

A STUDY ON TWO-LAYER CODING FOR ANIMATION IMAGES

Ouji NAKAGAMI[†], Toshinori MIYAZAWA[‡], Hiroshi WATANABE[†], Hideyoshi TOMINAGA^{†,‡}

[†] Graduate School of GITS
Waseda University

1-3-10, Nishiwaseda, Shinjuku-ku, Tokyo, Japan

[‡] Dept. of Elec. Info. and Comm. Eng.
Waseda University

3-4-1, Okubo, Shinjuku-ku, Tokyo, Japan

ABSTRACT

A coding scheme specifically designed for animation images is proposed. Taking characteristics of animation images into account, lines and homogeneous color regions are extracted from animation images. Lines and contours of the homogeneous color regions are approximated by straight-line and spline functions. We found that smoothing operations are effective to extract homogeneous regions from background images. Connected Filter is used for smoothing operation while keeping edges of the original image correctly. Precise approximation of the contours is also required. For this purpose, we propose to use Dynamic Programming with a feedback algorithm. As a result, an animation image can be represented by two layers; significant points as a base layer, and DCT blocks for an additional layer. For compensation scheme to the loss caused by smoothing operation, DCT is applied to the difference between the original images and lines and homogeneous color regions images. High-quality images can be obtained by adding this differential data to the approximated images.

1. INTRODUCTION

Discrete Cosine Transform (DCT) based coding, such as MPEG and JPEG, has been widely used for image compression. However, DCT-based approach may not be suitable for every kind of images. Some sort of images are often encoded efficiently by taking characteristics of the image content into account (Content-oriented Coding Scheme). As an example of this, we have been studying animation image coding.

Natural images mostly have smooth fluctuations in luminance, thus powers of DCT coefficient concentrate at low frequencies. On the other hand, animation images have many impulsive changes in luminance especially around line drawings, and for this, powers of DCT coefficient disperse from low to high frequencies. This effect makes coding efficiency worse in a DCT-based image coding, especially by causing many mosquito noises near edges.

Therefore, we have been proposing a coding scheme specifically designed for animation images [1]. Animation images are constructed from three components, such as (1) line drawings, (2) homogeneous color regions and (3) background regions (Fig.1). Line drawings, which an animation cartoonist creates, are usually black and serve as contour lines of homogeneous color regions. Homogeneous color regions can be defined as an area where pixel values of luminance and chrominance are almost the same. These regions can be seen on a cartoon character's face for example.

Background regions are similar to natural images since they usually have many colors and details without obvious edges.

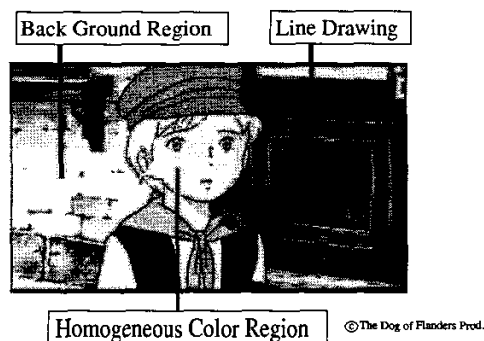


Figure 1: Three components of an animation image

In the method that we propose in this paper, first, line drawings and homogeneous color regions are extracted from an animation images. Then, all lines and contours of homogeneous color regions are approximated by some significant points. An advantage of this coding scheme is that it achieves high compression ratio by allowing some distortion of lines and edges. Computational complexity required for decoding process can be decreased, since only a few points are enough to represent lines. Furthermore, the size of the decoded images can be changed in any resolution by a number same as that of significant points. However, an encoder is estimated as 100 times more complex compared to that of DCT-based approach while a decoder has almost the same order.

2. ANIMATION IMAGE CODING SCHEME

Animation images are coded in two layers; base layer and additional layer. Fig.2 illustrates the flowchart of an animation image coding. Most animation images include many homogeneous color regions. Therefore, we represent animation images with lines and homogeneous color regions as a base layer image, which is constructed from only the significant points extracted from line image and region image. First, line drawings are extracted from an original animation image by using combinations of spatial filters. Next, homogeneous color regions in the original animation image are identified, and then the representative color and the contour lines are obtained. Detected lines, line drawings and

contour lines of homogeneous color regions, are traced to find significant points along the lines. These points are used to approximate the original lines by spline curves.

