An Efficient Polygonal Approximation Method in the Rate-Distortion Sense

Byoung-Ju Yun[1], Jae-Soo Cho[2], and Yun-Ho Ko[3]

[1]School of EECS, Kyungpook National University, Daegu, South Korea
[2]School of IME, Korea University of Technology and Education, Cheonan, South Korea,
[3]Dept. of Mechatronics Engineering, Chungnam National University, Daejeon, South Korea
bjisyun@ee.knu.ac.kr

Abstract

This paper proposes an efficient method for encoding the shape information of the object in the image. The proposed method is based on the PVS (Progressive Vertex Selection) method and uses the new vertex selection condition. The proposed system selects less number of vertices than conventional methods when maximum distortion ($D^{*}_{\max}$) is given, so gives 5.5 ~ 17.2% bit reduction than PVS and 13.8 ~ 31.3 % bit reduction than IRM over test sequences and is also robust to the noise of the contour by using wider search range than PVS. Simulation results show that the proposed method has outstanding performance than other conventional vertex selection methods in the rate-distortion sense.
Keywords: Shape information coding, polygonal approximation, progressive vertex selection

I. Introduction

Region-based image coding method [1] and object-based image coding method [2] introduced shape coding into image and video coding. A region is defined by its homogeneous texture and described by texture and shape, and an object is defined by its uniform motion and described by motion, shape and color parameters. The purpose of using shape was to achieve better subjective picture quality and increased coding efficiency as well as an object-based video representation. MPEG-4 visual is the first international standard allowing the transmission of arbitrarily shaped video objects (VO's) [3]. MPEG-4 visual transmits motion, texture, and shape information of one VO within one bitstream. By MPEG-4 visual, the importance of shape for video objects is recognized. In MPEG-7, with other information such as color, texture, and motion of an object, shape is also used as a Descriptor (D), which describes an object.

Shape information is very important in many applications, thus several researchers have been studying to efficiently encode shape information [4] ~ [7]. There are two major classes of shape coders: bitmap-based coders and contour-based coders. The former encodes each pixel whether it belongs to the object or not and the latter encodes the outline of the object. The proposed method belongs to the second type. It approximates the boundary by a polygon, which has merits in inherent quality control [8], and considers the problems of finding the polygon, which leads to the smallest bit-rates for a given distortion.

The most well-known method in the vertex-based contour coding is the polygon-based contour approximation. Two schemes are commonly employed for vertex selection in polygon-based contour approximation: the progressive vertex selection (PVS) method [9] and the iterated refinement method (IRM) [6].

PVS selects polygon's vertices progressively. It traces contour points from initial point and check whether the peak distance (d_P) between a straight line, whose two end points are initial point and current contour point, and the contour portion, which is approximated by the segment, is larger than given maximum distance criterion ($D^*_{\max}$). If d_P is larger than $D^*_{\max}$, the very previous point is selected as a polygon's vertex and the process is continued from this vertex. It selects minimum number of vertices under maximum distance criterion ($D^*_{\max}$) so it reduces the amount of data needed to encode polygon's vertices, but the size of error region between an object and a polygon is large.

On the other hand, IRM decides polygon's vertices by iteration. The polygon approximation is computed by using those two contour points with the maximum distance between them as the starting points. Then, additional points are added to the polygon where the approximation error between the polygon and the contour is maximum. This is repeated until the shape approximation error is less than $D^*_{\max}$. Because of its vertex selection condition, high curvature points are apt to be selected as polygon's vertices so the area of the error region is small. But the polygon, which IRM produces, has large number of vertices.

The latter has a drawback of selecting a large number of vertices when the maximum allowable distortion ($D^*_{\max}$) is given, so many number of bits for encoding them is needed. The former, however, has an advantage of small number of bits for encoding an approximating polygon because of the small number of vertices it selects, but has a shortcoming when there are some noises on the contour, it may select excessive vertices.

