Numerical smoothing and filtering in N dimensions

L. B. Hunt
Datatex, Hilltop, Whitebirk Industrial Estate, Blackburn, Lancs.

The processing of naturally occurring information (e.g. the examination of the visual field by a pattern recogniser) can be significantly simplified if the raw information is initially processed in such a way as to enhance those features that are important with respect to the given task whilst attenuating the spurious or undesired information.

A convenient method of specifying what information is to be removed and what preserved in, for example, the class of one-dimensional signals typified by the output from a microphone is to use the Laplace transform to generate a suitable realisable filter whose frequency characteristics can be easily and accurately specified.

A severe drawback associated with using these filters however is that, unavoidably with the class of signals mentioned above, the impulse response function is asymmetric. This asymmetry is responsible for the well-known phase shift effect produced by electrical filters. Such filters can be derived and analysed by means of the one-sided Laplace transform, the one-sidedness resulting from the limitations imposed by physical realisability. This note describes a method of using the one-sided Laplace transform to produce N-dimensional, approximately spherically symmetric filters. Although these are derived in numerical form they can be produced just as readily by means of actual hardware provided the appropriate data scanning is carried out. The method is compared with the mesh operator approach and its relative merits are enumerated.

The effect of these filters, which operate recursively, is demonstrated by their application to the enhancement of fingerprint images—a notoriously difficult cleaning-up problem. Fingerprints, by virtue of their almost constant width line structure have a very narrow two-dimensional band-width. Thus it is possible, by centring a band pass filter on the dominant frequency for a given fingerprint, to remove a great deal of the spurious data that is inevitably included by the data-capturing process.

(Received July 1970)

1. Introduction

This work was initiated in order to establish whether or not the vast body of standard filters developed by electrical, communications and control engineers could be readily extended to two (or more) dimensions. In particular, a desired objective was to employ the method of recursive or closed loop filtering in order to minimise the amount of computation required compared with that required to produce the same response using the open loop or mesh type filters which employ simple convolution with the data.

2. Frequency domain filters

An advantage inherent in using the standard types of filter mentioned above is that the filter transfer function, expressed in the frequency domain, can be obtained by mapping the filter's impulse response function from the spatial domain using the Laplace transform (Shinners, 1964). This considerably simplifies the synthesis of arbitrarily complex filters by virtue of the fact that such filters can be generated merely by forming the products and quotients of the required number of the basic standard forms.

The desired frequency response characteristic is obtained by adding together the logarithmic characteristics, or Bode diagrams, of the various standard forms used. As an example consider the standard spatial and frequency domain expressions for a simple exponential smoothing function in one dimension:

Spatial domain | *Frequency domain*

$$F(x) = Ke^{-x/T} \qquad F(S) = \int_0^\infty e^{-Sx} f(x)\, dx \quad (2.1)$$

$$= \frac{K}{1 + TS}$$

where $S = jw$ is the Laplace operator and w is the spatial frequency in radians/unit of distance.

Table 1 indicates the frequency domain transfer functions and the associated gain characteristics for the simplest possible forms of the three basic types of filter.

3. Convolution or mesh filters

The notes in this section are included in order to indicate the relative merits of the recursive scheme which will be developed in the next section. The result of operating on a one-dimensional data set $f(x)$ with a filter, whose weighting function (impulse response) is represented by $W(x)$, can be expressed as the convolution of these two functions taken over the interval of interest, i.e.

$$f^*(x) = \int_{-\infty}^{\infty} W(\alpha) \,.\, f(x - \alpha)\, d\alpha$$

or, for the discrete case:

$$f_i^* = \sum_{j=-n}^{n} W_j \cdot f_{i-j}$$

where $\pm n$ is the range of W and is determined from considerations of stability and accuracy of the result and of the computational load involved.

For the two-dimensional case the convolution integral assumes the form

$$f^*(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\alpha, \beta) \,.\, f(x - \alpha, y - \beta)\, d\alpha\, d\beta$$

or, for the discrete case

$$f_{ij}^* = \sum_{l=-n}^{n} \sum_{m=-n}^{n} W_{l,m} \cdot f_{i-l,\, j-m}$$

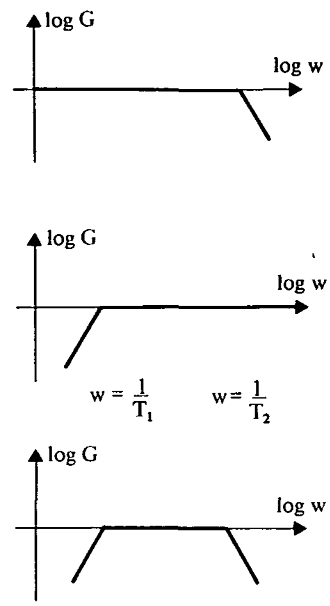
and the expressions for data fields of higher dimension follow naturally.

Mesh filter stability

Apart from the usual problems of the accumulation of errors

Table 1. Simple transfer functions

Transfer Function	Filter type	Gain Characteristics
$\frac{K}{(1+T_2S)}$	Low pass	
$\frac{KT_1S}{(1+T_1S)}$	High pass	
$\frac{KT_1S}{(1+T_1S)(1+T_2S)}$	Band pass	



