Variable Width Admissible Control Point Band for Vertex Based Operational-Rate-Distortion Optimal Shape Coding Algorithms

Ferdous Ahmed Sohel, Laurence S. Dooley, and Gour C. Karmakar

Gippsland School of Information Technology
Monash University, Churchill, Victoria – 3842, Australia.
{Ferdous.Sohel, Laurence.Dooley, Gour.Karmakar}@infotech.monash.edu.au

ABSTRACT

Existing vertex-based operational-rate-distortion (ORD) optimal shape coding algorithms use a *fixed width admissible control point band* (FCB) around the shape boundary as the search space for possible control points. The width of the band however, is fixed and arbitrarily chosen independent of the admissible distortion and shape contour, so it fails to fully exploit the admissible control point band to reduce the bit-rate. This paper proposes a *variable width admissible control point band* (VCB) where the width associated to each boundary point is dynamically determined from the admissible peak distortion and shape information. In addition, the paper uses an accurate distortion measurement method to overcome a key limitation of existing distortion and tolerance band based methods. Experimental results reveal that both the qualitative and quantitative performance of the existing ORD algorithms are improved by seamlessly integrating the VCB and accurate distortion measuring approach.

Keywords: Video coding, image coding, video signal processing.

1. INTRODUCTION

Despite facilitating increasingly effective retrieval, manipulation and interactive editing functionality for both natural and synthetic video sequences, object-oriented video coding using shape information remains a challenging research topic [1]-[6]. The inherent bandwidth limitations of the existing communication technologies mean that applications such as video streaming over the Internet, video-on-demand and mobile video transmission for handheld devices will benefit significantly from efficient shape coding strategies.

The aim of vertex-based operational rate-distortion (ORD) optimal shape coding algorithms [1]-[6], is that for some prescribed distortion, a shape contour is optimally encoded in terms of the number of bits, by selecting the set of control points (CP) that generate the lowest bit rate, and vice versa. In the basic ORD framework, the CP were selected from only boundary points, though this was subsequently relaxed by forming a *fixed-width admissible control point band* (FCB) around the shape contour to reduce the bit-rate [1], with the width of the band being the maximum permitted distortion. While FCB provides a bounded set of potential significant points for polygonal approximations, from a B-spline approximation perspective, it is not necessarily a bounded source of CP as there can be points that lie beyond the band that estimate a shape at a lower bit-rate, for the same peak distortion.

Katsaggelos et al. [1] restricted the peak admissible distortion to one value, which was generalised by dynamically determining the admissible distortion at each shape point from shape contour information [4]-[6]. These algorithms however, still used a FCB so the concept of variable admissible distortion was not fully exploited even in polygonal approximations, to reduce further the bit-rate. Moreover, none of these approaches focused on determining the dynamic width of the admissible control point band within the shape coding framework.

This paper addresses this hiatus by proposing a *variable-width admissible control point band* (VCB), where the width for an individual boundary point is determined from both the maximum admissible distortion and shape contour information.

The ORD framework [4]-[6] employed a *tolerance band* (TB) as its distortion measuring technique, which was in fact a generalisation of the *distortion band* (DB) [1]. In certain cases, the DB and by implication also the TB, have been shown to ignore particular parts of a shape leading to erroneous distortion measures [7], so for this reason in the proposed solution, the *accurate distortion* (AD) measure [7] is employed instead of the TB. Both the subjective and numerical performance of seamlessly embedding VCB into the ORD optimal shape coding framework has been analysed for a wide range of shapes, with superior rate-distortion results achieved with consistently improved distortion accuracy compared with [4]-[5].

The remainder of this paper is organised as follows: Section 2 provides a short overview of the existing vertex based ORD shape coding algorithms, while Section 3 presents the new VCB algorithm together with the bounds on the width of the band and a brief expose on the philosophy behind the AD measuring method. Section 4 analyses the experimental results which confirm the improved rate-distortion performance achieved using both the VCB and AD measure. Some concluding remarks are given in Section 5.

