Geometric distortion measurement for shape coding: a contemporary review

Journal Item

Geometric distortion measurement for shape coding: a contemporary review

Ferdous Ahmed Sohel[1]
The University of Western Australia, Australia

Gour Chandra Karmakar[2]
Monash University, Australia

Laurence Sean Dooley[3]
The Open University, United Kingdom

Mohammed Bennamoun[4]
The University of Western Australia, Australia

Abstract
Geometric distortion measurement and the associated metrics involved are integral to the rate-distortion (RD) shape coding framework, with importantly the efficacy of the metrics being strongly influenced by the underlying measurement strategy. This has been the catalyst for many different techniques with this paper presenting a comprehensive review of geometric distortion measurement, the diverse metrics applied and their impact on shape coding. The respective performance of these measuring strategies is analysed from both a RD and complexity perspective, with a recent distortion measurement technique based on arc-length-parameterisation being comparatively evaluated. Some contemporary research challenges are also investigated, including schemes to effectively quantify shape deformation.

Content indicator: Image processing/ coding.

Keywords: Object based video coding, shape coding, geometric distortion measurement.

1. Introduction

Advances in object-oriented video coding using shape information [Aghito and Forchhammer 2004; Aghito and Forchhammer 2006; Brady et al. 1997; Freeman 1961; Richardson 2003] are increasingly facilitating more efficient retrieval, manipulation and interactive editing functionality for both natural and synthetic sequences. The ubiquitous pursuit for greater coding efficiency coupled with the inherent

This work is partially supported by a Monash University Post Publications Award, ARC discovery projects (DP0664228 and DP0771294), the University of Western Australia (UWA) Postdoctoral Fellowship and a UWA Research Development Award.

[1] School of Computer Science and Software Engineering, The University of Western Australia, WA 6009, Australia. Email: Ferdous.Sohel@csse.uwa.edu.au, (Corresponding author)

[2] Gippsland School of Information Technology, Monash University, Victoria – 3842, Australia. Email: Gour.Karmakar@infotech.monash.edu.au

[3] Department of Communication and Systems, The Open University, Milton Keynes, MK7 6AA, United Kingdom. Email: L.S.Dooley@open.ac.uk

[4] School of Computer Science and Software Engineering, The University of Western Australia, WA 6009, Australia. Email: m.bennamoun@csse.uwa.edu.au

bandwidth limitations of existing communication technologies mean a wide range of diverse applications from medical imaging and patient monitoring, video-on-demand and Internet streaming of multimedia content, through to biometric authentication systems, mobile video transmissions for handheld devices and hyperlinked video/television, will all significantly benefit from more effectual shape coding strategies.

As video objects are defined by their shape, as well as texture and motion [Katsaggelos et al. 1998], shape coding has become an integral part of object-oriented video coding. Shape coders have evolved into two distinct classes [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998]; i) *bitmap-based* which encode every pixel within the shape and ii) *contour-based* which focus on just the object shape outline [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998]. Contour-based shape coding can best be illustrated by means of the example in Figure 1.

(a) (b) (c) (d)

Figure 1: Shape coding example – (a) the 30th frame of the *Miss America* sequence, (b) the segmented object shape, (c) the shape contour, and (d) shape coding – the solid line is the reconstructed contour using (lossy) encoded information, while the dotted line is the original contour (c).

Figure 1(a) is a single frame from the *Miss America* video sequence, with the corresponding extracted shape, commonly referred to as the binary alpha-plane, being shown in Figure 1 (b). The resulting shape contour is displayed in Figure 1 (c). The aim of all contour-based shape coding algorithms is that for some prescribed quality (distortion) threshold, a contour can be represented by a lower number of vertices than the original and vice versa. For example, the decoded shape (solid line) in Figure 1(d) requires only 13 vertices compared with the original 297 vertices (dotted line) for a distortion (quality) value of 3.8dB. The review of shape coding techniques in [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] draws the conclusion that the vertex-based polynomial shape coding framework is optimal in an *operational-rate-distortion* (ORD) sense. The term *rate* refers to the number of bits required to encode a shape while the quality of the reconstructed shape is usually measured in terms of geometric distortion. Interestingly, no cognisance is taken of any perceptual shape deformation in this *rate-distortion* (RD) nexus, raising the issue over the importance of *structural consistency* in the reconstruction of decoded shapes. If for example, all circular objects are decoded as squares, even though the relevant distortion criterion is upheld, the corresponding subjective impact could be confusing and even disturbing.