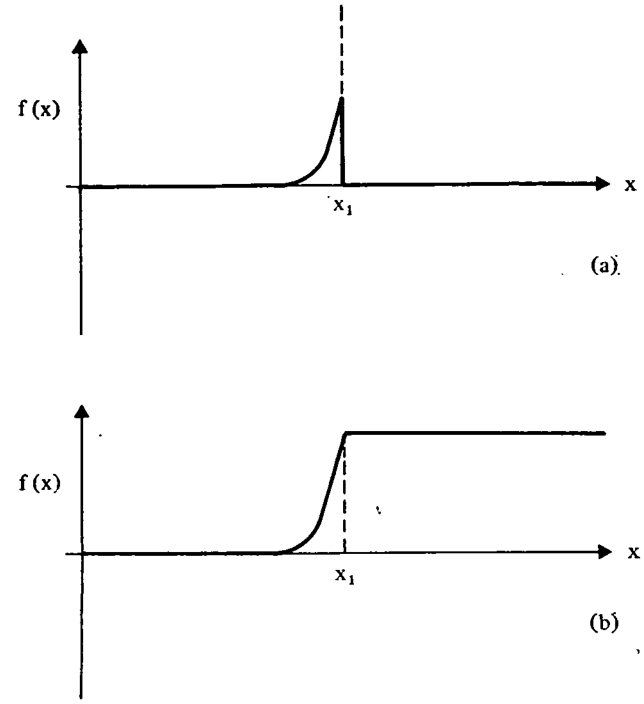


Fig. 2 Backward smoothing

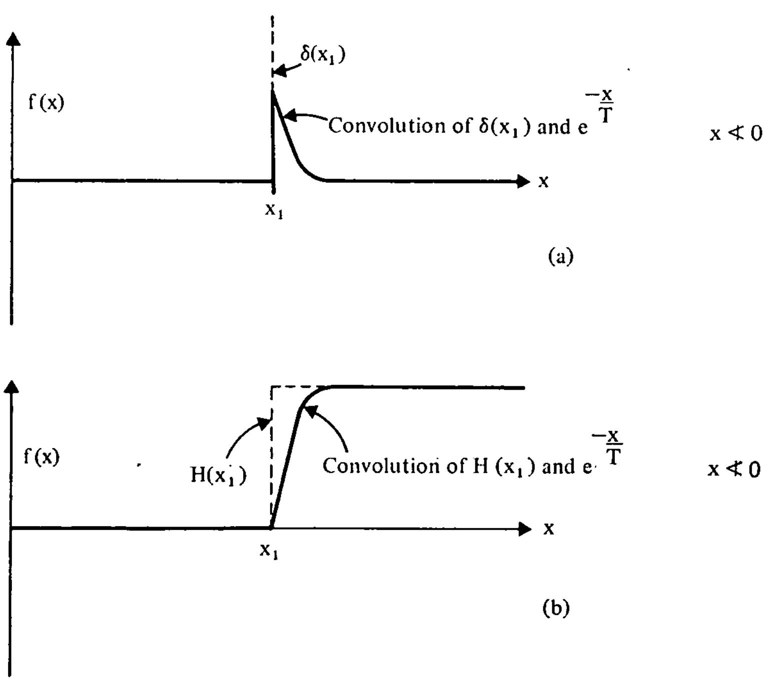


Fig. 1 Forward smoothing

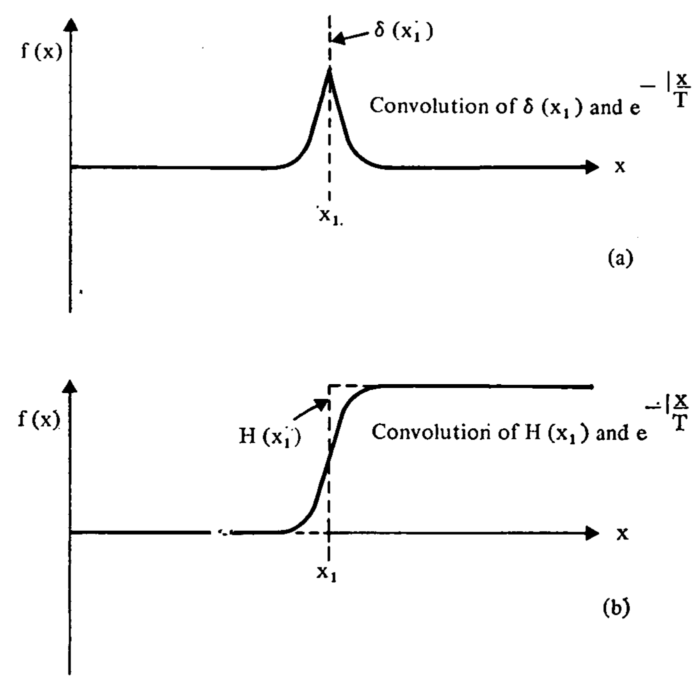


Fig. 3 Combined forward and backward smoothing

due to insufficient guarding figures being used in the calculations the following mutually exclusive criteria must also be satisfied in the appropriate case.

(*a*) For zero frequency pass filters we must have:

$$1 - \left| \sum_{l=-n}^{n} \sum_{m=-n}^{n} W_{l,m} \right| \leqslant \varepsilon$$

(*b*) For zero frequency stop filters we must have:

$$\left| \sum_{l=-n}^{n} \sum_{m=-n}^{n} W_{l,m} \right| \leqslant \varepsilon$$

The reasons for these conditions become obvious when one considers the fact that in the steady state (i.e. constant data) type (*a*) filters should have no effect whilst type (*b*) filters should reduce the amplitude of the data to zero. If very low frequencies are involved in type (*b*) filters then n becomes very large.

4. Recursive filters

The transfer functions listed in Table 1 can be used to generate the equivalent recursive filter by means of the substitutions:

$$S \equiv d/dx \text{ and } d/dx\, f_i = f_i - f_{i-1}$$

assuming that f_i and f_{i-1} are separated by unit distance. For example, considering the low-pass case we have:

$$f_i^* = \left[\frac{K}{1 + T_2S}\right] \cdot f_i \tag{4.1}$$

or

$$[1 + T_2S] f_i^* = K f_i$$

or

$$f_i^* + T_2 f_i^* - T_2 f_{i-1}^* = K f_i$$

or

$$f_i^* = \frac{1}{(1 + T_2)} [K f_i + T_2 f_{i-1}^*] \quad (4.2)$$

Note

It is usual to assume, at the start of the processing of a data stream, that $f_1^* = 0$. Any perturbations, which result from this assumption being in error, decay at a rate proportional to T_2. Errors of this kind give rise to typical Fresnel-like diffraction patterns near the boundaries of the data array when band pass filters are used.

We now consider the inherent difficulty associated with the use of the normal form of the Laplace transform. As indicated in equation (2.1) the standard expression is:

$$F(S) = \int_0^\infty e^{-Sx} f(x)\, dx \quad (4.3)$$

This expression, known as the one-sided Laplace transform because of the zero limit, is used in most practical applications. Its use is justified by the assumption that $f(x) = 0$ for all $x < 0$. This assumption is evidently inherent in the recursive filter indicated above in that it only makes use of backward difference approximations for the derivatives involved in the equations instead of the more accurate central difference formulae. Also implicit in these equations is the assumption that no knowledge can be obtained about data not yet encountered, i.e. we are unable to look forward to the next data point.

However, in the present application, we assume that we are dealing with recorded information and therefore we *do* have access to all points of the data stream.