2. OPERATIONAL-RATE-DISTORTION SHAPE CODING ALGORITHMS

Existing ORD shape coding algorithms seek to determine a set of significant points to represent a particular boundary and then encode these rather than the actual boundary points. Let the boundary $B = \{b_0, b_1, \cdots, b_{N_B-1}\}$ be an ordered set of shape points, where N_B is the total number of boundary points and $b_0 = b_{N_B-1}$ for a closed boundary. $P = \{p_0, p_1, \cdots, p_{N_p-1}\}$ is an ordered set of CP used to approximate B. The existing ORD optimal shape coding algorithms then seek to find P from a set of admissible vertices which fall within the FCB using either the DB or the TB as the distortion measure, each of which has some significant limitations, which are now discussed in the following sections.

2.1. Fixed width admissible control point band

In [1]-[3], a single admissible peak distortion ($D_{\max}$) is considered as the width ($W_{\max}$) of the FCB for the whole shape. In this case, the FCB is optimal for polygonal approximations, but may not be so for B-spline based approximations, since a CP lying outside this band can still produce a shape-approximating curve, that maintains the maximum admissible distortion. The example in Figure 1(a) illustrates this scenario with one CP being located outside the band, yet the resulting approximating curve still maintains the admissible peak distortion. The algorithms in [4]-[6] use two admissible peak distortion bounds ($T_{\max}$ and $T_{\min}$) to ensure efficient coding, thereby permitting a lower distortion at sharper boundary edges and a higher value during smoother and gradually changing portions of a shape contour. The admissible distortion for each boundary point ($T[j]$) is determined using a linear mapping between the curvature and distortion bounds. However, a FCB is again used, with for example $W_{\max} = 1pel$ in [4], and is invariant of the admissible distortion value, and as a consequence the FCB is no longer optimal for either polygonal or curve-based approximations. To illustrate this effect, consider the Figure 1(b) example where the encoder requires 4 CP for a polygonal shape approximation using a FCB, while the corresponding VCB solution requires just 3 CP (Figure 1(c)).

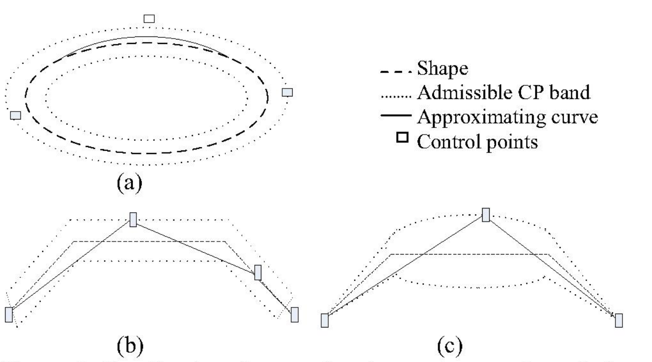


Figure 1: B-spline based approximation – a) approximated shape maintains the admissible distortion though one CP lies outside the FCB. Polygonal approximation – b) 4 CP required for FCB, c) 3 CP required for VCB.

2.2. Distortion and Tolerance Bands

As alluded to in Section 1, the fixed width DB was used in [1] as the distortion measurement technique, with the variable width TB [4]-[6] as its generalisation. The TB operates as follows: if each point on a candidate edge (curve) is a member of the TB, it is considered that the candidate edge (curve) maintains the distortion criteria. However, since the distortion for all boundary points associated with the candidate edge is not considered in the TB, it can lead to an erroneous distortion measurement as illustrated in Figure 2. All the boundary points in this example, with the exception of G, have an admissible peak distortion of 2 *pel,* while G has a distortion of only 1 *pel*. Assuming the radii for these two different admissible peak distortions are also respectively 2 *pel* and 1 *pel*, the TB can be formed. While the piecewise-edge EF of a curve lies entirely inside the TB, it still generates a distortion of 2.83 *pel* at boundary point G rather than the requisite 1 *pel* and even using a sliding window [1] of width 5 *pel*, the TB fails to constrain the admissible distortion for the curve-edge connecting E' and F'. Moreover, this limitation can increase the quantisation noise as the approximating curve points must be quantised in order to fit into the (preferably sub-pel) TB-grid.

