Sādhanā, Vol. 18, Part 2, June 1993, pp. 349–364. © Printed in India.

A generalized approach to the reconstruction of a restricted class of digitized planar curves

S CHATTOPADHYAY+ and P P DAS*

+ Department of Computer Science & Engineering, Jadavpur University, Calcutta 700 032, India

* Department of Computer Science & Engineering, Indian Institute of Technology, Kharagpur 721 302, India

Abstract. Reconstruction of an original continuous curve and the estimation of its parameters from the digitized version of the curve is a challenging problem, as quantization always causes some loss of information. In this paper, we have developed a scheme for reconstruction which is applicable to a class of curves having at the most two parameters. The class of curves for which the scheme works has also been characterized. We have shown that for one-parameter curves the exact domain of values of the parameter can be obtained. But in the two-parameter case, only the smallest rectangle containing the domain can be realised. The distinctive feature of our scheme is that it provides a unified approach to solve the reconstruction and the domain-finding problem for a class of curves.

Keywords. Discrete geometry; digitization; reconstruction; domain of digitization.

1. Introduction

Reconstruction of the original continuous curve from a given set of digital points, representing its discretization, is an important problem in pattern recognition and image processing. Since discretization inevitably causes some loss of information, the exact reconstruction of the original curve from its digitization is, in general, impossible. This loss of information opens up a number of related questions, most important among them being:

> "Is it possible to obtain the given digital image from a particular kind of curve?"

Consider the class of planar curves algebraically characterized by an equation $f(x, y, P) = 0$ where x, y are spatial variables, $P = (p_1, \ldots, p_k)$ is a vector of control parameters and f is the function relating x, y and P. Also let D_o be the given digital image data. Then the above question can be restated as:

> "Does there exist a vector $P_1 \in R^k$ (R denotes the set of real numbers) so that the digitization of $f(x, y, P_1) = 0$ yields the given digital data D_o?"

This problem is known as the *reconstruction problem* of a quantized image D_o with respect to a given function f. A solution to this problem finds out *at least* one P such that $D(f(x,y,P)=0)=D_o$, where D is the digitization operator.

As a generalization of the reconstruction problem we may also want to find out the set of *all* vectors $P, P \in R^k$ so that digitization of $f(x,y,P)=0$ gives the input image. The solution to this latter problem, called the *domain construction problem*, is a region S in the k-dimensional parametric space such that

$$\forall P \in S, D(f(x,y,P)=0)=D_o.$$

We may view the discretization procedure D as a transformation of a curve $f(x,y,P_o)=0$ to the corresponding quantized data D_o. Then, the problem is one of finding an inverse mapping from quantization to the specification of the curve. Formally, if $D(f(x,y,P_o)=0)=D_o$ then the *domain of digitization* is defined as

$$\text{Domain}(D_o,f)=\{P \mid P \in R^k \text{ and } D(f(x,y,P)=0)=D_o\}.$$

Note that when we talk about the domain of digitization or the domain of a discrete image, we have a particular type of curve in mind and the meaning of domain has to be understood in that context only. Let us illustrate this point. Let D_o be the following set of points

$$D_o=\{(0,0),(1,0),(2,1),(3,1),(4,1),(5,2),(6,2),(7,3),(8,3)\},$$

and $f(x,y,m,c)$ be $y-mx-c$. Then Domain(D_o,f) means the set of all continuous line segments which yield D_o on digitization. Again if $g(x,y,a,b)$ is $x^2/a^2+y^2/b^2-1$, then Domain(D_o,g) denotes the set of all ellipses with origin at centre and axes parallel to the coordinate axes, which produce the same quantization as D_o.

The reconstruction and domain construction problems for various curves and figures have been addressed in the literature over the last decade. In Anderson & Kim (1983) and Dorst & Smeulders (1984), a discrete representation of straight lines has been investigated in depth and a methodology is presented to find the set of all straight lines yielding a given set of discrete points. More results on straight lines and other simple linear figures may be found in Kim (1982), Kim & Rosenfeld (1982), Nakamura & Aizawa (1985), and Dorst (1986). The reconstruction of circles has been dealt with in Kim (1984) and Nakamura & Aizawa (1984). Most of these papers emphasize the reconstruction problem and the techniques used depend heavily on the geometry of the particular curve under consideration. In Anderson & Kim (1983) and Dorst & Smeulders (1984), there are algorithms to detect whether D_o, a given set of digital points, can be obtained by the digitization of a continuous line segment. But these algorithms cannot be adapted to recognise whether D_o may be regarded as the discrete version of, say, a parabola or an ellipse. Chattopadhyay & Das (1991, 1992) have solved the domain construction problems for straight lines and for a specialized class of conics using a novel and mostly uniform approach. This forms the major motivation of the present work. In this paper we develop an algorithmic scheme which utilises some general properties of the function representing the curve to solve the reconstruction problem. Moreover we also characterize the class of curves that are amenable to our unifying approach of analysis. This renders our approach applicable not only to one type of curve but to a rich class of curves having some common characterization.

The paper is organized as follows. We start with a preliminary discussion on the monotonicity and the digitization process in § 2. In the next section we take up curves with one parameter and develop our methodology to derive the domain of digitization. We treat two parameter curves in § 4. We show that for a restricted class of monotone curves we can find the smallest rectangle enclosing the domain in the parametric space. We finally conclude in § 5.

2. Mathematical preliminaries and digitization

In this section, we clarify the notation used in this paper, define the notion of monotonicity and discuss the digitization scheme. A planar curve with one or two parameters may be represented by the equation $f(x, y, a) = 0$ or $f(x, y, a, b) = 0$, respectively, where x, y denote the spatial variables and a, b denote the unknown parameters. Without any loss of generality, we can assume that the segment of the curve is contained in the first quadrant in a rectangular mesh defined by $0 \leqslant x \leqslant m$, $0 \leqslant y \leqslant n$. Moreover there is no grid line which does not intersect the curve. To be precise we assume that

Assumption 0. The curve is continuous and differentiable in the region of interest.