So, in this paper, an efficient vertex selection method is proposed, which is strong to the contour noise and selects small number of vertices satisfying the given maximum distortion ($D^*_{\max}$). The proposed method is based on the progressive vertex selection method and suggests a new vertex selection condition, in which all possible contour points are selected as the candidate vertices and among them the furthest point is selected as the vertex of a polygon, so can approximate an object by using small number of bits.

This paper is organized as follows: In section II, general structure of the vertex-based shape information encoder is presented. In section III, the proposed vertex selection method is presented. The experimental results are shown in section IV and conclusions are in section V.

II. Vertex-based shape information encoder

Fig. 1 shows a general framework for shape coding methods that relies on polygonal approximations of the shape [4].

Vertex-based shape information encoder consists of several modules such as shape

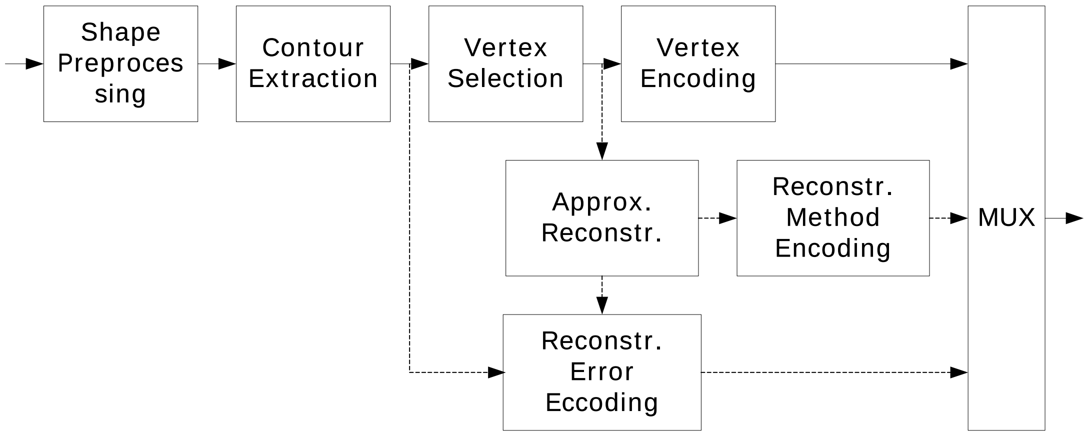


Fig. 1 General block diagram of vertex-based shape information encoder.

preprocessing, contour extraction, vertex selection, vertex encoding, approximation reconstruction, reconstruction method encoding, and reconstruction error encoding. At first, an input segmentation mask is preprocessed to decrease spatial resolution to the desired level and reduce unwanted noise and edge jaggedness. Then, a contour is determined and extracted from the input mask. Vertex selection (VS) module selects vertices defining a polygonal approximation of the contour subject to some distortion criteria or a limit of the number of vertices. Finally, vertex encoding (VE) module encodes the lists of vertices for each contour approximation. Some vertex-based coding methods reconstruct an approximation of the object contour and encode the approximation error in the dashed elements in Fig. 1.

Previous works in the conventional vertex-based encoding scheme focused on reducing the number of bits needed for encoding the given vertices by using the relations of successive vertices. We, in this paper, propose an efficient polygonal approximation scheme that selects less vertices of a polygon satisfying the maximum distortion criterion to improve the performance of the shape information coder.

III. Proposed vertex selection method

PVS method progressively searches the contour from the starting point (c_0) and selects the first contour point, which satisfies the vertex selection condition, as a vertex of a polygon as shown in Fig. 2, where the vertex selection condition of the conventional PVS method is given as follows:

Select a contour point c_n as the vertex
If $d(c_{\text{start}}, c_n) \le D^*_{\max}$ and $d(c_{\text{start}}, c_{n+1}) > D^*_{\max}$

Where, $d(c_a, c_b)$ is the peak distance between the partial contour (contour points from c_a to c_b) and the polygon line $\overline{(c_a, c_b)}$ and defined as follows.

$$d(c_a, c_b) = \max_{s \in \{c_a, c_{a+1}, \cdots, c_b\}} d'(c_a, c_b, s) \tag{1}$$