The problem that we are now faced with is that the use of data not yet processed is incompatible with the principle of feedback as illustrated by equation (4.2). Further, the use of simple feedback introduces asymmetry into the process.

This is illustrated by operating with a simple smoothing filter $f(x) = e^{-x/T}$ on a unit impulse (**Fig. 1**(*a*)) and on a unit step (**Fig. 1**(*b*)). If we regard the filter as a black box then these responses are produced by passing the data through it from right to left. Alternatively the filter can be regarded as passing over the data in the positive x direction. Analytically these responses are due to the zero lower limit of equation (4.3). If we pass the data through the filter from left to right then the responses illustrated in **Figs. 2**(*a*) and **2**(*b*) are obtained. We thus require a method for combining these two types of response in order to produce the symmetric forms illustrated in **Fig. 3.** As indicated in Fig. 3 the filter $f(x) = e^{-|x/T|}$ would in fact be suitable. However, it is analytically difficult to deal with. Using the following approach and by recourse to the two-sided Laplace transform, symmetric responses can in fact be produced.

Let

$$F(S) = \int_{-\infty}^{\infty} e^{-Sx} . f(x) . dx \quad (4.4)$$

$$= \int_{-\infty}^{0} e^{-Sx} . f(x)\, dx + \int_0^\infty e^{-Sx} . f(x)\, dx \quad (4.5)$$

$$= F_1(S) + F_2(S)$$

By making the substitutions $S = -\bar{S}$ and $x = -\bar{x}$ in the first of the two integrals on the right-hand side of equation (4.5) we obtain

$$F_1(S) = \int_0^\infty e^{-\bar{S}\bar{x}} . f(\bar{x})\, d\bar{x} \quad (4.6)$$

and noting that if $f(x) = e^{-|x/T|}$ then $f(x) = f(\bar{x})$.

Thus substituting $f(x) = e^{-|x/T|}$ into equation (4.5) we obtain

$$F(S) = \frac{1}{1 + T\bar{S}} + \frac{1}{1 + TS}$$

$$= \frac{1}{1 - TS} + \frac{1}{1 + TS} = F_1(S) + F_2(S) \quad (4.7)$$

where the $-S$ indicates that the data is to be passed through the filter in the negative direction. Now the two expressions on the right-hand side of equation (4.7) produce the two kinds of response indicated in Figs. 1 and 2 respectively and we note that the sum of say Figs. 1(*a*) and 2(*a*) does not produce Fig. 3(*a*). This is due to the fact that the limits of integration in equation (4.5) should be as follows:

$$F(S) = \int_{-\infty}^{x<0} e^{-Sx} . f(x) . dx + \int_{x \geq 0}^{\infty} e^{-Sx} . f(x) . dx \quad (4.8)$$

i.e. the data point $x = 0$ should not occur in both integrals. This difficulty can be avoided by observing that equation (4.7) can be rewritten in the form

$$F(S) = \frac{\sqrt{2}}{1 - TS} \left(\frac{\sqrt{2}}{1 + TS} \right) \quad (4.9)$$

This form implies that the data stream is first passed through the filter

$$\phi(S) = \frac{1}{1 + TS} \quad (4.10)$$

in the positive direction and then the resulting data stream is passed back through the same filter in the opposite direction. The scaling factor of 2 is in fact removed by normalisation. Using this technique the responses of Fig. 3 are produced.

Hence we see that by using quite ordinary filters, which may be designed using the normal one-sided Laplace transform, symmetric filtering operations (i.e. zero phase shift) can be obtained and, as the filters produced in this way always consist of sums of exponential terms, the substitution involving the modulus of the exponents is generally applicable.

It is worth noting at this point that filters with far steeper gain slopes at the cut-off frequencies may be generated using other techniques (Constantinides, 1968), the slopes of the Bode plots shown in Table 1 being only 6db/octave.

5. Extension to two (or more) dimensions

Having produced a class of filters with a symmetric response for recorded one-dimensional signals, the extension to two dimensions is almost automatic—but not quite.

The desired objective is to produce, at every point of the field, a circularly symmetric impulse of the kind which would be obtained by rotating the response shown in Fig. 3(*a*) about an axis aligned along the original impulse. (See **Fig.** (**4**(*b*).)

The reason for desiring this type of impulse response for any given filter is the fact that any departure from symmetricity will result in distortion of the visual field. It will be demonstrated that, although global circular symmetricity is not possible with this type of filter, the fact that the filter is approximately circularly symmetric about the axes is sufficient to prevent geometrical distortion. The aspherical shape of the actual filter does however cause the frequency response to vary with direction. This implies that for any filter the corner frequencies are greater along the diagonals than along the axes. Such variations can cause difficulties if very selective 'notch' filters are used but in the application illustrated here sufficient band-width could be allowed to reduce this effect to an acceptable level. The proposed method has the advantage of simplicity because of its recursive form and of speed during operation because of its simplicity.

We note that the expression which describes the response of the desired smoothing filter, operating on an impulse situated at (x_1, y_1) would assume the form

$$f(x, y) = K\, e^{-([x-x_1]^2 + [y-y_1]^2)^{\frac{1}{2}}/T} \quad (5.1)$$

Unfortunately, this function does not have a Laplace transform. This fact effectively prevents us from isolating the desired

