DBpia

論文 97-22-6-6

An Adaptive Approximation of Countours for a Region-based Image Sequence Coding

Chae Wook Lim*, Kang Hyuk Lee*, Kyeong Joong Kim*,
and Kyu Tae Park* *Regular Members*

※This study was funded in part by Korea Minitry of Information and Communication.

요 약

Encoding of segment contours is a critical part of a region-based coding system especially at low bit rates where the contour information occupies a majority of the bit rate. When approximating contours with polygons, a fixed upper bound on the distortion is set for the approximation process. Instead of using this fixed bound, adaptive approximation bound for a lossy coding of contourts is proposed in this paper. A function representing the relative importance of the contour segment is defined to take into account the spatial content of the image. By using this function, the contour can be approximated adaptively. This allows a more general approach than the methods with the fixed distortion measure. The effectiveness of the adaptive contour coding approach is verified through experiments.

Ⅰ. Introduction

Most traditional image coding schemes are block-based algorithms where an image is partitioned into blocks for processing. Owing to its efficiency, the block-based image coding approach has been incorporated into several international standards such as JPEG, H. 26x, and MPEG [1]. Despite this wide-spread use, the performance of block-based coding systems falls off considerably at low bit rates, exhibiting blocking and mosquito effects that are very annoying to the observers. These effects are caused by splitting the image into artificial blocks without any regard to the homogeneity of texture or motion.

To circumvent these shortcomings of block-based algorithms, the region-based (also refer to as the object-oriented) image coding approach was introduced [2]. The biggest advantage of this approach lies in the fact that the image is segmented into regions with similar texture and/or motion rather than artificial blocks, thus yielding a more natural looking image compared to the block-based schemes, particularly at low bit rates.

The major drawback of the region-based coding, however, is that the shape of the segments (called the contour) has to be explicitly encoded for each region whereas in the fixed-size block-based coding, no overhead is needed since every block has the same shape. The amount of contour information occupies a

---

*Dept. of Electronic Engineering, Yonsei University
論文番號:97085-0304
接受日字:1997年 3月 4日

www.dbpia.co.kr

major portion of the coding rate, and the proportion becomes even larger as the required rates decrease. In this respect, the coding of contour information has attracted much attention with greater emphasis on the lossy coding of contour points for a higher coding gain [3][4][5][6].

For the lossy coding of contour points, a fixed bound on the tolerable error is set for the approximation. This distortion criterion must be satisfied for the entire segment contour. However, not every portion of the contour requires the same degree of approximation accuracy. That is, by taking the inexact nature of the segmentation process into account, the maximum allowable distortion for a single segment can be varied. In this respect, for a more efficient region-based coding, we propose a novel contour coding scheme using an adaptive distortion measure. This paper is organized as follows; we review the lossy coding of contours in Section 2; in Section 3, a new distortion measure with adaptive bounds is presented; in Section 4, the contour encoding algorithm is described; the results of the experiments are given and discussed in Section 5; finally, the concluding remarks are given in Section 6.

Ⅱ. Lossy Coding of Contours

Coding of contour points has attracted much attention with emphasis on the lossy coding of the points for a higher coding gain. Though widely used, it is often impractical to use lossless coding of contour points such as the chain code [7] because of the high coding cost and the difficulties in rate control.

There are many approaches to contour approximation [8][9]. One of the most popular coding method is the polygon approximation. In this method, a contour is approximated by straight lines, and only those points corresponding to the endpoints of the lines (called vertices) need to be coded. There are many methods (e.g. iterative endpoint fit and split, uniform bounded-error approximation, etc.) available for finding the positions of the vertices, but regardless of the method employed an upper bound on the tolerable error, called the maximum allowable distortion, has to be set to control the degree of approximation accuracy.