Several distortion metrics have been adopted within existing shape coding frameworks, including *peak signal-to-noise ratio* (PSNR) [Richardson 2003], the MPEG-4 distortion parameter D_n [Aghito and Forchhammer 2004; Aghito and Forchhammer 2006; Brady, Bossen and Murphy 1997; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Kondi et al. 2001] and a gamut of geometric distortion metrics [Bandyopadhyay and Kondi 2005; Chen and Ngan 2004; Hötter 1990; Hötter 1994; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Kondi et al. 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi et al. 2004; Meier et al. 2000; Melnikov et al. 2000; O'Connell 1997; Schuster and Katsaggelos 1997; Schuster et al. 1998; Wang et al. 2003], which each have differing intrinsic qualities. In many cases the foundations of the coding model are based upon a specific distortion metric, such as in the vertex-based shape coding and Polygon/B-spline (BS) frameworks. Moreover, the reconstruction quality of an object shape is highly dependent on the distortion metric employed with the way the distortion is

measured influencing the performance of the metric and by implication, the encoder. Geometric distortion metric has a wide range of application domains including for example, medical imaging for patient monitoring where accurate measurement of the 2D/3D deformation in digital mammograms [Matsubaraa et al. 2005] aids diagnosis and discrimination between cancerous and breast tissue structures, and in automatic damage assessment and structural deformation monitoring [Qiaoa et al. 2007]. To clarify the terminology used in this paper, the *distortion metric* is a parameter that objectively reflects the quality of a measurement, while *distortion measurement* refers to the underlying algorithm used to calculate the distortion component of a distortion metric. In this context, it is clear *distortion measurement techniques* play a crucial role in both distortion metrics and the underlying shape coding process.

Existing geometric distortion measurement techniques include the *shortest absolute distance* (SAD) [Hötter 1990; Hötter 1994; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998], *distortion band* (DB) [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] and *tolerance band* (TB) [Kondi, Meier, Schuster and Katsaggelos 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi, Melnikov and Katsaggelos 2004], *accurate distortion for shape coding* (ADMSC) [Sohel et al. 2006], and *chord-length-parameterisation based distortion measurement for faster encoding* (DMCLP) [Sohel et al. 2007]. This paper aims to provide a contemporary treatise on the performance of these distortion metrics and measurement techniques from a shape coding perspective, allied with an investigation into current research challenges in this field including, ways of characterising *shape deformation* within the encoding paradigm. The paper also examines the recently introduced approach *distortion measurement based on arc-length-parameterisation* (DMALP) [Sohel and Bennamoun 2008], which exhibits enhanced RD performance.

The remainder of this paper is organised as follows: Section 2 presents a brief description of the most popular distortion metrics, while Section 3 provides a comprehensive review of geometric distortion measurement techniques. An experimental results analysis is presented in Section 4, with some future research directions discussed in Section 5. Finally, some concluding remarks are given in Section 6.

2. Distortion metrics

To specify, analyse and contrast various shape coding systems, it is necessary to determine the quality of the resulting decoded shapes. Visual quality is inherently subjective and influenced by many factors that make it difficult to obtain accurate, consistent and repeatable measures of the perceived quality [Richardson 2003]. A viewer's opinion of visual quality is often influenced on factors like their psychophysical state or the nature of the task at hand. Examples include passively watching a movie, keenly watching the closing moments of an exciting sporting event, participating in a video-conference session, or trying to identify either a person or objects in a video surveillance scene. Furthermore, subjective measures are time consuming and manually intensive [Richardson 2003]. Conversely, measuring visual quality with objective criteria does afford accurate and repeatable results at much lower cost, though as the *video quality experts group* (VQEG) [VQEG 1999] reported, there is no unified quantitative measurement system that faithfully reproduces the perceptual experience of a human observer and no single metric that consistently outperforms all other techniques from a subjective viewpoint [Wu and Rao 2006]. As a consequence, various numerical techniques for objective quality measurement have evolved for shape coding, a few of which will now be reviewed.