Our next assumption about the class of curves is that the equation $f(x, y, a, b) = 0$ may be rearranged to write the following four equivalent equations, $x = f_x(y, a, b)$, $y = f_y(x, a, b)$, $a = f_a(x, y, b)$ and $b = f_b(x, y, a)$. We call it the *separability property* of the curve. It is important to note that to solve the reconstruction problem using our scheme the curve has to possess this *separability property*.

Assumption 1. The curve possesses the separability property.

Now let us formalize the concept of monotone curves in our context. In the following definition $\mathbf{x} = (x_1, \ldots, x_n)$ denotes a vector of n variables. Also, $\mathbf{x}^i$ denotes the vector of $(n-1)$ variables $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

DEFINITION

A function $g(\mathbf{x})$ is *monotone decreasing with respect to* x_i if and only if $x_i > x_i'$ implies $g(x_1, \ldots, x_i, \ldots, x_n) < g(x_1, \ldots, x_i', \ldots, x_n)$.

Similarly, $g(\mathbf{x})$ may be defined to be monotone increasing w.r.t. x_i. In other words, $g(\mathbf{x})$ is monotone decreasing if and only if $\partial g/\partial x_i < 0$ over our interval of attention.

DEFINITION

A function $g(\mathbf{x})$ is *monotone* if and only if $\forall x_i, 1 \leqslant i \leqslant n, g(\mathbf{x})$ is either monotone increasing or monotone decreasing w.r.t. x_i. Henceforth, we refer to g as an increasing (a decreasing) function if g is monotone increasing (decreasing).

DEFINITION

We say that a curve given by $f(\mathbf{x}) = 0$ is monotone if and only if the function f is monotone.

Assumption 2. The curve is monotone.

As we shall see in theorem 1, if f is monotone and $x_i = f_i(\mathbf{x}^i)$, then the function f_i is also monotone for all $i, 1 \leqslant i \leqslant n$ and vice versa. Then it follows from the above discussions that we can capture the monotonicity of a particular curve $f(\mathbf{x}) = 0$, if we know the nature of the monotonicity of the individual functions f_i, $1 \leqslant i \leqslant n$. This can be easily depicted by a matrix, which we shall call the *monotonicity matrix*.

DEFINITION

We define a monotonicity matrix (MM) as an $n \times n$ matrix whose rows correspond to the functions $f_i, 1 \leqslant i \leqslant n$, and the columns correspond to variables $x_i, 1 \leqslant i \leqslant n$, respectively. The entry in a cell for such a matrix is either I or D. $\mathrm{MM}(i,j) = I$ (or D) indicates that the function f_i is increasing (decreasing) with respect to the variable x_j.

Although the monotonicity matrix fully describes the nature of monotonicity of a curve, the enormous number of MM depicting different kinds of monotone curves is a severe drawback for the easy handling of such matrices. Moreover such a characterization of a monotone curve is not obtained directly from the equation of the curve. The next theorem, however, offers a solution.

Theorem 1. *For any separable function* $f(\mathbf{x})$ *we have* $(\partial f_i/\partial x_j) > 0$, *if and only if* $(\partial f/\partial x_i)\cdot(\partial f/\partial x_j) < 0$ *for* $1 \leqslant i, j \leqslant n$, $i \neq j$.

Proof. Since all other variables are treated as constants, we may write $f(\mathbf{x}) = 0$ as $f(x_i, x_j) = 0$. Restated differently, $x_i = f_i(x_j)$. We know that $(\partial f_i/\partial x_j) = -(\partial f/\partial x_j)/(\partial f/\partial x_i)$.

Hence $(\partial f_i/\partial x_j) > 0$, iff $(\partial f/\partial x_i)$ and $(\partial f/\partial x_j)$ have opposite signs. Q.E.D.

The above theorem helps us to construct the monotonicity matrix if we define a partial derivative sign vector (PDSV).

DEFINITION

A PDSV of a function $f(\mathbf{x})$ is an n-tuple where the ith element stands for the sign of the partial derivative $(\partial f/\partial x_i)$. If π is a PDSV then π_i, the ith component of π, is '+' if $(\partial f/\partial x_i)$ is positive and '−' if it is negative.

If we consider a function $f(x, y, a)$ then every PDSV π of f is a three-tuple. π_1, π_2, π_3 denote the signs of $(\partial f/\partial x)$, $(\partial f/\partial y)$ and $(\partial f/\partial a)$ respectively. Since each of these π_i may be either a '+' or a '−', there may be at most eight PDS vectors for the given function as are listed in figure 1.

Every PDSV corresponds to a unique MM in the following manner. Let π be the given PDSV.

$$\mathrm{MM}(i,j) = I, \text{ if } \pi_i \neq \pi_j$$
$$= D, \text{ if } \pi_i = \pi_j.$$

This is a direct consequence of the definition of $\mathrm{MM}(i,j)$ and theorem 1.

$PDSV_1 = (+, +, +)$

$PDSV_2 = (+, +, -)$

$PDSV_3 = (-, +, -)$

$PDSV_4 = (+, -, -)$

$PDSV_5 = (-, -, -)$

$PDSV_6 = (-, -, +)$

$PDSV_7 = (+, -, +)$

$PDSV_8 = (-, +, +)$

Figure 1. Listing of all PDSV for one-parameter planar curves.

Clearly then, there can be no more than 2^n (number of n-tuple PDSV) distinct monotonicity matrices defining the monotone functions. We use the following properties of PDS vectors to identify equivalent MM (with respect to our analysis). We assume that f is a d-dimensional curve with $k = n - d$ unknown parameters, i.e. $f(x_1, x_2, \ldots, x_d, a_1, \ldots, a_{n-d}) = 0$.

Property 1. (Complementation) The complement π_c of a π, where a '+' (or '−') in π is changed to a '−' (or '+') in π_c, denotes the same monotonicity matrix as π.