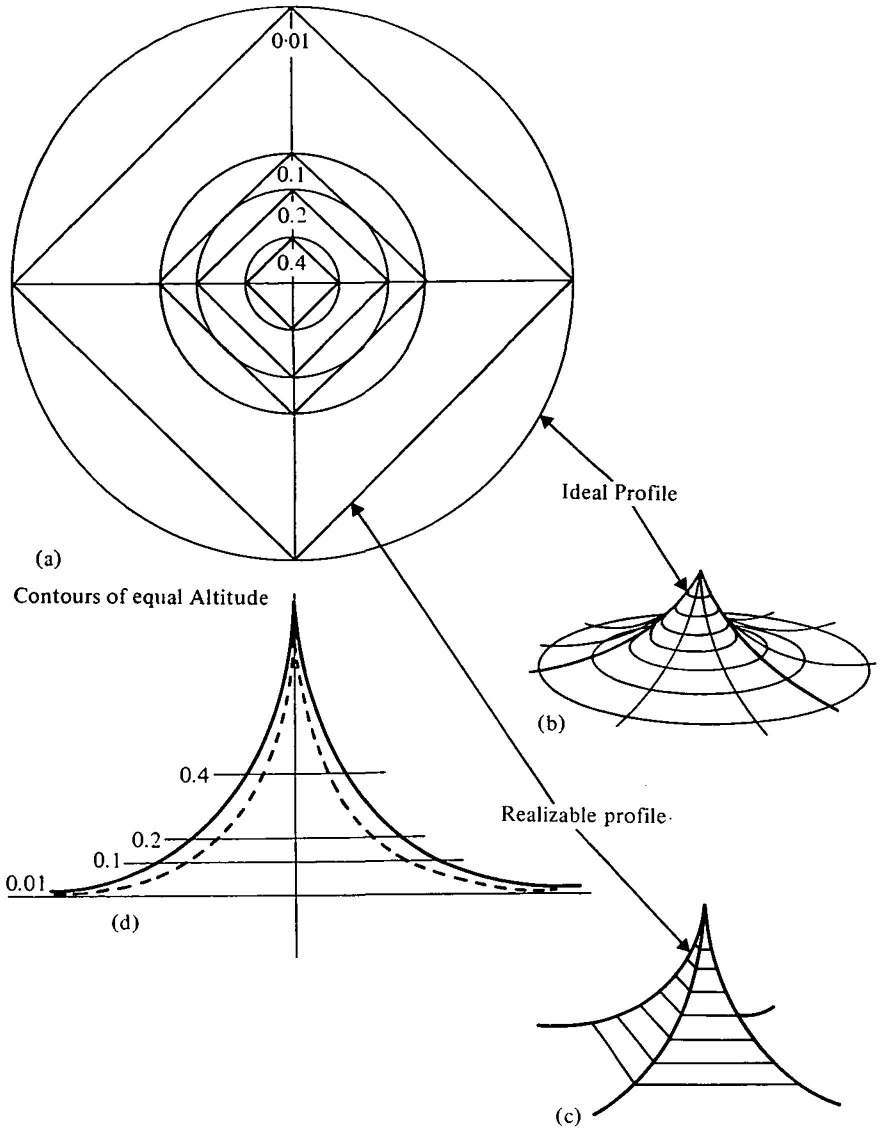


Fig. 4. Impulse responses of $F(S) = \dfrac{1}{(1 + 0{\cdot}16S)}$

two-dimensional filter using standard techniques. We are obliged therefore to seek a suitable approximation to equation (5.1) which *does* have a Laplace transform. Bearing the earlier symmetric filtering expression in mind we note that

$$e^{-([x-x_1]^2+[y-y_1]^2)^{1/2}} \doteqdot e^{-|x-x_1|/T} \,.\, e^{-|y-y_1|/T} \qquad (5.2)$$

The left-hand side of equation (5.2) is of course perfectly symmetrical about any point and therefore contours of equal altitude will be concentric circles (see Fig. 4). The right-hand side, however, varies by a factor of $e^{-\sqrt{2}} = 0{\cdot}66$ according to direction if a circular path is traversed about the point of attention, i.e. the value of the function on the right-hand side of equation (5.2), at any distance, 'r' along a diagonal, is 0·66 of the value at the same distance 'r' along any axis.

The contours of equal altitude for this case are therefore concentric diamonds (see Fig. 4). It is worth noting at this point that these contours are produced by the simplest possible smoothing filter. It is in fact possible to produce sets of contours which tend toward circles as r tends towards zero by suitable choice of filter type. In particular the quadratic smoothing filter $1/(a + bS + S^2)$ produces this effect (see **Fig. 5**).

Having stated this possibility the rest of the paper will be confined to the simplest types of filter which have impulse responses with diamond symmetricity at all values of r. Equation (5.2) thus provides us with an approximation to the desired circularly symmetric impulse response of equation (5.1) which we can implement in the following way.

For exponential smoothing the two-dimensional convolution integral now assumes the form

$$\begin{aligned} f^*(x, y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-[|\alpha|+|\beta|]/T} \,.\, f(x-\alpha, y-\beta)\, d\alpha\, d\beta \\ &= \int_{-\infty}^{\infty} e^{-|\beta|/T} \int_{-\infty}^{\infty} e^{-|\alpha|/T} \,.\, f(x-\alpha, y-\beta)\, d\alpha\, d\beta \end{aligned} \qquad (5.3)$$

or in operator form:

$$f^*(x, y) = \phi_y[\phi_x \,.\, f(x, y)] \qquad (5.4)$$

Hence we see that, given a two-dimensional array of data the desired filter can be approximated by two separate transformations. Firstly each row of the array is processed in the manner described by equation (4.9) and secondly each column of the resulting array is operated on in the same way. Thus every point of the array is operated on by the same expression four times, albeit the sense of the operation is different each time.

It is again evident that this process can be generalised to the whole range of filtering operations and the following section will illustrate the effects of band pass as well as smoothing filters on two-dimensional data. The extension to more dimen-