Where, $d'(c_a, c_b, s)$ is the distance from a point s to the polygon line $\overline{(c_a, c_b)}$ and defined as follows.

$$d'(c_a, c_b, s) = \frac{|(s_x - c_{a,x})(c_{b,y} - c_{a,y}) - (s_y - c_{a,y})(c_{b,x} - c_{a,x})|}{\sqrt{(c_{b,x} - c_{a,x})^2 + (c_{b,y} - c_{a,y})^2}} \tag{2}$$

Where, s_x and s_y are the x- and y-address of s, $c_{a,x}$ and $c_{a,y}$ are the x- and y-address of c_a, and $c_{b,x}$ and $c_{b,y}$ are the x- and y-address of c_b, respectively.

Fig. 2 shows an example of vertex selection process. Because the peak distance $d(c_0, c_{10})$ is larger than $D^*_{\max}$, PVS method selects the very previous contour point (c_9) as the vertex of a polygon. And from the contour point c_9, the process continues until the search process meets the last contour point. In this example, PVS method selects four contour points as the vertices of a polygon $(c_0, c_9, c_{21}$, and $c_{30})$ (see Fig. 2 (a)).

The proposed method selects the contour point, whose distance on the contour from the start point is longest, as the vertex of a polygon among the candidate points, which satisfy the vertex selection condition. The candidate vertex selection condition is defined as follows:

Select a contour point c_n as the candidate vertex
1. If $d(c_{\text{start}}, c_n) \le D^*_{\max}$ and $d(c_{\text{start}}, c_{n+1}) > D^*_{\max}$
2. If $d(c_{\text{start}}, c_n) \le D^*_{\max}$ after the contour point meets condition 1

That is, the first candidate vertex is the same point as that of the conventional PVS method selects. After that, the proposed method continues the contour search to find the vertices that satisfies the second condition. From this, the proposed method found 6 candidates $(c_9, c_{15}, c_{16}, c_{17}, c_{18}, c_{19})$ as shown in Fig. 2 (b). The proposed method selects the last candidate vertex as the vertex of a polygon, which has the longest distance from the previous vertex. As shown in Fig. 2 (b), using the new condition, there are just three vertices for approximating the open contour.

From the observation, we know that the proposed method can approximate an object using smaller number of vertices than the conventional PVS. This means that it can approximate an object by using small number of bits when the maximum distortion measure is given.

To find all the points that satisfy the maximum distortion criterion, it should search whole contour points until it meets the end point. This, however, causes the computational burden. To reduce it, the search range is restricted until the peak distance from the polygon line to the partial contour is not more than the two times of the given maximum distortion $(d(c_{\text{start}}, c_{\text{end}}) < 2 \cdot D^*_{\max})$. After the peak distance exceeds $2 \cdot D^*_{\max}$, there is not much possibility that the peak distance is less than $D^*_{\max}$ again, so we restrict the search range to $2 \cdot D^*_{\max}$. We can reduce the computational burden by restricting the search range.