2.1. Peak signal to noise ratio (PSNR)

This ubiquitous distortion metric is normally expressed on a logarithmic scale by:

$$PSNR_{dB} = 10 \cdot \log_{10} \frac{\left(2^n - 1\right)^2}{MSE} \quad (1)$$

where MSE is the *mean-squared-error* between an original and approximated signal and *n* is the sample size in bits.

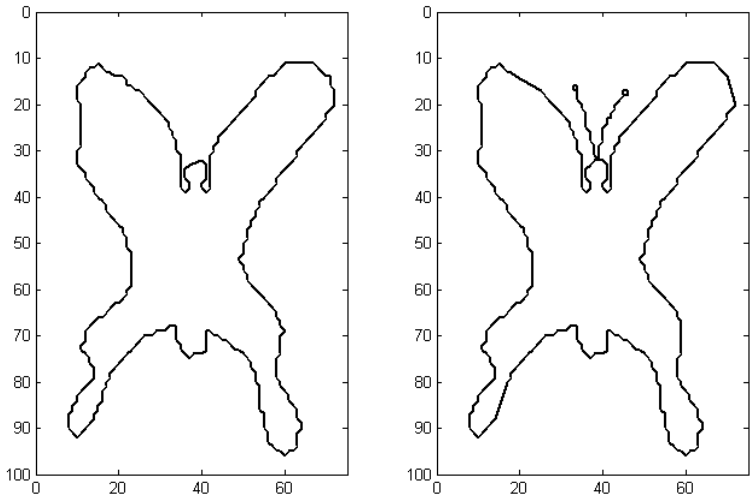
Despite its broad appeal and application, PSNR has not gained popularity in (binary) shape coding [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] primarily because the numerator value in (1) is 1, so it mainly depends on the denominator, which is the *mean-squared* (MS) distortion. PSNR also has some fundamental limitations [Richardson 2003], most notably requiring an unimpaired video as a reference, which may not always be readily available. It is also not easy to verify the original video had perfect fidelity, so it does not necessarily equate to an absolute subjective measure, i.e., a decoded video with higher PSNR value can produce poorer subjective quality than one with lower PSNR. These factors have led to the development of quality measurement techniques more appropriate to the shape coding domain, with two such metrics being extensively adopted.

2.2. The MPEG-4 metric

This metric (D_n) is employed in the MPEG-4 standard and computes the ratio of the number of erroneous pixels in the approximated shape to the total number of pixels in the original to represent the shape distortion [Brady 1999]. It is formally defined as:

$$D_n = \frac{\text{number of pixels mismatched in the approximated shape}}{\text{number of pixels in the original shape}} \quad (2)$$

with D_n usually represented in percentile form [Wang et al. 2005]. The total number of erroneous pixels is the *absolute error* (AE) and this has been used as a quality metric in many applications [Schuster et al. 2004; Soares and Pereira 2004]. D_n provides a broader estimate of the reconstruction quality, and since its inclusion within the MPEG-4 standard, it has become widely applied [Aghito and Forchhammer 2004; Aghito and Forchhammer 2006; Brady, Bossen and Murphy 1997; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Kondi, Melnikov and Katsaggelos 2001].



(a) (b)

Figure 2: *Butterfly* shape with the same $D_n = 0.05\%$ – (a) with the antenna lost and (b) with the antenna preserved.

It needs to be emphasised that as varying shapes will have different ratios of contour pixels to shape pixels, D_n only truly has a physical interpretation when different approximations of the same shape are compared [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998]. The corollary of this is that D_n may not sufficiently represent the actual distortion scenario for all cases. In the two *Butterfly*[5] examples in Figure 2 for instance, though both shape contours have the same $D_n = 0.05\%$, the appearance of the Figure 2(a) object looks subjectively dissimilar since the antenna has been completely lost, while it is fully preserved in Figure 2(b) (note unless otherwise specified, the numerals along the axes in the Figures throughout this paper are the respective Cartesian coordinate values). This implies the former contour will have a higher perceived distortion than the latter shape approximation. In contrast, geometric distortion measurement affords a direct representation of the quality of a shape approximation around the entire contour.

