x',y')\right] H(x-x',y-y') \\ &= \sum_{r=0}^{R-1}\sum_{s=0}^{S-1}\left[\sum_{x'=0}^{RL_1-1}\sum_{y'=0}^{SM_1-1} E_{rs}(x',y')H(x-x',y-y')\right] \\ &= \sum_{r=0}^{R-1}\sum_{s=0}^{S-1} G_{rs}(x-rL_1, y-sM_1), \end{aligned} \quad (9)$$

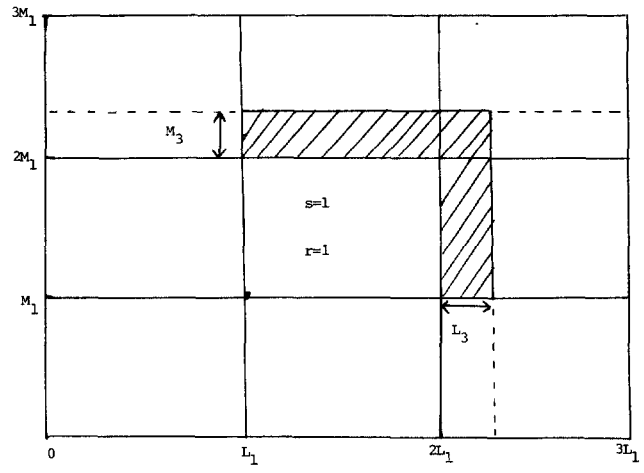


FIG. 1. Floor plan for subdividing the overall pattern into $S \times R$ regions or "tiles." Stitching regions for the overlap-and-add method are shown hatched for the tile with $s = 1$ and $r = 1$.

where, recalling that $E_{rs}(x, y)$ is nonzero only as shown in Eq. (8), we can write

$$G_{rs}(x, y) = \sum_{x'=rL_1}^{(r+1)L_1-1} \sum_{y'=sM_1}^{(s+1)M_1-1} E_{rs}(x', y')H(x + rL_1 - x', y + sM_1 - y')$$
$$= \sum_{x''=0}^{L_1-1} \sum_{y''=0}^{M_1-1} E_{rs}^{T}(x'', y'')H(x - x'', y - y''). \quad (10)$$

Here,

$$E_r^T(x, y) = E_{rs}(x + rL_1, y + sM_1) \quad (11)$$

is a copy of $E(x, y)$ for the rs^{-th} mosaic subregion (or tile) translated down to the origin of the x, y plane. Equation (10) is a conventional linear convolution upon the translated tile whose nonzero extent is $[0, L_1 - 1] \times [0, M_1 - 1]$.

The nonzero extent of H will also be assumed finite and confined to the region $[0, L_3 - 1] \times [0, M_3 - 1]$. Note that M_3 might be different from M_2. Following the previous discussion for Eqs. (5) and (6) we see that each $G_{rs}(x, y)$ must be calculated in a region

$$[0, L_1 + L_3 - 2] \times [0, M_1 + M_3 - 2],$$

in order to accomodate all the nonzero results which can potentially arise.

When the data on these tiles are retranslated by Eq. (11) back to the proper position we find that the new tiles are bigger than the original ones, and they will not fit in a nonoverlapped manner into the same spaces. Hence, the new tiles must overlap as shown in Fig. 1. In the overlapped regions the results from each tile simply add together as shown in Eq. (9). Therefore, the stitching of the tiles is straightforward in the FP method, and the convolutions for each of the tile regions of Eq. (10) can be executed with fast Fourier transform (FFT) hardware. Observe that one must recheck the positivity of the FP dose in the stitching regions. Note also that several tile overlap regions may contribute to the stitching corrections at a given point when these overlap regions intersect. It is essential that these stitching corrections be included since features in these regions might otherwise become distorted or lost. This is because of the "spill" of the proximity effects from one tile to the next.

It is normally important to make the tiles of the original sections for $E(x, y)$ in Eq. (8) have as *large* a dimension as fast memory storage permits using a given computing system, since the FFT operations will then be more efficient. However, in highly periodic cellular spatial patterns such as are found in memories there are obvious advantages to using the method with smaller tiles (which are the cells) since after FP correction the results are simply superposed using Eq. (9) taking into account the proper stitching corrections. There is also no reason why the tiles all have to be of the same size.

IV. HAAR THINNING

The data base for $G(x, y)$ produced by the FP method contains far too much detail to be practical, and a method to reduce this complexity (or "thin" it) is needed. The process of simplifying the data base is equivalent to approximating the FP dose while controlling the error in the exposure $E(x, y)$. For lithography the pointwise error made in $E(x, y)$ during this approximation process is crucial since an error at any point (x, y) which exceeds the developer threshold will obviously change the final result. We also note that the concept of a threshold (which is nearly universally used in most commercial two-dimensional proximity correction codes) may itself be an imperfect model of the behavior of the development process, especially at submicron dimensions. Nevertheless, the uniformity produced by the FP method is still advantageous.

Unfortunately, a fast, convenient method for monitoring this pointwise error during the approximation process is unknown. Instead, the mean-square error is monitored. This is because approximations involving the deletion of terms in an orthonormal expansion introduce a mean-square error which can be directly computed from the expansion coefficients of the deleted terms.