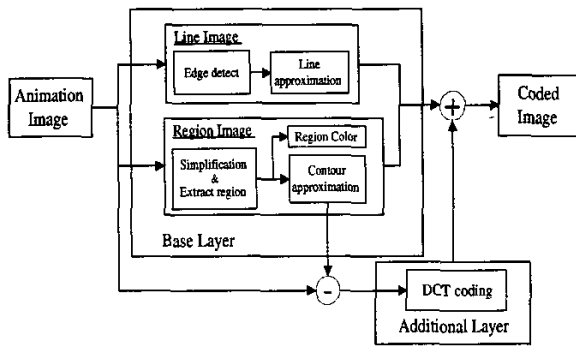


Figure 2: Flowchart of animation image coding

The extraction of a base layer has two issues to be solved. Firstly, in the animation image coding [1], background regions, which were defined as neither back ground regions nor line drawing in an animation images, were coded by DCT since they are close to natural images. A homogeneous color region is detected by region-growing algorithm [2]. However, it fails in some cases because of the mixture of the homogeneous and the background regions. Thus, it is necessary to improve such extraction process. We propose to introduce a joint method of smoothing operator and region-growing algorithm for this purpose. A flowchart of the proposed animation image coding is shown in Fig.2.

Secondly, line approximation algorithm fitting for various shapes to the contours is needed. Ramer algorithm [5] has been used in the conventional approach. It is suitable for straight-line approximation. Thus, significant points determined by the Ramer algorithm are not always proper for spline curve function. Therefore, Dynamic Programming with feedback algorithm is developed.

As an additional layer, a compensation method to the result of smoothing and approximation operation is proposed, and it enables to acquire a high quality coded image. DCT is used for the difference between the original images and the approximated images because luminance signals of the difference contain smooth changes. In addition, each region does not contain edges that are not suitable for DCT-based coding. High-quality images can be obtained by adding this differential data to approximated images that are created by homogeneous color regions and line drawings.

2.1. Region Extraction with Smoothing Operator

We use smoothing operator and region-growing algorithm jointly to extract homogeneous color regions from background regions. Smoothing operations by traditional linear filter loses edges in original images, that they are not suitable for animation images with many edges. Connected Filter [3], which is used in this method, is one of smoothing filters for the simplification of images. Since this filter keeps high frequency components unlike linear low-pass filters, the edges in the filtered image are kept. Fig.3 shows the

result of the smoothing by Connected Filter. Then, region-growing algorithm extracts homogeneous color regions more correctly than the conventional approach.

Fig.4 and Fig.5 show the process of the base layer coding. Fig.4(a) ("dog1"), which does not include background regions, indicate that the edges in the image are kept in spite of smoothing operation by Connected Filter. Fig.5(a) ("dog2") indicates that the homogeneous color regions are extracted by region-growing algorithm from simplified background regions.

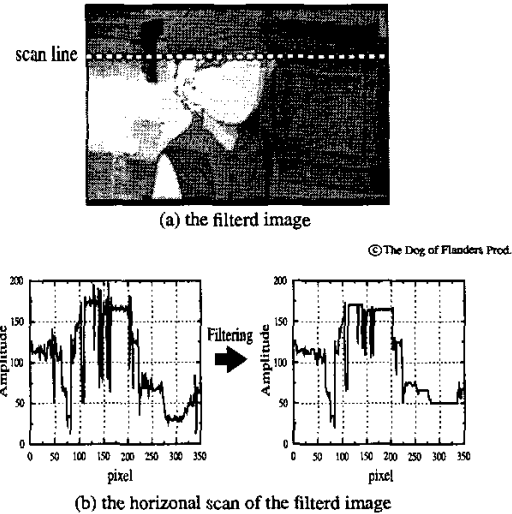


Figure 3: The result of Connected Filter

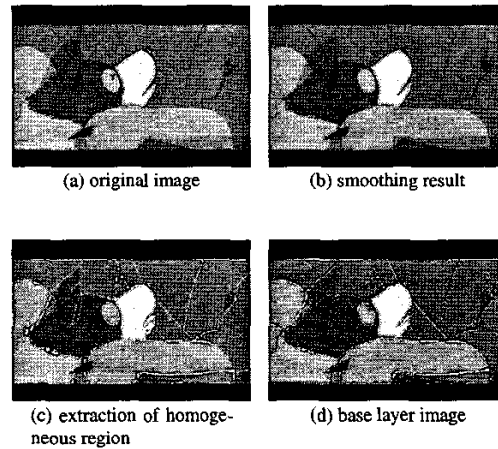
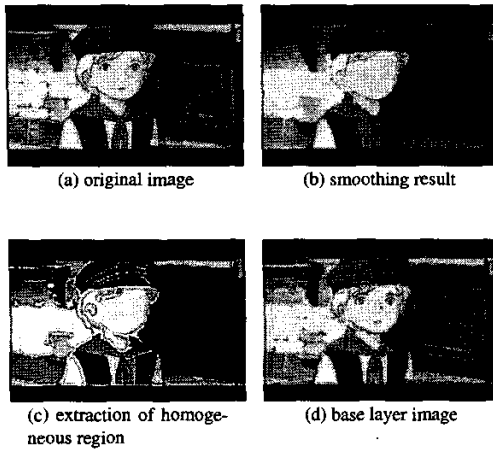


