Optimal Digitization of 2-D Images

Nielsen, Lars; Åström, Karl Johan; Jury, Eliahu I

1983

Document Version:
Publisher's PDF, also known as Version of record

CODEN:LUTFD2/(TFRT-7265)/1-023/(1983)

OPTIMAL DIGITIZATION OF 2-D IMAGES

L NIELSEN
K J ÅSTRÖM
E I JURY

DEPARTMENT OF AUTOMATIC CONTROL
LUND INSTITUTE OF TECHNOLOGY
NOVEMBER 1983

LUND INSTITUTE OF TECHNOLOGY
DEPARTMENT OF AUTOMATIC CONTROL
Box 725
S 220 07 Lund 7 Sweden

Document name
REPORT

Date of issue
November 1983

Document number
CODEN:LUTFD2/(TFRT-7265)/1-023/(1983)

Author(s)
L Nielsen*, K J Åström*, E I Jury**

Supervisor

Sponsoring organization
The Swedish Board of Technical Development

Title and subtitle
OPTIMAL DIGITIZATION OF 2-D IMAGES

Abstract
A theoretical formulation of the optimal digitization problem is given. The problem is to represent an image by M x N samples and b bits per sample. The constraint is a fixed number of bits, i.e. $M \cdot N \cdot b$ = constant. In this paper the formulas for the individual numbers M, N, and b are obtained as solutions to an optimization problem. The solution is tested experimentally and agrees well with human visual quality. An advantage is that the solution is given in closed form. This makes it easy to use as a rule of thumb. It also clearly points out the dependence of image characteristics. This dependence explains and agrees with what is found in other subjective tests.

* Department of Automatic Control, Lund Institute of Technology, P.O. Box 725, S-220 07 Lund 7, Sweden.

** Research Professor, University of Miami, Coral Gables, Florida, USA.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
23

Recipient's notes

Security classification

1. INTRODUCTION

During the last two decades vigorous research activities have been devoted to image processing and related fields. The interest in these problems stem from the practical applications of digital signal processing as applied to images. Several books have been written on several aspects of the theory and practical applications [1-3]. The problem of optimal digitization of 2-D images have been sporadically mentioned in several texts, but has not been addressed in full details [2-4]. In the works of Abdou and Wong [5] experimental study of the effect of coarse scan/fine print for bilevel images has been initiated. However, a theory is still lacking in these investigations. For the 1-D case, Steiglitz [6] has presented a detailed theory of transmission of an analog signal over a fixed bit-rate channel. This work has motivated the extension to the 2-D case in this paper.

In this paper a definition of optimal digitization (quantization and sampling) is given for the first time. This definition is also meaningful in practical applications. Other forms of optimization based on filtering of images have been investigated in the literature. Among such methods are the works of Lebedev and Markin [7] which are based on multicategory Wienerfilter and the works of Woods [8-9] based on the variation of Lebedev-Markin process and using 2-D Kalman filters. Both the objectives and the imposed constraints of these optimization procedures are different from the one introduced in this paper.

Having introduced the optimal definition, the theoretical formulation of the problem is discussed in Section 2. The optimal solution is derived in Section 3. The theoretical studies are illustrated in Section 4 by a computer experiment which gives insight into the properties of the optimization procedure. These experiments are also used to verify that the optimization algorithm is somewhat suitable for this problem. The results are summarized in the conclusions in Section 5.

2. PROBLEM FORMULATION

Let the original image f(x,y) be a function defined on

$$\Omega = [\,0,\ L_x] \times [\,0,\ L_y] \subset R^2$$

and with values in $V = [\min f,\ \max f] \subset R^+$.

The image is registered by a video camera and then sampled to give a sampled image $\hat{f}(x_i,y_j)$ defined on a M x N rectangular grid G with values in $V \subset R^+$. Ideal sampling is assumed, i. e.

$$\hat{f}(x_i, y_j) = f(x_i, y_j) \qquad \text{for } x_i, y_j \in G.$$

In addition to sampling the values of $\hat{f}(x_i,y_j)$ are also quantized so that $V \subset R^+$ is represented by b bits i.e. 2^b quantization levels. The quantization of $\hat{f}(x_i,y_j)$ is denoted by $Q\hat{f}(x_i,y_j)$. This is the digitized image defined on a M x N grid with discrete values.

We want to reconstruct the new function $\tilde{f}(x,y)$ defined on Ω with values in V. From the function $Q\hat{f}(x_i,y_j)$ the function $\tilde{f}(x,y)$ can be obtained using many different interpolation schemes [1,5,6].

The optimal digitization problem can be simply formulated as follows:

Assume that $M \cdot N \cdot b$ is constant. That is there exists a fixed number of bits to represent an image. Determine the individual values of M, N, b so that the error

$$E = \iint_{\Omega} g[f(x,y) - \tilde{f}(x,y)]\ dxdy \ /\ \iint_{\Omega} dxdy \qquad (2.1)$$

is minimum.

Special assumptions

The above formulation is quite general and in order to make the problem analytically tractable we impose the following specifications:

1. The function f(x,y) is characterized by:

   a) The value range i.e. min f and max f is known.

   b) The fluctuation rates is known in terms of their mean square variations

   $$\sigma_x^2 = \overline{(f'_x)^2} \quad \text{and}$$

   $$\sigma_y^2 = \overline{(f'_y)^2}$$

2. A quantization error n(x,y) is obtained when $\hat{f}(x,y)$ is represented by $Q\hat{f}(x,y)$. The value range V of $\hat{f}$ is divided in 2^b equal intervals and $Q\hat{f}$ is chosen as the midpoint in the intervals.

   $$Q\hat{f}(x,y) = \hat{f}(x,y) + n(x,y)$$

   This error will be assumed independent of the value $\hat{f}(x,y)$. Errors at two different points in the image are also assumed to be independent.

3. Zero order hold interpolation is used. This means that $\tilde{f}(x,y) = Q\hat{f}(x_i,y_j)$ for x,y around x_i,y_j.

4. A quadratic function g is used, i.e.

   $$g(u) = u^2$$

   Other criteria can be considered. The quadratic has the advantage that an analytical solution may be obtained.

3. SOLUTION

3.1 Expanding the criterion

The criterion will be expanded and evaluated by dividing the image in M x N cells. The contributions from all cells will then be summed up.

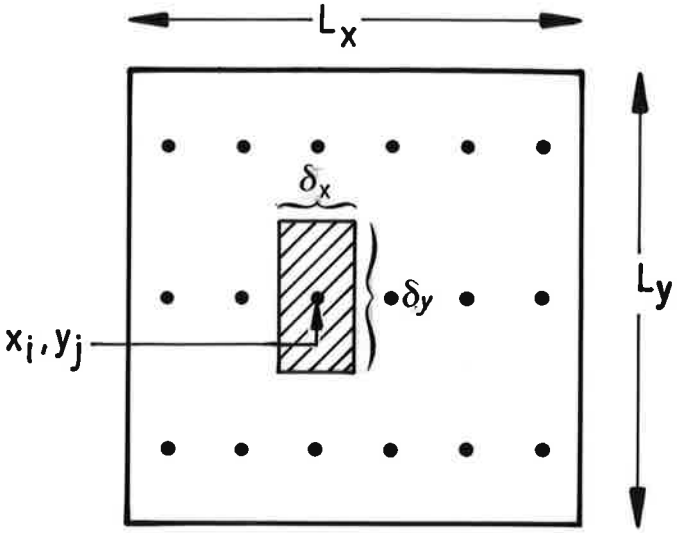

Fig. 3.1 The width and height of the image are denoted L_x and L_y. The point x_i,y_j belongs to the grid G and is surrounded by a cell with sides $\delta_x=L_x/M$ and $\delta_y=L_y/N$.

Let $\iint_{\square}$ denote integration over one cell.