Walsh functions have been studied for thinning[1] but a comparison with the Haar transform suggests that the Haar transform produces superior results.[8] In either case the "thinning" process consists of deleting some of the basis functions in the transforms. The orthonormal pulselike Haar basis functions $H(m, x/L)$ are defined by[9]

$$\mathrm{HAR}(0, x/L) = 1 \quad 0 \leqslant x/L \leqslant 1 \quad (12)$$

$$\mathrm{HAR}(1, x/L) = \begin{cases} 1 & 0 \leqslant x/L < 1/2 \\ -1 & 1/2 \leqslant x/L \leqslant 1 \end{cases} \quad (13)$$

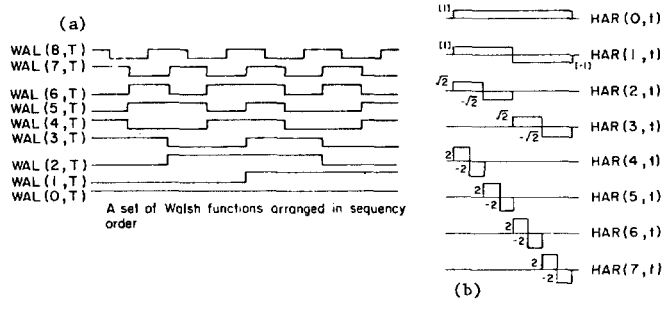


FIG. 2. The first few (a) Walsh and (b) Haar basis eigenfunctions.

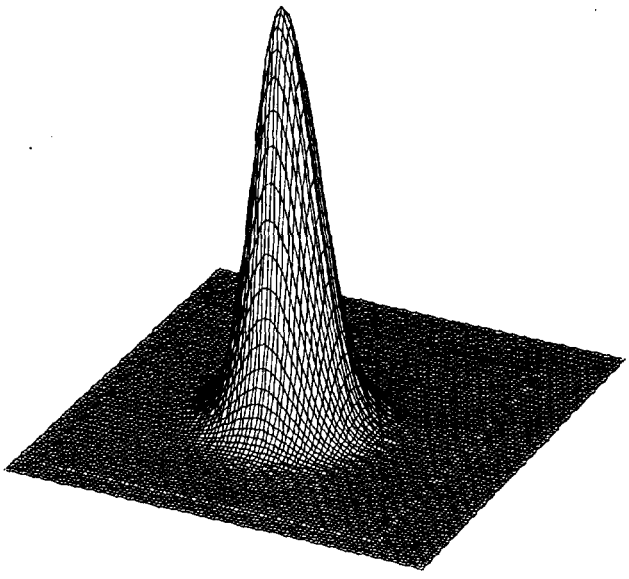


FIG. 3. Point spread function with $B_f = 0.10\,\mu$, $B_b = 2.00\,\mu$, and $\eta = 0.95$. The square base of the figure is $1.0\,\mu$ on a side.

and

$$\mathrm{HAR}(2^\sigma + n, x/L), \quad n = 0, \ldots, 2^\sigma - 1$$
$$= \begin{cases} 2^\sigma & (n/2^\sigma) \leqslant x/L < (n + 1/2)/2^\sigma \\ -2^\sigma & (n + 1/2)/2^\sigma \leqslant x/L < (n + 1)/2^\sigma, \\ 0 & \text{elsewhere} \end{cases}$$

for x taking discrete values in the interval $[0, L]$. The Haar basis functions are collected together by this notation into groups fixed by the size parameter σ, which sets the width of the pulses in a given group to $1/2^\sigma$ (see Fig. 2). Deleting all basis functions with a pulse group size index greater than σ_1 limits the spatial feature resolution of the overall result to $1/2^{\sigma_1}$. This provides one measure of the degree of thinning accomplished by deleting an entire group. Of course, if deletion of a specific basis function introduces a large error one must retain it regardless of its group size σ. The index $2^\sigma + n$ is termed the "sequency" number for a given Haar basis function.

The two-dimensional Haar transform pair can be written as follows:

$$G(x, y) = \sum_{p=1}^{L-1} \sum_{q=1}^{M-1} G_{pq}\, \mathrm{HAR}(p, x/L)\, \mathrm{HAR}(q, y/M), \tag{14}$$

$$G_{pq} = \frac{1}{LM} \sum_{x=1}^{L-1} \sum_{y=1}^{M-1} G(x, y)\, \mathrm{HAR}(p, x/L)\, \mathrm{HAR}(q, y/M). \tag{15}$$

The most likely point in time for application of this transformation would be the precompensation for a tile after overlap-and-add stitching corrections have been performed. For best results it has been found that the transitions or step discontinuities should be spatially synchronized with the transitions in the desired pattern as much as possible.[8] A Haar function which "straddles" one of the pattern transitions creates large errors if deleted. On the other hand retaining these straddle basis functions causes the data base to enlarge. Of course, not all transitions will be synchronized and a shift of the transform origin becomes part of the procedural adjustments to maximize synchronization.

V. THE EFFECTIVE EXPOSURE ERROR

The Haar functions are orthogonal and form a complete basis for square integrable functions in the Lebesgue sense. Consequently, the derivation provided in Ref. 1 for Walsh thinning still applies and the mean-square $\bar{\epsilon}_N^2$ induced in the final exposure by the thinning operation is bounded by

$$\bar{\epsilon}_N^2 \leqslant \Big(\sum_{p,q \in \Psi} \sum |G_{pq}| \Big)^2, \tag{16}$$

where Ψ is the list of indices corresponding to the deleted terms. Because of the pulselike character of the Haar basis functions the error caused by a deletion is confined to the interval where the pulse is nonzero. By comparison, all of the Walsh basis functions are nonzero everywhere and deleting any one of them from the complete expansion introduces errors throughout the entire data base. The data compression observed using the Walsh transform occurs primarily in the sequency transform space. The data base compression is less powerful in the spatial domain using Walsh functions than with the Haar functions. For this reason the short duration pulselike character of the Haar basis functions gives a more nearly pointwise interpretation of the mean-square approximation error in $E(x, y)$ and the data base compaction in the spatial domain is more evident.

