SEGMENTATION OF COLOR STILL IMAGES USING VORONOI DIAGRAMS

Susumu Itoh and Ichiro Matsuda
Science University of Tokyo
2641 Yamazaki Noda-shi, Chiba Prefecture, 278 JAPAN
Tel: +81 471 24 1501, ext. 3711; fax: +81 471 24 9367
e-mail: itoh@itohws01.ee.noda.sut.ac.jp

ABSTRACT

This paper proposes a new segmentation method based on Voronoi diagrams in order to develop efficient region-oriented coding for color still images. The method disposes generators according to local activity of a color image, and modifies their positions so that boundaries between Voronoi regions can run parallel to the principal contours in the image. Since a Voronoi diagram is uniquely determined by only positions of generators, the method can efficiently represent region-shapes. Moreover it can segment images quite freely, because there is in general no limitation about disposition of generators. Simulation results indicate that the method realizes better segmentation even at a low coding rate than a conventional method.

1 INTRODUCTION

DCT coding has become the mainstream in the field of efficient coding for still images, and JPEG is well-known as a typical example. However DCT coding usually partitions an image into square blocks independent of its contents, therefore interferences such as mosquito-noise and blocking-effects deteriorate reproduced image quality. Moreover, interblock correlation which remains in most of square-block-based coding schemes decreases coding efficiency.

As a promising coding technique to overcome these problems in essence, region-oriented coding which utilizes two-dimensional structure of still images has recently attracted special interest [1]. Region-oriented coding first segments an image into homogeneous regions, then encodes region-shapes and region-contents separately. Accordingly, it is necessary to represent efficiently these two kinds of information, especially region-shapes for the purpose of attaining high coding performance.

From this point of view, we formerly proposed adaptive transform coding based on variable-shape-blocks [2]–[4]. The coding scheme, which we call VSB hereafter, partitions an image into square blocks and transfigures these blocks into appropriate variable-shape-blocks so that their boundaries can run parallel to the principal contours in the image. Since VSB limits every block-shape to a quadrilateral, it can encode block-shapes, or region-shapes at the low coding rate of about 0.12 bits/pel for any image [4].

However, the above-mentioned limitation about block-shapes as well as preservation of connective relations between adjacent vertices of each block reduces degree of freedom related to segmentation of an image. In other words, methods which can more freely segment an image are required. And moreover in order to realize well-balanced bit allocation to coding of region-shapes and region-contents, it is necessary to develop such segmentation methods as not only encode region-shapes at a low coding rate but can control the rate freely. To meet the requirements, we previously reported on a new segmentation method of monochrome still images using Voronoi diagrams [5]. This paper describes application of the method to segmentation of color still images and evaluation of its performance.

2 SEGMENTATION USING A VORONOI DIAGRAM

Let $\boldsymbol{g}_i = (x_i, y_i)$ $(i = 1, 2, \ldots, N)$ be one of N points called generators on an image. The set $V(\boldsymbol{g}_i)$ of all points $(\boldsymbol{p} = (x, y))$ which satisfy the following equation is called the Voronoi region of $\boldsymbol{g}_i$:

$$V(\boldsymbol{g}_i) = \{\boldsymbol{p} |\, d(\boldsymbol{p}, \boldsymbol{g}_i) < d(\boldsymbol{p}, \boldsymbol{g}_j) \quad \text{for} \quad j = 1, 2, \cdots, N \quad \text{and} \quad j \neq i\}, \tag{1}$$

where $d(\boldsymbol{p}, \boldsymbol{g})$ represents a distance between two points $\boldsymbol{p}$ and $\boldsymbol{g}$. There are several kinds of definitions with regard to the distance; we adopt the Euclidean distance as an example in this paper.

A Voronoi diagram is then expressed as the union of Voronoi regions of all the generators $\boldsymbol{g}_i$s, and it is uniquely determined by only positions of the N generators. Therefore we can utilize a Voronoi diagram to represent efficiently region-shapes. Since there is generally no limitation with respect to disposition of generators, it is possible to segment an image quite freely. In addition, such a segmentation method can almost freely control the coding rate of region-shapes by changing a value of N. These facts indicate that a Voronoi diagram has properties suitable for representation of region-shapes in region-oriented coding.