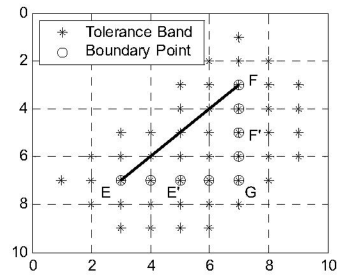


Figure 2: Distortion measure using the TB.

3. VARIABLE WIDTH ADMISSIBLE CONTROL POINT BAND FUNDAMENTALS

This section firstly presents the theoretical basis and implementation process of VCB, before formally developing the bound for the width of the admissible control point band for each individual boundary point. Subsequently, a brief discussion upon the AD measurement technique is provided.

3.1. Variable width admissible control point band

VCB is an ordered set of vertices formed around the shape, with each boundary point has some associated vertices in the band. As each boundary point may have a different VCB width, so the number of VCB points associated with each point will vary. Let $W[j]$ be the width of VCB for the boundary vertex b_j and $W_{\max}$ be the maximum allowable width for any point. The VCB, along with the proper ordering of the vertices is then formed by using Algorithm 1.

Algorithm 1: Variable width admissible control point band (VCB) formation.

Inputs: $B = \{b_0, b_1, \cdots, b_{N_B-1}\}$, $W[j]$ - VCB width around b_j.

Variables: $d(v_0, v_i)$ - Euclidean distance between v_0 and v_i.

Outputs: C -ordered set in VCB, $L[j]$ - number of points in C associated with b_j.

1. $C = B$; $i = j = 1$; $L[j] = 1, \quad 0 \le j \le N_{B-1}$;
2. WHILE $(d(v_0, v_i) \le W_{\max})$
3. FOR $1 \le j \le N_B - 2$
4. IF $(d(v_0, v_i) \le W[j])$ AND $(\{b_j + v_i\} \notin C)$
5. $C = C \cup \{b_j + v_i\}$; $L[j] = L[j] + 1$;
6. Assign vertex $b_j + v_i$ to boundary point b_j.
7. $i = i + 1$

The vector increments v_i used when ordering the admissible vertices are those used in [2].

3.2. Bounds for the width of the admissible control point band

Lemma 1 and Lemma 2 focus on the width of admissible control point band for polygonal and quadratic B-spline based approximations respectively.

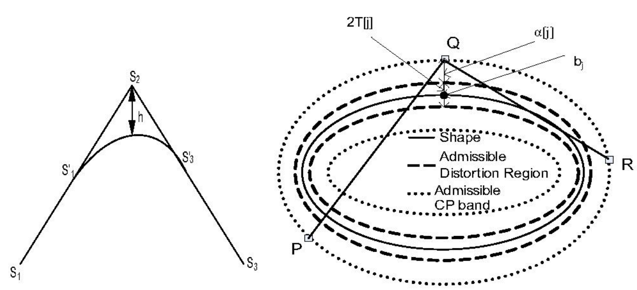
***Lemma 1*:** For a polygonal approximation, the admissible CP band width for any boundary vertex is bounded by the peak admissible distortion for that point, i.e., $W[j] \le T[j]$ for b_j.

Proof (by contradiction): Let there be such a vertex u associated with boundary vertex b_j where $u \in C$ and the distance of u from b_j is greater than $T[j]$, i.e., $W[j] > T[j]$. Assume u is now selected as a CP. Since it is a polygonal approximation, the approximated shape will pass through this vertex (in fact, it will be an end point of one edge), so the distortion at this vertex will exceed the peak admissible distortion for b_j. Thus for b_j, such a u can never be selected as a CP and be within the associated VCB. □

As discussed in Section 2.1, B-spline CPs can lie outside the admissible distortion, yet still provide approximating curves which uphold the distortion criteria. The maximal width of the admissible CP band depends not only upon the admissible distortion, but also the shape. Moreover, the encoding strategy used can limit the size of the range of the distance between consecutive CPs.