Equation (2.1) becomes

$$L_xL_yE = \iint_{\Omega} [f(x,y) - \tilde{f}(x,y)]^2 \, dxdy =$$

$$= \sum_{i,j} \iint_{\square} [f(x,y) - Q\hat{f}(x_i,y_j)]^2 \, dxdy =$$

$$= \sum_{i,j} \iint_{\square} [f(x,y) - (f(x_i,y_j) + n(x_i,y_j))]^2 dxdy =$$

$$= \sum_{i,j} \{\iint_{\square} [f(x,y) - f(x_i,y_j)]^2 \, dxdy -$$

$$- \iint_{\square} 2\,(f(x,y) - f(x_i,y_j))\, n(x_i,y_j)\, dxdy +$$

$$+ \iint_{\square} [n(x_i,y_j)]^2\, dxdy)$$

Calculating the mean square error now gives

$$\bar{E} = \frac{1}{L_x L_y} \sum_{i,j} \{ \iint_{\square} \overline{[f(x,y) - f(x_i,y_j)]^2}\, dxdy +$$

$$+ \iint_{\square} \overline{n(x_i,y_j)^2}\, dxdy) \qquad (3.1)$$

Compared to Steiglitz notation the first term is the reconstruction error which depends only on the sampling fineness δ_x and δ_y. The second term is the quantization error which depends only on the number of quantization levels 2^b. However, recall that these numbers are coupled via the constraint $M \cdot N \cdot b$ constant.

A. The reconstruction error will be expanded using Taylor series excluding high order terms.

$$\iint_{\square} \overline{[f(x,y) - f(x_i,y_j)]^2}\, dx\, dy$$

then gives

$$= \iint_{\square} \overline{[x\, f_x'(x_i,y_j) + y\, f_y'(x_i,y_j)]^2}\, dxdy =$$

$$= \iint_{\square} [x^2\, \overline{(f_x')^2} + 2xy\, \overline{(f_x' \cdot f_y')} + y^2\, \overline{(f_y')^2}]\, dxdy$$

Each term is now evaluated.

$$\iint_{\square} x^2\, dxdy = \int_{-\delta_x/2}^{\delta_x/2} \int_{-\delta_y/2}^{\delta_y/2} x^2\, dxdy = \delta_y \cdot \frac{1}{3} \cdot \frac{\delta_x^3}{4}$$

$$\iint_{\square} xy\, dx\, dy = 0$$

$$\iint_{\square} y^2 \, dxdy = \delta_x \cdot \frac{1}{3} \cdot \frac{\delta_y^3}{4}$$

These expressions and the definitions of σ_x^2 and σ_y^2 from assumption 1 in section 2 are inserted. This gives the following expression for the reconstruction error in one cell.

$$\frac{1}{12} \cdot (\delta_x^3 \, \delta_y \, \sigma_x^2 + \delta_x \, \delta_y^3 \, \sigma_y^2) \qquad (3.2)$$

B. The quantization error will now be evaluated.

$$\iint_{\square} \overline{n(x_i, y_j)^2} \, dx \, dy = \delta_x \delta_y \, \overline{n(x_i, y_j)^2}$$

From the assumptions introduce R = max f - min f, let the number of quantization levels be 2^b and assume equidistant quantization. The quantization grain then becomes

$$\delta = \frac{R}{2^b}$$

which gives

$$\overline{n^2} = \frac{1}{\delta} \int_{-\delta/2}^{\delta/2} x^2 \, dx = \frac{1}{12} \delta^2 = \frac{1}{12} \cdot \frac{R^2}{2^{2b}}$$

The quantization error in one cell is thus

$$\delta_x \cdot \delta_y \cdot \frac{1}{12} \cdot \frac{R^2}{2^{2b}} \qquad (3.3)$$

C. The total mean square error.

By using $M \, \delta_x = L_x$ and $N \, \delta_y = L_y$ the criteria is expressed in M and N. The total mean square error (3.1) is summed up using (3.2) and (3.3), and the fact that there are $M \cdot N$ cells. The optimal digitization problem (both sampling and quantization) can now be expressed as follows.

Optimal digitization

Minimize the criterion

$$\bar{E} = \frac{1}{12} \cdot \left[\frac{L_x^2 \sigma_x^2}{M^2} + \frac{L_y^2 \sigma_y^2}{N^2} + \frac{R^2}{2^{2b}} \right] \tag{3.4}$$

under the constraint

$$M \cdot N \cdot b = C. \tag{3.5}$$

Here L_x, σ_x, L_y, σ_y, R and C are known constants.

3.2 Optimization

The solution to (3.4) and (3.5) will now be given.

Denote

$$J = 12\,\bar{E} = \frac{L_x^2\sigma_x^2}{M^2} + \frac{L_y^2\sigma_y^2}{N^2} + \frac{R^2}{2^{2b}}$$

Complete the squares

$$J = \left[\frac{L_x\sigma_x}{M} - \frac{L_y\sigma_y}{N}\right]^2 + \frac{2L_x\sigma_x L_y\sigma_y}{M \cdot N} + \frac{R^2}{2^{2b}}$$

From the constraint (3.5) it follows that

$$M \cdot N = \frac{C}{b}$$

which gives

$$J = \left[\frac{L_x\sigma_x}{M} - \frac{L_y\sigma_y}{N}\right]^2 + \frac{2L_x\sigma_x L_y\sigma_y}{C} \cdot b + \frac{R^2}{2^{2b}}$$

Notice that J is expressed as a sum of two parts. One is a square dependent only on M and N, and the other is a function only of b. Denote this function as h(b). The minimum of J is thus obtained for

$$\begin{cases} \dfrac{L_x\sigma_x}{M} = \dfrac{L_y\sigma_y}{N} & (3.6) \\ h'(b) = 0 & (3.7) \end{cases}$$

Equation (3.7) gives

$$\frac{2L_x\sigma_x L_y\sigma_y}{C} - \frac{R^2 2 \ln 2}{2^{2b}} = 0$$

$$b = \frac{1}{2 \ln 2} \cdot \ln\left[\frac{C R^2 \ln 2}{L_x\sigma_x L_y\sigma_y}\right]$$

Solving for M and N using (3.6) leads to the following result.

$$b = \frac{1}{2 \ln 2} \ln \left[C \cdot \ln 2 \cdot \frac{R^2}{L_x \sigma_x \, L_y \sigma_y} \right] \tag{3.8}$$

$$M = \sqrt{\frac{L_x \sigma_x}{L_y \sigma_y}} \cdot \sqrt{\frac{C}{b}} \tag{3.9}$$

$$N = \sqrt{\frac{L_y \sigma_y}{L_x \sigma_x}} \cdot \sqrt{\frac{C}{b}} \tag{3.10}$$

It is interesting to see how the image properties influences the solution. The expression

$$\frac{R^2}{L_x \sigma_x L_y \sigma_y}$$

contains the image characteristics in the form of a relation between the value range and the fluctuation rate. More fluctuation (i.e. more image detail) leads to fewer bits and more samples for image resolution. Less fluctuation needs more bits for intensity resolution, and fewer samples. This agrees with earlier subjective tests [3,4]. The formulas for M and N also explicitly gives the tradeoff between sampling rates in respective direction due to fluctuation in each direction.

4. THE EXPERIMENTS

Two experiments will be presented. First a computer generated image where the image characteristics are calculated from the function. Second a real life image where the characteristics are estimated from image statistics.

The original image and some digitized versions (sampled and quantized) will be presented in both cases. The individual numbers, M, N and b must be integers. This means that the product $C = M \cdot N \cdot b$ cannot be kept constant. There are also a hardware limitation in the hardware on M and N. For equidistant sampling the only values possible are 512, 256, 170, 128, 102, 85... (=trunc. (512/k)). The numbers actually used in each experiment are given.

4.1 Rotated trianglewave image

An image is generated by rotating a triangle wave with period T and amplitude R/2 around the image midpoint. This image is seen in 1 A.

