PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 59 (73), 1996, 169-176

A CONSTANT SPACE REPRESENTATION OF DIGITAL CUBIC PARABOLAS

Dragan M. Acketa and **Joviša Žunić**

Communicated by Gradimir Milovanović

Abstract. The concept of "noisy" straight line introduced by Melter and Rosenfeld is generalized and applied to digital cubic parabolas. It is proved that digital cubic parabola segments and their least square cubic parabola fits are in one-to-one correspondence. This leads to a constant space representation of a digital cubic parabola segment. One such representation is (x_1, n, a, b, c, d), where x_1 and n are the left endpoint and the number of digital points, respectively, while a, b, c and d are the coefficients of the least square cubic parabola fit $Y = aX^3 + bX^2 + cX + d$ for the given cubic parabola segment.

1. Introduction. Let p be a cubic parabola $y = \alpha x^3 + \beta x^2 + \gamma x + \delta$ in the Euclidean plane. A digital curve is defined to be the result of subjecting a continuous curve to a particular digitalization process. The cubic parabola p will be digitalized by a digitalizing method in which the first digital points (points with integer coordinates, often called "pixels") below a given cubic parabola are taken. This is obviously equivalent to translating the cubic parabola by -0.5 in the vertical direction and rounding.

So, the associated set of digital points for the cubic parabola p, called a *digital cubic parabola*, is defined by

$$C(p) = \{(j, \lfloor \alpha j^3 + \beta j^2 + \gamma j + \delta \rfloor),\ j \text{ is an integer}\}$$

where $\lfloor k \rfloor$ denotes the largest integer not larger than k.

In general, we will operate with finite subsets of $C(p)$, or, more precisely, with digital cubic parabola segments obtained by digitalizing the parts of cubic parabolas lying between the lines $x = x_1$ and $x = x_{\text{end}}$ for some numbers x_1 and

---

AMS Subject Classification (1991): Primary 65D10; Secondary 68G10

Key Words and Phrases: coding scheme, shape representation, image vision, digital geometry.

x_{end} (where $x_1 < x_{\mathrm{end}}$). Without loss of generality, one can assume that x_1 and x_{end} are integers. If $n = x_{\mathrm{end}} - x_1 + 1$, then the digital cubic parabola segment $C_n(p)$ for the considered cubic parabola p is defined as:

$$C_n(p) = \{(j, \lfloor \alpha j^3 + \beta j^2 + \gamma j + \delta \rfloor),\ j = x_1, x_1+1, \ldots, x_1 + (n-1) = x_{\mathrm{end}}\}.$$

Obviously, n is the number of digital points in the digital cubic parabola segment $C_n(p)$. This number is called the *length* of the segment.

Our main goal is to give a constant space representation for a digital cubic parabola segment of arbitrary length. Six parameters (x_1, n, a, b, c, d) are used for such a representation, where x_1 and n denote the left endpoint and the length of the digital parabola segment, while a, b, c and d are the coefficients of its least square cubic parabola fit $Y = aX^3 + bX^2 + cX + d$.

The idea for using the least squares fitting technique for representation of digital curves is not new. Namely, the concept of a "noisy" straight line segment, based on least squares line fitting and defined in terms of bounds on correlation coefficients, was introduced in [**5**]; it was also shown that this is a generalization of a digital straight line. The following question was posed in the same paper: if a continuous line is digitalized and the least squares fitting is applied to the points of the digital line segment, can the original line be recovered? The positive answer was given for line segments in [**6**]. More precisely, it was proved that the least square line fit uniquely determines the digital line on a segment. Thus any digital line segment can be uniquely coded by four numbers (x_1, n, b_0, b_1), where x_1 is x-coordinate of the left endpoint, the integer n is the number of digital points, and b_0 and b_1 are the coefficients of the least square line fit $Y = b_0 + b_1 X$ for the given digital line segment. This representation is an alternative for the well-known representation of digital lines by adjacent pairs given in [**3**], as well as for the one suggested in [**2**].

So, while there exist several constant space representations for digital lines, currently no such representation for other kinds of digital curves is known. This paper solves the problem for digital cubic parabolas.