Property 2a. (Permutation of spatial variables) Since x_i and x_j can be interchanged (with proper adjustment of the coordinate system) $\pi = (\pi_1, \ldots, \pi_{i-1}, \pi_i, \pi_{i+1}, \ldots, \pi_{j-1}, \pi_j, \pi_{j+1}, \ldots, \pi_d, \ldots, \pi_n)$ and $\pi' = (\pi_1, \ldots, \pi_{i-1}, \pi_j, \pi_{i+1}, \ldots, \pi_{j-1}, \pi_i, \pi_{j+1}, \ldots, \pi_d, \ldots, \pi_n)$ define equivalent monotone classes of curves.

Property 2b. (Permutation of non-spatial variables) As a_i and a_j can be interchanged by renaming, their order also becomes immaterial in a PDSV.

In this paper we have treated only one ($k = 1$) or two ($k = 2$) parameter planar ($d = 2$ and $n = 3$ or 4) curves. Thus at most eight PDS vectors may arise in the first case while their number may go up to sixteen in the latter. In figure 1, the last four PDS vectors are complements of the first four. Therefore by property 1 we consider the first four PDS vectors only. Again using property 2a the third PDSV and the fourth PDSV are equivalent. Hence there are three PDS vectors to consider for a planar curve with one parameter. Similarly, the number of distinct PDS vectors reduces to five for curves with two parameters. The different PDS vectors are listed in figure 2. The MM for $PDSV_1$ in figure 2a and $PDSV_3$ in figure 2b are shown in figures 3 and 4 respectively.

Next let us discuss the digitization scheme used. While many different schemes like object boundary quantization (OBQ) (Dorst & Smeulders 1984), grid intersection quantization (GIQ) (Freeman 1970, pp. 241–66), Kim's (1982) digitization etc. are in use for different tasks, OBQ is specially suitable for image analysis purposes (Dorst 1986). Hence, we have selected the OBQ scheme for our approach. The OBQ image of a curve is the set of grid (digital) points nearest to the curve that lie consistently on one side of the curve. Whenever the curve is closed (a part of which is under consideration), we can uniquely define the inside or outside of the curve using Jordan's theorem. But

$\text{PDSV}_1 = (-, +, -)$

$\text{PDSV}_2 = (+, +, -)$

$\text{PDSV}_3 = (+, +, +)$

(a)

$\text{PDSV}_1 = (+, -, -, -)$

$\text{PDSV}_2 = (+, -, -, +)$

$\text{PDSV}_3 = (+, +, -, -)$

$\text{PDSV}_4 = (+, +, +, -)$

$\text{PDSV}_5 = (+, +, +, +)$

(b)

Figure 2. Listing of distinct PDSV for (a) one- and (b) two-parameter planar curves in the first quadrant.

inside or outside of open curves (like straight lines, parabolas etc.) is not defined at all. For closed curves, therefore, when we say that we travel the curve in a clockwise direction it has a meaning. The OBQ and GIQ image of a closed contour is shown in figure 5. However, for open curves such a sense of traversal does not exist. So in case of open curves, we shall deliberately close it by connecting it to the x-axis by vertical lines. After doing so we can traverse the curve clockwise. Now let us formally define the OBQ technique.

DEFINITION

$I(f)$, the OBQ image of a curve given by $f(x, y, a, b) = 0$, is the set of digital points obtained as follows.

While traversing f *clockwise* (i) whenever f passes through a grid point P, then P belongs to $I(f)$, (ii) whenever f crosses a grid line L but not a grid point, the nearest grid point to the *right* of the curve and on L is a point of $I(f)$, (iii) no other point is included in $I(f)$. Thus for a closed contour $f = 0$, $I(f)$ is the set of grid points inside f and "nearest" to its boundary. Note also that we should collect the points on the left of f if we traverse it anti-clockwise.

	x	y	a
f_x	-	I	D
f_y	I	-	I
f_a	D	I	-

Figure 3. Monotonicity matrix corresponding to PDSV_1 of figure 2a.

	x	y	a	b
f_x	-	D	I	I
f_y	D	-	I	I
f_a	I	I	-	D
f_b	I	I	D	-

Figure 4. Monotonicity matrix corresponding to $PSDV_3$ of figure 2b.

This definition, though precise, is not algebraic in nature. Fortunately, under the assumptions of monotonicity, $I(f)$ can be expressed through easy formulas also. To see this, we first define another set of digital points as D from $f = 0$ and subsequently prove its equivalence with $I(f)$.

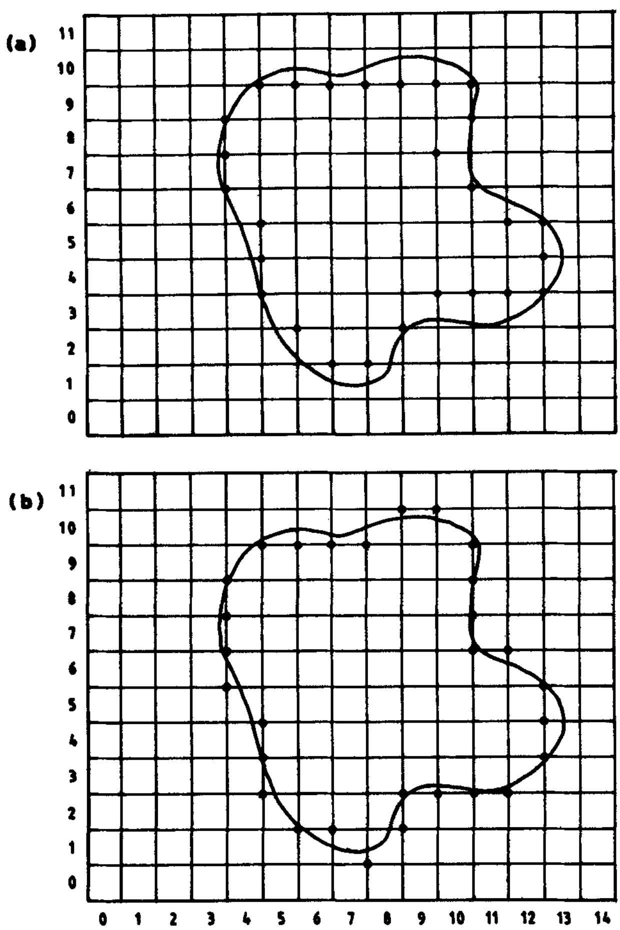

Figure 5. Illustration of (**a**) OBQ and (**b**) GIQ digitization schemes for an arbitrary closed contour.