By integrating the square of the derivatives it is found that

$$\sigma_x = \sigma_y = \sqrt{2} \cdot \frac{R}{T}$$

and hence

$$\frac{R^2}{L_x\sigma_x \cdot L_y\sigma_y} = \frac{1}{2} \cdot \frac{T}{L_x} \cdot \frac{T}{L_y} \cdot$$

The following numbers were used in the experiments

$$\frac{T}{L_x} = \frac{T}{L_y} = 0.12$$

$$C = [256]^2 = 65536.$$

Table 1. The rotated square wave image experiment. The used numbers M, N, b, the total of bits C, and the resulting value of the criteria J.

Image	M	N	b	C	J ($\cdot 10^{-3}$)
1 A	Original test image				
1 B	256	256	1	65536	16.5
1 C	170	170	2	57800	4.69
1 D	128	170	3	65280	1.88
1 E	128	128	4	65536	1.36
1 F	102	128	5	65280	1.48
1 G	85	102	7	60690	2.12

4.2 Authors' image

An image of the authors was used in the second experiment. The original image is seen in 2 A. The image characteristics are estimated directly from the image itself. A scan of the image gives

$$\frac{R}{L_x \sigma_x} = 4.06 \cdot 10^{-2} \qquad (4.1)$$

$$\frac{R}{L_y \sigma_y} = 9.68 \cdot 10^{-2} \qquad (4.2)$$

It is seen from the expressions (4.1-2) that there are more fluctuation in the horizontal direction than in the vertical direction. The number of bits used for digitization is $C = [256]^2 = 65536$.

Table 2. The authors image experiment. The used numbers M, N, b, the total of bits C, and the resulting value of the criteria J.

Image	M	N	b	C	J ($\cdot 10^{-3}$)
2 A	Original authors' image				
2 B	256	256	1	65536	16.0
2 C	256	128	2	65536	4.81
2 D	170	102	3	52020	2.88
2 E	170	85	4	57800	2.44
2 F	128	102	4	52224	3.15
2 G	170	73	5	62050	2.58
2 H	128	64	7	57344	3.89

Discussion

The integers which are closest to the optimal are M = 170, N = 85 and b = 4. If fewer bits/pixel than the optimal (b < 4) are used the image looks either blurred (Image 2 B) or banded (Image 2 C). If more bits/pixel (b > 4) and fewer sampling points are used some detail is lost, e.g. in image 2 H the spectacles of Karl Johan have disappeared.

In image 2 F the optimal number of bits (b = 4) are used but less care is taken to the difference in the two directions (M = 128 and N = 102). This gives a bit "edgier" appearance because some detail is lost horizontally but not much is gained vertically.

5. CONCLUSIONS

A theoretical formulation of the optimal digitization problem is given. The solution is obtained from an optimization of a criterion due to the constraint of fixed number of bits. The solution is tested experimentally and agrees well with human visual quality.

An advantage is that the solution is given in closed form (eq. 3.8 - 10). This makes it easy to use as a rule of thumb. It also clearly points out the dependence of image characteristics. This dependence explains and agrees with what is found in other subjective tests.

In future works the assumptions made in this paper will be somewhat relaxed. Also, other optimization criteria will be attempted in order to obtain the best one suitable for this problem. It is known that for biological systems the optimization criteria are "Application-Dependent" [10] and the visual quality of the image data involving the human visual system is, of course, of no exception.

6. ACKNOWLEDGEMENTS

L. Nielsen is supported by the Swedish Board of Technical Development in the project "Control Based on Image Information" (STU-82-3429). This contract has also supported the visit of E. I. Jury in Lund the summer 1983. The research efforts by E. I. Jury are also partially supported by NSF grant ECS-8116847.

REFERENCES

[1] Pratt, W.K. (1978): Digital Image Processing. John Wiley & Sons, Inc., New York.

[2] Pavlidis, T. (1982): Algorithms for Graphics and Image Processing. Computer Science Press. (p.39).

[3] Rosenfeld, A., A.C. Kak (1982): Digital Picture Processing. Academic Press. (p.111)

[4] Huang, T.S., O.J. Tretiak, B. Prasada, Y. Yamaguchi (1967): Design considerations in PCM transmission of low resolution monochrome still pictures. Proc. IEEE 55, 331-335.

[5] Abdou, I.E., K.Y. Wong (1982): Analysis of Linear Interpolation Schemes for Bi-Level Image Applications. IBM J. Res. Develop. 26, 6.

[6] Steiglitz, K. (1966): Transmission of an Analog Signal Over a Fixed Bit-Rate Channel. IEEE Trans. on Information Theory. IT-12, 4.

[7] Lebedev, D.S. and L.T. Mirkin (1975): Smoothing of two-dimension images using the "composite" model of a fragment. Inconics, Digital Holography, and image processing. Institute for Problems in Information Transmission, Academy of Sciences, U.S.S.R. 57-62.

[8] Ingle, V.K. and J.W. Woods (1979): Multiple model recursive estimation of images. ICASSP 79 Conf. Rec. 642-645.

[9] Woods, J.W. (1981): Two-Dimensional Kalman Filtering. Topics in Applied Physics. 42. Editor: T.S. Huang, Springer-Verlag.

[10] Krishnan, V.V., E.I. Jury and L. Stark (1982): Biological Optimality: Comparison of Generalized Feasibility and Optimality Conditions for Linear Error Minimizing Systems. Bulletin of Matematical Biology, vol. 44, 6, pp. 777-791.

