(15)

"A Method for Compressing Character or Pictorial Image Data".

This invention relates to the compression of character or pictorial image data, and particularly, though not exclusively to a data compression methods which establish sampling points on a character outline, and which establish blocks and store information for specifying outlines of a character, pictorial image or the like (hereinafter referred to simply as "character"), and wherein the outline of each character is approximated by a set of functional curves or straight lines.

First of all, the establishment of sampling points will be described. In a method for compressing character data wherein an outline of a character is split up between sampling points, and each of the outline segments thus split off is approximated by means of functional curves or straight lines, it is important to establish as small a number of sampling points as possible, at a high rate and in such a way that the characteristics of the shape of the character outline do not suffer.

Heretofore, as one of the simplest methods for establishing such sampling points, the sampling points have been established by incrementing a variable with respect to the character outline which has been developed on X- and Y-coordinates. However, the establishment of sampling points in such a method depends only on the variable increment so that a prescribed precision is not guaranteed for a relationship between a curve (or a straight line) to be approximated and an actual character outline. As a result, it has been difficult to establish the most suitable sampling points.

Furthermore, there has been proposed such a method wherein the fit between a curve (or a straight line) to be approximated and an actual character outline is determined in respect of candidate sampling points, then, the candidate sampling points are moved by trial and error in response to the precision

determined, and the candidate sampling points become established sampling points when a prescribed precision is obtained. However, since this method requires a trial and error process for establishing one sampling point, the processing therefor is slow.

According to the present invention there is provided, a method of compressing data representing a character outline wherein the outline is initially divided into a set of blocks by split points and each block is sampled at a plurality of sample points, the method comprising: establishing a new split point subdividing the initial block into two new blocks, by computing for each sample point the density of sample points in its vicinity and establishing one such sample point as said new split point on the outline when the density of sample points about the said such sample point exceeds a prescribed threshold.

A preferred embodiment of the invention will now be described with reference to the accompanying drawings, wherein:

Figure 1 is an diagram showing how the outline of a character is divided into a set of blocks;

Figures 2A - 2C are diagrams illustrating processes for establishing sampling points on a block B of a character outline in accordance with the present invention;

Figure 3 is a block diagram illustrating an embodiment of the method for establishing sampling points;

Figure 4 is a flow chart illustrating the operational process of Figure 3;

Figure 5 is an diagram illustrating an example wherein the method for establishing sampling points in accordance with the construction of Figure 3 or the process of Figure 4 is effected;

Figure 6 is an diagram illustrating an embodiment of a method for establishing blocks;

Figure 7 is a diagram illustrating an example wherein outlines of a character are actually processed by combining a method for establishing sampling points in accordance to the present invention with a method forestablishing blocks in accordance with the invention.

First of all, a pretreatment for character outlines which is a premise of establishing sampling points will be described hereinbelow.

Figure 1 is a typical explanatory view illustrating a situation in which a character outline is split into plural blocks (B_1, B_2, B_3, ...), and the character outline is grapsed as a set of the respective blocks B wherein mark "o" and mark "•" indicate each terminal outline point in these blocks, respectively.

Namely, as a pretreatment for establishing sampling points, split points are suitably established on the character outline developed on X- and Y-coordinates, and the outline is split in an interval of one valued function involving x as its variable, whereby the blocks have previously been determined.

The splitting of these blocks may generally be processed by establishing an arbitrary interval in which a series of X-coordinate values of the outline increase or decrease monotonously as a block.

Furthermore, as a more suitable block establishing method, there is such a method that, as illustrated in Fig. 1, positions at which a series of X-coordinate values on the outline become extreme values are determined, thereafter, positions corresponding to the aforesaid extreme values are established as initial split points D (mark "o" in Fig. 1), and further, as fully described in conjunction with the undermentioned block establishing method, intervals split with each split point (D or D') are established as blocks B, respectively, by establishing second split points D' (mark "•" in Fig. 1) on the respective outline blocks split with the aforesaid split points D with a prescribed relationship.

Physical meaning of each block B established as described above is in that each block B corresponds to a set of points making bright points to be scanned in Y-direction turn-on or turn-off in case of reproducing character image on a raster scanning type character display unit such as cathode ray tube (CRT).