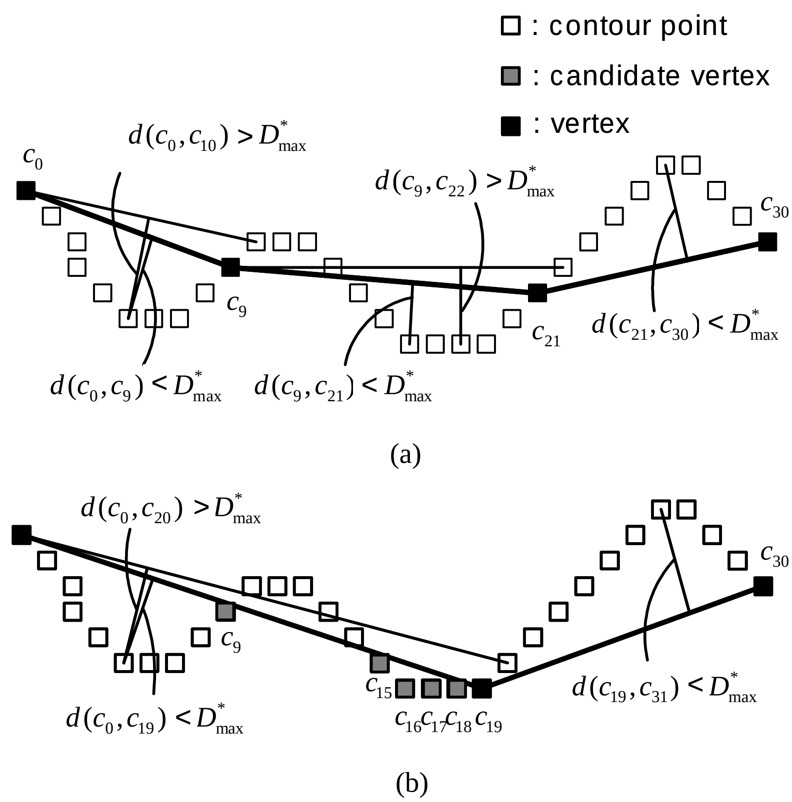
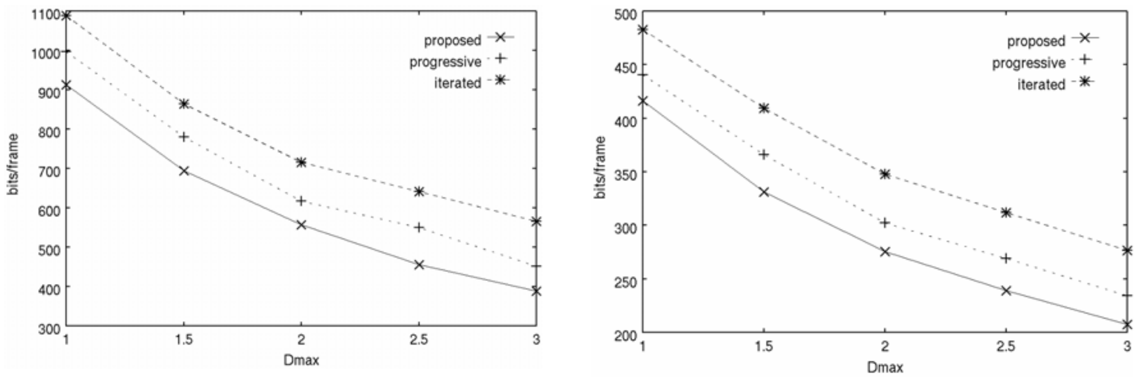


Fig. 2 Progressive vertex selection method
(a) conventional selection method (b) proposed selection method.

IV. Experimental Results

Experiments were conducted to compare the performance of the proposed vertex selection method with other methods, especially PVS (Progressive Vertex Selection) and IRM (Iterated Refinement Method), in the rate-distortion sense. Many kinds of MPEG-4 test sequences were used in the experiments, and, among them, the results of two sequences (Children kids (352×288) and Stefan (352×240)) are presented here.

The vertex of a polygon is selected by using three different vertex selection methods, as described above. For calculating the bit rates, the entropy coding of the relative address of two successive vertices is used, except the first vertex which is encoded by using the fixed length coding of the absolute addressing.



(a) (b)

Fig. 3 Rate-distortion curve (a) Children-kids image (b) Stefan image.