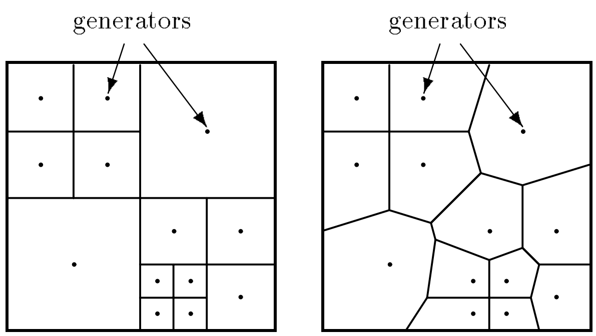
In this paper we assume that an image is a digitized one composed of $M \times M$ pels. Accordingly a Voronoi diagram is constructed by applying Eq.(1) to each pel $\boldsymbol{p} = (x, y)$ $(x, y = 1, 2, \ldots, M)$ and by giving $\boldsymbol{p}$ a label which shows that $\boldsymbol{p}$ is included in $V(\boldsymbol{g}_i)$. Furthermore we assume that a color image consists of Y, C_R and C_B signals. In order to reduce the amount of information on region-shapes as well as the amount of computation required for giving the label to each pel, the region-shapes for the three signals accord with one another. In other words the segmentation method constructs only one Voronoi diagram for one color image.

3 INITIAL DISPOSITION OF GENERATORS

To dispose appropriately generators at variable density according to local activity of image signals, initial positions of generators are determined as follows. First of all the color image is divided into four square blocks composed of $M/2 \times M/2$ pels. In each block AC_Y, AC_{CR} and AC_{CB}, which represent AC energy of Y, C_R and C_B signals respectively, are calculated and the following parameter T_{RGB} which approximately represents the total AC energy in the RGB signal space [4] is also calculated as the local activity:

$$T_{RGB} = 3.0 \cdot AC_Y + 1.26 \cdot AC_{CR} + 1.04 \cdot AC_{CB}. \quad (2)$$

This equation is derived from the color coordinate transform matrix on the assumption that AC components of Y, C_R and C_B signals at the same pel are statistically noncorrelated with one another. Then the block where T_{RGB} is the largest is redivided into four square blocks. The same operation for dividing the color image is recursively iterated until the number of blocks becomes N. Finally a generator is placed in the center of each block as shown in Fig.1 (a). The segmentation method always places each generator just on some pel, however, so as to prevent the amount of information on its position from increasing. Fig.1 (b) indicates the initial Voronoi diagram determined by the generators.



(a) Initial disposition of generators ($N = 13$) (b) Initial Voronoi diagram

Fig.1 An example of initial disposition of generators and the initial Voronoi diagram determined by these generators.