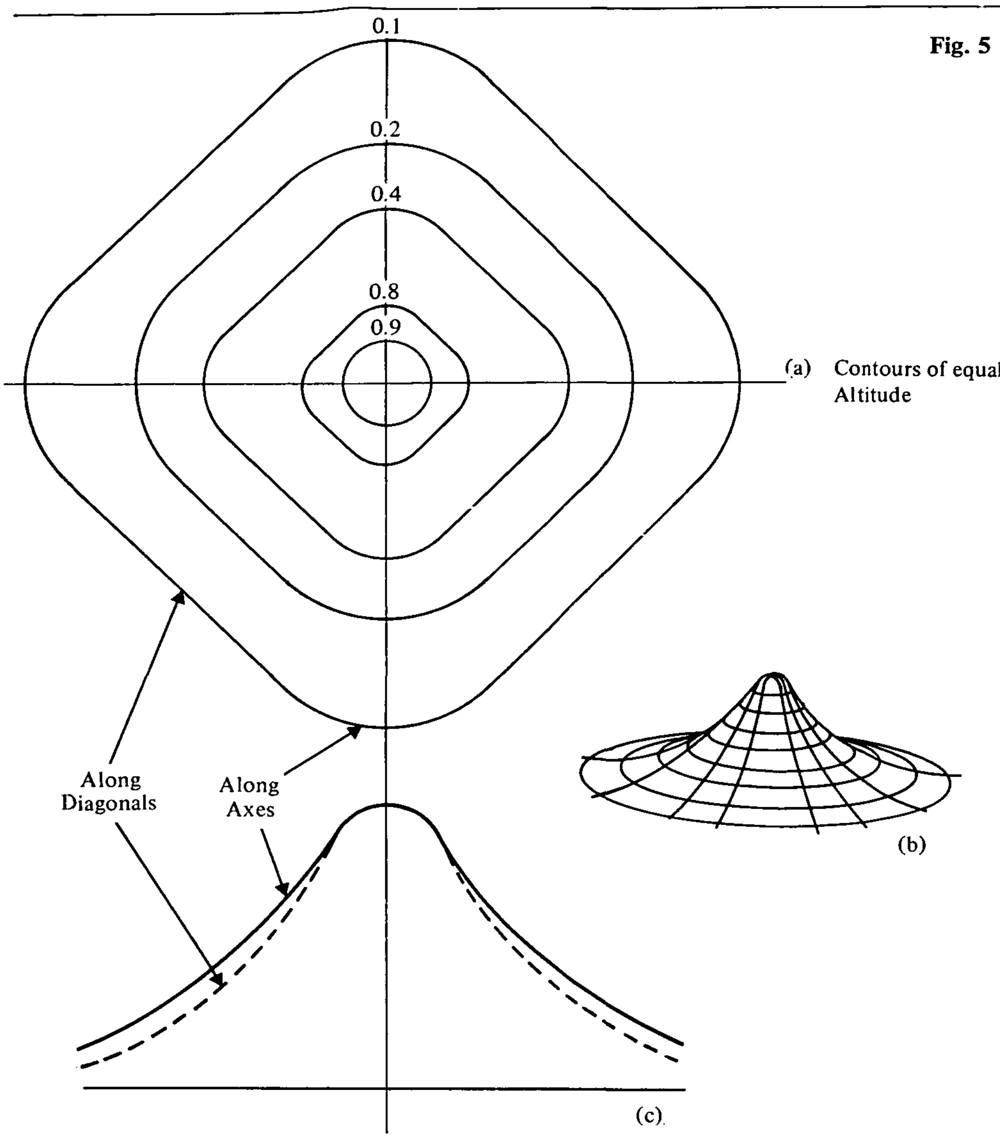
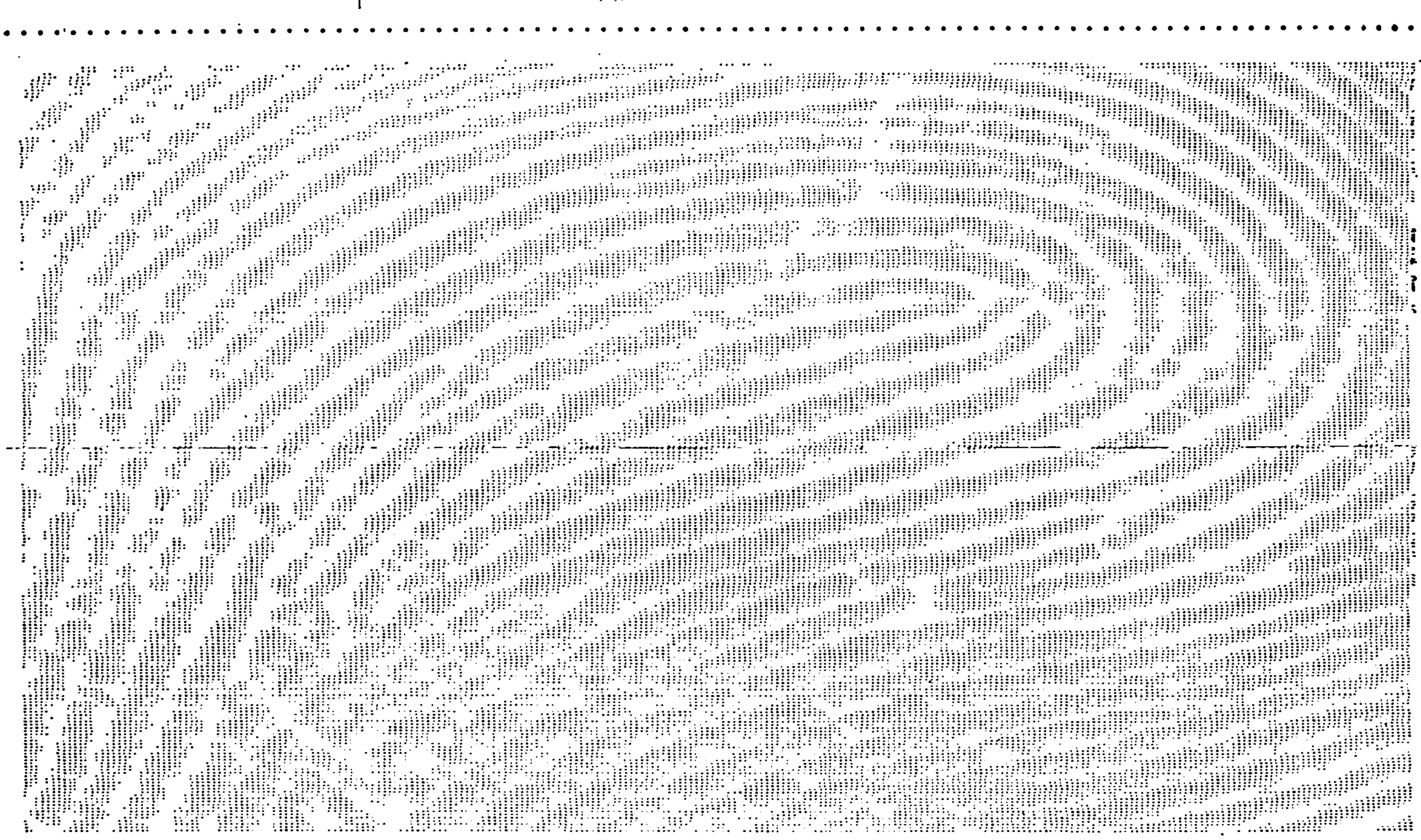


Fig. 5 left. Impulse responses of $F(S) = \frac{1}{(2+3S+)S^2}$

Fig. 6 below. Digitised fingerprint coded on an eight level grey scale



sions obviously follows the same process of cascading the filters for each dimension as indicated in equation (5.4).