DEFINITION

Given f, $D = D(f)$ is a set of grid points defined as follows:

$$D = D_x \cup D_y \text{ where}$$
$$D_x = \{(x_i, i) | x_i = \lfloor f_x(i, a, b) \rfloor, 0 \leqslant i \leqslant n\} \text{ if } M(f_x, y) = D.$$
$$= \{(x_i, i) | x_i = \lceil f_x(i, a, b) \rceil, 0 \leqslant i \leqslant n\} \text{ if } M(f_x, y) = I.$$
$$D_y = \{(i, y_i) | y_i = \lfloor f_y(i, a, b) \rfloor, 0 \leqslant i \leqslant m\}.$$

The above expressions can easily be justified for the given monotonicity. We can informally say that if a curve is decreasing in the first quadrant then it bends down (or bends to the left) as its x-coordinate (or y-coordinate) increases. The converse holds if the curve is increasing in the first quadrant. The following theorem presents the equivalence between D and $I(f)$.

Theorem 2. *For a monotone curve f, $I(f) = D(f)$.*

Proof. Directly follows from the definitions of $D(f)$ and $I(f)$.
We can also easily separate out the sets D_x and D_y from $I(f)$. So while $I(f)$ may be obtained experimentally from the acquired image data, it suffices to carry out the theoretical analysis using D_x and D_y only. The reader may note that the digitization for a one-parameter curve can be analogously defined.

3. One parameter case

As already mentioned, the equation of the curve is given by $f(x, y, a) = 0$ which may be rewritten as $x = f_x(y, a)$, $y = f_y(x, a)$ and $a = f_a(x, y)$. Henceforth, we shall denote the original value of a by a_o, and D_o will denote the digitization of $f(x, y, a_o) = 0$ i.e., $D_o = D(f(x, y, a_o))$. For the ease of analysis, let us consider one particular monotonicity matrix which is given in figure 3. In this case, $x_i = \lceil f_x(i, a_o) \rceil$ and $y_i = \lfloor f_y(i, a_o) \rfloor$. Using the properties of $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ functions, we can write,

$$f_x(i, a_o) \leqslant x_i < f_x(i, a_o) + 1 \text{ and}$$
$$f_y(i, a_o) - 1 < y_i \leqslant f_y(i, a_o).$$

As f_a is decreasing in x,

$$f_a(x_i, i) \leqslant a_o = f_a(f_x(i, a_o), i) < f_a(x_i - 1, i). \tag{1}$$

Similarly as f_a is increasing in y,

$$f_a(i, y_i) \leqslant a_o = f_a(i, f_y(i, a_o)) < f_a(i, y_i + 1). \tag{2}$$

From (1) and (2) we obtain two bounds of a_o as follows:

$$a_l = \text{Max}(\text{Max}_i(f_a(x_i, i)), \text{Max}_i(f_a(i, y_i))),$$
$$a_u = \text{Min}(\text{Min}_i(f_a(x_i - 1, i)), \text{Min}_i(f_a(i, y_i + 1))).$$

For any value of a such that $a_l \leqslant a < a_u$,

$$f_a(x_i - 1, i) \geqslant a_u > a \geqslant a_l \geqslant f_a(x_i, i).$$

That is,

$$f_a(x_i - 1, i) > a \geqslant f_a(x_i, i).$$

Hence

$$x_i = f_x(i, f_a(x_i, i)) \geqslant f_x(i, a) > f_x(i, f_a(x_i - 1, i)) = x_i - 1.$$

Rearranging,

$$f_x(i, a) \leqslant x_i < f_x(i, a) + 1.$$

Consequently,

$$x_i = \lceil f_x(i, a) \rceil.$$

Similarly

$$\forall a, a_l \leqslant a < a_u, y_i = \lfloor f_y(i, a) \rfloor.$$

In fact, if $a < a_l$, it is easy to see that $D(f(x, y, a))$ will miss some point (x_i, i) or (i, y_i) of D_o. Also, if $a > a_u$ then some point is included in $D(f(x, y, a))$ which is not in D_o. Thus, we have the following theorem.

Theorem 3. *$a_l \leqslant a < a_u$ iff $D(f(x, y, a)) = D_o$. In other words, Domain of $D_o = (a_l, a_u)$.*

We conclude this section by giving one example from Chattopadhyay & Das (1992).

Example 1. Let $f \equiv y^2 - 4ax = 0$. Here, the monotonicity matrix is the same as given in figure 3. The different functions are listed below.

$$x = f_x(y, a) = y^2/4a,$$

$$y = f_y(x, a) = (4ax)^{1/2},$$

$$a = f_a(x, y) = y^2/4x.$$

It has been shown (Chattopadhyay & Das 1992) that

$$a_l = \max(\max_i i^2/4x_i, \max_i y_i^2/4i), \text{ and}$$

$$a_u = \min(\min_i i^2/4(x_i - 1), \min_i (y_i + 1)^2/4i).$$

These formulae are the same as the ones given in this section.

4. Two-parameter case

In contrast to the one-parameter case, we cannot formulate a straightforward analysis scheme of the discrete point inequalities like (1) or (2) to determine the possible domain of a and b values. However, we construct an iterative refinement technique using a similar approach. Again we shall consider one particular MM to highlight the theme of our analysis. The MM we are considering is given in figure 4.

As in the last section, let a_o, b_o denote the original values of a and b from which the given digitization D_o has been derived. Therefore, $x_i = \lfloor f_x(i, a_o, b_o) \rfloor$ and $y_i = \lfloor f_y(i, a_o, b_o) \rfloor$. We present the main result in the following theorem.