4 RENEWAL OF POSITIONS OF GENERATORS

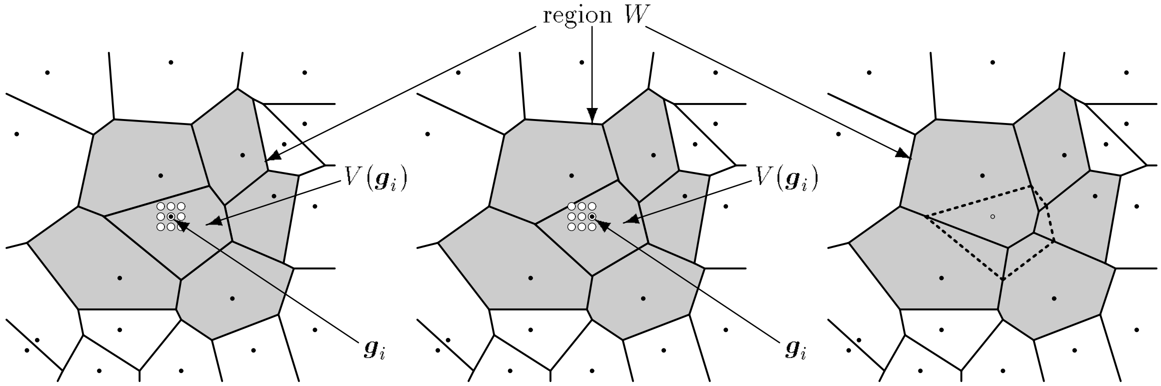
4.1 Basic Idea for Renewal

In order to segment a color image into homogeneous regions, boundaries between Voronoi regions need to run parallel to the principal contours in the image. Thereupon the segmentation method repeatedly renews a position of each generator and gradually transfigures the Voronoi diagram so that the following evaluation parameter can have a minimum. The evaluation parameter is also defined by Eq.(2). In this case, however, AC_Y, AC_{CR} and AC_{CB} represent energy of residual signals concerned with Y, C_R and C_B signals respectively [4]. The residual signals concerned with Y signals, for example, are produced by subtracting an interpolation image from the original monochrome image, and the interpolation image is constructed by feeding an only mean value of Y signals in each region to the adaptive-smoothing filter [3]. Such an interpolation image and residual signals are individually produced in the same way for C_R and C_B signals as well. Minimization of the parameter T_{RGB}, which represents approximately the total energy of the residual signals in the RGB signal space in this case, brings a good segmentation result suitable for color images.

4.2 Concrete Procedures for Renewal

Construction of a Voronoi diagram as well as the three interpolation images (color interpolation image) requires a lot of computation. Therefore it is difficult in practice to calculate repeatedly a value of the above evaluation parameter in the whole image. To cope with this problem, the parameter is redefined on the assumption that the only shapes of $V(\boldsymbol{g}_i)$ and its neighboring Voronoi regions change with the position of the generator $\boldsymbol{g}_i$. In other words, it is assumed that the external shape of the region W, which is the union of these Voronoi regions and is painted gray in Fig.2, does not depend on the position of $\boldsymbol{g}_i$. The new evaluation parameter E_k $(k = 0, 1, \ldots, 9)$ is defined as the total energy of the three residual signals in the region W. If $\boldsymbol{g}_i$ is moved through a very short distance, the above assumption is satisfied, as a result E_k is equivalent to the foregoing parameter T_{RGB}.

Concrete procedures for renewing the position of $\boldsymbol{g}_i$ are as follows. First of all, a value of E_0 is calculated when $\boldsymbol{g}_i$ is located on its initial position as shown in Figs.2 (a) and 3. Next, moving temporarily $\boldsymbol{g}_i$ to the



(a) Initial position of the generator $\boldsymbol{g}_i$ $(k=0)$

(b) Movement of $\boldsymbol{g}_i$ to the pel $\boldsymbol{p}_1$ $(k=1)$

(c) Removal of $\boldsymbol{g}_i$ from the image $(k=9)$

Fig.2 Renewal of the position of the generator $\boldsymbol{g}_i$ and reconstruction of Voronoi diagrams.

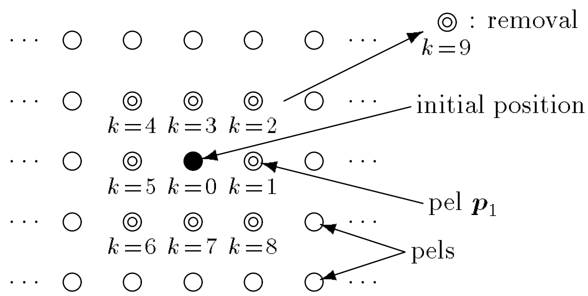


Fig.3 Initial position of the generator $\boldsymbol{g}_i$ and candidates (●, ◎) for its next position.

pel $\boldsymbol{p}_1$ in Fig.3, and reconstructing a Voronoi diagram in the only region W, we obtain Fig.2 (b). Then a value of E_1 is calculated with respect to the Voronoi diagram. Besides, we calculate values of $E_2 \sim E_8$ in the same way as E_1 after moving temporarily $\boldsymbol{g}_i$ to each of the seven remaining pels shown in Fig.3. On the other hand, Fig.2 (c) indicates the Voronoi diagram constructed by removing temporarily $\boldsymbol{g}_i$ from the image, and a value of E_9 is calculated in this case. Finally the generator $\boldsymbol{g}_i$ is moved to the pel where the parameter E_k is the smallest. If E_9 is the smallest, however, $\boldsymbol{g}_i$ is removed.