Lemma 2: For quadratic B-spline based rate-distortion constrained approximation of a shape, $\alpha[j] \le \min\left\{\frac{3\delta + 4T_{\max} + 2T[j]}{6}, \frac{\rho\sqrt{2}}{4}\right\}$, where δ and ρ are respectively the largest chord length of the boundary and the largest run-length possible for the code used [2], $\alpha[j]$ is the difference between the corresponding admissible distortion and width of the admissible CP band, i.e., $W[j] = \alpha[j] + T[j]$.

Proof: Figure 3(a) shows a uniform quadratic B-spline curve produced by CP $S_1, S_2 \,\&\, S_3$. This is in fact, a Bezier curve (BC) generated by $S_1', S_2 \,\&\, S_3'$, where $S_1' = \frac{1}{2}(S_1 + S_2)$ and $S_3' = \frac{1}{2}(S_2 + S_3)$, with h being the minimum distance of the middle CP S_2 from the BC. It thus follows from [8], that $2h \le \max\left\{\left|S_1'S_2\right|, \left|S_2S_3'\right|\right\}$, where $\left|S_2S_3'\right|$ is the length of edge S_2S_3', so $4h \le \max\left\{\left|S_1S_2\right|, \left|S_2S_3\right|\right\}$.



(a) (b)

Figure 3: a) Distance between a quadratic B-spline curve and its control point, b) Maximal width of admissible control point band calculation.

In the example shown in Figure 3(b), three CPs $P, Q \,\&\, R$ are employed to encode a shape segment that includes the boundary point b_j which has an admissible distortion $T[j]$. Assuming $PQ \ge QR$, the distance of the B-spline curve from Q is always $\le \frac{1}{4}|PQ|$, so the maximum length of PQ is:-

$$\delta + T_{\max} + T_{\max} + \alpha_{\max} + \alpha_{\max} = \delta + 2T_{\max} + 2\alpha_{\max}$$

where $\alpha_{\max}$ is the maximum value of α so $\delta + 2T_{\max} + 2\alpha_{\max} \ge 4\alpha_{\max}$.

Hence $\alpha_{\max} \le \frac{\delta}{2} + T_{\max}$ (1)

The corresponding $\alpha[j]$ for boundary point b_j;

$$4\alpha[j] \le \delta + T_{\max} + \alpha_{\max} + T[j] + \alpha[j] \quad (2)$$

$$\alpha[j] \le \tfrac{1}{6}\left(3\delta + 4T_{\max} + 2T[j]\right) \quad (3)$$

The encoding strategy adopted can limit the length of an edge, since for example, the logarithmic code [1] can support a maximum length of $\rho = 15$, while using a 3-connected chain as the direction encoder, it is able to encode a maximum length of $\rho\sqrt{2}$ (through the diagonal) so that;

$$\alpha[j] \le \frac{\rho\sqrt{2}}{4} \quad (4)$$

From (3) and (4) $\alpha[j] \le \min\left\{\frac{3\delta + 4T_{\max} + 2T[j]}{6}, \frac{\rho\sqrt{2}}{4}\right\}$ (5)

□

3.3. Modification of the accurate distortion measure

In order to overcome the limitations of the TB as discussed in Section 2.2, the edge distortion is measured by using the accurate distortion (AD) measurement technique described in [7], though this was proposed for fixed admissible distortion. Therefore, it is required to be modified in order to be embedded into the variable admissible peak distortion framework. This is achieved in two steps. Firstly, the distortions for all associated boundary points with respect to the candidate edge (curve) are to be determined and secondly, these distortions are checked with the corresponding admissible distortions. This candidate edge (curve) will be further considered only if the distortion criterion for all the boundary points is upheld.