A distortion measure used for the polygon approximation is,

$$d(v_i, v_{i+1}) = \max_{(x,y)\in S_i} \frac{(c_{i+1}-c_i)x + (r_i - r_{i+1})y + (r_{i+1}c_i - r_i c_{i+1})}{\sqrt{(r_{i+1}-r_i)^2 + (c_{i+1}-c_i)^2}} \qquad (1)$$

where,

i : $0, 1, \cdots, N_v - 1$
N_v : number of vertices
v_i : i-th vertex
S_i : ordered of contour points covered by a straight line between v_i and v_{i+1}
(r_i, c_i) : coordinate of v_i
(x, y) : coordinate of a contour point in S_i.

N_v line segments are needed for approximating the segment, and the distortion for a partial contour S_i is the longest distance between the approximated line and the contour points in S_i. Here, the distortion must be less than the given maximum allowable threshold $d_{\max}$ for the entire segment. Then, (1) can be rewritten as,

$$D = \max_{i\in\{0,1,\ldots,N_v-1\}} \{d(v_i, v_{i+1})\} \le d_{\max}. \qquad (2)$$

In an attempt to approximate the curve in a more natural looking way, a combined polygon-spline approximation can be used [4], but the distortion criterion remains the same.

Ⅲ. Adaptive Approximation Bound

1. New Distortion Measure

www.dbpia.co.kr

Unless we have an exact segmentation and the exact reconstruction of its boundary points is critical[1], we can alter the contour of the segment to reduce the coding rate without noticeable damage to the image quality. This assumption justifies the use of contour approximation schemes. In existing region-based coding schemes, the encoding of contour points is considered to be independent of the segmentation process. Though this is a valid assumption, the coding efficiency can be improved by taking into account the inexact nature of the segmentation process.

In this paper, instead of setting a constant distortion bound $d_{\max}$ for the entire segment contour, we propose an adaptive approximation bound where the distortion bound changes within the segment contour according to the nature of the segment. The following adaptive distortion measure is defined for the maximum approximation bound,

$$d_{\max}(v_i, v_{i+1}) = d_0 + \beta(v_i, v_{i+1}) \tag{3}$$

where d_0 is the base distortion bound and $\beta(v_i, v_{i+1})$ is the tolerance level for the approximation interval which will be denoted simply by β. By adjusting the value of β, the approximation constraint becomes tighter for critical portions of the segment and less so for other portions. Hence, an adaptive approximation of the contour points is possible. Note that when β is constant for all i, the problem reduces to that of a simple contour approximation with the constant maximum distortion of $d_0 + \beta$. Thus, the proposed adaptive bound allows a more general encoding of segment contours. The question now is the determination of β. In this paper, we define β as a function of the luminance of the image. The determination of β is described in detail in the next section.

2. Determination of β

One of the most widely used features in image processing literature is edge properties. It is enticing to use edge properties in applications involving boundaries or contours due to the significance it has in human vision [10]. By defining β as a function of the edge strength, its value can be adjusted to reflect the relative importance of a given partial contour. That is, those portions with greater visual importance are more accurately approximated. In this paper, Canny operator is used to extract the edges [11].

Suppose $G_n(x, y)$ is a first derivative of a two-dimensional Gaussian smoothing operator in the direction n that is normal to the edge.

$$G_n(x, y) = \frac{\partial G}{\partial \mathbf{n}} = \mathbf{n} \cdot \nabla G \tag{4}$$

where,

$$G(x, y) = \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right].$$

σ is the standard deviation of the associated probability distribution.

The edge position is then at the local maximum in the direction n of the operator convolved with the image,

$$\frac{\partial}{\partial \mathbf{n}} G_n(x, y) * I(x, y) = 0 \tag{5}$$

Substituting in (5) for $G_n(x, y)$ from (4) we have,

$$\frac{\partial^2}{\partial \mathbf{n}^2} G(x, y) * I(x, y) = 0 \tag{6}$$

The edge strength is evaluated when (6) is satisfied. In other words, the directional edge strength is evaluated only where a zero-crossing occurs. Let us denote the directional edge strength by $E(x, y)$. Then,