As in the case of the Walsh transform, we note that for the Haar transform (or, in fact, any orthogonal expansion) the

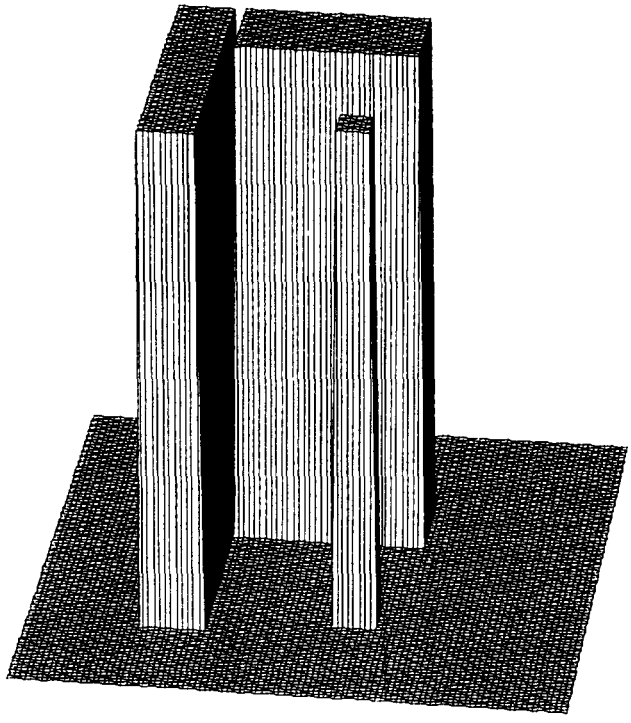


FIG. 4. Two-dimensional desired net exposure pattern for comparison with other *space domain* figures. The narrow gap is $0.20\,\mu$ in width, and the small square is $0.50\,\mu$ on each side. The square base in this and all subsequent 3D perspective *space domain* plots is $10.0\,\mu$ on a side. That is, the horizontal scale is ten times that shown in Fig. 3.

coefficients remaining after the deletion of some of the terms need not be readjusted. They are already optimal under the least mean-square error criterion and any adjustment will only increase the mean-square error further.

To avoid problems in the vicinity of the stitching regions the thinning process should be made less aggressive where the tiles abut to insure more accuracy. This assures that any differences in the independent thinning processes which occur in two adjacent tiles won't clash at the boundaries of the tiles.

Additional errors arise if the applied dose is quantized. To be absolutely certain that the deviations of the net exposure after thinning (and dose quantization) do not cause an actual crossing of the developer threshold, a simulation using Eq. (1) is advisable. If, in fact, the developer margin is too small in some region then some of the Haar coefficients may have to be reintroduced in the vicinity of the occurrence of the insufficient margin.

The preference for the Haar transform here is motivated in part by the fact that fast Haar transform (FHT) algorithms usually lend themselves to implementation on the same computing hardware which accelerates the fast Fourier transform (FFT). In addition, whereas the Fourier transform for a block of data of length n requires $n \log_2(n)$ operations, the Haar transform is linear in n and therefore, even faster than the FFT, leading to extremely fast thinning times.[8]

One final comment is in order concerning the amount of working disk storage required to implement the FP method followed by thinning. The largest amount of storage is required when executing the FP calculations for neighboring tiles. Pixel level representations for the pattern are only required during these computations. If thinning or data base compaction is performed immediately, it is never necessary to store the *entire* pattern at the pixel level (which would certainly require hundreds of megabytes of disk storage).

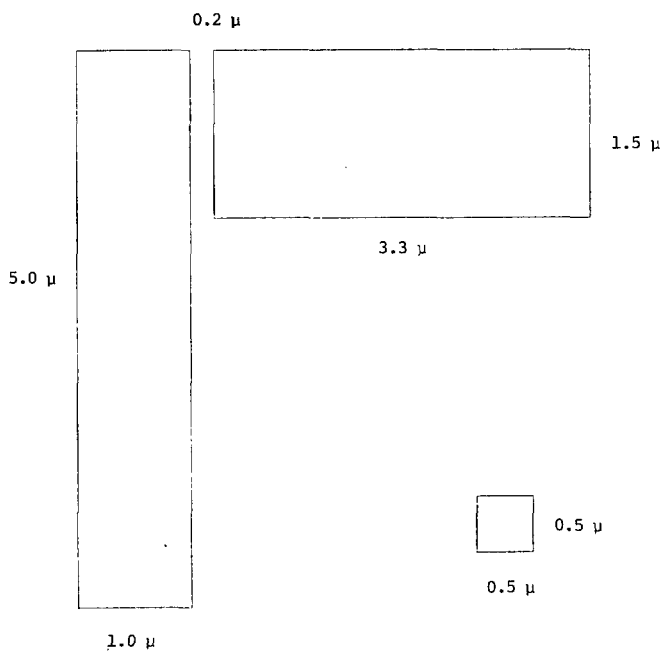


FIG. 5. Plan view projection of the pattern features corresponding to Fig. 4.

VI. COMPUTED RESULTS

For the purposes of discussion we have arbitrarily adopted parameters for a point spread function measured at IBM for the EL-2 using 0.6 μ of polymethyl methacrylate (PMMA) on a silicon substrate with a beam acceleration voltage of 25 keV. Here[11,12]

$$F(x,y) = \exp[-(x^2+y^2)/B_f^2] + \eta(\alpha^2/\beta^2)\exp[-(x^2+y^2)/B_b^2], \quad (17)$$