The sampling point establishing method relates to the one in which an outline corresponding to any one of the respective blocks B established by means of the above described pretreatment is split into a larger number of outline segments.

Figs. 2A - 2C are typical explanatory views showing processes for establishing sampling points on any one of the blocks B in accordance with the method of the present

invention. Fig. 3 is a block diagram illustrating an embodiment of the sampling point establishing method according to the present invention.

In Fig. 3, reference character 31 designates an outline information storing part for storing coordinate information of the outline corresponding to any one block B represented by a solid line curve in Fig. 2 and outputting such information at need, 32 a sampling point coordinate storing part for successively storing coordinate information of the sampling points established in accordance with the method of the present invention and outputting such information as occasion demands, 33 a distance computing part for computing a distance L between adjacent sampling points, respectively, in respect of the sampling points which have already been established on the outline, that is, which have been stored in the aforesaid sampling point coordinate storing part, 34 an approximate curve computing part which will be described hereunder, 35 an area computing part for computing an area S enclosed by outline segments cut out by adjacent sampling points as well as an approximate curve obtained by connecting the respective sampling points which have already been established on the outline with each other, 36 an appreciation quantity computing part for computing an appreciation quantity ξ being a ratio S/L, in respect of each segment on the basis of the distance L computed in the aforesaid distance computing part 33 and the area S determined in

the area computing part 35, 37 a maximum appreciation quantity detecting part, 38 a comparing part for comparing magnitude of the maximum value ξ_{max} of the appreciation quantity determined by the appreciation quantity computing part 36 with that of a predetermined allowable value ξ' supplied from an allowable value establishing part 39, and 40 a newly-established sampling point computing part for computing new sampling points on the basis of a prescribed relationship in the case when the result compared in the comparing part 38 is in $\xi_{max} > \xi'$ and the newly-established sampling point computing part 40 comprises, for example, a deviation computing part 41 and a maximum deviated position detecting part 42.

Operations of the construction illustrated in Fig. 3 will be described hereinbelow in connection with the illustrations of Fig. 2.

First, the coordinate information at both terminal outline points P_1 and P_2 (Each of these terminal outline points corresponds to the split point D or D' of Fig. 1 obtained by the aforesaid pretreatment.) of a block B in Fig. 2A is stored in the sampling point coordinate storing part 32 as the coordinate information of the sampling points which have been preset to establish an initial condition of the operations.

When establishment of the aforesaid initial condition is completed, a distance L (Fig. 2A) between the sampling points P_1 and P_2 which have already been established and

being read out from the sampling point coordinate storing part 32 is computed by the distance computing part 33 on the basis of the coordinate information of both the sampling points, and the result computed is outputted to the appreciation quantity computing part 36.

On the other hand, at the same time of the above described operation, the approximate curve computing part 34 determines an approximate curve connecting the sampling point P_1 with the point P_2 which have already been established and being similarly read out from the sampling point coordinate storing part 32 on the basis of the coordinate information of both the sampling points, and the approximate curve thus determined is outputted to the area computing part 35.

However, the present case wherein an operation is started from the initial condition involves two established sampling points so that the aforesaid approximate curve comes to be straight line C_1 as shown in Fig. 2A.

Then, an area $S = S' + S''$ illustrated in Fig. 2A is computed by means of the area computing part 35 on the basis of the straight line C_1 determined by the aforesaid approximate curve computing part 34 and the outline coordinate information stored in the outline information storing part 31, the result thus computed is outputted to the appreciation quantity computing part 36.

The appreciation quantity computing part 36 computes an appreciation quantity ξ being S/L on the basis of the

distance L determined by the aforesaid distance computing part 33 and the area S determined by the area computing part 35.

Since this case involves one appreciation quantity to be calculated, the above appreciation quantity is supplied to the comparing part 38 as the maximum value ξ_{max} of appreciation quantity through the maximum appreciation quantity detecting part 37.

To the comparing part 38, a predetermined allowable value ξ' has been given from the allowable value setting part 39 as other input so that the newly-established sampling point computing part 40 is actuated in case of $\xi_{max} > \xi'$.