6. Application to fingerprint enhancement

The fingerprint data for these tests was obtained by means of a flying spot scanner of the reflectance type. The encoded data consists of a 256 × 256 matrix of pattern points. Each point consists of a 6-bit code representing the intensity of the reflected light at the corresponding point on the fingerprint image thus providing 64 levels of intensity ranging from 0 black) to 63 (white). The sampling frequency is approximately 17 points/mm which is rather on the low side. 20 to 30 points/mm would have been better in order to keep the dominant frequency (corresponding to the average ridge spacing) well away from the sampling frequency cut-off point. This margin is desirable in order to observe the smoothing effect of the filters clearly. **Fig. 6** is a line printer representation of the basic input data coded on an 8-level grey scale. The eight characters were chosen to provide a coarse shading effect corresponding to the light intensity on the pattern in the following way:

0 to 7	@
8 to 15	#
16 to 23	*
24 to 31	+
32 to 39	–
40 to 47	'
48 to 55	.
56 to 63	space

The other fingerprint representations (**Figs. 7-9**) have all been reduced to binary patterns by a simple thresholding process at level 30. Fig. 7 is the thresholded version of Fig. 6. We are immediately able to appreciate the problems associated with simple thresholding in this case. There is a large amount of 'salt and pepper' noise present and in several areas the ridges have been removed entirely by the threshold—due to a general lightening in that region.

Fig. 8 illustrates the effect of a two-dimensional smoothing filter prior to the thresholding process. In this case the filter was

$$F(S) = \frac{2}{1 + 0{\cdot}3S}$$

As can be seen by comparison with Fig. 7 the edges of the lines are more regular and most of the 'salt and pepper' noise has been removed. However, a lot of the fine detail has also been lost and the residual lines in the lower left corner and elsewhere have been almost completely suppressed.

Fig. 9 demonstrates the advantages of using a band pass filter, as opposed to a simple smoothing filter, prior to the thresholding process. The filter used has the basic form

$$F(S) = 2 \, . \, \frac{0{\cdot}3S}{(1 + 0{\cdot}01S)\,(1 + 0{\cdot}3S)}$$

In this case we see that, in addition to smoothing the edges of the lines and removing the 'salt and pepper' noise, the removal of the low frequencies has preserved those line details in the lower left corner, and elsewhere, that were removed by the simpler processes.

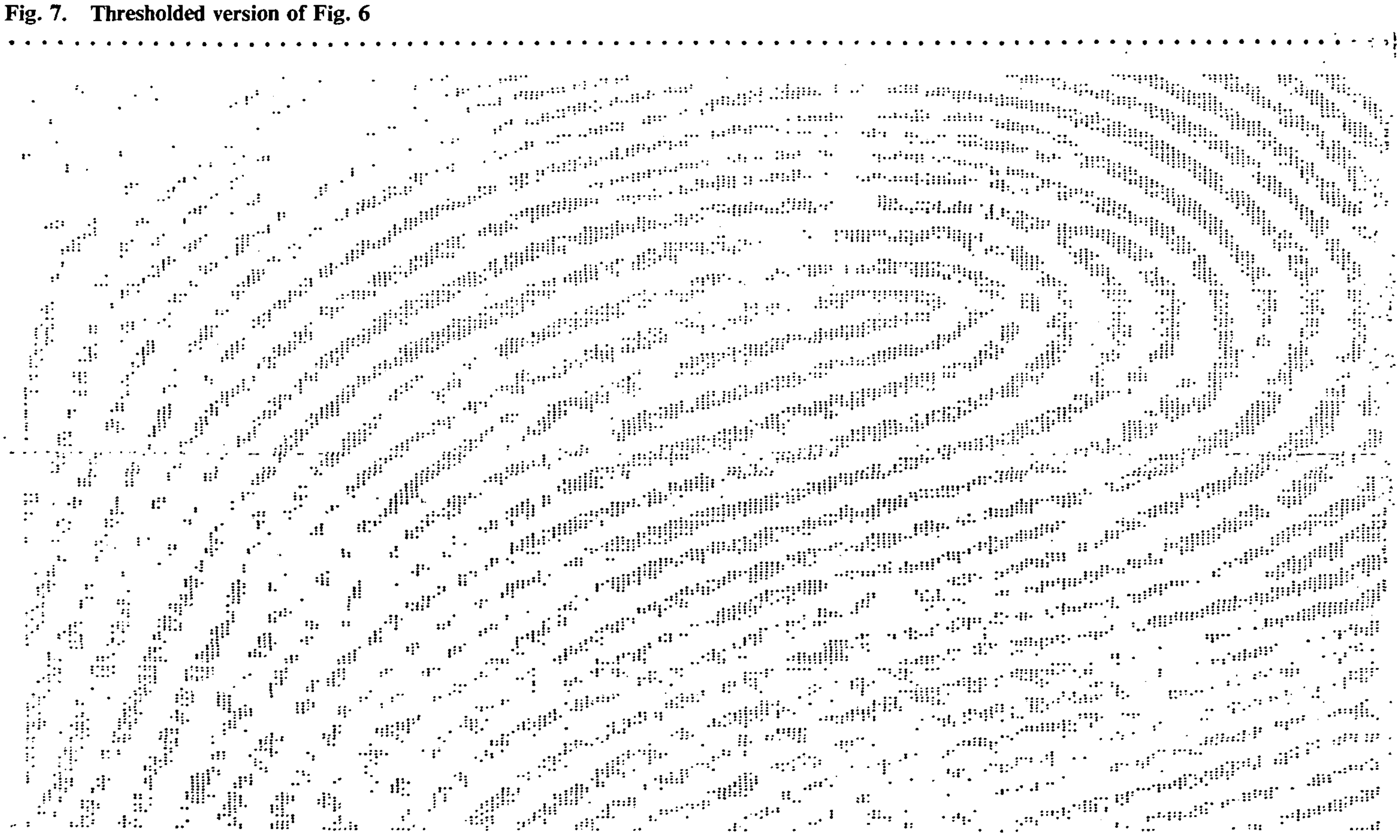
A high pass filter was not used on the fingerprint data because of the degradation it produces by accentuating the 'salt and pepper' type noise.

Errors

The sources of numerical error to which the scheme is subject are:

1. Truncation effects of the arithmetic operations involved. In the examples illustrated the data is stored in character form and therefore, because the numbers are only represented by 6 bits during the various transformations, rounding errors became significant. However, despite this, the filters effect has been clearly demonstrated.