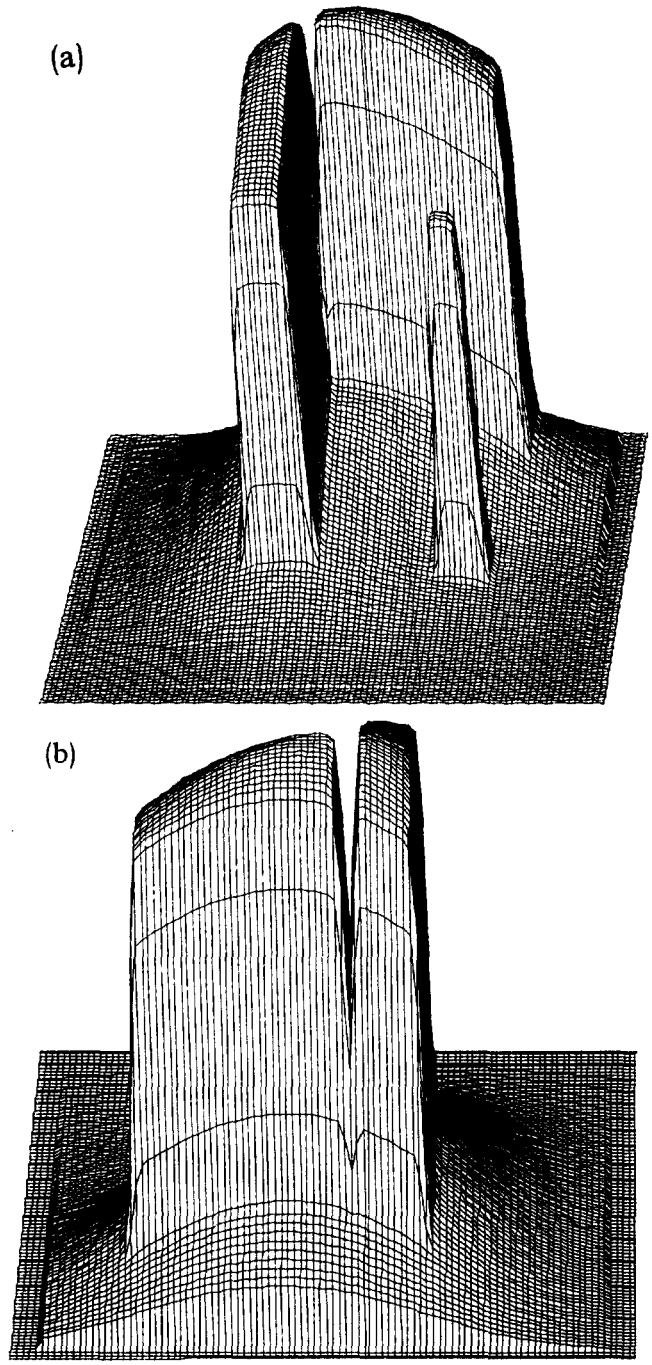


FIG. 6. (a) Unprecompensated net exposure obtained by convolving Fig. 4 with Fig. 3 according to Eq. (1) (front view). Note that the small 0.50 μ square feature is underdosed relative to the other features primarily due to *back* scattering near the larger features which tends to increase their local exposure levels. (b) Rear view of (a) showing the narrow 0.20 μ gap has nearly disappeared primarily due to *forward* scattering.

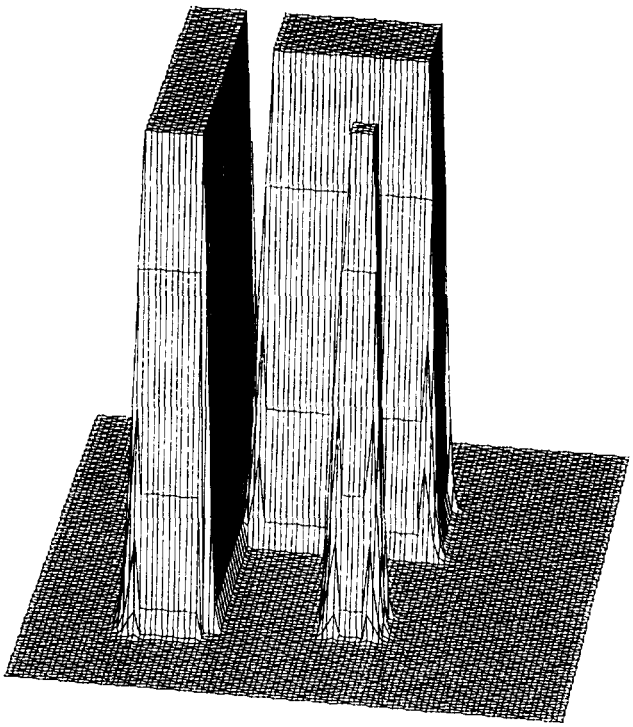


FIG. 7. Original desired dose of Fig. 4 with tapers and aprons affixed.

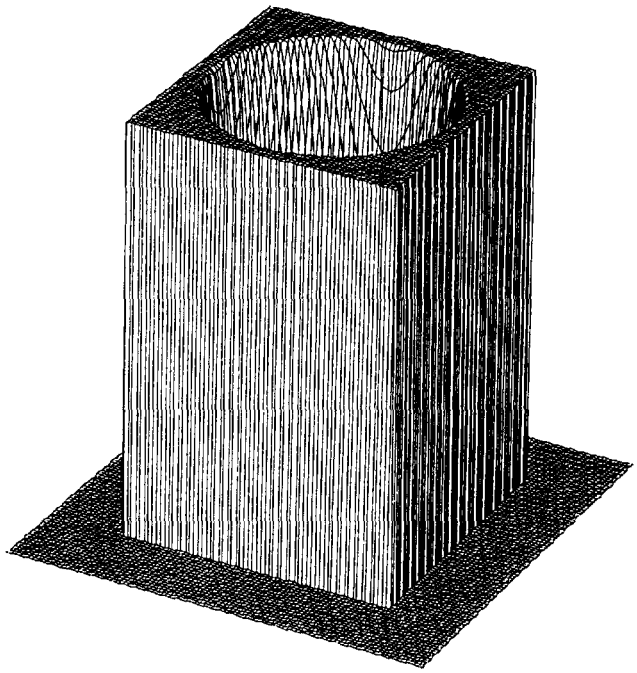


FIG. 8. Low-pass spatial filter used to partially suppress some of the more extreme Gibbs oscillations. The *spatial frequency domain* origin may be taken to be at the lower left corner. The highest spatial frequencies in both the x and y direction are located at the center of the hole and are $10.0\,\mu^{-1}$, respectively. Negative spatial frequencies are reflected symmetrically in the figure about its center as is the usual case when using FFT software.

where the first term is due to the forward scattering in the resist and the second is due to backscattering in the substrate. (We have taken[15] $B_f = 0.10\ \mu$, $B_b = 2.00\ \mu$, and $\eta = 0.95$). This is plotted in Fig. 3.