The deviation computing part 41 being a component of the newly-established sampling point computing part 40 calculates deviation Δ illustrated in Fig. 2A with respect to whole the interval extending over the points $P_1 - P_2$ on the basis of the curve (in this case, the same is particularly the straight line C_1) obtained by means of the aforesaid approximate curve computing part 34 and the outline coordinate information stored in the outline information storing part 31, and a positional information at point P_3 being the maximum value on the block B is determined by the maximum deviated position detecting part 42. As a result, the aforesaid point P_3 is stored and established in the sampling point coordinate storing part 32 as a newly-established sampling point.

As described above, when the newly-established sampling point P_3 is established on the outline of the block B, the construction of Fig. 3 executes again the operations as described above. Such operation processes will be described hereinbelow in conjunction with Fig. 2B. In this case, the newly-established sampling point P_3 is renewed as an established sampling point.

Namely, in this case, the distance computing part 33 determines first distances L_1 and L_2 illustrated in Fig. 2B. At the same time, an approximate curve C_2 (represented by each dotted line in Fig. 2) passing through the established sampling points P_1, P_2 and P_3 on the block B is determined by the approximate curve computing part 34.

The area computing part 35 computes individually areas S_1 and S_2 (indicated by slant lines in Fig. 2) enclosed by the respective outline segments cut out with sampling points as well as the aforesaid approximate curve C_2.

Then, appreciation quantities ξ_1 and ξ_2 relating to which $S/L = \xi$ are determined in the appreciation quantity computing part 36 on the basis of the areas S_1 and S_2 as well as the distances L_1 and L_2 corresponding thereto, respectively.

Now, in case of Fig. 2B, since a relationship between both the appreciation quantities is such that $\xi_1 < \xi_2$, value of ξ_2 is supplied from the maximum appreciation quantity detecting part 37 to the comparing part 38 as the maximum value ξ_{max} of appreciation quantity.

In case of ξ_{max} ($\simeq \xi_2$) > ξ', the newly-established sampling point computing part 40 is again actuated by means of a command derived from the comparing part 38 so that the appreciation quantity at point P_4 at which the quantity becomes Δ_{max} similarly as described in the case of Fig. 2A is established in the sampling point coordinate storing part 32 as a newly-established sampling point.

The processing situation after the newly-established sampling point P_4 is determined as mentioned above is as illustrated in Fig. 2C, but the operations themselves of the construction in Fig. 3 at that time are the ones of repetition which have already been determined. Accordingly, the description therefor will be omitted hereinbelow.

And when the aforesaid comparing part 38 detects a relationship of $\xi_{max} \leq \xi'$ in such processes wherein newly-established sampling points P_i are successively determined, operations of the construction illustrated in Fig. 3 relating to any one block B are completed.

Of course, it means that the sampling points P_1 to P_i stored in the sampling point coordinate storing part 32 have been established at that time as desired sampling points concerning the outline of block B.

Furthermore, a value of the distance L in the aforesaid distance computing part 33 can very easily be determined through calculation of an equation $L = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. Moreover, a desired area S can be comparatively easily determined by integrating differences of two curves in

y-direction towards x-direction in the aforesaid area computing part 35. Therefore, computation upon a value of the appreciation quantity ξ being given as the ratio S/L as mentioned before may also be comparatively simply effected. And such appreciation quantity ξ is obtained by dividing the area S by the length L so that the appreciation quantity is the one involving dimension of length. As a result, such quantity comes to be an index for indicating a discrete relationship between two curves (outline and approximate curve) defining the area S so far as these both curves do not involve a peculiar irregularity.

Fig. 4 is a flow chart illustrating operational processes of the construction of Fig. 3.

Next, a method for establishing blocks by which the method for establishing sampling points in accordance with the construction of Fig. 3 or the processes of Fig. 4 is made more effective will be described hereinbelow in connection with a case wherein a character outline involves the peculiar irregularity as described above.

Namely, when the method for establishing sampling points in accordance with the construction of Fig. 3 or the processes of Fig. 4 is practiced, there is such a tendency that sampling points P concentrate in a portion where the character outline has a small curvature radius as typically illustrated in Fig. 5. As a result, there causes such a case wherein improvement of data compressibility being dependent on the number of sampling points is prevented.