2.3. L_p norms

This family of geometric distortion metrics comprises the L_p norm category [Topiwala 1998], where the error term ε_p is defined as:

$$\varepsilon_p = \frac{1}{N}\left\|x - \tilde{x}\right\|_p^p = \frac{1}{N}\sum_{i=0}^{N-1}\left|x_i - \tilde{x}_i\right|^p \qquad (3)$$

where x and $\tilde{x}$ are the original and approximated shapes respectively, N is the number of pixels in the shape, x_i and $\tilde{x}_i$ the i^{th} pixels of the original and approximated shapes respectively, and $\left|x_i - \tilde{x}_i\right|$ is the distance between them based upon the measurement criteria. Various weightings derived from the L_p norm definition in (3) can be used as quality measures. For example, *p=2* refers to the MS distortion,

[5] IMSI's Master Photo Collection, 1895 Francisco Blvd. East, San Rafael, CA 94901-5506, USA.

while the L_∞ ($p=\infty$) and L_1 (*p=1*) norms correspond to the *peak absolute distortion* ($D_{\max}$) and *sum-of-distortion-magnitudes* (*average distortion-magnitudes*) respectively. The peak distortion is formally defined as:

$$D_{\max} = \max_{i=0,1,\cdots,N-1}\left\{\left|x_i - \tilde{x}_i\right|\right\} \tag{4}$$

while the MS distortion is given by:

$$D_{MS} = \frac{1}{N}\sum_{i=0}^{N-1}\left|x_i - \tilde{x}_i\right|^2 \tag{5}$$

As this pair of metrics provide the approximated shape quality around the whole contour, so efficiently reflecting dominant geometric features like sharp edges and corners, it has been broadly embraced by the shape coding research community [Bandyopadhyay and Kondi 2005; Chen and Ngan 2004; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Kondi, Meier, Schuster and Katsaggelos 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi, Melnikov and Katsaggelos 2004; Meier, Schuster and Katsaggelos 2000; Schuster and Katsaggelos 1997; Sohel et al. 2007; Wang, Schuster and Katsaggelos 2005].

While the PSNR and D_n calculations in (1) and (2) are both straightforward, differing approaches have been employed to compute the geometric distortion metrics. Their efficacy greatly depends on the distortion measurement techniques, so in the following section an overview of the most popular geometric shape distortion measurement techniques will be elucidated.

3. Geometric distortion measurement techniques

The simplest and most widely employed measurement technique is the SAD [Hötter 1990; Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Koplowitz 1981; Schuster and Katsaggelos 1997; Sohel, Dooley and Karmakar 2006], which calculates the shortest absolute distance between the original and approximating shape contours. Other approaches that have been applied within the vertex-based polynomial shape coding framework, include the DB [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] and TB [Kondi, Meier, Schuster and Katsaggelos 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi, Melnikov and Katsaggelos 2004] methods, while more recently two intuitive measuring techniques in [Sohel, Dooley and Karmakar 2006] and [Sohel, Karmakar and Dooley 2007] have been proposed, which respectively focus upon the accuracy and computational speed of the underlying distortion measurement process. A detailed delineation of these techniques follows.

3.1. The Shortest Absolute Distance (SAD)

In this technique the distortion at a contour point is defined as the perpendicular distance of that point from the corresponding edge of the approximating polygon. The SAD at an arbitrary contour point b_t with respect to an approximating polygon edge with endpoints s_{k-1} and s_k is given by:

$$d\left(s_{k-1}, s_k, b_t\right) = \frac{abs\left(\left(b_{t,x} - s_{k-1,x}\right)\left(s_{k,y} - s_{k-1,y}\right) - \left(b_{t,y} - s_{k-1,y}\right)\left(s_{k,x} - s_{k-1,x}\right)\right)}{\sqrt{\left(s_{k,x} - s_{k-1,x}\right)^2 + \left(s_{k,y} - s_{k-1,y}\right)^2}} \quad (6)$$

where *x* and *y* are the corresponding Cartesian coordinate values and $abs(\cdot)$ is the *absolute value* function. SAD calculates the distance from either the edge or its extensions, so in the Figure 3 example, segments *GK*, *IM* and *HJ* correspond to the respective SAD of line *EF* from contour points *G*, *I* and *H*.