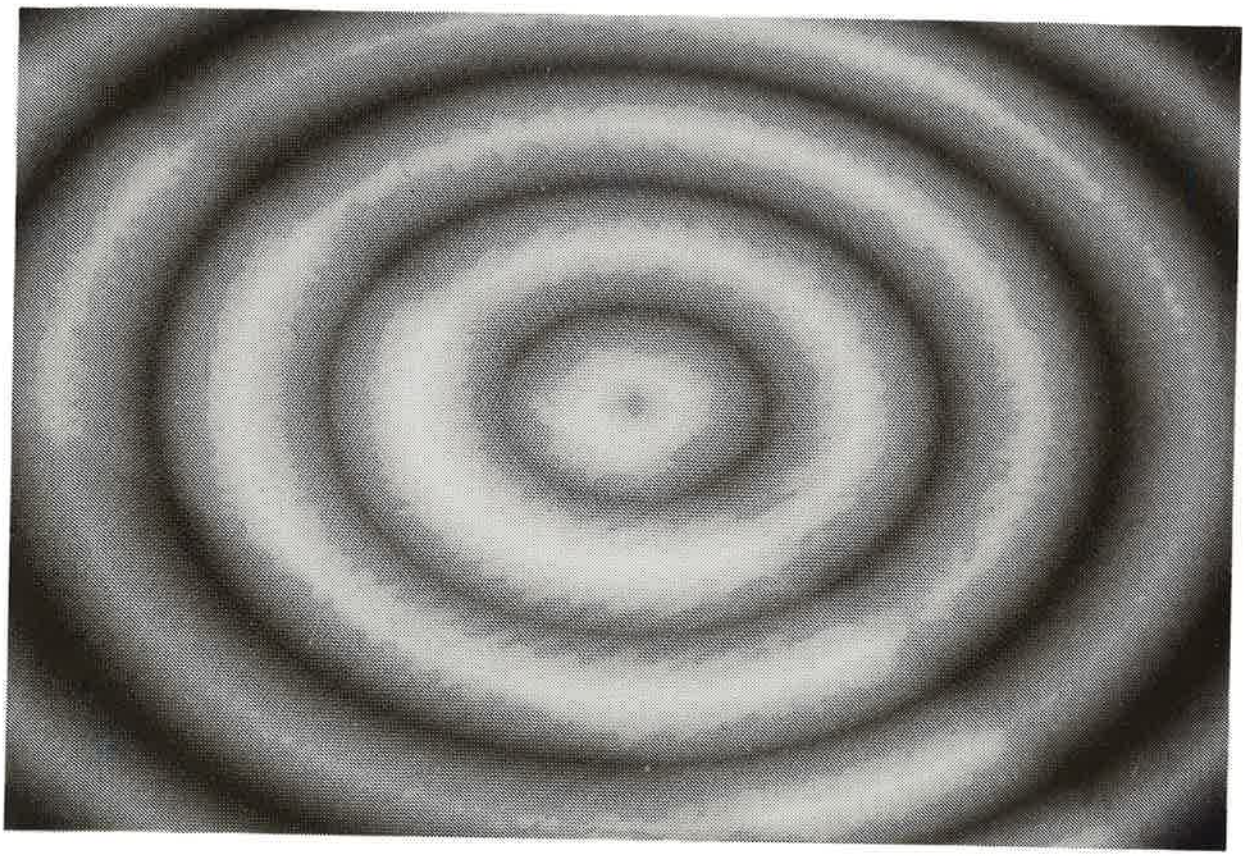

Image 1 A. Original test image

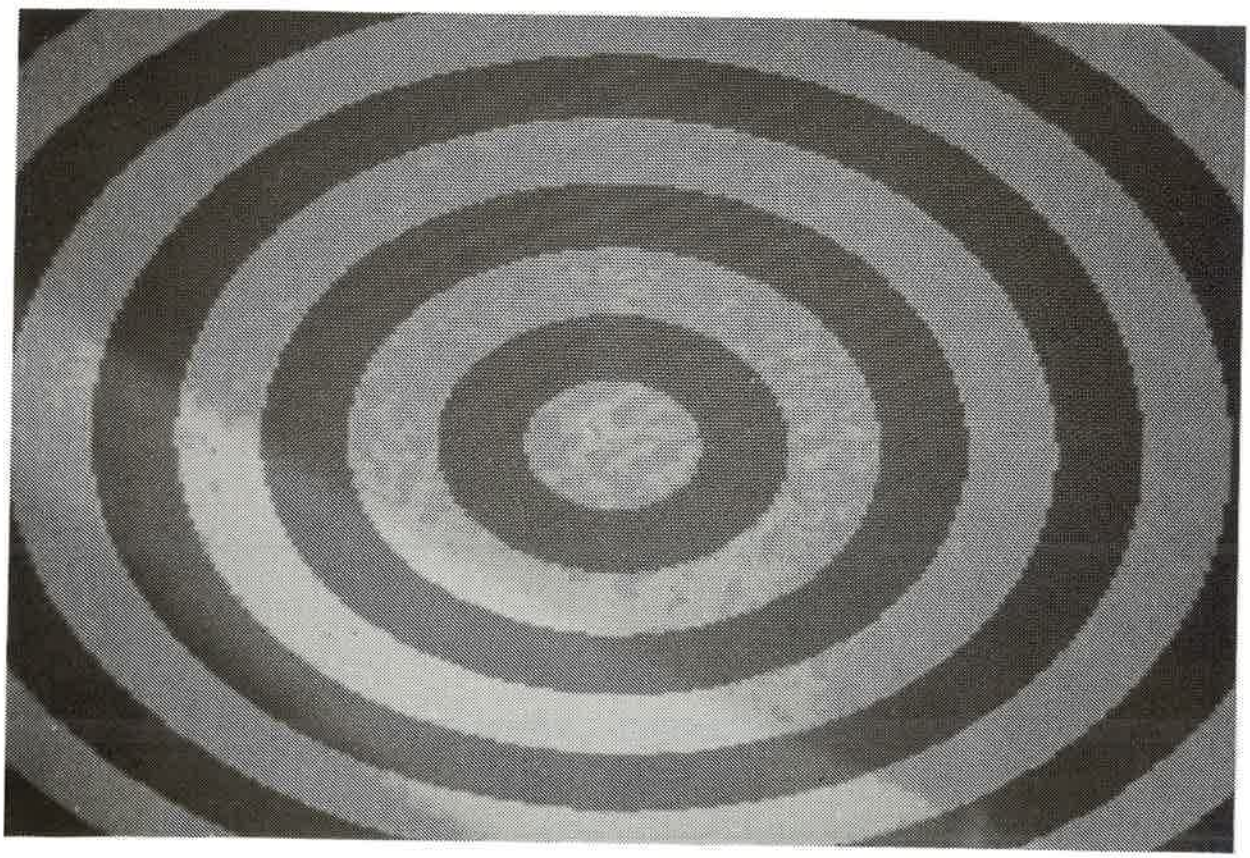

Image 1 B. M=256 N=256 b=1 M·N·b=65536

Image 1 C. M=170 N=170 b=2 M·N·b=57800

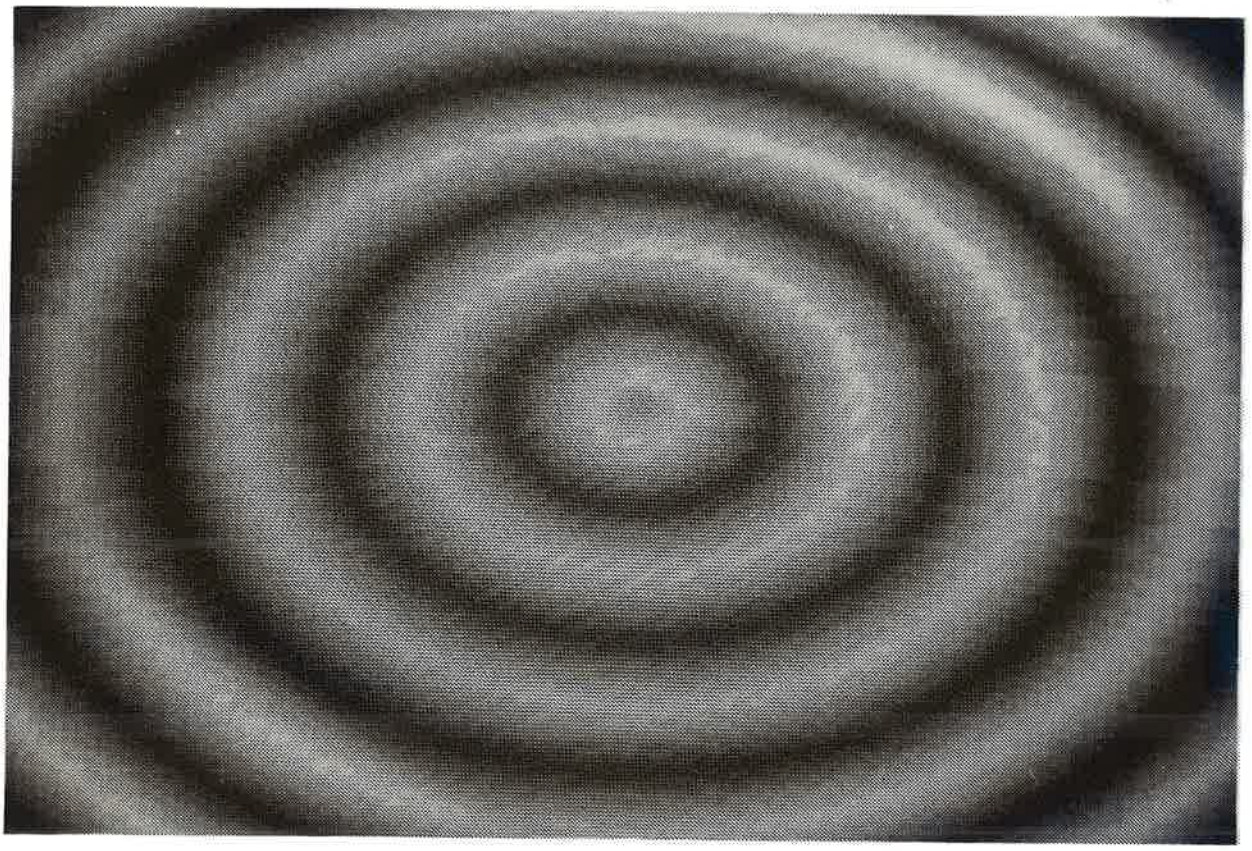