2. Preliminaries. A finite set of points in the plane is sometimes called a scatter diagram. The least square curve for a scatter diagram is a curve which minimizes the sum of the squares of the vertical distances to the data points. The method for determining such curves is well known from statistics [**1**].

If a scatter diagram is given by $\{(x_i, y_i), i = 1, 2, \ldots, n\}$, and the equation of its least squares cubic parabola is $y = ax^3 + bx^2 + cx + d$, then the function $F(a,b,c,d) = \sum_{i=1}^{n} (ax_i^3 + bx_i^2 + cx_i + d - y_i)^2$ should be minimized. So, the equations $\frac{\partial F}{\partial a} = 0$, $\frac{\partial F}{\partial b} = 0$, $\frac{\partial F}{\partial c} = 0$, $\frac{\partial F}{\partial d} = 0$ must be satisfied. This gives:

$$(\#) \qquad S_{k+3}a + S_{k+2}b + S_{k+1}c + S_k d = M_k \quad (k = 0, 1, 2, 3),$$

where $S_k = \sum_{i=1}^{n} x_i^k$ and $M_k = \sum_{i=1}^{n} y_i x_i^k$.

If the previous scatter diagram is a digital cubic parabola segment $C_n(p)$, then $x_i = x_{i-1} + 1$ for $i = 2, 3, \dots, n$. For the sake of convenience and without loss of generality, it will be assumed that $x_1 = 0$ in the rest of the paper. Then (see e.g. [**4**, p. 87, Problem 29]:

$$S_0 = n, \quad S_1 = \frac{1}{2}n(n-1), \quad S_2 = \frac{1}{6}n(n-1)(2n-1), \quad S_3 = \frac{1}{4}n^2(n-1)^2,$$
$$S_4 = \frac{1}{30}n(n-1)(2n-1)(3n^2-3n-1), \quad S_5 = \frac{1}{12}n^2(n-1)^2(2n^2-2n-1),$$
$$S_6 = \frac{1}{42}n(n-1)(2n-1)(3n^4-6n^3+3n+1).$$

The coefficients a, b, c, d of the least square parabola fits can be determined by solving the above system. The determinant of the system is

$$\frac{1}{6048000}n^4(n+1)^3(n-1)^3(n+2)^2(n-2)^2(n+3)(n-3),$$

so for $n > 3$ the system has a unique solution $a(p)$, $b(p)$, $c(p)$, $d(p)$.

3. Constant space representation. A key question is whether there exist two different digital cubic parabola segments $C_n(p)$, $C_n(q)$ with the same least square cubic parabola fit, i.e. $a(p) = a(q)$, $b(p) = b(q)$, $c(p) = c(q)$, $d(p) = d(q)$. The answer is negative. This means that the digital parabola segments and their least square parabola fits are in one-to-one correspondence. Two auxiliary lemmas will be proved first.

LEMMA 1. *If the sequences $\{a_n\}$ and $\{b_n\}$ of non-negative integers satisfy*

$$0 \le a_1 \le \cdots \le a_k < b_1 \le \cdots \le b_n < a_{k+1} \le \cdots \le a_n,$$
$$a_1 + \cdots + a_n = b_1 + \cdots + b_n,$$

then $a_1^2 + \cdots + a_n^2 > b_1^2 + \cdots + b_n^2$.

Proof. Elementary transformations give that

$$\sum_{i=1}^{k}(b_i^2 - a_i^2) = \sum_{i=1}^{k}(b_i - a_i)(b_i + a_i) \le (b_k + a_k)\sum_{i=1}^{k}(b_i - a_i)$$
$$= (b_k + a_k)\sum_{i=k+1}^{n}(a_i - b_i) < \sum_{i=k+1}^{n}(a_i - b_i)(a_i + b_i) = \sum_{i=k+1}^{n}(a_i^2 - b_i^2).$$

The strict inequality follows from $a_k < a_{k+i}, i = 1, 2, \dots, n-k$. $\square$