Figure 4: Extraction of the base layer for the image with no background regions ("dog01")

2.2. Dynamic Programming with feedback algorithm

Dynamic Programming examines all points on the curve since they may be chosen as significant points. It deter-



©The Dog of Flanders Prod.

Figure 5: Extraction of the base layer for the image including complexity background regions ("dog02")

mines a set of the significant points that minimize the total error between original and approximated line. Although this method requires high computational complexity, the optimized approximation can be obtained.

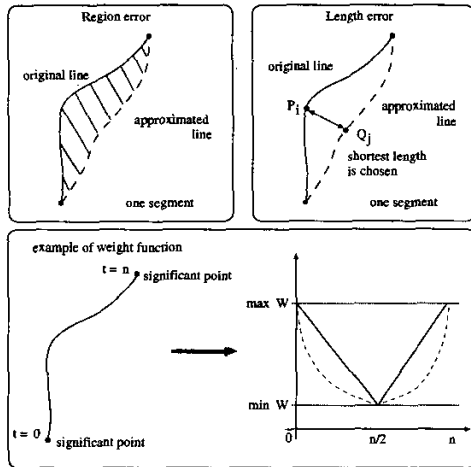


Figure 6: Estimation of the approximation error

There are two ways of estimating the total error in the conventional study[5], [6]; by region error algorithm and by minimizing the length error algorithm (Fig.6). The former estimates an error area between an original line and an approximation line. The latter does a summation of a length from points on an original line to the nearest points on an approximation line. The equation in the latter algorithm enables each approximation points to have a weight based on the characteristic of original lines (Eqn.(1)).

$$E = \frac{1}{l} \sum_{i=1}^l w_i |P_i - Q_j|^2 \quad (1)$$

where, E is an approximation error, l is the number of contour points, and w_i is the functional weight about point i on an original line. P_i are the coordinates of point i , and Q_j are the coordinates of point j , which is the nearest from point i . The set of significant points is calculated by the following steps.

1. Choose several significant points at high curvature points on original line.
2. Divide an original line into segments between significant points. The value of E is estimated by normalization of the length. Therefore, if the segment length is too long, DP algorithm cannot acquire a precise approximation result.
3. Check the value of E for each segment. If the error is larger than a threshold T_e , try other set of significant points in the segment.
4. Determine the points which minimize E of each segment as the significant points.

There are various kinds of contours in animation images. Since DP algorithm requires the threshold T_e to determine the best set of significant points, we must set suitable T_e for each contours. If the threshold is too small, many determined significant points concentrate at high curvature points on an original line. On the contrary, a large T_e increases the approximation error.

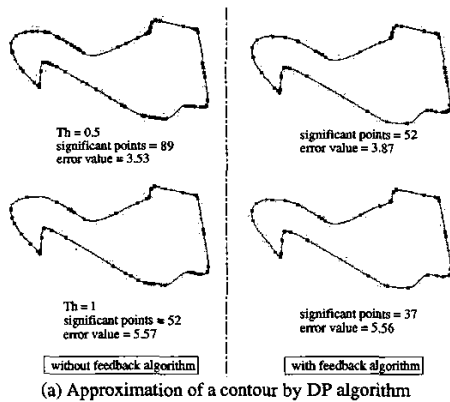
Therefore, we propose feedback algorithm to eliminate the useless significant points. First, a sufficient small T_e is set, and DP acquires enough significant points. Next, DP is applied again to eliminate significant points that do not reduce approximation error. Fig.7 shows the result of DP with and without feedback algorithm. Fig.7(a) illustrates the relation between the number of significant points and approximation error. DP with feedback algorithm enables more precise approximation in spite of the small number of significant points than that of conventional method (Fig.7(b)).