Figure 4 shows the two-dimensional desired pattern which will be examined throughout the rest of the figures. The two critical features consist of the narrow $0.20\ \mu$ gap between the two large rectangles, and the small $0.50\ \mu$ square. This pattern contains features which are extremely challenging to reproduce. Figure 5 contains the plan projection view giving detailed dimensions. With no precompensation the convolution of Eq. (1) produces the net exposure shown in Figs. 6(a) and 6(b). The narrow gap has nearly disappeared, and the small square is insufficiently dosed to have these two features develop to proper size simultaneously.

Next, following the conclusions of Ref. 1 the edges and corners of the patterns have a taper or apron affixed to all edges to minimize the Gibbs oscillations, especially the negative ones. This is shown in Fig. 7. The dose taper consists of a fillet whose shape has been taken as

$$F(\xi) = \left\{\frac{1}{2} + \frac{1}{2}\cos[(\xi - \xi_0)/D]\right\}^{\rho}, \tag{18}$$

where ξ is the distance from the onset of the ideal dose downward transition, ξ_0 is an offset into the feature designed to place the maximum downward slope at the location of the transition, D is the taper width, and ρ is typically 3 for the feature sizes discussed here.

Next, the inverse precompensation filter h_{pq} is computed and further multiplied by the low-pass characteristic[13,14] shown in Fig. 8. The resulting low-passed precompensating filter is shown as a magnitude plot in the space frequency domain in Fig. 9. The corresponding picture of $H(x, y)$ for Fig. 9 is shown in Fig. 10.

Figure 11 shows the FP corrected applied dose (without the additive offset). After Haar processing we obtain Fig. 12 for the thinned precompensated dose. The plan views of these two pictures showing the regions of constant dose are shown in Figs. 13 and 14, illustrating the amount of data

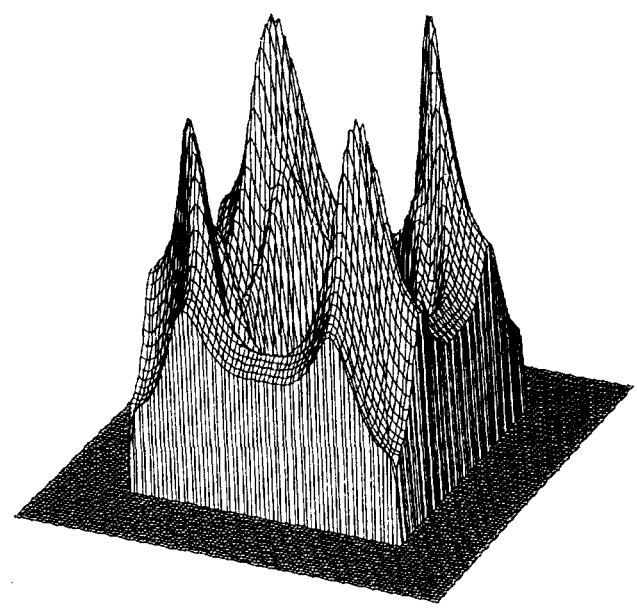


FIG. 9. Low-pass filtered precompensation filter $H(\omega_1,\omega_2)$ corresponding to $H(x, y)$ shown in the *spatial frequency domain* using the same coordinate system used in Fig. 8.

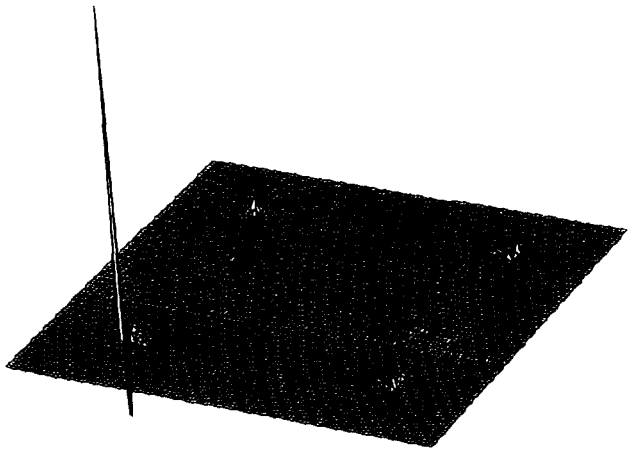


FIG. 10. Two-dimensional inverse FFT of Fig. 9 showing $H(x, y)$. Note the large dynamic range implied.

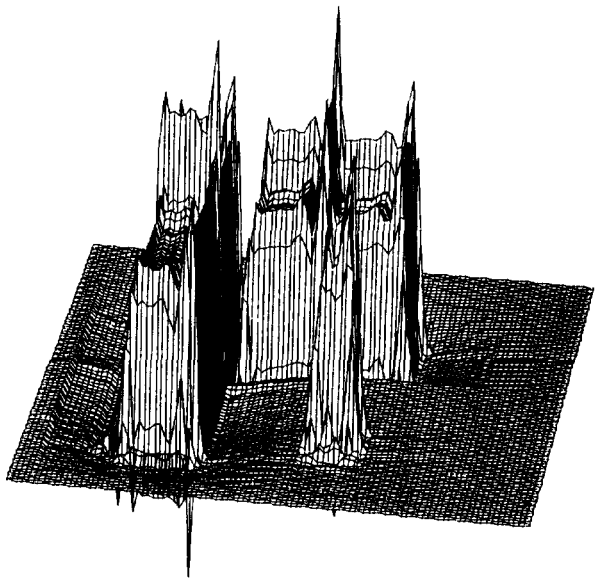


FIG. 11. The FP corrected applied dose shown in the *spatial domain* corresponding to Fig. 4 (with no additive offset or thinning).