4. RESULTS AND ANALYSIS

To qualitatively and quantitatively analyse the performance of the vertex based ORD shape coding algorithms with the VCB and AD measurement technique embedded, they were both implemented in Matlab 6.1 and applied to a number of natural and synthetically generated arbitrary shapes with all object shapes being manually segmented. The subjective results produced by different techniques upon the *lip* region of the 30th frame of the *Miss America* video sequence for $T_{\max} = 2, T_{\min} = 1pel$ are presented in Figure 4(a)-(f) while the corresponding numerical results for this and other various admissible distortion combinations are summarised in Table 1.

The results in Table 1 reveal that for a polygonal approximation with $T_{\max} = 3, T_{\min} = 1pel$, the original ORD algorithms involving FCB and TB (FCB-TB) required 68 bits to encode the shape while in comparison, the new VCB and AD measure combination (VCB-AD) required only 63 bits. Interestingly, the combination FCB-AD mandated 74 bits, which would in fact be the actual bit rate if the peak admissible distortion was uphold for this shape, since in the TB example, the maximum distortion was *3.2 pel* rather than the prescribed maximum of *3 pel*. This means the TB erroneously ignored some parts of the shape, thereby requiring fewer bits than using AD for the same FCB, though the key feature of the AD measure [7] is that it always ensures the distortion is bounded to the admissible distortion. Clearly when using the same AD measure, VCB always requires fewer bits than FCB because it provides a more efficient and dynamic search space for potential CP, so directly reflecting upon the admissible distortion and enabling the encoder to fully utilise the concept of an admissible CP band to reduce the bit rate. The results in Figure 4(a) provide an example of this distortion inaccuracy for the TB, whereby the maximum distortion value is greater than the admissible distortion (*2 pel)* in the two highlighted regions. Note these inaccuracies do not appear in the corresponding results in Figure 4(c) and (e), where the AD measure is employed.

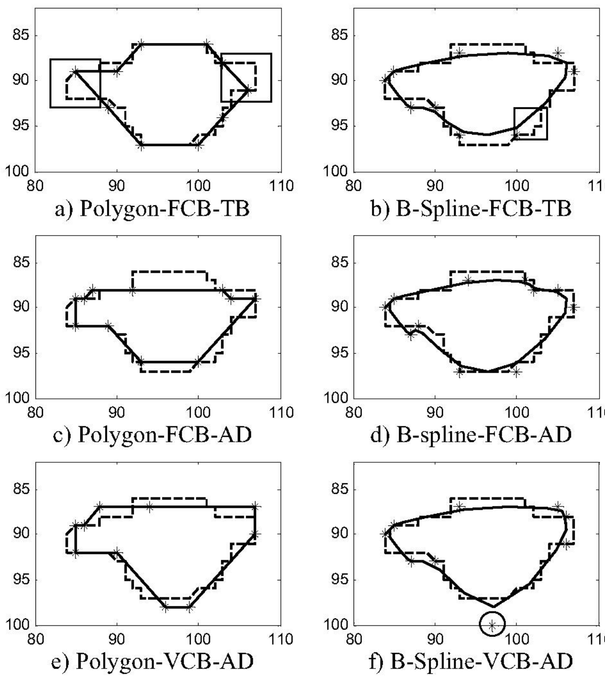


Figure 4: Results for the *lip* region of the 30th frame of the *Miss America* sequence for $T_{\max} = 2, T_{\min} = 1 pel$ (dashed line: shape, solid line: approximation, asterisk: control points).

Table 1: Bit-rates required for various admissible distortion sets (in *pel*) using different band-measure combinations – peak distortion values in parenthesis indicate when the admissible distortion bound has been exceeded.