Image 1 D. M=128 N=170 b=3 M·N·b=65280

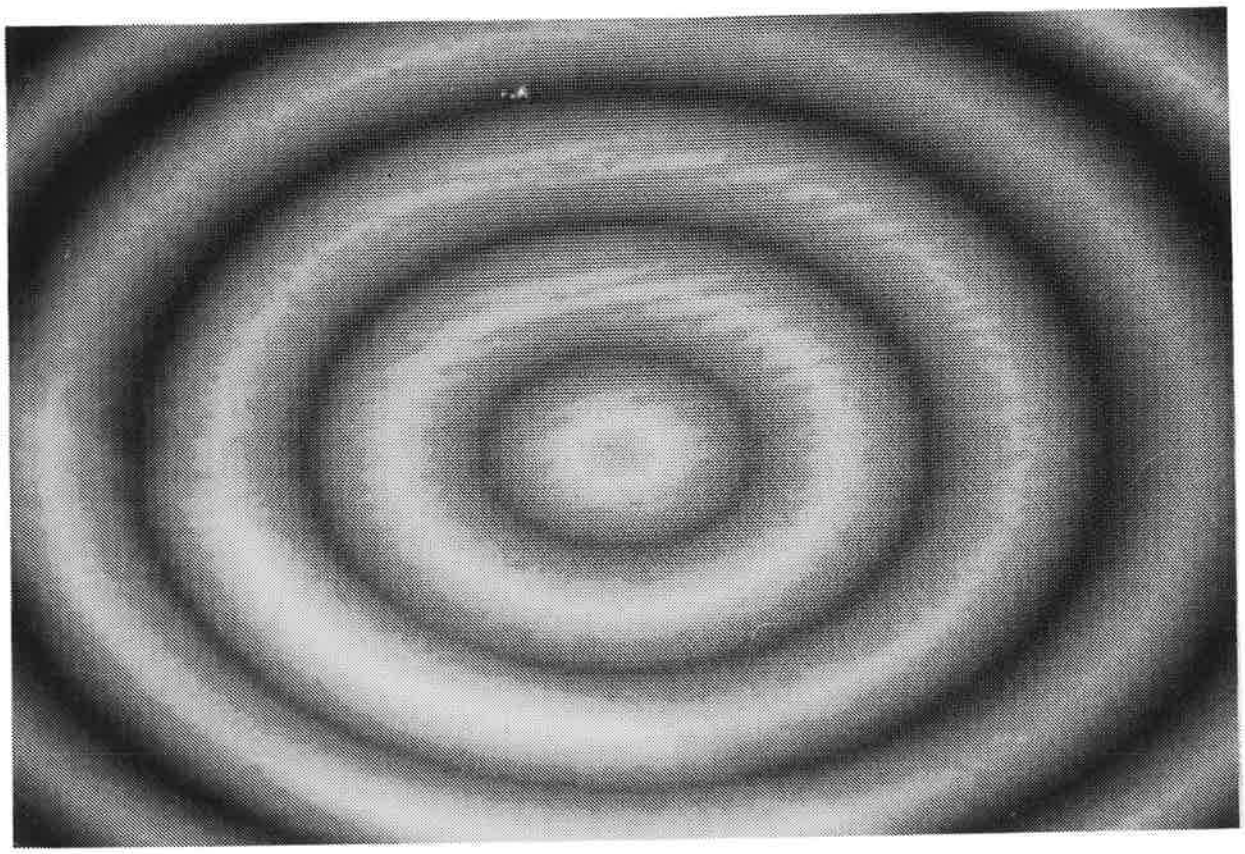

Image 1 E. M=128 N=128 b=4 M·N·b=65536

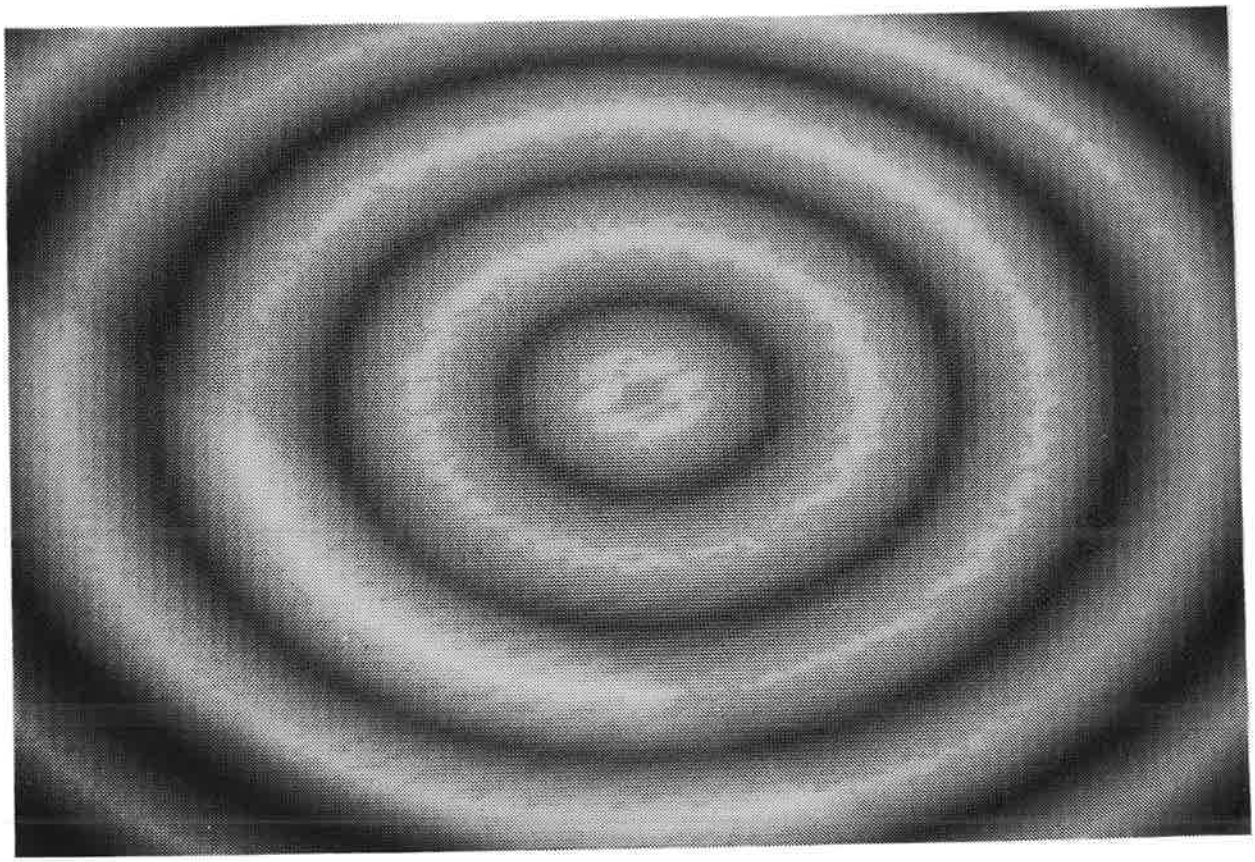

Image 1 F. M=102 N=128 b=5 M·N·b=65280