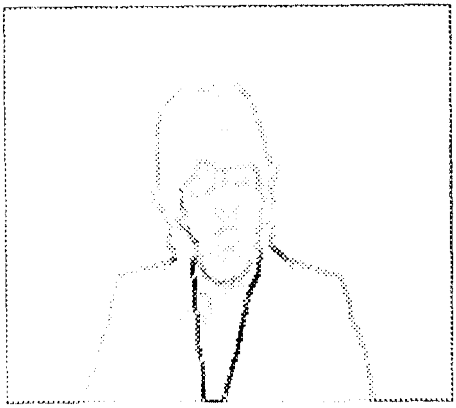
$$E(x, y) = \begin{cases} |G_n(x, y) * I(x, y)|, \text{ where } \frac{\partial^2}{\partial \mathbf{n}^2} G(x, y) * (x, y) = 0 \\ 0 \qquad , \text{elsewhere} \end{cases} \tag{7}$$

---

[1] Note that this requires a lossles coding which is not the focus of our study.

www.dbpia.co.kr

Note that there is no thresholding process to find single-pixel-width binary edges. A normalized $E(x, y)$ of an image is shown in Fig. 1.



(a)

(b)

Fig. 1 (a) $E(x, y)$ of an image ($\sigma = 1.0$) and (b) the orginal image.

For a practical implementation of the presented scheme, β should be set according to the nature of the segment and at the same time it should not deviate too far from d_0. Otherwise an excessive geometrical distortion may result. In this paper, the bound on $C \cdot d_0$, i.e., β cannot differ C times the value of the base bound.

We now define the relationship between the normalized $G_n(x, y)$ and β as,

$$\beta(v_i, v_{i+1}) = d_0 \cdot \left(1 - \frac{\mu(v_i, v_{i+1})}{\mu_E}\right) \tag{8}$$

where,

$$\begin{aligned} \mu_E &= \frac{1}{N_{S^\Theta}} \sum_{(x,y) \in S^\Theta} E(x, y), \\ \mu(v_i, v_{i+1}) &= \frac{1}{N_{S_i^\Theta}} \sum_{(x,y) \in S_i^\Theta} E(x, y) \end{aligned} \tag{9}$$

Here, S is the set of all contour points. The superscript Θ means that the set is extended to include the second-order neighborhood of the points in the set, and $N_{S_i^\Theta}$ is the number of elements in S_i^Θ. In other words, the ratio of local mean and the global mean determines the value β.

Ⅳ. Implementation

1. Object Extraction

A spatio-temporal segmentation is performed to extract the segments needed for the experiment [15]. The motion of the extracted objects are estimated using the 6-parameter model [16]. The method involved in extracting the segments needed for the contour coding is not of primary concern and is not addressed in detail. The main focus, rather, is on the effect of approximated contours on the decoded image quality. So, with a given set of segments and their motion parameters, we reconstruct images using the motion parameters and the approximated contours.

2. Vertex Encoding

In any lossy coding system, the goal is to minimize the distortion at a given rate or minimize the rate given an upper distortion bound. An optimal solution in the rate-distortion (RD) sense for finding the polygon vertices is given in [12]. In this scheme, an iterative scheme which converges to the optimal solution is given. This scheme employs a shortest path algorithm for a weighted directed acyclic graph.

www.dbpia.co.kr

We can directly apply the optimal procedure with the new adaptive distortion measure. Note that with this scheme, we can obtain an optimal set of vertices for any given vertex encoding scheme. For the experiment we assume that the vertices must lie on the initial given contour[2]. Doing so ensures the stability of the adaptive scheme, i.e., the vertices themselves have the distortion of zero. The vertices are encoded using a relative address coding scheme [14]. Temporal prediction of contour points is not performed.

V. Experimental Results

The new contour coding algorithm using the adaptive approximation bound is compared to that using the fixed distortion bound for 30 frames of QCIF clair sequence. To find the edge strength, the standard deviation of the Gaussian filter is set to 1. Edge thinning is not necessary since the goal is not on finding the exact location of the edges but on finding the relative edge strength. For the fixed bound scheme, $d_{\max}$ is set to 2.0, and for the proposed method d_0 is adjusted to match the rate of the fixed distortion bound scheme. Fig. 2 shows the results compared to that of using the fixed distortion bound.