	Polygonal encoding			B-spline based encoding		
Admissible peak distortion bounds	FCB-TB	FCB-AD	VCB-AD	FCB-TB	FCB-AD	VCB-AD
T_{max}=2, T_{min}= 1	75 (2.8)	79	75	61 (2.2)	62	59
T_{max}=3, T_{min}=1	68 (3.2)	74	63	57(5.0)	61	55
T_{max}=3, T_{min}=2	56 (4.0)	62	55	56(5.0)	57	52

The results for quadratic B-spline based encoding using the VCB always require fewer bits than FCB for the same admissible peak distortion. For example, with $T_{\max} = 3, T_{\min} = 2 pel$ the proposed VCB-AD method required 52 bits compared with 56 and 57 bits respectively for FCB-TB and FCB-AD. The reason for VCB-AD requiring fewer bits is that the variable band affords a larger search space in the selection of control points. For example, the encircled CP in Figure 4(f) is outside the $W_{\max} = 1$ *pel* FCB and approximates a significant portion of the shape so lowering the bit-rate. A similar trend is observed in the distortion results, with existing ORD algorithms (FCB-TB) often failing to maintain the admissible distortion bound. For example, with $T_{\max} = 3, T_{\min} = 1 pel$ the distortion for FCB-TB was *5 pel* which was larger than the maximum admissible distortion of *3 pel*. Moreover, as shown in Figure 4(b), FCB-TB produced a peak distortion of *2.2 pel* compared with the prescribed maximum of *2 pel*. The results in Table 1 also confirm that there was no inaccuracy in distortion measurements when the AD technique was used, a fact endorsed in Figure 4(d) and (f).

It is noteworthy to mention that [1] provided a rigorous review of shape coding techniques, where it is established that the ORD algorithms [1]-[6] are optimal and this particular paper proposes some seamless enhancements to all these algorithms.

5. CONCLUSIONS

This paper has presented a *variable-width admissible control point band* (VCB) used in combination with an accurate distortion (AD) measure to improve the rate-distortion performance of existing vertex-based operational-rate-distortion (ORD) optimal shape coding algorithms. For both polygonal and B-spline based approximations, bounds of the VCB-width have been established along with the procedure of implementation. The AD has been modified for variable admissible distortion. Both the VCB and modified AD have been seamlessly embedded into the ORD optimal shape coding framework to improve the bit-rate performance and overcome inaccuracies in distortion measurement. Both qualitative and quantitative results have endorsed the improvements in the rate distortion performance and accuracy in distortion measurement over the existing shape coding algorithms.

6. REFERENCES

[1] A.K. Katsaggelos, L.P. Kondi, F.W. Meier, J. Ostermann, and G.M. Schuster, "MPEG-4 and rate-distortion-based shape-coding techniques," Proc. IEEE, vol.86, no.6, pp.1126-1154, 1998.

[2] G.M. Schuster and A.K. Katsaggelos, Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding, Kluwer Academic Publishers, 1997.

[3] G.M. Schuster, G. Melnikov, A.K. Katsaggelos, "Operationally optimal vertex-based shape coding," IEEE Sig. Proc. Mag., vol.15, no.6, pp.91-108, 1998.

[4] L.P. Kondi, F.W. Meier, G.M. Schuster, and A.K. Katsaggelos, "Joint optimal object shape estimation and encoding," in Proc. SPIE Conf. Visual Communi. and Image Proc., pp.14-25, 1998.

[5] L.P. Kondi, G. Melnikov, and A.K. Katsaggelos, "Joint optimal object shape estimation and encoding," IEEE Trans. Circuits Syst. Video Technol., vol.14, no.4, pp.528-533, 2004.

[6] S.K. Bandyopadhyay and L.P. Kondi, "Optimal bit allocation for joint contour-based shape coding and shape adaptive texture coding," in Proc. Int. Conf. Image Proc., vol.1, pp.589-592, 2005.

[7] F.A. Sohel, L.S. Dooley, and G.C. Karmakar, "Accurate distortion measurement for generic shape coding," Pattern Recognition Letters, vol.27, no.2, pp.133-142, 2006.

[8] D. Nairn, J. Peters, and D. Lutterkort, "Sharp, quantitative bounds on the distance between a polynomial piece and its Bezier control polygon," Computer Aided Geometric Design, vol.16, no.7, pp.613-631, 1999.