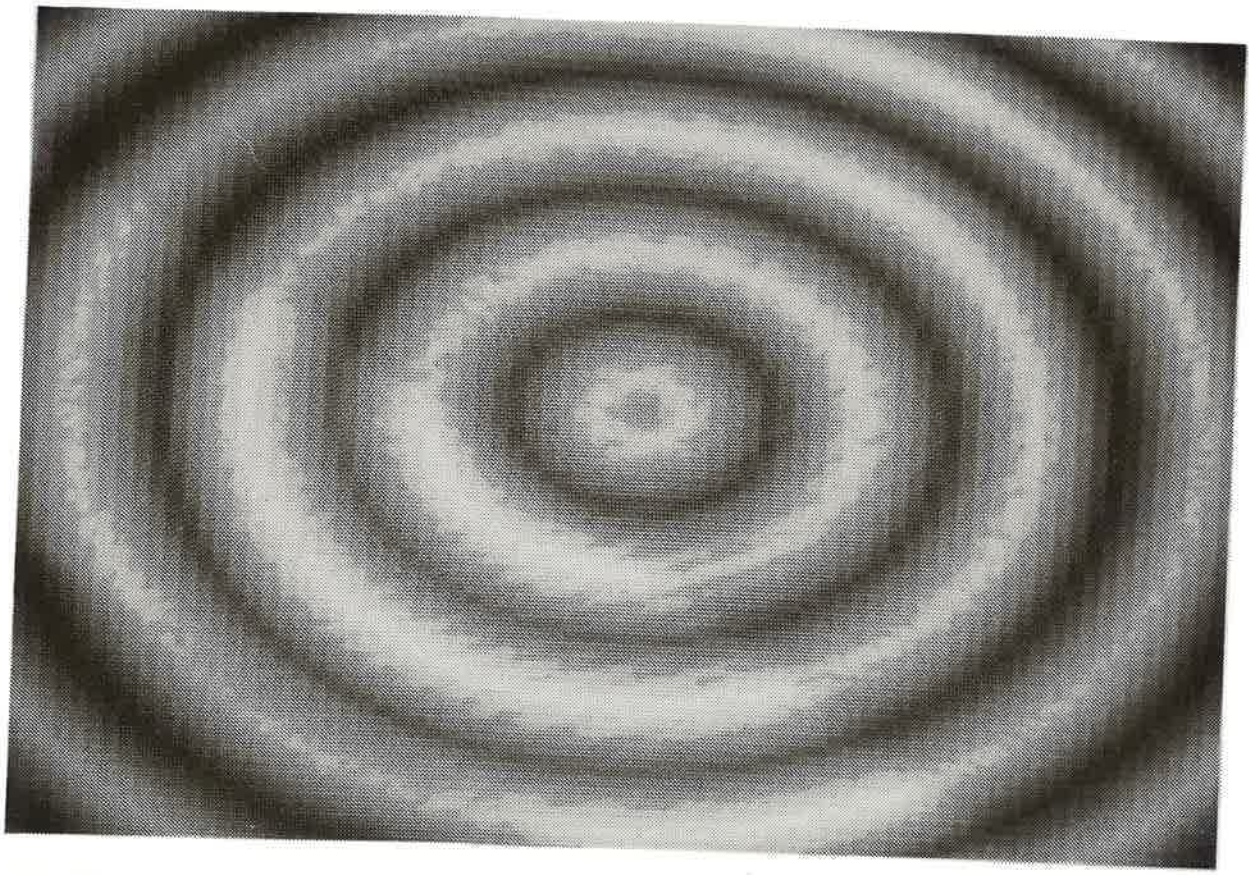

Image 1 G. M= 85 N=102 b=7 M · N · b=60690

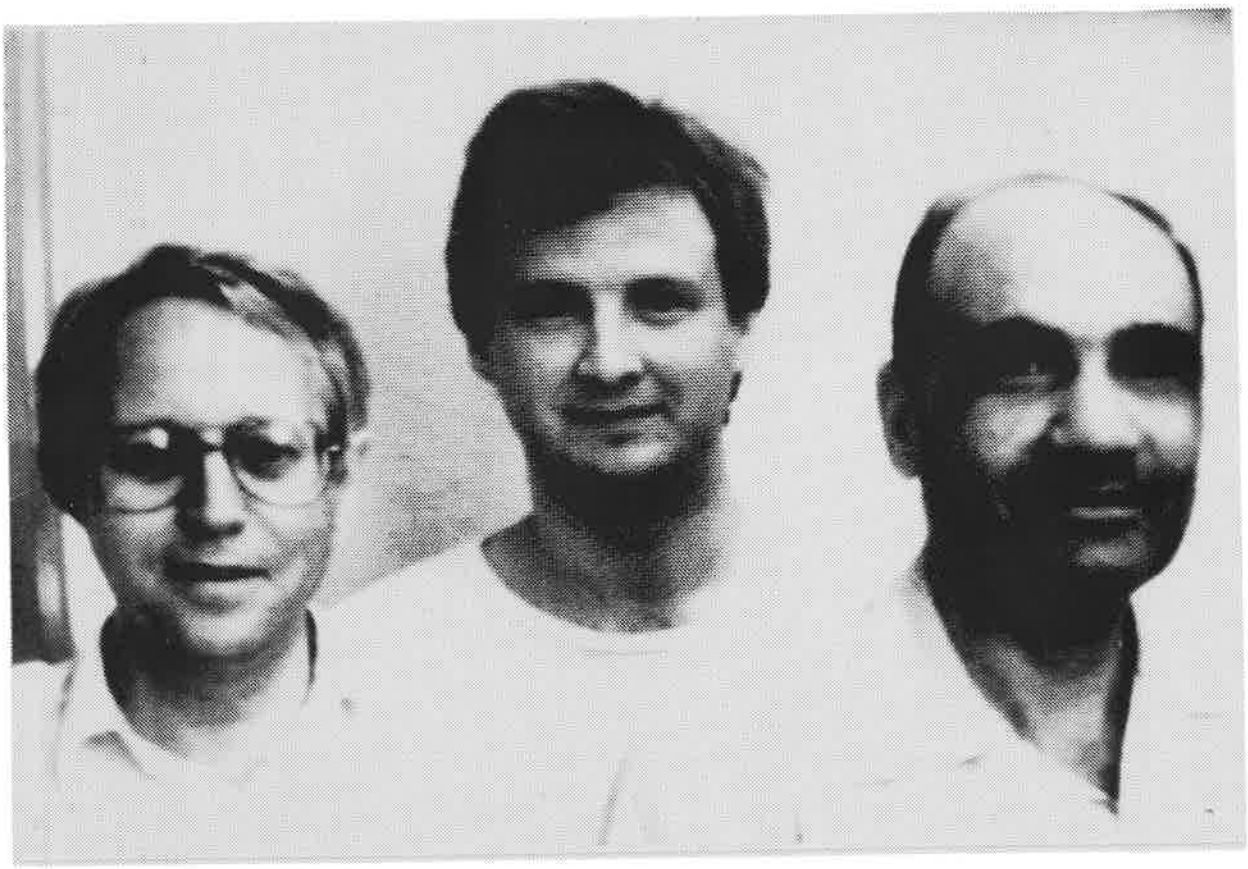

Image 2 A. Original authors' image

Image 2 B. M=256 N=256 b=1 M·N·b=65536