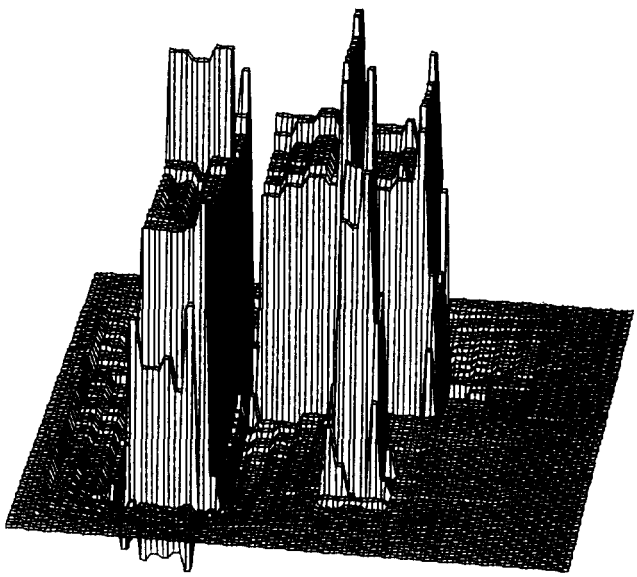


FIG. 12. Result of Haar thinning of the precompensated dose shown in Fig. 11.

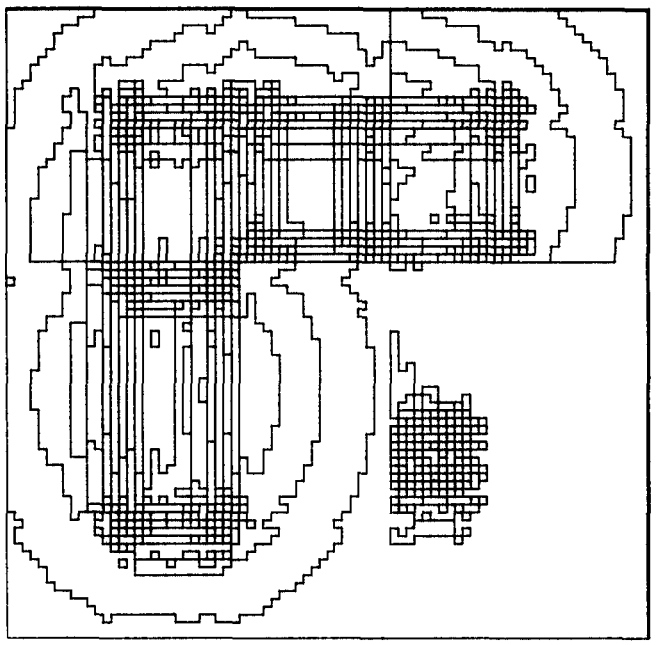


FIG. 13. The plan view of the unthinned FP applied dose of Fig. 11, showing regions of constant dose, 0.1 μ resolution. The outer square in this and all subsequent plan views is 8.00 μ on a side.

base reduction possible using the 2D Haar transform. Additional improvements in the thinning strategy may reduce this complexity even further. The net exposure resulting from the thinned precompensated dose is shown in Figs. 15(a) and 15(b). These can be compared with Figs. 6(a) and 6(b).

Now for many of the currently available commercial electron beam systems the applied dose must be further quantized to 4 or 5 bit numbers with a separate exposure pass required for each bit plane. Figure 16 shows the thinned, quantized precompensated applied dose using 4 bit quantization (16 possible dose levels). The plan view corresponding

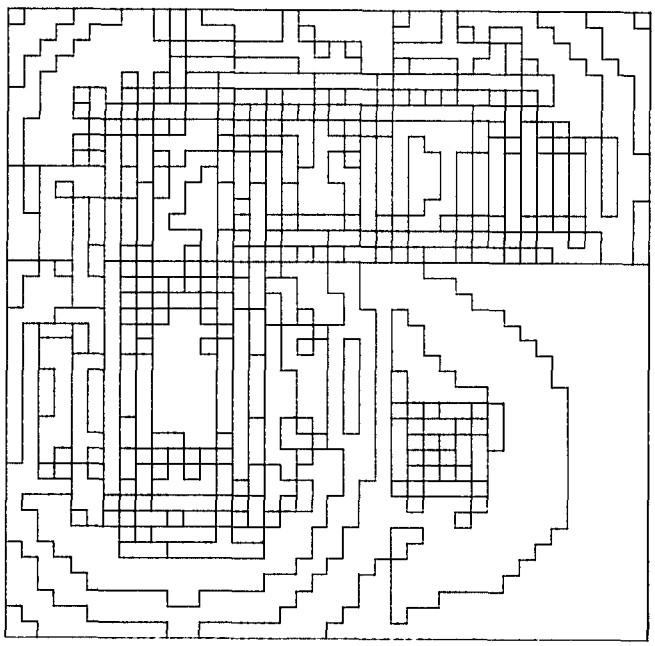
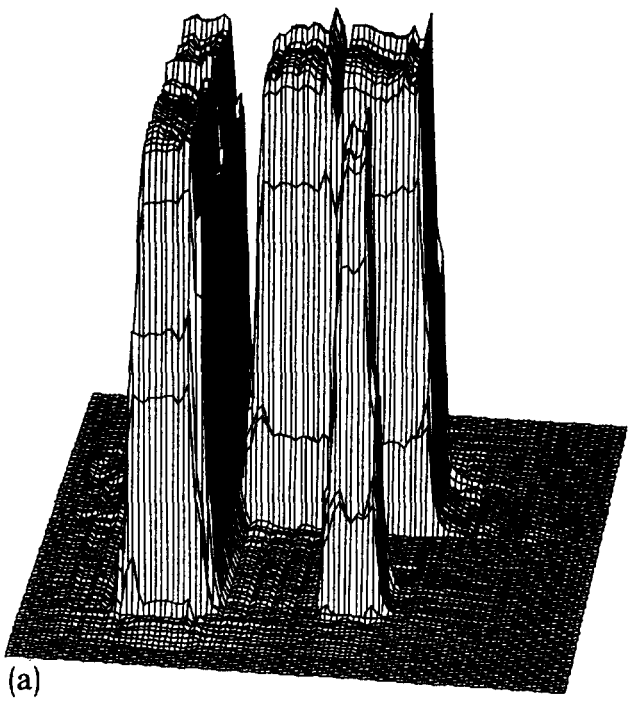


FIG. 14. The plan view of the thinned applied dose with σ_1 set to reject all but 0.2 μ sized basis functions.