Fig. 7. Thresholded version of Fig. 6



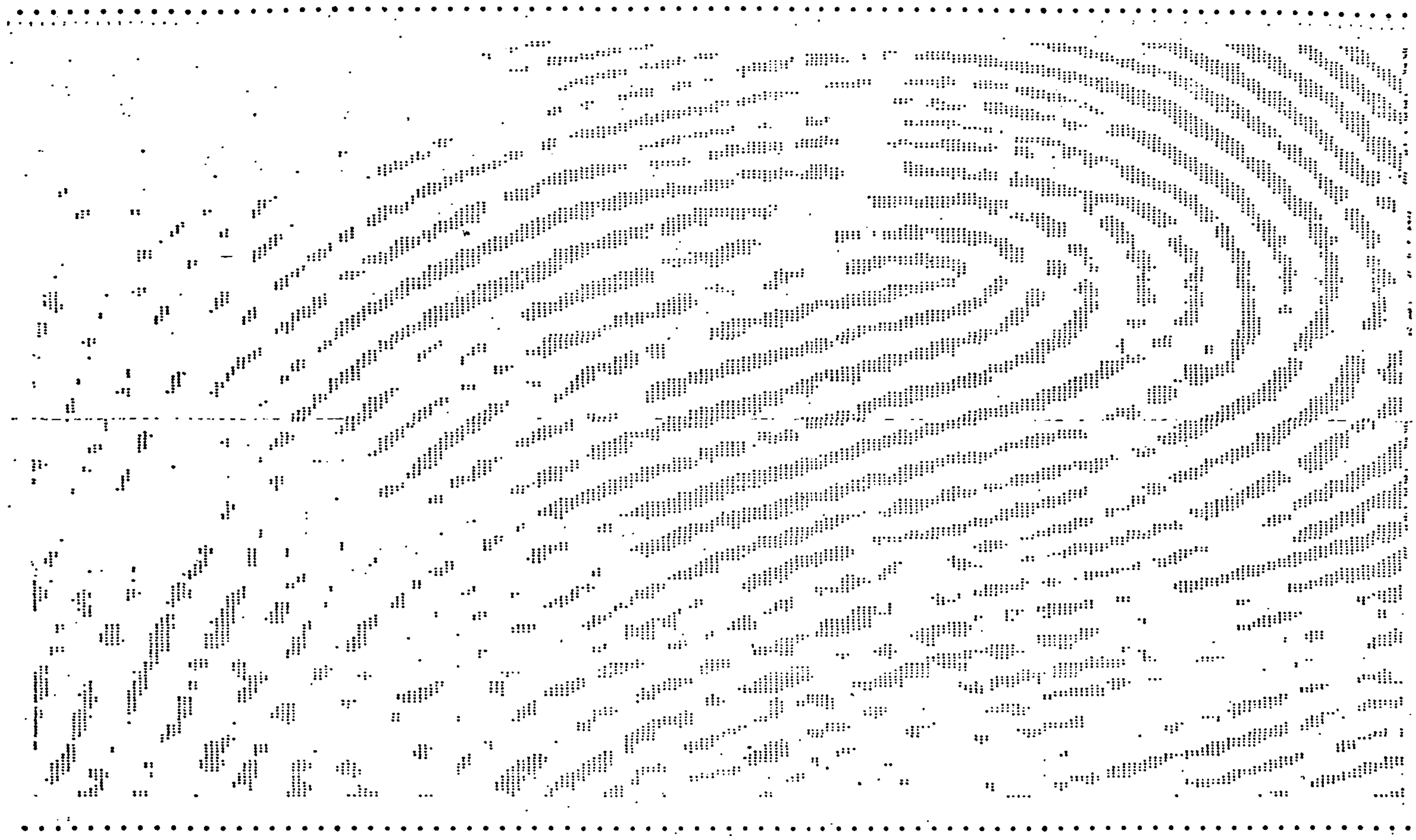


Fig. 8. Low pass filtered and thresholded version of Fig. 6

2. Sampling frequency cut off. This effectively limits the higher frequency effects.
3. Boundary effects. This effect limits the low frequency discrimination to frequencies whose period is of the same dimension as the pattern matrix.
4. Errors due to the low order difference equation used (see Appendix 1) are insignificant for this application but in any case they can be reduced further by using higher order approximations to the various derivatives involved.
5. Errors due to the aspherical nature of the filter, which in fact has triangular symmetry about any axis, vary from 0 to 0·3 between any axis and the diagonal through the adjacent quadrant. Such effects, however, are not pronounced in the immediate vicinity of a given point of operation as is illustrated in Fig. 5 and can be dramatically reduced by increasing the order of the filter. In fact, for all practical purposes, circularity can be obtained in the region of the operating point by using quadratic filters (see Fig. 5).

7. Conclusions

A method for affecting spatially symmetric filtering operations has been derived. The effectiveness of the filters has been demonstrated by their application to the enhancement of fingerprint images and their usefulness in smoothing and band-pass operations has been clearly shown. This paper reports on a preliminary study of the technique in which the data was held to an accuracy of 6 bits. Further work is being conducted using much greater precision which should improve the filter response significantly.

The theory produces an extremely simple implementation (compare for instance with Cole and Davie, 1969) which, in the case of two-dimensional data, is more flexible than the equivalent optical technique (though obviously slower in response) and has arbitrary accuracy.

The method employs a mapping from one data array to another using a one-dimensional recursive operator used repeatedly. Because recursive filters involve feedback they do not suffer from the usual problem, associated with the mesh variety, of steady state errors; the influence of any impulse being distributed over the entire array rather than over the range of the mesh.

Further, because the method employs well-known and well-tried standard forms with well-defined characteristics, precise filtering operations can be readily specified.

Acknowledgements

The programs for this study were written in the ICL PLAN language and executed on the ICL 1905 at the City University under the auspices of the Mathematics Department at the City University.

Access to the flying spot scanner at the National Physical Laboratory was kindly provided by Dr. John Parks of the Computer Sciences Division of the NPL.

Appendix 1 Generation of recursive numerical filters

As illustrative examples of the process we will consider the simplest possible smoothing filter and the simplest possible band pass filter. These filters belong to the general class

$$F(S) = \frac{a_0 + a_1 S + a_2 S^2 + \ldots a_n S^n}{b_0 + b_1 S + b_2 S^2 + \ldots b_m S^m}, \quad m \geqslant n$$

where we make the substitution $S = d/dx$.