Theorem 4. *Let $a_l^{k+1}, b_u^{k+1}, a_u^{k+1}$ and b_l^{k+1} be defined by the following iterative algorithm for $k \geqslant 0$.*

$$a_l^{k+1} = \operatorname*{Max}_i(f_a(x_i, i, b_u^k)),$$

$$b_u^{k+1} = \operatorname*{Min}_i(f_b(i, y_i + 1, a_l^k)),$$

$$a_u^{k+1} = \operatorname*{Min}_i(f_a(x_i + 1, i, b_l^k)), \text{ and}$$

$$b_l^{k+1} = \operatorname*{Max}_i(f_b(i, y_i, a_u^k)).$$

With a proper choice of a_l^0, b_u^0, a_u^0 and b_l^0 that satisfies

(a) $a_l^0 \leqslant a_o < a_u^0, b_l^0 \leqslant b_o < b_u^0$, and

(b) $a_l^1 \geqslant a_l^0, b_u^1 \leqslant b_u^0, a_u^1 \leqslant a_u^0, b_l^1 \geqslant b_l^0$,

there exists a_l, a_u, b_l, b_u such that

(A) $\lim_{k\to\infty} a_l^k = a_l$, $\lim_{k\to\infty} b_u^k = b_u$, $\lim_{k\to\infty} a_u^k = a_u$ and $\lim_{k\to\infty} b_l^k = b_l$ and

(B) $a_l \leqslant a_o < a_u, b_l \leqslant b_o < b_u$.

Proof. First, let us show that $a_l^k \leqslant a_o$ implies $b_u^{k+1} > b_o$ for all k.
Since

$$y_i = \lfloor f_y(i, a_o, b_o) \rfloor, \text{ we can write}$$

$$f_y(i, a_o, b_o) - 1 < y_i \leqslant f_y(i, a_o, b_o).$$

Rearranging,

$$y_i + 1 > f_y(i, a_o, b_o) \geqslant y_i.$$

As f_b is increasing in y and $y_i + 1 > f_y(i, a_o, b_o)$, we get,

$$b_o = f_b(i, f_y(i, a_o, b_o), a_o) < f_b(i, y_i + 1, a_o).$$

Further, f_b being a decreasing function in a, we obtain from $a_l^k \leqslant a_o, b_o < \text{Min}_i(f_b(i, y_i + 1, a_o)) < \text{Min}_i(f_b(i, y_i + 1, a_l^k)) = b_u^{k+1}$. Similarly, $\forall k, b_u^k > b_o$ implies $a_l^{k+1} \leqslant a_o$. From the statement of the theorem, we know that $a_l^0 \leqslant a_o$ and $b_u^0 > b_o$. Therefore, $\forall k, k \geqslant 0, a_l^k \leqslant a_o$ and $b_u^{k+1} > b_o$.

Next, we prove that $a_l^{k+1} \geqslant a_l^k$ implies $b_u^{k+2} \leqslant b_u^{k+1}$ for $k \geqslant 0$.

$$a_l^{k+1} \geqslant a_l^k \Rightarrow \forall i[(f_b(i, y_i + 1, a_l^k)) \geqslant f_b(i, y_i + 1, a_l^{k+1})]/$$

$$(\text{since } M(f_b, a) = \text{'}D\text{'})$$

$$\Rightarrow \text{Min}_i(f_b(i, y_i + 1, a_l^k)) \geqslant \text{Min}_i(f_b(i, y_i + 1, a_l^{k+1}))$$

$$\Rightarrow b_u^{k+1} \geqslant b_u^{k+2}.$$

Again, as $a_l^1 \geqslant a_l^0$ and $b_u^1 \leqslant b_u^0$, we have the following two monotonic sequences upper- or lower-bounded by a fixed value.

$$a_l^0 \leqslant a_l^1 \leqslant a_l^2 \leqslant \ldots \leqslant a_l^k \leqslant \ldots \leqslant a_o,$$

$$b_u^0 \geqslant b_u^1 \geqslant b_u^2 \geqslant \ldots \geqslant b_u^k \geqslant \ldots > b_o.$$

This shows that there exist limiting values of a_l^k, b_u^k as k tends to infinity, which may be denoted by a_l and b_u. Clearly then, $a_l \leqslant a_o$ and $b_u > b_o$. The remaining part of the proof follows similarly. Q.E.D.

Now, if we consider the curves $f(x, y, a_l, b_u) = 0$ and $f(x, y, a_u, b_l) = 0$, we can find a few interesting properties which are summarized as in the following theorems and lemmas.

The following theorem is required to prove the tightness of the bounds of a_o and b_o.

Theorem 5. *The curves* $f(x, y, a_l, b_u) = 0$ *and* $f(x, y, a_u, b_l) = 0$ *intersect.*

Proof. The proof is by contradiction.

Let $f_{lu} \equiv f(x, y, a_l, b_u) = 0$ be below and to the left of the curve $f_{ul} \equiv f(x, y, a_u, b_l) = 0$ as shown in figure 6.

From the definition of a_l, $a_l \geqslant f_a(x_j, j, b_u)$, and since f_x is increasing w.r.t. a,

$$f_x(j, a_l, b_u) \geqslant f_x(j, f_a(x_j, j, b_u), b_u) = x_j.$$

In other words, $x_j \leqslant f_x(j, a_l, b_u)$. Similarly, $x_j + 1 \geqslant f_x(j, a_u, b_l)$. But as f_{lu} is to the left of f_{ul} at $y = j$, $f_x(j, a_u, b_l) > f_x(j, a_l, b_u)$. Combining,

$$f_x(j, a_l, b_u) - 1 < f_x(j, a_u, b_l) - 1 \leqslant x_j \leqslant f_x(j, a_l, b_u).$$

Now, let us say that (i, j) is a grid point lying in between f_{lu} and f_{ul} as shown in figure 6.

Therefore $x_j + 1 \geqslant f_x(j, a_u, b_l) > i > f_x(j, a_l, b_u) \geqslant x_j$, which is impossible as i is an integer. Hence, no grid point may exist between f_{lu} and f_{ul}.