In addition when E_0 is the smallest, the following probabilistic method similar to the simulated annealing is carried out. First, an integer is chosen from among $1 \sim 9$ at random as a value of the subscript k, and the probability $P(k)$ that we regard E_k as smallest instead of E_0 is defined by:

$$P(k) = \exp\left\{-(E_k - E_0)/T\right\}, \qquad (3)$$

where T represents temperature in the annealing. Then we decide according to the value of $P(k)$ whether the generator $\boldsymbol{g}_i$ is moved to the pel corresponding to the chosen value of k in Fig.3 or not (in the case of $k = 9$, $\boldsymbol{g}_i$ is removed or not). Accordingly when E_0 is the smallest, the probability that $\boldsymbol{g}_i$ does not move anywhere but stays at the present position is $1 - P(k)$. Introduction of such a probabilistic method diminishes the harmful influence of local minima and improves remarkably performance of segmentation [5].

The above-mentioned procedures are carried out once for each generator in turn. Thus the first operation for transfiguration of the Voronoi diagram in the whole image comes to an end. Then we newly append generators, the number of which is equal to that of the removed generators, to the Voronoi regions where the total energy of the three residual signals is relatively large. The operation including re-disposition of the removed generators is iterated until the Voronoi diagram converges. After conducting segmentation of the color image in this way, the final position of each generator and a mean value of each of Y, C_R and C_B signals in each Voronoi region are encoded using run-length coding and predictive coding, respectively. We can reconstruct the Voronoi diagram and the color interpolation image by utilizing only these two kinds of information.

5 SIMULATION RESULTS

Fig.4 indicates relationship between signal-to-noise ratios (SNR) and coding rates of color interpolation images, which are obtained through computer simulation using the original color image of Lenna shown in Fig.5 (a). In this paper SNR is defined by:

$$\mathrm{SNR(dB)} = 20 \cdot \log_{10}(255/\sqrt{\mathrm{MSE}}). \qquad (4)$$

The proposed method provides about 2 dB additional improvement in SNR of Y signals over VSB. With regard to C_R and C_B signals, SNRs of the proposed method are about 1 dB higher than those of VSB. This is mainly because the former which is based on a Voronoi diagram has greater degree of freedom concerning segmentation of an image than the latter. Besides by changing a value of N, the proposed method can easily and freely control

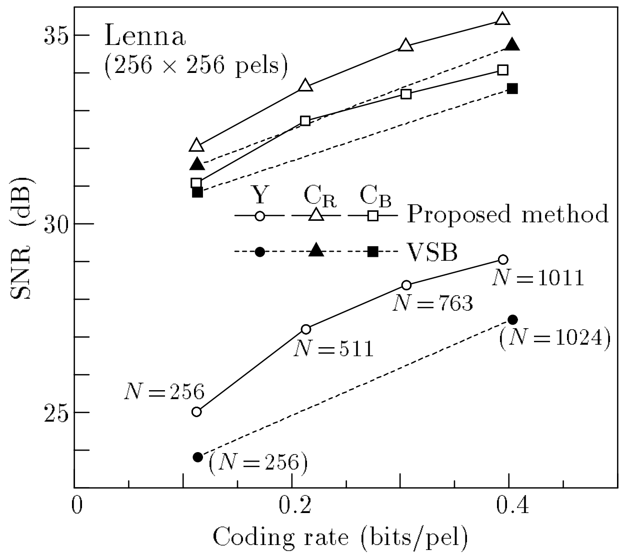


Fig.4 Signal-to-noise ratios vs. coding rates of color interpolation images.

a coding rate of a color interpolation image as shown in Fig.4. On the contrary, since VSB must generally employ uniform square blocks in an initial state of segmentation [3], [4], it can not control the coding rate so easily and freely as the proposed method can.