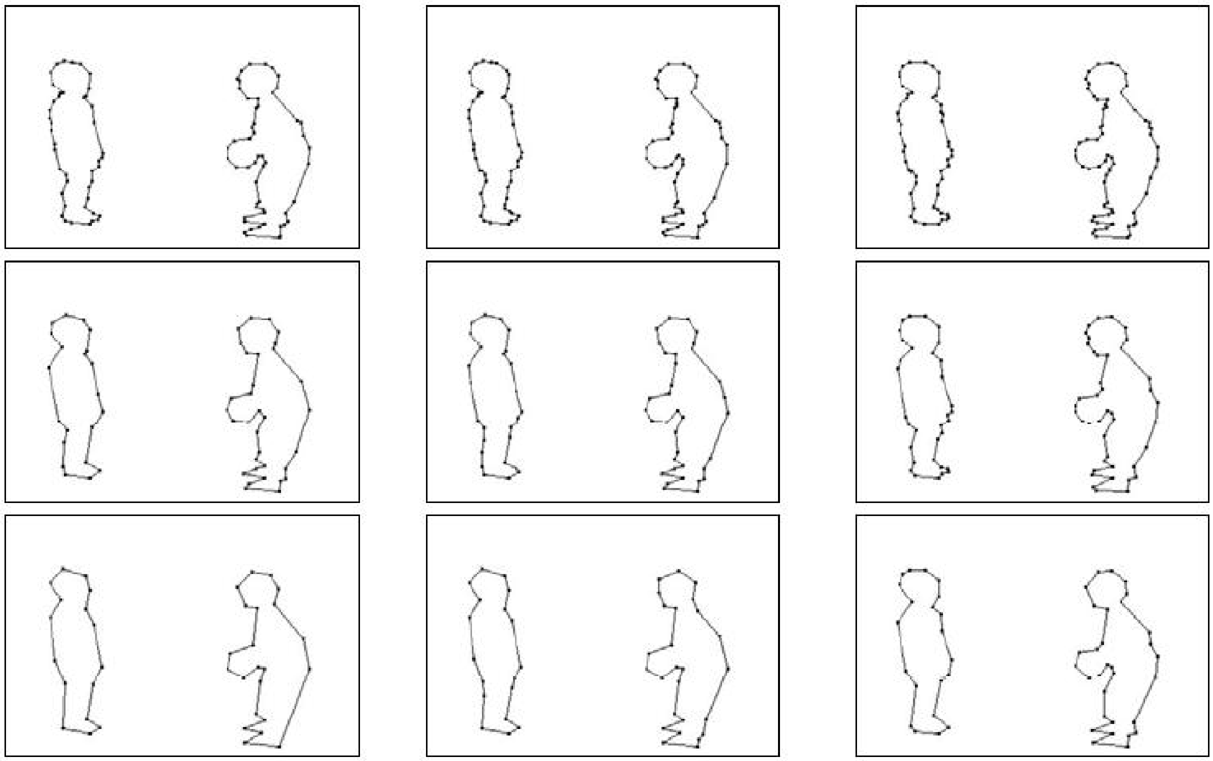
Fig. 3 (a) and (b) show the rate-distortion curves of 'Children-kids' and 'Stefan' sequences, respectively. The proposed method showed outstanding performance compared with the conventional vertex selection methods over all distortion levels.

Fig. 4 shows the reconstructed shape images with vertices of 'Children-Kids' image; (a) is the reconstructed shape image by the proposed method, (b) is the reconstructed one by PVS, and (c) is the reconstructed one by IRM. As shown in the top row of Fig. 4, the subjective qualities of the reconstructed images are almost same in the lower distortion region ($D^*_{\max} = 1.0$). As the $D^*_{\max}$ increases the subjective quality of the reconstructed shape image by the proposed method is almost same with that by PVS and is slightly worse than that by IRM. The proposed method, however, saves bits considerably in the vertex coding step, so can use them in the approximation error encoding by using VQ (vector quantization) or DST (Discrete Sine Transform) to enhance the subjective quality.

V. Conclusions

This paper proposed an efficient method for encoding the shape information of the object in the image. The proposed method adopted conventional PVS as the basic vertex selection scheme and proposed a new vertex selection condition when the maximum distortion is given. To make more efficient polygon in the rate-distortion sense, the proposed method searched the contour up to the peak distance between the polygon and the partial condition is not more than $2 \cdot D^*_{\max}$.

The proposed method gave 5.5 ~ 17.2% bit reduction than PVS and 13.8 ~ 31.3% bit reduction than IRM over test sequences. The subjective quality of the reconstructed shape image, however, is slightly poorer than IRM. But it can be recovered if some of the saved bits are used for the approximation error encoding by VQ or DST. This can provide the function of the bit-stream scalability to the shape information encoder.