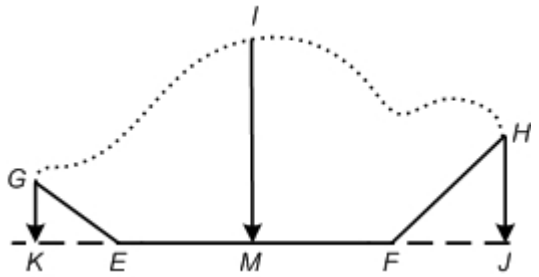
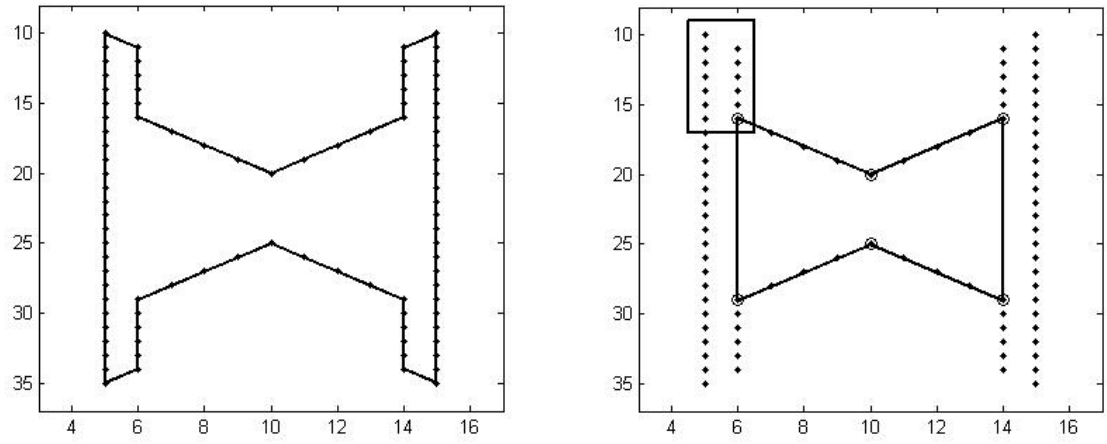


Figure 3: Illustration of the SAD and its limitations when the distance is measured from an extended line.

The SAD measure is attractive due to its computational simplicity and minimal number of arithmetic operations [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Schuster and Katsaggelos 1997; Sohel, Dooley and Karmakar 2006], though it may not always accurately reflect the true distortion since the perpendicular distance does not necessarily correspond to the actual distance for all contour points [Sohel, Dooley and Karmakar 2006]. For instance, when the SAD from an extended part of a line segment is considered to be the minimum distance between a contour point and the line, it clearly fails to represent the actual distance. In these circumstances, the actual distance is measured from the contour point and the closer endpoint of that line segment which leads to discrepancies between the calculated and perceived distortion values, especially for contours with sharp edges and corners. The reason of this anomaly is illustrated in Figure 3. Using SAD, segments *GK*, *IM* and *HJ* correspond respectively to the shortest distances of line *EF* from contour points *G*, *I* and *H*. Of these three distances, both *GK* and *HJ* are measured from the extended lines *EK* and *FJ* respectively, though from a perceptual viewpoint this misinterprets the actual distortion, since *GE* is the distance of *G* from *EF* rather than *GK* as implied by (6). A similar observation applies to point *H*, which leads to the conclusion that as *GK* < *GE*, SAD-based ORD algorithms fail to take cognisance of the perceptual distance by excluding this significant distortion component. This is especially relevant when the shape contour has distinct geometric features such as the sharp edges and corners in Figure 4.

Figures 4(a) and (b) respectively show an arbitrarily shaped object and its corresponding decoded contour using SAD-based ORD optimal algorithm for a peak distortion (4) of $D_{\max} = 1$ *pixel*. The decoded contour has a peak distortion of 6.08 *pixels* at each of the four extrema object points, despite the distortion supposedly being bounded by $D_{\max} = 1$ *pixel*. This is because at each corner point, the distortion has been measured by SAD with a value of 1 *pixel* from the extended line segments and as a consequence, has

generated an inaccurate measurement. Another shortcoming of the SAD technique is the computational speed incurred for BS-based encoding.