Accordingly, an improved method for establishing blocks as described hereinbelow notices a property in concentration of the aforesaid sampling points to utilize such property so that a second split point (corresponding to the split point D' in Fig. 1) is newly-organized at a portion where a preset density of the sampling points exceeds a prescribed threshold to divide a block into two sections. The establishment of sampling points in accordance with the construction of Fig. 3 is again practiced with respect to the respective blocks thus divided, whereby the improved method contemplates to contribute to decrease in the number of sampling points established in the whole character outline.

An embodiment of the improved method for establishing blocks will be described hereinbelow on the basis of the illustration in Fig. 6.

In Fig. 6, P_1 and P_n designate terminal outline points of any one of blocks as well as P_{i-1} to P_{i+3} designate sampling points determined by applying the aforesaid method for establishing sampling points to block B.

In case of practicing the improved method for establishing blocks, for example, such a value $\overline{P_i, P_{i+1}} + \overline{P_{i+1}, P_{i+2}} = \rho_i$ is first computed in respect of all the cases of $i = 2$ to $(n - 3)$ as densities of the sampling points, and the values thus computed are compared with threshold ρ', respectively.

And such a case wherein $\rho_i < \rho'$ is extracted as a concentrated interval of sampling points.

Then, the farthest point Q with respect to a straight line $\overline{P_{i-1}, P_{i+3}}$ in this concentrated interval of sampling points, that is, the interval $\left[P_{i-1}; P_{i+3}\right]$ is determined.

In addition, an angle $(P_{i-1}, Q, P_{i+3}) = \theta_q$ defined by straight lines $\overline{P_{i-1}, Q}$ and $\overline{Q, P_{i+3}}$ is determined, and then, the angle θ_q is compared with threshold θ', whereby one block is split into two blocks by adopting the point Q in the case where $\theta_q < \theta'$ as a fresh split point D'.

The improved method for establishing blocks is as mentioned above. And when the method for establishing sampling points in the construction of Fig. 3 is again applied to the respective blocks thus split into two sections, the number of sampling points decreases in the vicinity of peculiar points such as, particularly, linear intersections or an extreme end portion of so-called "hane" in Chinese or katakana character of the character outline, and as a result, it contributes to effects for data compression.

One example in the case where a character outline is actually processed on the basis of the method for establishing sampling points as well as the method for establishing blocks which have been described in detail hereinbefore will be illustrated in Fig. 7 for reference.

In Fig. 7, mark "o" designates an initial split point being described in conjunction with the illustration of

Fig. 1, mark "•" designates a second split point determined by means of the improved method for establishing blocks, and mark " " designates a sampling point which was determined after establishing the aforesaid second split point, and in case of Fig. 7, the number of these points is 40, 9 and 99, respectively.

Attention is drawn to our co-pending application No. 8303086, which is directed to the method of establishing sampling points described herein.

CLAIMS:

1. A method of compressing data representing a character outline wherein the outline is initially divided into a set of blocks by split points and each block is sampled at a plurality of sample points, the method comprising: establishing a new split point subdividing the initial block into two new blocks, by computing for each sample point the density of sample points in its vicinity and establishing one such sample point as said new split point on the outline when the density of sample points about the said such sample point exceeds a prescribed threshold.

2. A method of compressing data as claimed in claim 1, wherein the initial block is sampled at sample points P_1 to P_n, including points P_{i-1} to P_{i+3}, and the density of sample points for the point P_i is given as ρ_i where

$$\rho_i = \overline{P_i, P_{i+1}} + \overline{P_{i+1}, P_{i+2}}$$

and when ρ_i is smaller than a threshold ρ' establishing a point Q on the initial block at a maximum perpendicular distance from a line joining points P_{i-1} and P_{i+3} and where an angle Θq given by

$$\Theta q = P_{i-1}, Q, P_{i+3}$$

is smaller than a prescribed value Θ' establishing the said point Q as the new split point.

3. A method of establishing blocks for compressing character data on a character outline, substantially as herein described, with reference to the accompanying drawings.