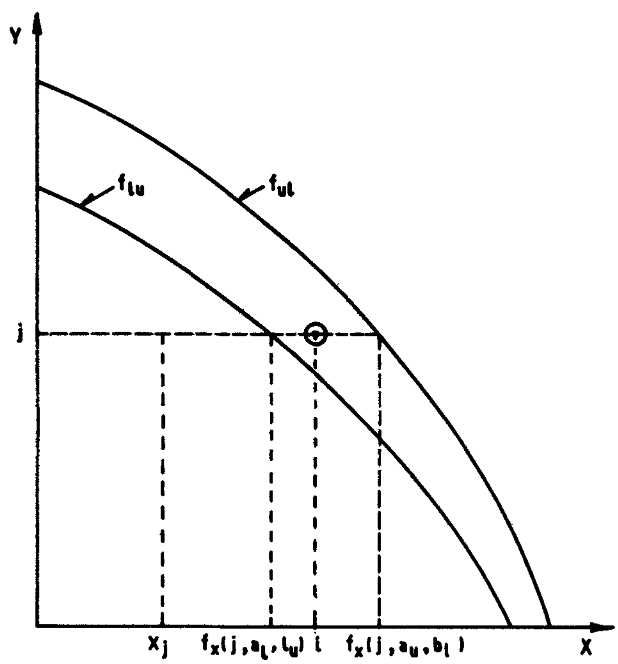

Figure 6. Illustration for the proof of theorem 5.

We can also derive the following inequality

$$f_y(i, a_l, b_u) - 1 < y_i \leqslant f_y(i, a_u, b_l).$$

As no grid point exists in between f_{ul} and f_{lu}, either $y_i \leqslant f_y(i, a_l, b_u)$ or $y_i = f_y(i, a_u, b_l)$. But $y_i = f_y(i, a_u, b_l)$ means that if $y_i = k$ then $x_k = i > f_x(y_i = k, a_l, b_u) \geqslant x_k$. Hence $y_i \neq f_y(i, a_u, b_l)$ and consequently $y_i \leqslant f_y(i, a_l, b_u)$. So, $x_i \leqslant f_x(i, a_l, b_u)$ and $y_i \leqslant f_y(i, a_l, b_u)$. Thus all grid points of D_o lie strictly inside f_{ul} (i.e. no grid points lie on it) which cannot happen for at least one point of D_y. Hence, our assumption is contradicted. A similar contradiction is arrived at if we start with the assumption the f_{ul} is below f_{lu}. Therefore f_{ul} must intersect f_{lu}. Q.E.D.

COROLLARY 1

The open region between f_{lu} and f_{ul} (i.e. excluding the arcs of the curves) do not contain any grid point.

Let us define p, p', q, q' such that $a_l = f_a(x_p, p, b_u), b_l = f_b(p', y_{p'}, a_u)$, $b_u = f_b(q, y_q + 1, a_l)$ and $a_u = f_a(x_{q'} + 1, q', b_l)$.

COROLLARY 2

$D(f_{lu})$ or $D(f_{ul})$ differs from D_o only at point(s) like $(q, y_q + 1)$ or $(x_{q'} + 1, q')$ which define b_u and a_u.

From the above corollary we can say that f_{lu} or f_{ul} are very close approximations of the original curve represented by the function f. If we assume one more property about the curve then we can prove that

(i) the Domain of D_o in the $a - b$ parametric space is contained in the rectangle R_{ul} whose diagonally opposite vertices are (a_l, b_u) and (a_u, b_l); and
(ii) R_{ul} is the smallest rectangle containing the domain.

This assumption is stated below.

Assumption 3. Let, $(\partial f_y / \partial x) = g(x, a, b)$. Then either

$$\forall x, g(x, a_1, b_1) < g(x, a_2, b_2), \text{ for } a_1 < a_2 \text{ and } b_1 > b_2, \text{ or}$$

$$\forall x, g(x, a_1, b_1) > g(x, a_2, b_2), \text{ for } a_1 < a_2 \text{ and } b_1 > b_2.$$

In terms of the limiting and original curves the above assumption helps us in proving the following properties.

Lemma 1. *If f satisfies assumption 3 then $f_1 \equiv f(x, y, a_1, b_1) = 0$ and $f_2 \equiv f(x, y, a_2, b_2) = 0$ cannot intersect more than once for $a_1 < a_2$ and $b_1 > b_2$.*

Proof. If f_1 and f_2 do not intersect at all then the lemma trivially holds. So we assume that f_1 and f_2 intersect at least twice. Let any two consecutive intersections be $P_1(\equiv (u, v))$ and $P_2(\equiv (u', v'))$. Clearly the curve intersecting the other from above at P_1 will be below it at P_2. Hence, the more oblique curve at P_1 becomes steeper at P_2. This means that if $g(u, a_1, b_1) < g(u, a_2, b_2)$ then $g(u', a_1, b_1) > g(u', a_2, b_2)$ or vice versa. Both contradict assumption 3. Hence f_1 and f_2 will meet at the most once.

Q.E.D.

In the sequel we shall use the first part of the last assumption only. Symmetric results may be obtained using the other part.

Lemma 2. *If p and q are defined such that* $a_l = f_a(x_p, p, b_u)$ *and* $b_u = f_b(q, y_q + 1, a_l)$ *then* $x_p > q$ *and* $p < y_q + 1$.

Proof. Using lemma 1, we say that the curves $f_o \equiv f(x, y, a_o, b_o) = 0$ and f_{lu} intersect at $P = (u, v)$. (Note that if f_o and f_{lu} do not intersect, then either (x_p, p) is outside f_o or $(q, y_q + 1)$ is inside f_o – both leading to contradictions). As $a_l \leqslant a_o$ and $b_u > b_o$, $g(x, a_l, b_u) < g(x, a_o, b_o)$. So, to the right of P, f_o is above f_{lu} and to the left of P, f_o is below f_{lu}. Since f_{lu} passes through (x_p, p) and (x_p, p) is inside $f_o, x_p \geqslant u$ and $p \leqslant v$. Similarly, $q < u$ and $y_q + 1 > v$. Combining, we get, $x_p > q$ and $p < y_q + 1$. Q.E.D.