(a) (b) (c)

Fig. 4 Comparison of the reconstructed shape of Children-Kids image (a) Proposed method (b) PVS (c) IRM (from top to bottom $D^*_{\max} = 1.0, 2.0, 3.0$).

References

[1] M. Kunt,A. Ikonomopoulos, and M. Kocher, "Second-generation image coding techniques," *Proceedings of the IEEE*, vol. 73, no. 4, pp. 549 – 574, April 1985.

[2] H. G. Musmann, M. Hötter, and J. Ostermann, "Object-Oriented Analysis-Synthesis Coding of Moving Images," *Signal Processing: Image Communication*, vol. 1, pp. 117 – 138, October 1989.

[3] R. Koenen, Ed., "Overview of the MPEG-4 standard," International Standards Organization, Stockholm meeting, ISO/IEC/JTC1/SC29/WG11 N1730, July 1997.

[4] Kevin J. O'Connell, "Object-Adaptive Vertex-Based Shape Coding Method," *IEEE Transactions on Cicuits and Systems for Video Technology*, vol. 7, no. 1, pp. 251 – 255, February 1997.

[5] J. Ostermann, E. S. Jang, J. S. Shin, and T. Chen, "Coding of Arbitrarily Shaped Video Objects in MPEG-4," *Proceedings of International Conference on Image Processing, ICIP97*, vol. 1, pp. 496 – 499, 1997.

[6] A. K. Katsaggelos, L. P. Kondi, F. W. Meier, J. Ostermann, and G. M. Schuster, "MPEG-4 and Rate-Distortion-Based Shape-Coding Techniques," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1126 – 1154, June 1998.

[7] B. J. Yun, S. W. Lee, and S. D. Kim, "Vertex adjustment method using geometric constraint for polygon-based shape coding," *Electronics Letters*, 7th, vol. 37, no. 12, pp. 754 – 755, June 2001.

[8] P. Gerken, "Object-Based Analysis-Synthesis Coding of Image Sequences at Very Low Bit Rates," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.4, pp. 228 – 235, June 1994.

[9] Si-Woong Lee, et al., "Adaptive vertex selection based-on interregion contrast for polygon-based contour approximation," *Optical Engineering*, vol. 43, no. 1, pp. 44 – 50, Jan. 2004.

Byoung-Ju Yun received the B.S. degree in electronics engineering from Kyungpook National University, Daegu, Korea, in 1993 and the M.S. and Ph.D. degrees in electrical engineering and computer science from KAIST (Korea Advanced Institute of Science and Technology), Daejeon, Korea, 1996 and 2002, respectively. From 1996 to May 2003, he has been with Hynix Semiconductor Inc., where he worked as a senior engineer. From June 2003 to February 2005, he has been with the Center for Next Generation Information Technology Kyungpook National University, where he worked as assistant professor. Since June 2003, he has been with the school of Electrical Engineering and Computer Science, where he works as assistant professor.

His current research interests include image processing, image compression, H.264, digital image watermarking, multimedia communication system and HCI.

Jae-Soo Cho received the B.S. degree in electronic and electrical engineering from the Kyungpook National University, Republic of Korea in 1993 and the M.S. and the Ph.D. degrees in electrical and electronic engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 1996 and 2001, respectively. He was a senior researcher of Samsung Electronics, Suwon, Korea from 2001 to 2003. He is currently an assistant professor in the school of Internet-Media Engineering of Korea University of Technology and Education.

His research interests include target tracking, computer vision, pattern recognition, and intelligent information systems such as neural networks and fuzzy logic.

Yun-Ho Ko received the B.S. degree in electronics engineering from Pusan National University, Pusan, Korea in 1995 and the M.S. and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, in 1997 and 2002, respectively. From 2002 to 2004, he has been a senior engineer of MCU Design Team in Hynix Semiconductor Incorporated Company. Since 2004, he has been with the department of mechatronics engineering, Chungnam National University, where he is currently a full-time instructor.

His current research interests are in image processing, computer vision, pattern recognition, image coding, image processing system, and digital watermarking.