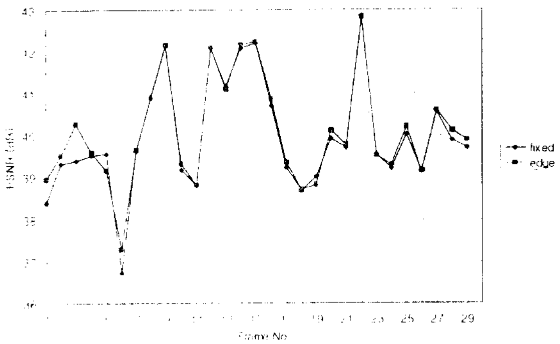


Fig. 2 Performance comparison.

The absolute PSNR values of decoded images themselves are not of great significance since the purpose of the experiment is not on proving the effectiveness of the entire coding system, but on investigating the performance of the adaptive contour coding scheme compared to the one with the conventional fixed distortion bound. In Fig. 3 the achieved gain using the adaptive scheme is shown.

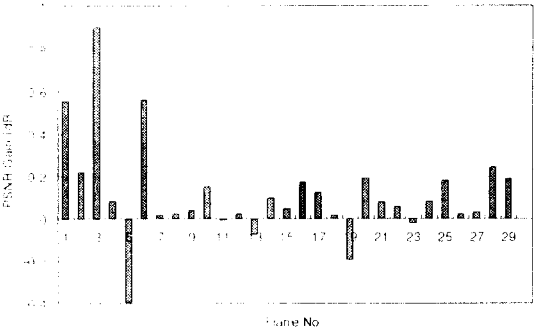
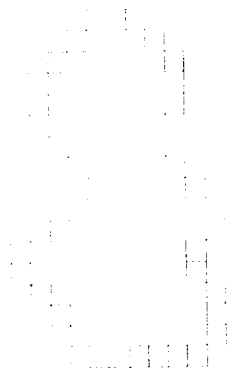


Fig. 3 Achieved PSNR gain.

An example of the shape approximation using the adaptive bound in comparison with the fixed bound is shown in Fig. 4. The region corresponds to the nose and mouth area. This figure clearly shows the advantage of using the adaptive bound. The dots represent the locations of the vertices. As can be seen,

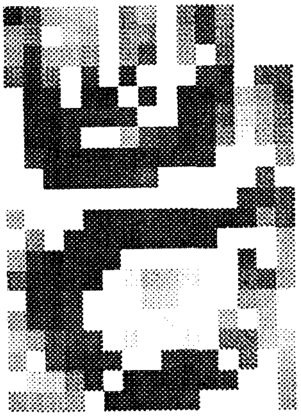


(a) original contour

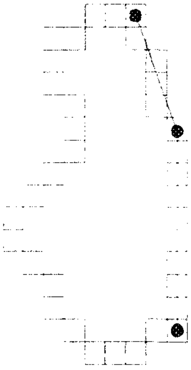
---

[2] This is not a necessary condition [13].

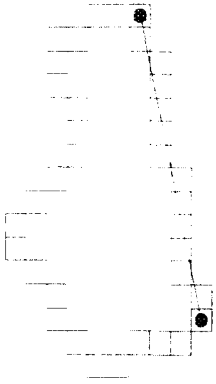
www.dbpia.co.kr



(b) $E(x,y)$



(c) fixed bound



(d) adaptive bound

Fig. 4 Shape approximation using the adaptive bound.

where the edge strength is relatively weak a vertex can be removed without any noticeable difference, and around the region with stronger edge strengths the approximation results are the same. Therefore, by using the proposed shape approximation scheme the object shape can be approximated more efficiently.

Ⅵ. Conclusion

An adaptive approximation bound for a lossy coding of segments has been presented. With the assertion that a fixed contour distortion criterion does not necessarily lead to a better image quality, the effectiveness of the new adaptive approximation bound was verified. Experimental results showed that at the same rate the adaptive bound leads to better image quality in terms of PSNR. For further studies, improvements in determining β which can reflect the relative importance of contour points is necessary.