Now, we are ready to claim that the domain of D_o is properly contained in the rectangle R_{ul} defined by the diagonally opposite points (a_l, b_u) and (a_u, b_l) in the parametric space. This is formally proved in the next theorem.

Theorem 6. $D(f(x, y, a, b)) = D_o$ *implies that* $a_u > a \geqslant a_l$ *and* $b_u > b \geqslant b_l$.

Proof. We first show that $D = D(f(x, y, a, b)) = D_o$ implies $a \geqslant a_l$ and $b < b_u$. We proceed by contradiction in three cases. Note that the first two cases may be treated without the last assumption.

Case 1. $a < a_l$ and $b \leqslant b_u$.

As $x_p = f_x(p, a_l, b_u)$ is an integer and $M(f_x, a) = M(f_x, b) =$ 'I', we get $[f_x(p, a, b)] < [f_x(p, a_l, b_u)] = x_p$. Thus $(x_p, p) \notin D$.

Case 2. $a \geqslant a_l$ and $b \geqslant b_u$.

Again $y_q + 1 = f_y(q, a_l, b_u)$ being an integer, $[f_y(q, a, b)] \geqslant [f_y(q, a_l, b_u)] = y_q + 1$. Hence $(q, y_q + 1) \in D$ but $\notin D_o$.

Case 3. $a < a_l$ and $b > b_u$.

Consider the curves $f \equiv f(x, y, a, b) = 0$ and f_{lu}. If f and f_{lu} do not intersect then either f is always below f_{lu} or always above f_{lu}. In the first case, $(x_p, p) \notin D$ and in the second $(q, y_q + 1) \in D$. This is a contradiction. Hence, f and f_{lu} must meet exactly once (lemma 1). Let us call this point of intersection $P \equiv (u, v)$. Now, since $a < a_l$ and $b > b_u, g(u, a, b) < g(u, a_l, b_u)$. Therefore, f_{lu} intersects f from below and thus, to include the point (x_p, p), $u \geqslant x_p$ and $v \leqslant p$. But then f also includes $(q, y_q + 1)$ as $u \geqslant x_p > q$ and $v \leqslant p < y_q + 1$ (from lemma 2). Hence $D \neq D_o$. Q.E.D.

In the next theorem, we proceed one step further to show that R_{ul} not only contains the domain of D_o but R_{ul} is the smallest such rectangle enclosing the domain of D_o.

Theorem 7. *If we select some 'a' so that* $a_l \leqslant a < a_u$ *then there exists some b for which* $D(f(x, y, a, b)) = D_o$.

Proof. Let us say that f_{ul} and f_{lu} intersect (theorem 5) at $P \equiv (u, v)$. Consider the curve $f \equiv f(x, y, a, b) = 0$ which passes through P. As $f(u, v, a, b) = 0$, we can say $b = f_b(u, v, a)$. Since f_b is decreasing in a and $a_l \leqslant a < a_u$ we have,

$$b_u = f_b(u, v, a_l) \geqslant b = f_b(u, v, a) > f_b(u, v, a_u) = b_l.$$

Thus the curve f touches P and passes through the region between f_{lu} and f_{ul} in accordance with our last assumption. Using corollary 1 of theorem 5, $D(f)$ may not match with D_o only at P if P is either $(x_{q'} + 1, q')$ or $(q, y_q + 1)$ for some q. In that case, we can shift the curve f vertically downwards by an infinitesimally small amount so that f goes out of P without disturbing other grid points of D_o. Q.E.D.

We complete this section by considering the following two examples.

Example 2. Let f be an ellipse with its centre at the origin and its axes parallel to the coordinate axes. So $f \equiv x^2/a^2 + y^2/b^2 - 1 = 0$ and the parameters to estimate are 'a' and 'b'. Clearly f is continuous and its MM is the same as given in figure 4. f is also separable and the separated functions are:

$$x = f_x = a(1 - y^2/b^2)^{1/2}, \; y = f_y = b(1 - x^2/a^2)^{1/2},$$
$$a = f_a = x/(1 - y^2/b^2)^{1/2} \text{ and } b = f_b = y/(1 - x^2/a^2)^{1/2}.$$

Further $g(x, a, b) = \partial f_y/\partial x = -bx/(a^2(1 - x^2/a^2)^{1/2})$. Therefore $g(x, a_1, b_1) < g(x, a_2, b_2)$ for $a_1 < a_2$ and $b_1 > b_2$ and f satisfies assumption 3 as well.

We have dealt thoroughly with this curve in Chattopadhyay & Das (1992). It is

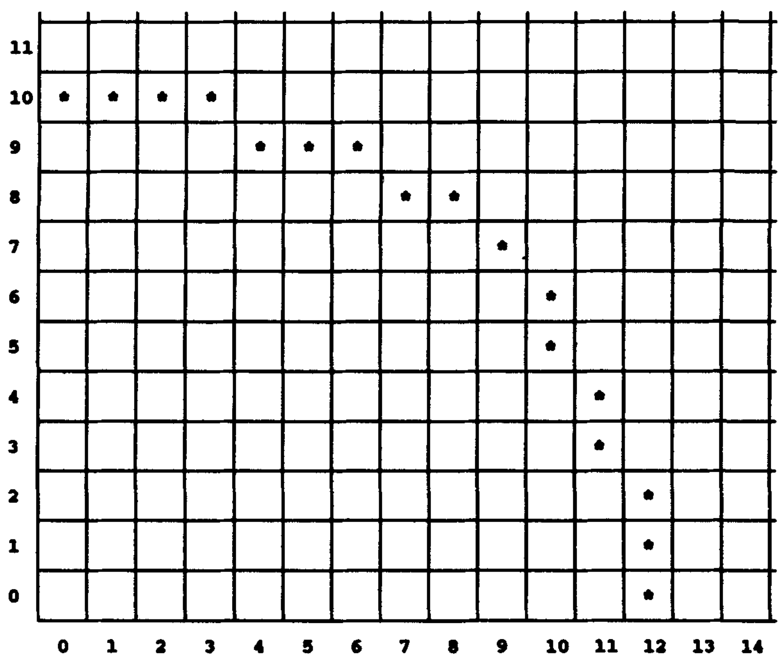

Figure 7. OBQ image of an ellipse with $a_o = 12{\cdot}5$ and $b_o = 10{\cdot}5$. Grid points belonging to D_o have been marked with '$*$'.