Lemma 2. *If the sequences* $\{a_n\}$ *and* $\{b_n\}$ *of non-negative integers satisfy:*

$$(1) \qquad a_1 \le \cdots \le a_k < b_1 \le \cdots \le b_l < a_{k+1} \le \cdots \le a_n < b_{l+1} \le \cdots \le b_n,$$

$$(2) \qquad a_1 + \cdots + a_n = b_1 + \cdots + b_n, \qquad (3) \quad a_1^2 + \cdots + a_n^2 = b_1^2 + \cdots + b_n^2,$$

then: (*a*) $k < l$, (*b*) $b_1^3 + \cdots + b_n^3 > a_1^3 + \cdots + a_n^3$.

Proof. (a) Suppose that $k \ge l$. Then

$$\begin{aligned} (b_i - a_{k-l+i}) &> 0 \quad \text{for} \quad 1 \le i \le l, \\ (b_{l+i} - a_{k+i}) &> 0 \quad \text{for} \quad 1 \le i \le n-k, \\ (b_{l+n-k+i} - a_i) &> 0 \quad \text{for} \quad 1 \le i \le k-l. \end{aligned}$$

Summing up these inequalities, one obtains $\sum\limits_{i=1}^{n}(b_i - a_i) > 0$, a contradiction with (2).

(b) Using (a), the conditions (2) and (3) can be written in the form:

$$(2') \qquad \sum_{i=1}^{k}(b_i - a_i) + \sum_{i=l+1}^{n}(b_i - a_i) = \sum_{i=k+1}^{l}(a_i - b_i)$$

$$(3') \qquad \sum_{i=1}^{k}(b_i^2 - a_i^2) + \sum_{i=l+1}^{n}(b_i^2 - a_i^2) = \sum_{i=k+1}^{l}(a_i^2 - b_i^2),$$

(the common value of both sides of $(2')$ will be denoted by s) while the requirement of the theorem can be formulated as

$$\sum_{i=1}^{k}(b_i^3 - a_i^3) + \sum_{i=l+1}^{n}(b_i^3 - a_i^3) > \sum_{i=k+1}^{l}(a_i^3 - b_i^3).$$

All the expressions in brackets are natural numbers. Let *BA-difference* and *AB-difference* denote a difference of the form $(b_i^t - a_i^t)$ and $(a_i^t - b_i^t)$, respectively. The equalities $(2')$ and $(3')$ imply the balance of BA-differences (on the left-hand side) and AB-differences (on the right-hand side), for degrees 1 and 2. An *elementary difference of degree* t is a difference of the form $(x+1)^t - x^t$, for some natural numbers x and t. Thus, elementary differences of degrees 1, 2 and 3 are equal to 1,$2x+1$ and $3x^2+3x+1$ respectively.

Replace now each (BA- or AB-) difference $v^t - u^t$ of degree t (where $v > u$, $t = 1, 2, 3$) by a sum

$$(v^t - (v-1)^t) + ((v-1)^t - (v-2)^t) + \cdots + ((u+2)^t - (u+1)^t) + ((u+1)^t - u^t)$$

of elementary differences of degree t.

The condition (2′) implies that the numbers of elementary AB-differences and elementary BA-differences, (of degree 1 and consequently of any degree) are the same and equal to the common value s of the sums on the left-hand side and right-hand side of the equality (2′).

Let $x_1 \leq \cdots \leq x_s$ denote the *sorted* values of u, which are used in the elementary BA-differences. Similarly, let $y_1 \leq \cdots \leq y_s$ denote the *sorted* values of u, which are used in the elementary AB-differences.