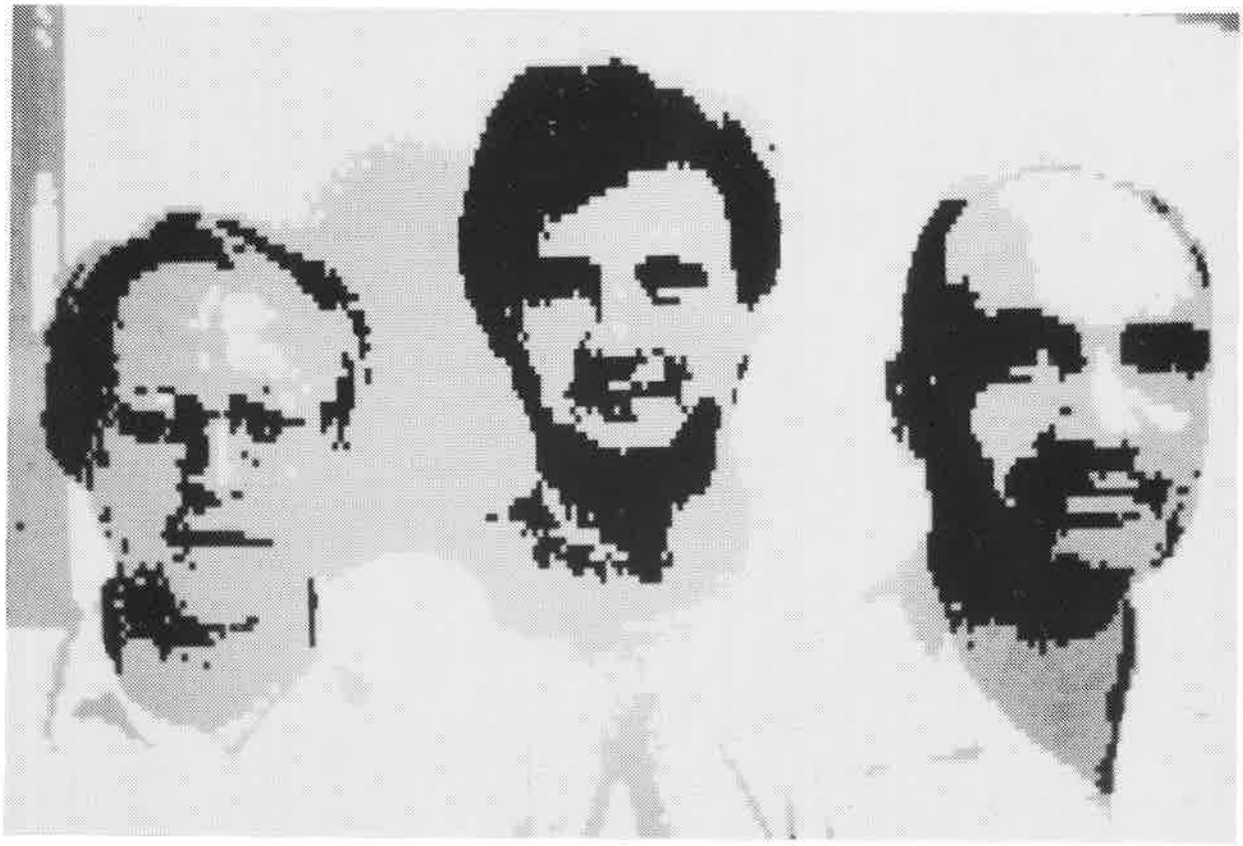

Image 2 C. M=256 N=128 b=2 M·N·b=65536

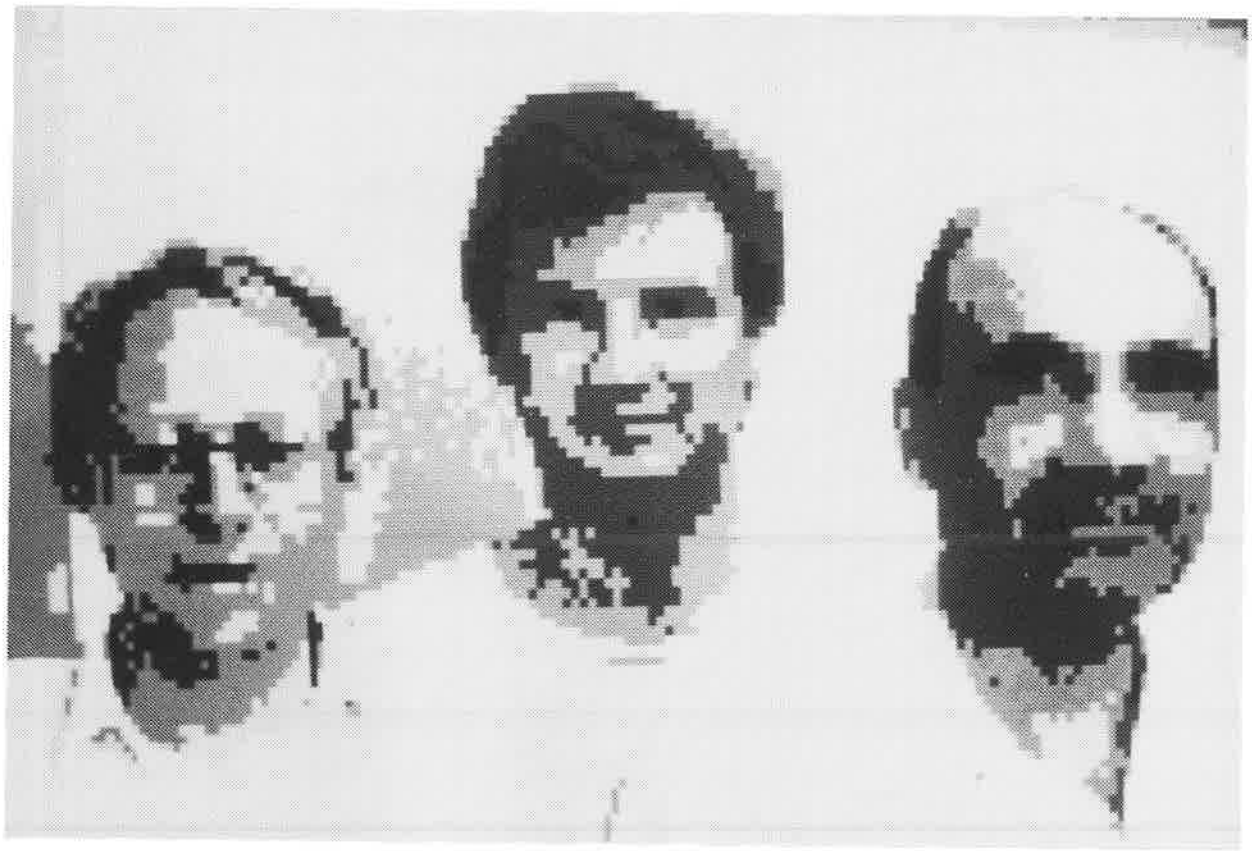

Image 2 D. M=170 N=102 b=3 M·N·b=52020

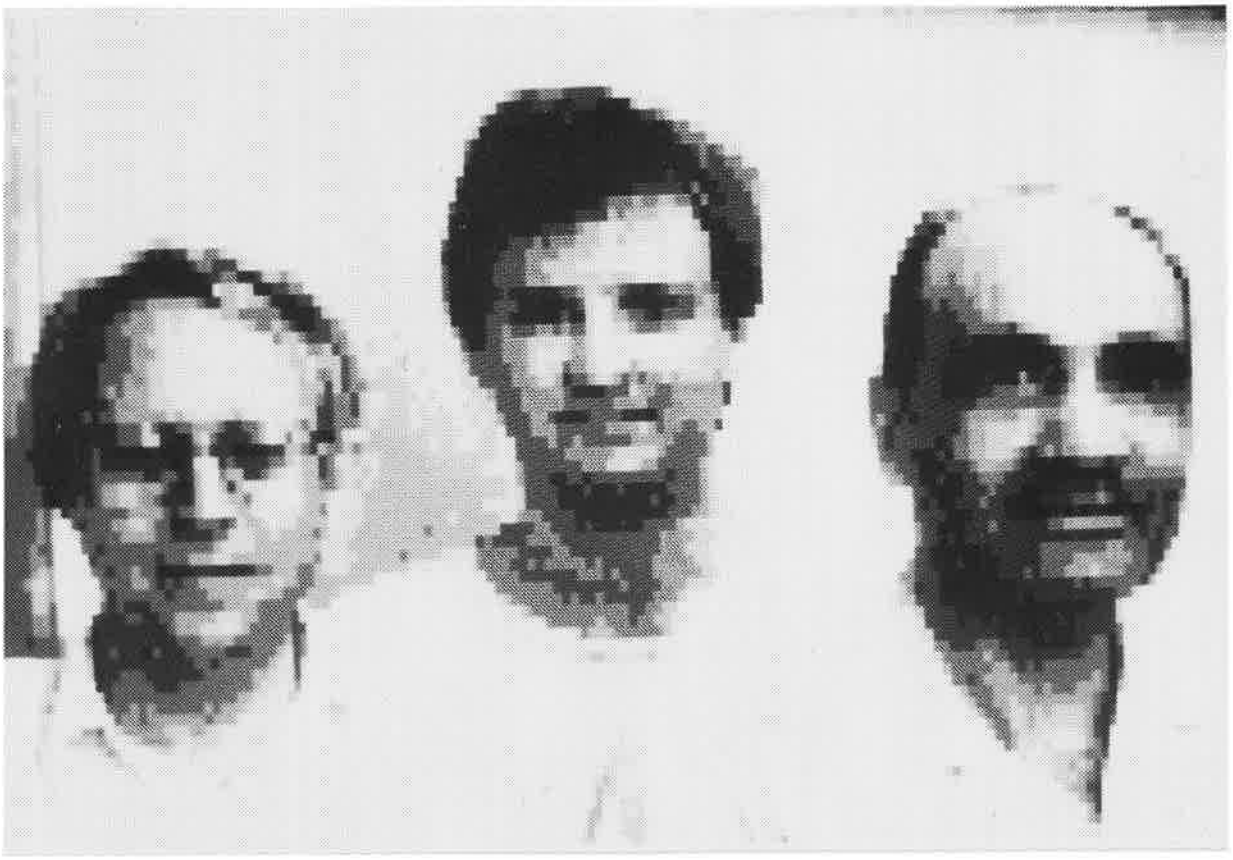

Image 2 E. M=170 N= 85 b=4 M·N·b=57800