2.3. DCT Coding of Differential Image

We propose a compensational method to the result of smoothing and approximation operation as an additional layer. First, the difference between each approximated homogeneous color region and the original image is calculated. It is divided into 8×8 pixel blocks and coded by DCT. Some DCT blocks include edges when they locate on the contours of homogeneous color regions. To avoid mosquito noise caused by DCT coding, dummy data is padded to pixels not existing on homogeneous color regions (Fig.8).

Connected Filter is effective for not giving influence to the homogeneous color regions. The differential value is calculated at each point in the DCT block using Eqn.(2). D is the sum of $D(i, j)$ of the block. A block on the background regions has a bigger D than that homogeneous color regions because of the smoothing operation. Therefore, it is effective to prioritize the processing order of block according to the D value.

$$D(i, j) = P_o(i, j) - P_h(i, j) \quad (2)$$



(a) Approximation of a contour by DP algorithm

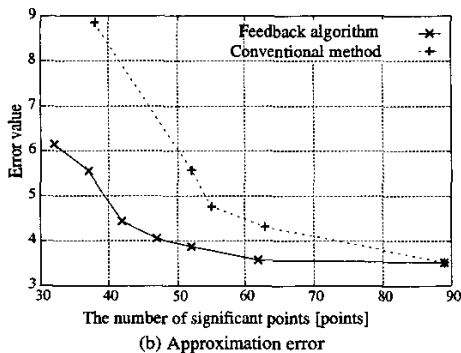


Figure 7: The result of DP with the feedback algorithm

$$D = \sum_{i=0}^7 \sum_{j=0}^7 D(i, j) \quad (3)$$

Fig.9 shows the effect of the additional layer. In the conventional method, the regions that can not be extracted as homogeneous color regions or line drawings were encoded by DCT. The homogeneous color regions detected in failure caused much error. On the contrary, the proposed method enables to control the number of DCT blocks to code. The result of Fig.9(b) shows that the additional layer is necessary for animation images including complicated background regions. When noisy image is input, noise component is once removed by Connected Filter, then it appears

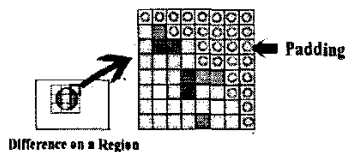


Figure 8: Padding dummy data

again in a differential image.

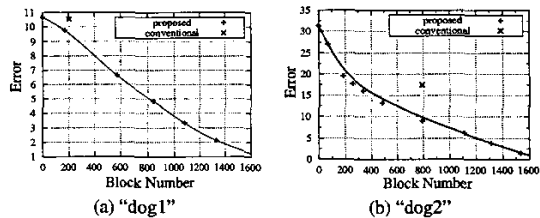


Figure 9: The effect of the additional layer

3. CONCLUSION

In this paper, we proposed the two-layer coding scheme designed for animation images. On the base layer, an animation image can be encoded at low bit rate since line drawings and contours of each homogeneous color region is expressed by significant points used for the approximation of the contour line. DP with feedback algorithm enables precise approximation for the contours by a controllable number of significant points. The lost data caused by smoothing operation is compensated by any number of DCT blocks on the additional layer. This enables smaller coding error than the conventional method. Currently, the proposed coding scheme is only designed for still images. Thus, this scheme can be used for the Intra frame coding of video coding.

4. REFERENCES

- [1] T. Miyazawa, W. Kameyama, H. Watanabe, T. Sakatani and H. Tominaga, "Animation Image Coding," in *Proc. World Multiconference on Systemics, Cybernetics and Informatics*, Vol.XIII, pp.404-409, July 2001.
- [2] R. Adams and L. Bischof, "Seeded Region Growing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(6), pp 641-647, Jun. 1994.
- [3] P. Salembier and J. Serra, "Flat Zones Filtering, Connected Operators, and Filters by Reconstruction," *IEEE Transactions on image processing*, Vol.4, No.8, pp.1153-1160, Aug. 1995.
- [4] C. A. Christopoulos, W. Philips, A. N. Skodras and J. Cornelis, "Segmented image coding: Techniques and experimental results," *Signal Processing:Image Communication*, Vol.11, pp.63-80, Nov. 1997.
- [5] U. Ramer "An iterative procedure for the polygonal approximation of planar curves" *Computer Graphics and Image Processing*, Vol.1, pp.244-256, 1972.
- [6] F. Lu, E. E. Milios, "Optimal Spline Fitting to Planar Shape," *Signal Processing*, vol.37, pp.129-140, 1994.