The condition (3′), i.e., the equality of the sums of elementary AB- and BA-differences of degree 2, has the following consequence of degree 1:

$$(2x_1 + 1) + \cdots + (2x_s + 1) = (2y_1 + 1) + \cdots + (2y_s + 1),$$

which implies

$$(*) \qquad x_1 + \cdots + x_s = y_1 + \cdots + y_s$$

The value of u in an elementary difference $D = (u+1)^2 - u^2$ is chosen from the following intervals:

$a_1 \leq b_k - 1$, if D is a BA-difference $b_i^2 - a_i^2$ for some i satisfying $1 \leq i \leq k$,
$b_{k+1} \leq a_l - 1$, if D is an AB-difference $a_i^2 - b_i^2$ for some i satisfying $k+1 \leq i \leq l$,
$a_{l+1} \leq b_n - 1$, if D is a BA-difference $b_i^2 - a_i^2$ for some i satisfying $l+1 \leq i \leq n$.
Since it is necessarily true that $b_k - 1 < b_{k+1}$ and $a_l - 1 < a_{l+1}$, one can derive that the following inequality holds:

$$x_1 \leq \cdots \leq x_p < y_1 \leq \cdots \leq y_s < x_{p+1} \leq \cdots \leq x_s, \text{ where } p = \sum_{i=1}^{k} (b_i - a_i).$$

In other words, the x-sequence and the y-sequence satisfy the order condition of Lemma 1. When joined with the equality condition (*), Lemma 1 (for degree 2) can be applied to derive the conclusion

$$(**) \qquad x_1^2 + \cdots + x_s^2 > y_1^2 + \cdots + y_s^2.$$

The condition (2) of Lemma 2 is equivalent to the following inequality for elementary differences of degree 3:

$$(3x_1^2 + 3x_1 + 1) + \cdots + (3x_s^2 + 3x_s + 1) > (3y_1^2 + 3y_1 + 1) + \cdots + (3y_s^2 + 3y_s + 1).$$

Taking into account (*), this inequality is equivalent to the already proved inequality (**). □

THEOREM 1. *Let $C_n(p)$ and $C_n(q)$ be two digital cubic parabola segments, the left endpoints of which have abscissas equal to 0. If $a(p)$, $b(p)$, $c(p)$, $d(p)$ and $a(q)$, $b(q)$, $c(q)$, $d(q)$ are coefficients of the least square parabola fit associated to $C_n(p)$ and $C_n(q)$ respectively, then ($a(p) = a(q)$ and $b(p) = b(q)$ and $c(p) = c(q)$ and $d(p) = d(q)$) is equivalent to $C_n(p) = C_n(q)$.*

Proof. The direction $\Leftarrow$ follows from the definitions. The opposite direction will be proved by a contradiction.

An interpretation of the sums appearing on the right hand sides of equalities of the system (#), will be primarily given, under the assumption that all the values y_i and y'_i are greater than or equal to 0 (otherwise one can apply the translation in the vertical direction by $-\min\{y_i, y'_i\}$, for $i = 1, 2, \ldots, n$).

(1) $\sum_{i=1}^{n} y_i$ denotes the number of digital points below the digital cubic parabola and above the x-axis.
(2) $\sum_{i=1}^{n} x_i y_i$ can be understood as the sum of the abscissa values of all digital points lying between the digital cubic parabola and the x-axis.
(3) $\sum x_i^2 y_i$ is the sum of squares of the abscissa-values of all digital points lying between the digital cubic parabola and the x-axis.
(4) $\sum x_i^3 y_i$ is the sum of cubes of the abscissa-values of all digital points lying between the digital cubic parabola and the x-axis.

Assume, on the contrary, that there exist two digital cubic parabola segments $C_n(p)$ and $C_n(q)$ with the same associated coefficients of least square cubic parabola fits, i.e. $a(p) = a(q)$, $b(p) = b(q)$, $c(p) = c(q)$ and $d(p) = d(q)$. Also, let the digital cubic parabola segments $C_n(p)$ and $C_n(q)$ be obtained by digitalization of the cubic parabola segments p and q respectively.

Let the scatter diagrams of digital cubic parabola segments $C_n(p)$ and $C_n(q)$ be $C_n(p) = \{(x_i, y_i), i = 1, \ldots, n\}$ and $C_n(q) = \{(x_i, y_i^{'}), i = 1, \ldots, n\}$, respectively.