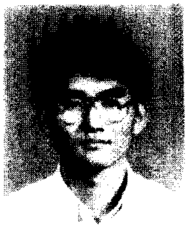
References

1. R. Clarke, *Digital Compression of Still Images and Video*, Academic Press, 1995.
2. H. Musmann, M. H tter, and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images," *Signal Processing : Image Communication*, vol. 1, pp.117-128, Oct. 1989.
3. M. H tter, "Object-oriented analysis-synthesis coding based on moving two-dimensional objects," *Signal Processing : Image Communication*, vol. 2, pp. 409-428, 1990.
4. P. Gerken, "Object-based analysis-synthesis coding of image sequences at very low bit-rates," *IEEE Trans. Circ. Syst. Video Technol.*, Vol.4, No.3, pp. 228-235, June 1994.
5. M. H tter, "Optimization and efficiency of an object-based analysis-synthesis coder," *IEEE Trans. Circ. Syst. Video Technol.*, Vol.4, No.2, pp.181-194, April 1994.
6. G. Heckner, "Redundance reducing coding of moving

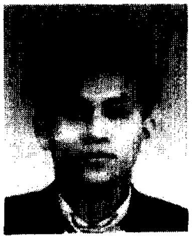
www.dbpia.co.kr

object shapes," *Signal Processing: Image Communication*, Vol. 9, No.1, pp.91-98, 1997.

7. R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley Publishing Co., 1993.
8. R. Haralick and L. Shapiro, *Computer and Robot Vision*, Vol.1, Addison-Wesley, 1992.
9. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*, Chapman & Hall, 1993.
10. D. Marr, *Vision-A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman and Co., 1982.
11. J. Canny, "A computational approach to edge detection," *IEEE Trans. PAMI*, Vol.8, No.6, pp. 679-698, 1986.
12. G. Schuster and A. Katsaggelos, "An optimal lossy segmentation encoding scheme," in *Proceedings of the Conference on Visual Communications and Image Processing*, SPIE, pp.1050-1061, Mar. 1996.
13. G. Schuster and A. Katsaggelos, "An efficient boundary encoding scheme which is optimal in the rate distortion sense," in *Proceedings of International Conference on Image Processing*, IEEE, Vol. 2, pp.77-80, 1996.
14. J. Chung, J. Kim, and J. Moon, "Shape information reduction based on contour prediction and shape coding type," ISO/IEC JTC1/SC29/WG11 Document MPEG95/0461, Nov. 1995.
15. L. Wu, J. Benois-Pineau, P. Delagnes, and D. Barba, "Spatio-temporal segmentation of image sequences for object-oriented low bit-rate coding," *Signal Processing: Image Communication*, Vol. 8, No.1, pp.513-543, 1996.
16. R. Tsai and T. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol. Assp-29, No.6, pp.1147-1152, Dec. 1981.



林 采 旭(Chae Wook Lim) 정회원
1968년 2월 20일생
1989년 2월 : 연세대학교 전자공학과 학사
1991년 8월 : 연세대학교 전자공학과 석사
1992년 9월~현재 : 연세대학교 전자공학과 박사과정
※주관심분야 : 영상 부호화, 시각 인식, 멀티미디어 시스템



李 康 赫(Kang Hyuk Lee) 정회원
1972년 4월 25일생
1995년 2월 : 연세대학교 전자공학과 학사
1995년 9월~현재 : 연세대학교 전자공학과 석사과정
※주관심분야 : 영상 부호화, 멀티미디어 시스템

金 慶 重(Kyeong Joong Kim) 정회원
1969년 7월 10일생
1992년 2월 : 연세대학교 전자공학과 학사
1994년 2월 : 연세대학교 전자공학과 석사
1994년 3월~현재 : 연세대학교 전자공학과 박사과정
※주관심분야 : 영상 부호화, 시각 인식

朴 圭 泰(Kyu Tae Park) 정회원
한국통신학회지 제21권 11호 참조

www.dbpia.co.kr