(a) (b)

Figure 4: (a) An arbitrary shaped object and (b) approximated contour (solid line) from the encoded data by the basic vertex-based ORD optimal shape coding framework using SAD with $D_{\max} = 1$ *pixel.*

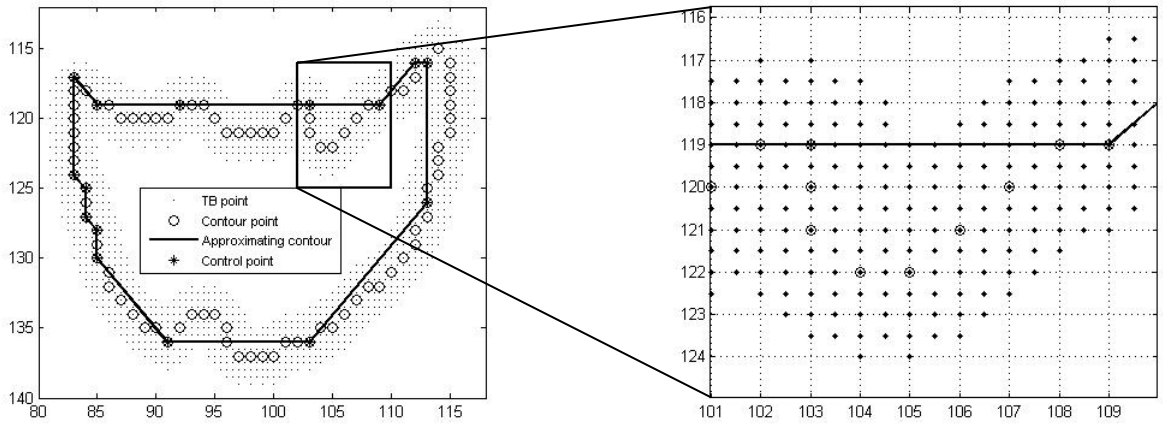
Computational complexity: If there are N_B points along the contour segment where the distortion is measured, SAD incurs $O(N_B)$ time when the approximated contour is represented by a single polygon edge. This is because the edge-distortion for all associated contour points is calculated from the candidate edge and checked against the corresponding admissible distortion. BS-based encoding in contrast requires $O(N_B^2)$ time to monitor the distortion, because a BS curve is in fact a concatenation of piecewise polygon-edges and so for each contour point associated with a candidate curve, the individual distortion has to be measured from all edges forming the approximating curve. The minimum edge-distortion value is then assumed as the distortion for that particular contour point and compared against the corresponding admissible distortion value.

3.2. The Distortion and Tolerance Bands (DB and TB)

The DB technique only considers the peak (L_∞ norm in (4)) admissible distortion within the vertex-based ORD optimal polynomial shape coding framework [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998]. A band of width equal to the fixed admissible distortion $(D_{\max})$ is drawn around the original contour, so it is then only required to detect whether either a candidate approximating polygon-edge or BS curve resides completely inside the band. The DB was in fact originally designed to support fixed admissible peak distortions [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998; Schuster and Katsaggelos 1997], with the philosophy being successfully extended in [Bandyopadhyay and Kondi 2005; Kondi, Meier, Schuster and Katsaggelos 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi, Melnikov and Katsaggelos 2004] to support variable admissible distortion based upon either the image intensity gradient of the object around the contour or the shape-curvature at the contour points [Kondi, Melnikov and Katsaggelos 2001]. This inherently affords the freedom to give greater emphasis to higher image gradient (or alternatively curvature) parts of a contour and *vise versa*. To support the notion of variable admissible distortion, the concept of DB was extended to the TB [Kondi, Meier, Schuster and Katsaggelos 1998]. TB-based models use two admissible peak distortion bounds ($T_{\max}$ and $T_{\min}$) for efficient coding

with an admissible distortion $T[j]$ at each individual contour point b_j being determined using a linear mapping between the image-intensity gradient (curvature) and admissible distortion bounds. The TB operates according to the following three steps [Kondi, Meier, Schuster and Katsaggelos 1998; Kondi, Melnikov and Katsaggelos 2001; Kondi, Melnikov and Katsaggelos 2004]: i) draw a circle around each contour point b_j of radius $T[j]$, so the TB consists of the set of all points that lie inside the circles; ii) Convert the candidate polygon-edge or BS curve into points compatible with the TB-grid, which preferably is a sub-pixel grid in order to provide high precision and accuracy; and iii) Check the distortion and if all points on a candidate polygon-edge (or BS curve) lie inside TB, it is considered the candidate edge (curve) upholds the requisite distortion criteria.