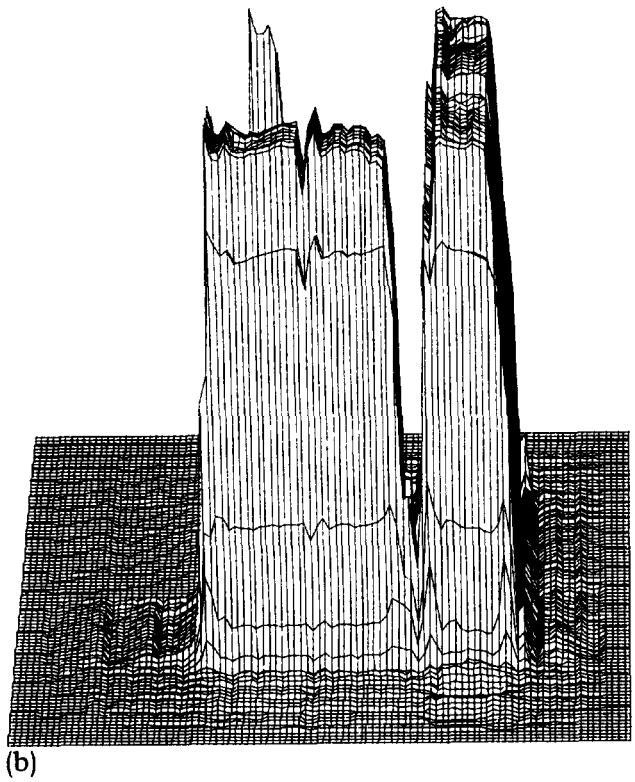


FIG. 15. (a) Net exposure resulting from the thinned FP applied dose shown in Fig. 12 (front view) corresponding to Fig. 6(a). (b) Net exposure resulting from the thinned FP applied dose shown in Fig. 12 (rear view) corresponding to Fig. 6(b).

to this figure is Fig. 17 which shows the additional dose simplification which can result from this quantization. Finally, Fig. 18 shows the degradation of the net exposure resulting from quantization. Figure 19 shows the most complex bit plane (2′) for Fig. 17.

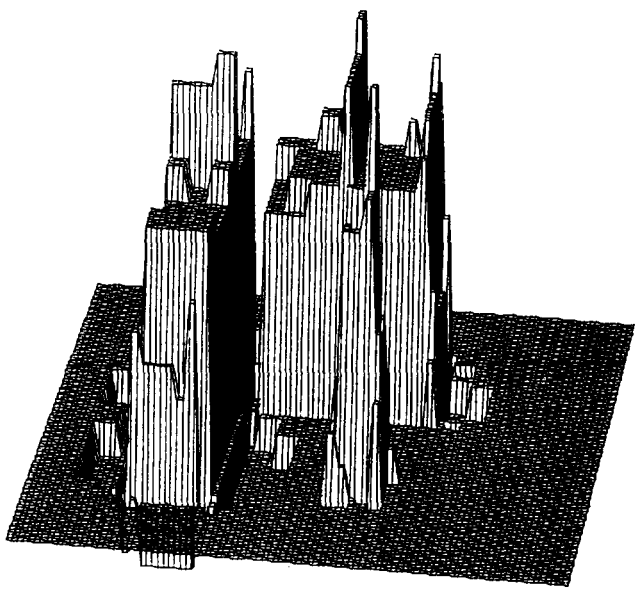


FIG. 16. Thinned precompensated applied dose quantized to 4 bits.

VII. CONCLUSIONS

A procedure for implementing the 2D FP proximity correction method for patterns of large size has been described which utilizes the principle of linear superposition and the overlap-and-add method for stitching. The precompensated dose is then thinned using the 2D Haar transform using the least mean-square error as an error monitor. An *ad hoc* edge transition taper had been introduced for minimizing Gibbs oscillations, especially the negative excursions. The design of a truly optimal edge taper and the optimization of the low-

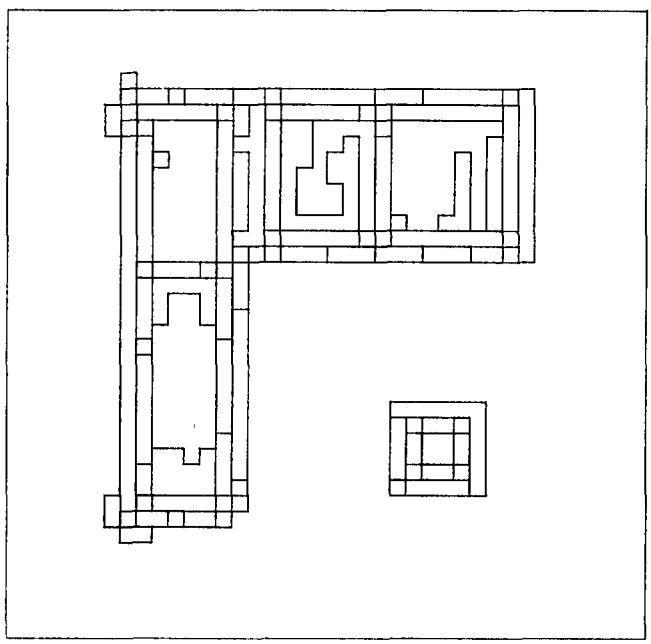


FIG. 17. Plan view corresponding to Fig. 16.