k	a_l^k	a_u^k	b_l^k	b_u^k
0	12.000000	13.000000	10.000000	11.000000
1	12.203404	12.579418	10.277402	10.606602
2	12.219196	12.546421	10.366421	10.584755
3	12.220127	12.536441	10.385010	10.583110
4	12.220197	12.534392	10.390681	10.583013
5	12.220202	12.533769	10.391848	10.583006
6	12.220202 *	12.533641	10.392203	10.583005
7	12.220202	12.533602	10.392276	10.583005 *
8	12.220202	12.533594	10.392299	10.583005
9	12.220202	12.533591	10.392303	10.583005
10	12.220202	12.533591 *	10.392305	10.583005
11	12.220202	12.533591	10.392305 *	10.583005

Figure 8. Iterative refinement of a_l, a_u, b_l and b_u for $a_o = 12{\cdot}5$ and $b_o = 10{\cdot}5$.

seen that the initial choice of the lower and upper bounds of a and b for theorem 4 can be given by the following set of equations:

$$a_l^0 = x_0, a_u^0 = x_0 + 1, b_l^0 = y_0, b_u^0 = y_0 + 1.$$

Also, theorems similar to 5, 6 and 7 have been proved in Chattopadhyay & Das (1992) from a different perspective using the specific properties of an ellipse. Let D_o be the set of points, which is obtained by digitizing an ellipse with $a_o = 12{\cdot}5$ and $b_o = 10{\cdot}5$, shown in figure 7. The convergence of the algorithm (theorem 4) producing lower and upper bounds of a_o and b_o, that is, R_{ul} is shown in figure 8.

We have presented an analysis of the class of curves whose nature of monotonicity is depicted by the particular MM as given in figure 4. Further the other four monotonicity matrices may be treated along similar lines to achieve similar results. In the next example we consider a straight line because it has a different MM from the earlier one.

Example 3. Let $f \equiv y - mx - c$. The MM for this f is different from that in figure 4. Its PDSV $= (-, +, -, -)$, i.e. a variant of PDSV_1 in figure 2b. Also we consider the segment from $x = 0$ to $x = n$. We can develop similar iterative refinement schemes using $m_l^0 = -1/n$, $m_u^0 = (n+1)/n, c_l^0 = y_0$, $c_u^0 = y_0 + 1$ as initial choices of the bounds of the parameters. It is easy to verify that f satisfies all assumptions. Hence the results of this paper hold well for it. They, however, have been proved independently in Chattopadhyay & Das (1991).

5. Conclusion

In this paper, we have addressed the reconstruction problem of a restricted class of one or two parameter curves. It is shown that the domain of the given digitization can be exactly formulated for *separable, continuous* and *monotone* one-parameter curves. In the case of two-parameter curves we have proved that the smallest rectangle containing the domain may be derived if the curve possesses another additional property in the form of assumption 3. If the curve under consideration satisfies all the properties as listed in the assumptions only, can our method be applied for reconstruction problems involving such curves.

The contribution of this paper is in the development of a unified methodology to solve the reconstruction problem for a class of curves. In separate papers (Chattopadhyay & Das 1991, 1992) we have carried out a detailed analysis of straight lines and a special class of conics using the same approach. An advantage of assuming particular functional forms is that we can algebraically characterize the domain. In fact, for straight lines the domain is completely formulated. In Chattopadhyay & Das (1992) the domain of digital conics in normal positions has been numerically computed. We hope that the technique of numerically computing the domain can be extended to curves satisfying the set of assumptions already mentioned. Hence, the strength of the scheme developed in this paper lies in its wider applicability to solve the reconstruction and domain construction problems.

The iterative refinement technique also has a drawback. It is not easy to extend this technique either for more general curves or for curves with more parameters. Whether this method may be adapted to three-dimensional surfaces is an unsolved problem. For 3-D planes which involve three unknown parameters, an iterative refinement algorithm can be devised to obtain the upper and lower bounds of the parameters.

The authors would like to thank the anonymous reviewer of our earlier paper (Chattopadhyay & Das 1992) for suggesting the problem solved in this paper. PPD also gratefully acknowledges the financial support of the Indian National Science Academy under a Young Scientist Award.

References

Anderson T A, Kim C E 1983 Representation of digital line segments and their preimages, Technical Report, CS-83-120, Washington State University

Chattopadhyay S, Das P P 1991 A new method of analysis for discrete straight lines. *Pattern Recogn. Lett.* 12: 747–755

Chattopadhyay S, Das P P 1992 Parameter estimation and reconstruction of digital conics in normal positions. *Comput. Vision Graph. Image Process.: Graph. Models Image Process.* 54: 385–395

Dorst L 1986 *Discrete straight line segments: Parameters, primitives and properties*, Ph D thesis, Delft University of Technology, The Netherlands

Dorst L, Smeulders A W M 1984 Discrete representation of straight lines. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6: 450–463

Freeman H 1970 Boundary encoding and processing. In *Picture processing and psychopictorics* (eds) B S Lipkin, A Rosenfeld (New York: Academic Press) pp. 241–266

Kim C E 1982 Digital convexity, straightness and convex polygons. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-4: 618–626

Kim C E 1984 Digital disks. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-6: 372–374

Kim C E, Rosenfeld A 1982 Digital straight lines and convexity of digital regions. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-4: 149–153

Nakamura A, Aizawa K 1984 Digital circles. *Comput. Vision Graph. Image Process.: Graph. Models Image Process.* 26: 246–255

Nakamura A, Aizawa K 1985 Digital images for geometric pictures. *Comput. Vision Graph. Image Process.: Graph. Models Image Process.* 30: 107–120