It follows from the uniqueness of the solution of the system (#) that the following equalities must be satisfied:

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} y_i^{'}, \tag{1}$$

$$\sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} x_i y_i^{'}, \tag{2}$$

$$\sum_{i=1}^{n} x_i^2 y_i = \sum_{i=1}^{n} x_i^2 y_i^{'}, \tag{3}$$

$$\sum_{i=1}^{n} x_i^3 y_i = \sum_{i=1}^{n} x_i^3 y_i^{'}. \tag{4}$$

Suppose that the cubic parabolas p and q determine different digital cubic parabola segments for $x \in [0, n-1]$. Depending on the number of their intersection points in the same interval, four possible cases should be considered:

Case 1. There are no intersection points of p and q. It follows that there is a digital point between them with the abscissa-value belonging to the interval $[0, n-1]$, but this is a contradiction with (1).

In the remaining cases, it will be assumed that $y_1 \geq y'_1$. The assumption that the two digital parabola segments are different implies that in the area A_1 below the parabola p and above the parabola q there is at least one digital point (with abscissa-value belonging to $[0, n-1]$), as well that there is at least one digital point belonging to the area A_2 above the parabola p and below the parabola q. From (1), it follows that the number of digital points inside the areas A_1 and A_2 are the same. Denote this common number by w, let the abscissa-values of the points from A_1 be denoted by $a_1, a_2, \ldots, a_w$ and let the abscissa values of the points from A_2 be denoted by $b_1, b_2, \ldots, b_w$.

Case 2. There is one intersection point of p and q. The following inequality is a consequence of this case:

$$0 \leq a_1 \leq \ldots \leq a_w < b_1 \leq \ldots \leq b_w,$$

which implies the strict inequality

$$a_1 + \ldots + a_w < b_1 + \ldots + b_w,$$

a contradiction with (2).

Case 3. There are two intersection points of p and q. It can be concluded from (2), (3) and the proposed case that the following relations are satisfied:

$$\begin{gathered} 0 \leq a_1 \leq \ldots \leq a_l < b_1 \leq \ldots \leq b_w < a_{l+1} \leq \ldots \leq a_w, \\ a_1 + \ldots + a_w = b_1 + \ldots + b_w, \quad a_1^2 + \ldots + a_w^2 = b_1^2 + \ldots + b_w^2, \end{gathered}$$

a contradiction with Lemma 1.

Case 4. There are three intersection points of p and q. It can be concluded from the conditions (2), (3), (4) and the proposed case that the folllowing relations are satisfied:

$$\begin{gathered} 0 \leq a_1 \leq \ldots \leq a_l < b_1 \leq \ldots \leq b_k < a_{l+1} \leq \ldots \leq a_w < b_{k+1} \leq \ldots b_w, \\ a_1 + \ldots + a_w = b_1 + \ldots + b_w, \\ a_1^2 + \ldots + a_w^2 = b_1^2 + \ldots + b_w^2, \\ a_1^3 + \ldots + a_w^3 = b_1^3 + \ldots + b_w^3, \end{gathered}$$

a contradiction with Lemma 2b. $\square$

References

1. I.W. Burr, *Applied Statistical Methods*, Academic Press, New York, 1974.
2. L. Dorst, A.W.M. Smeulders, *Discrete representation of straight lines*, IEEE Trans. PAMI **6** (1984), 450–463.

3. M. Lindenbaum, J. Koplowitz, *A new parametrization of digital straight lines*, IEEE Trans. PAMI **13** (1991), 847–852.
4. V.A. Krechmar, *A Problem Book in Algebra*, Mir Publishers, Moscow, 1974.
5. R.A. Melter, A. Rosenfeld, *New views of linearity and connectedness in digital geometry*, Pattern Recognition Letters **10** (1989), 9–16.
6. R.A. Melter, I. Stojmenović J. Žunić, *A new characterization of digital lines by least square fits*, Pattern Recognition Letters **14** (1993), 83–88.

Institut za matematiku
Prirodno-matematički fakultet
Trg D. Obradovića 4
21000 Novi Sad

(Received 06 10 1994)
(Revised 11 09 1995)

Institut za primenjene osnovne discipline
Fakultet tehničkih nauka,
Trg D. Obradovića 6
21000 Novi Sad