Given that the ith point in the discrete stream of input data to the filter is represented by $f_i(x)$, the associated output from the filter is given by

$$f_i^*(x) = F(d/dx) . f_i(x)$$

Using backward difference approximations we note that

$$Sf_i = d/dx(f_i(x)) = f_i(x) - f_{i-1}(x)$$

and

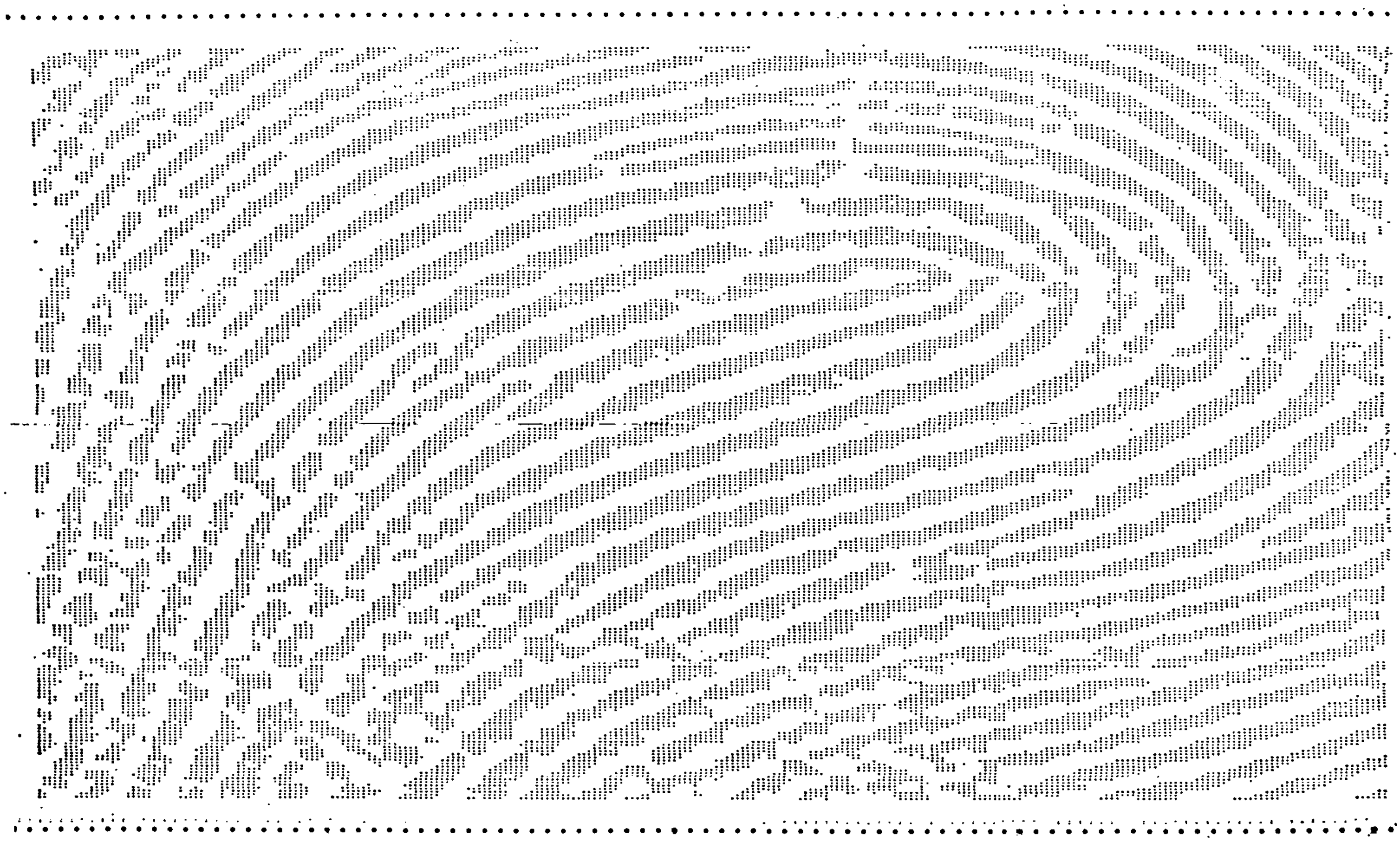


Fig. 9. Band pass filtered and thresholded version of Fig. 6

$$S^2 f_i = d^2/dx^2(f_i(x)) = f_i(x) - 2f_{i-1}(x) + f_{i-2}(x)$$

These expressions are used in the following two types of filter:

Low pass or smoothing filter

$$f_i^* = \left[\frac{K}{1 + TS}\right] f_i$$

or

$$[1 + TS] f_i^* = K f_i$$

or

$$[1 + T] f_i^* - T f_{i-1}^* = K f_i$$

or

$$f_i^* = \left[\frac{1}{1 + T}\right] [K f_i + T f_{i-1}^*]$$

and if $K = T = 1{\cdot}0$ we have the simple form

$$f_i^* = \tfrac{1}{2}[f_i + f_{i-1}^*]$$

which is very useful for test purposes.

Band pass filter

$$f_i^* = \left[\frac{K T_1 S}{(1 + T_1 S)(1 + T_2 S)}\right] f_i$$

or

$$[1 + (T_1 + T_2) S + T_1 T_2 S^2] f_i^* = K T_1 S f_i$$

or

$$f_i^* + (T_1 + T_2) [f_i^* - f_{i-1}^*] + T_1 T_2 [f_i^* - 2f_{i-1}^* + f_{i-2}^*] = K T_1 [f_i - f_{i-1}]$$

or

$$f_i^* = \left[\frac{1}{1 + T_1 + T_2 + T_1 T_2}\right] [K T_1 (f_i - f_{i-1}) + (T_1 + T_2 + 2T_1 T_2) f_{i-1}^* - T_1 T_2 f_{i-2}^*]$$

References

Cole, A. J., and Davie, A. J. T. (1969). Local Smoothing by Polynomials, *The Computer Journal*, Vol. 12, No. 1.

Constantinides, A. G. (1968). Synthesis of Recursive Digital Filters, University of London, Ph.D. Thesis, Electrical Engineering Department, The City University, St. John Street, London EC1.

Shinners, S. M. (1964). *Control System Design*, Wiley.