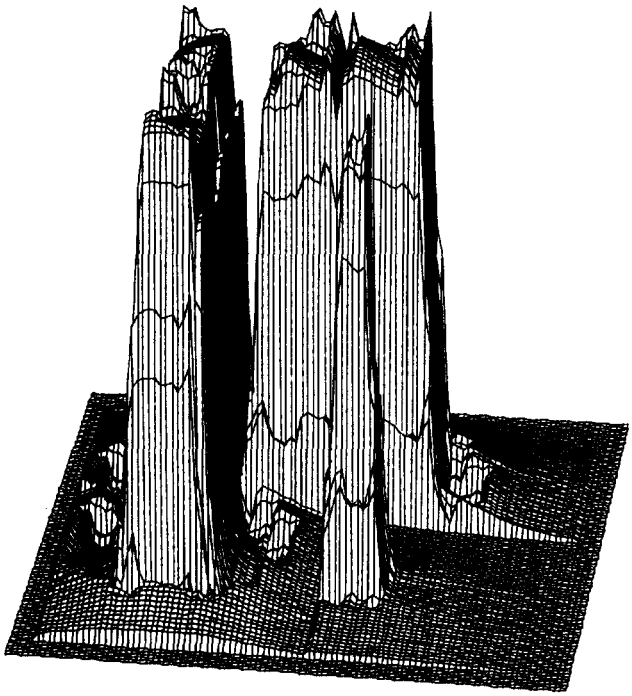


FIG. 18. Degradation caused by 4 bit quantization (front view).

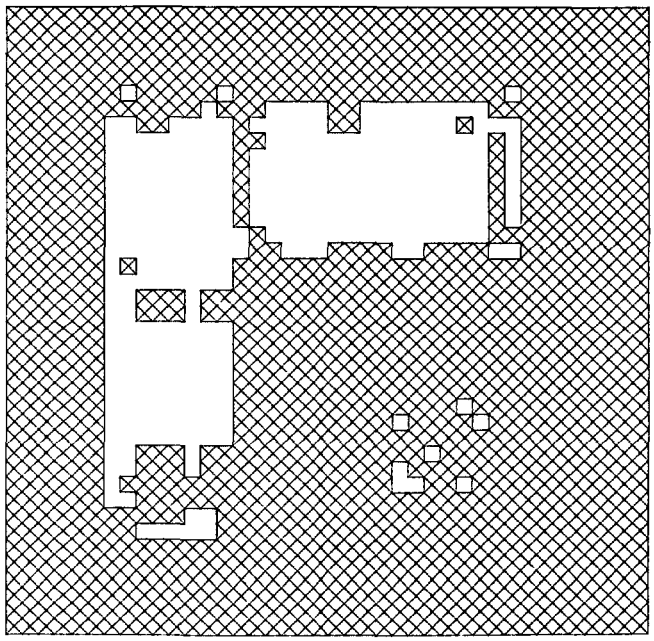


FIG. 19. Most complex bit plane—the second lowest bit (2^1)—of the four bit planes corresponding to the dose of Fig. 17 plus an offset for the negative dose.

pass characteristics of the spatial filter are now seen to be of paramount importance in controlling the residual Gibbs oscillation and hence the complexity of the resulting thinned and quantized data base. We have also observed that these techniques can be used effectively to deal with patterns containing highly periodic features such as memory cells.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to IBM, especially the East Fishkill division, for their donation of the EL-2 electron beam lithography machine to the Rensselaer Polytechnic Institute Center for Integrated Electronics (CIE) and for their constant assistance, encouragement, and support. Special mention of thanks is due to Mr. Ed Weber and Mr. Joe Giuffre at East Fishkill for providing us with accurate point spread function parameters for EL-2 in the context of the article. The authors also acknowledge the strong financial support of the Semiconductor Research Cooperative (SRC), Dr. B. Agusta, research monitor, under agreement No. 83-01-041. Finally, the authors wish to acknowledge the assistance of other colleagues at CIE, and the general support, assistance, and encouragement of the CIE corporate sponsors.

[1]D. G. L. Chow, J. F. McDonald, D. C. King, W. Smith, K. Molnar, and A. J. Steckl, J. Vac. Sci. Technol. B **1**, 1383 (1983).
[2]D. Kern, *Proceedings of the 9th International Conference and Ion Beam Science and Technology*, edited by R. Bakish (The Electrochemical Society, Princeton, 1980).
[3]H. Sewell, J. Vac. Sci. Technol. **15**, 927 (1978).
[4]M. Parikh, J. Appl. Phys. **50**, 4371 (1979).
[5]L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Prentice–Hall, Englewood Cliffs, New Jersey, 1973).
[6]A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice–Hall, Englewood Cliffs, New Jersey, 1975).
[7]R. C. Singleton, IEEE Trans. Audio Electroacoust., AU-**15**, 91 (1967).
[8]D. G. L. Chow, J. F. McDonald, D. C. King, and A. J. Steckl, *Proceedings of the 1983 Microcircuit Engineering Conference*, edited by H. Ahmed, J. R. A. Cleaver, and G. A. C. Jones (Academic, London, 1983), p. 65.
[9]K. G. Beauchamp, *Walsh Functions and Their Applications* (Academic, New York, 1975).
[10]J. E. Shore, IEEE Trans. Commun. (1973), p. 209.
[11]T. H. P. Chang, J. Vac. Sci. Technol. **12**, 1271 (1975).
[12]G. R. Brewer, *Electron-Beam Technology in Microelectronic Fabrication* (Academic, New York, 1980).
[13]D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing* (Prentice–Hall, Englewood Cliffs, New Jersey, 1984).
[14]V. Cappellini, A. Constantinides, and P. E. Miliani, *Digital Filters and their Applications* (Academic, New York, 1978).
[15]J. Giuffre, IBM East Fishkill Division (personal communication).