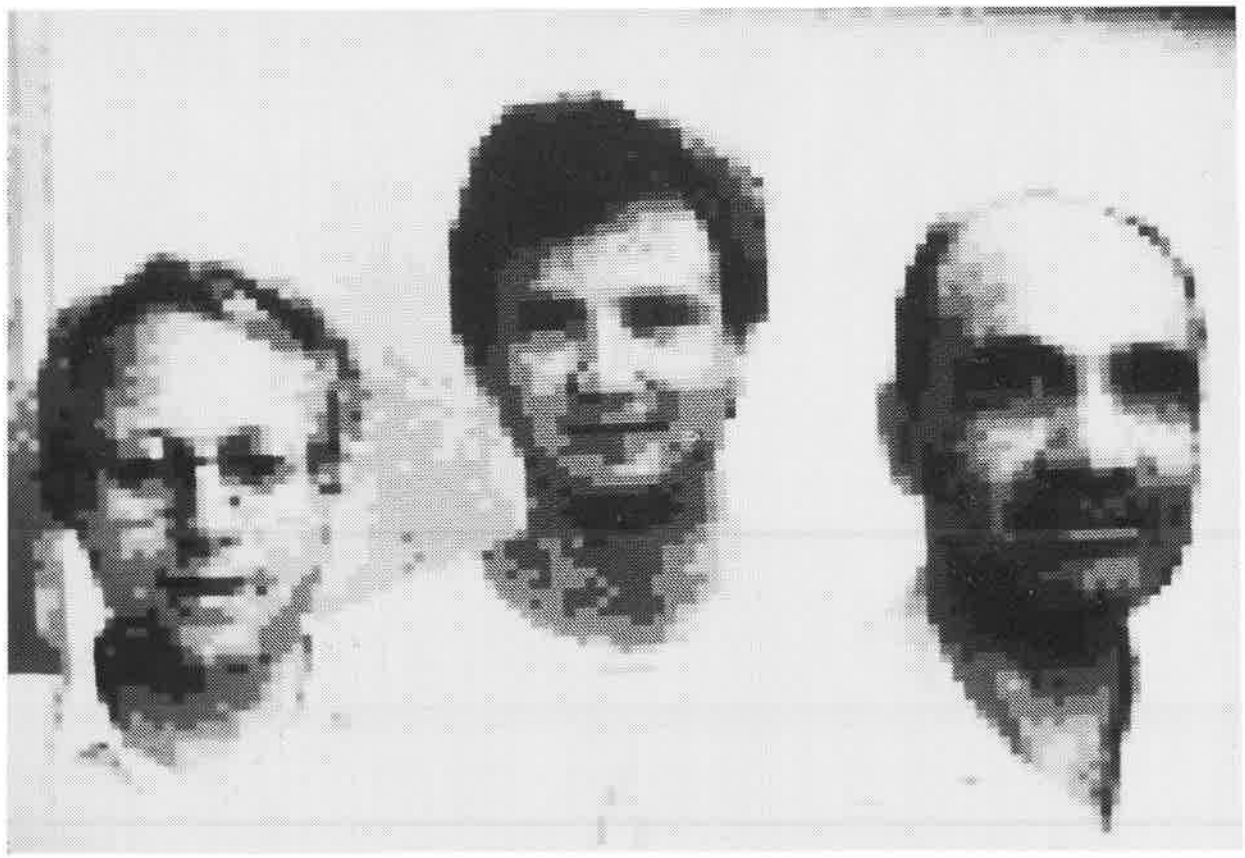

Image 2 F. M=128 N=102 b=4 M·N·b=52224

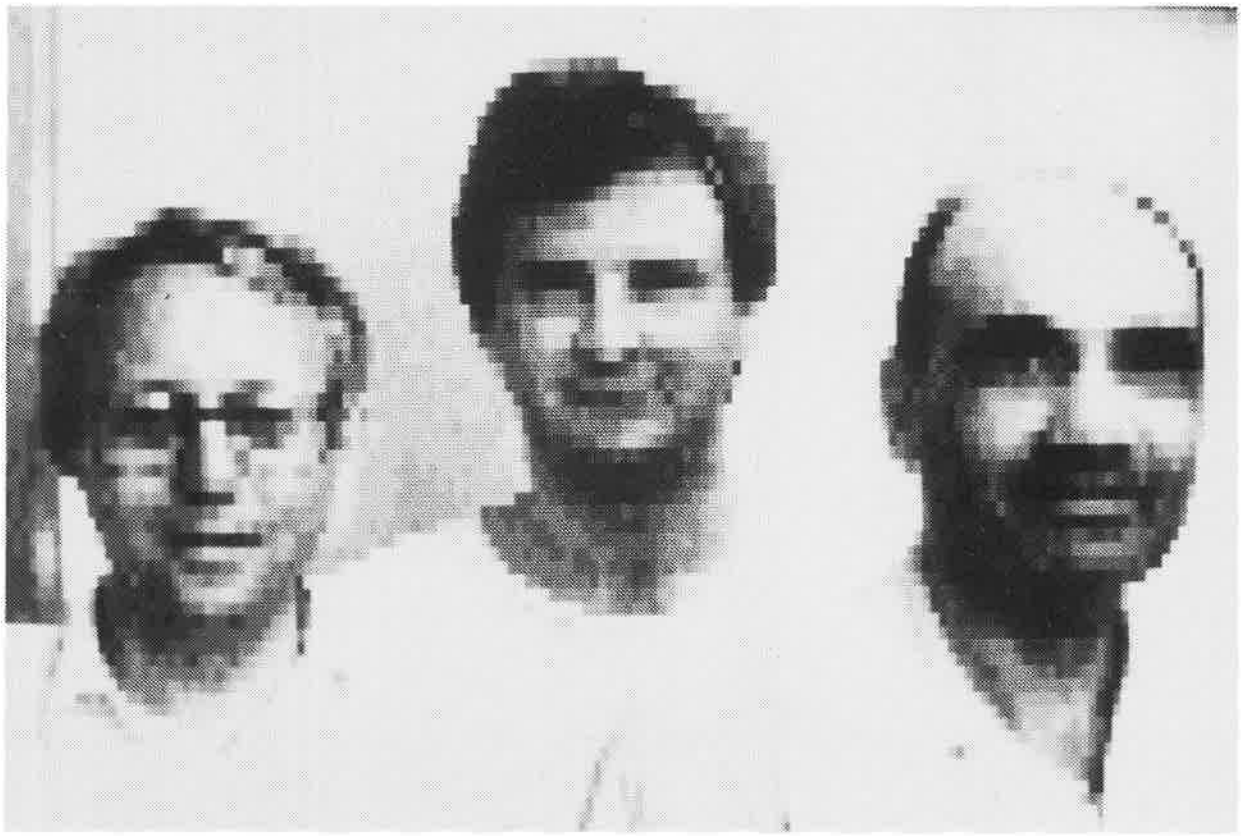

Image 2 G. M=170 N= 73 b=5 M·N·b=62050

Image 2 H. M=128 N= 64 b=7 M·N·b=57344