Fig.5 shows examples of color interpolation images which have been constructed at the almost identical coding rate. The figure demonstrates that the proposed method can represent outlines of the original image, especially the eyes, the lips and the accessories of the hat much more clearly than VSB. Consequently, the proposed method has much superiority over VSB as a segmentation method of color still images.

6 CONCLUSIONS

Toward the development of an efficient region-oriented coding scheme, this paper has proposed a new segmentation method of color still images using a Voronoi diagram. The method can quite freely segment an image, and moreover it can not only encode region-shapes at a low coding rate but control the rate easily. Simulation results indicate that the proposed method achieves high performance from the viewpoints of coding efficiency as well as quality of interpolation images. The reduction of the amount of computation required for segmentation and the development of a coding technique suitable for region-contents are the important problems for a future study.

REFERENCES

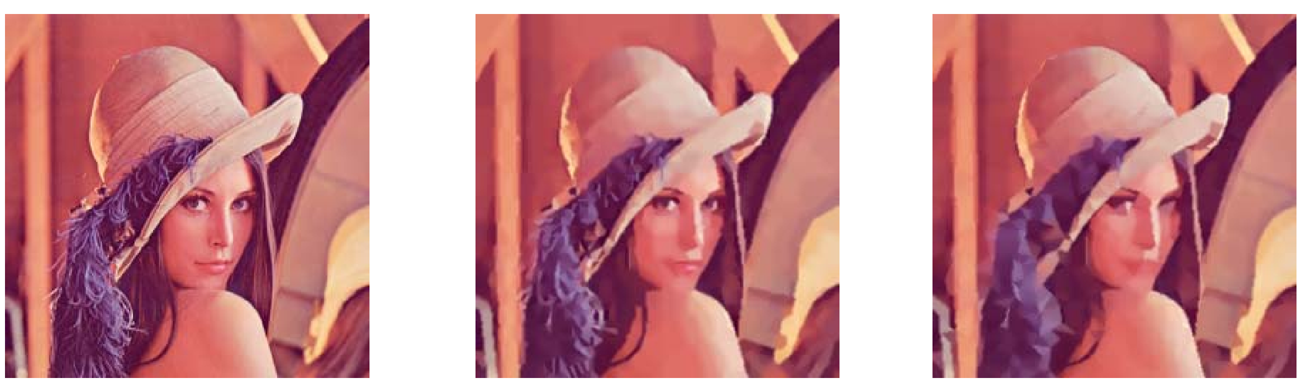
[1] M. Kunt, A. Ikonomopoulos and M. Kocher, “Second Generation Image Coding Techniques”, Proceedings of IEEE, Vol. 73, No. 4, pp. 549–574, Apr. 1985.

[2] S. Itoh, I. Matsuda and T. Utsunomiya, “Adaptive Transform Image Coding Based on Variable-Shape-Block-Segmentation with Smoothing Filter”, Proceedings of Visual Communications and Image Processing (VCIP’92), pp. 1338–1349, Boston, Nov. 1992.

[3] I. Matsuda, S. Itoh and T. Utsunomiya, “Adaptive Transform Image Coding Based on Variable-Shape-Blocks and Directional Autocorrelation Models”, Signal Processing VII, Theories and Applications (Proceedings of EUSIPCO–94), pp. 163–166, Edinburgh, Sep. 1994.

[4] S. Itoh and I. Matsuda, “New Transform Coding of Color Still Images Based on Variable-Shape Blocks”, Proceedings of IEEE Workshop on Nonlinear Signal and Image Processing (NSIP’95), pp. 344–347, Neos Marmaras (Greece), June 1995.

[5] S. Itoh and I. Matsuda, “A Segmentation Method Using Voronoi Diagrams for Still Image Coding”, Proceedings of International Picture Coding Symposium (PCS’96), pp. 535–540, Melbourne, March 1996.



(a) Original color image of Lenna (b) Proposed method ($N = 1011$) (c) VSB ($N = 1024$)

Fig.5 Original color image and color interpolation images.