Generally, the DB and TB perform well in the classical ORD optimal shape coding framework and as the creation of the TB is performed outside the main computing loops, it can be efficiently used in the core of the ORD algorithms at modest computational cost. The TB checking process ensures every point on the approximating curve lies within the admissible distortion bound and while this is a necessary condition for distortion maintenance, it alone is crucially not sufficient. This is because the distortion of all contour points associated with a candidate edge are not individually considered, so there is the risk of some points lying beyond the admissible distortion from the approximated contour, as illustrated in Figure 5.



(a) (b)

Figure 5: Distortion measurement using the tolerance band – (a) the TB technique and (b) a magnified version of the region indicated by the rectangle in (a).

Figure 5(a) shows the TB along with all its points (the half-pixel TB grid provides quarter-pixel accuracy) for the *Neck* region of the 31st frame of the *Miss America* video sequence with $T_{\max} = 2$ and $T_{\min} = 2$ *pixels* (i.e., $D_{\max} = 2$ *pixels*). Figure 5(b) displays the zoom-in portion of Figure 5(a) indicated by the rectangle, which reveals that despite the entire decoded contour lying inside the TB, two contour points with Cartesian coordinates $(104,122)$ and $(105,122)$ generate a peak distortion of 3 *pixels*.

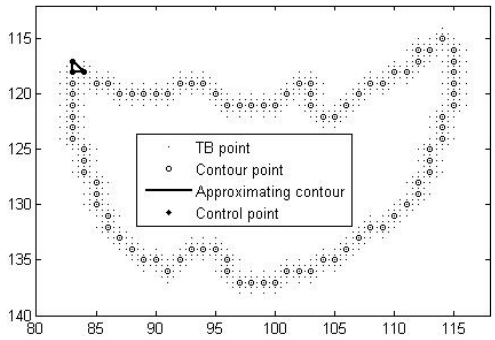


Figure 6: Example to illustrate trivial solution problem of TB/DB.

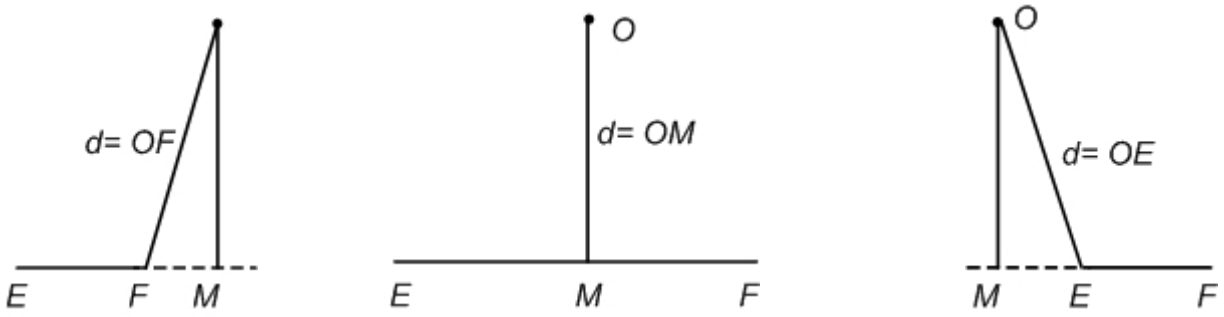
Vigilance is required whenever applying both TB/DB to avoid *trivial solutions* [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998], whereby the encoder only codes a small portion of the contour, whilst appearing to have accurately encoded the entire contour. By definition, the framework attempts to select a set of *control points* (CP), between the first and last contour points so the reconstructed contour lies entirely inside the DB/TB. For a closed contour, the first and last contour points will obviously be coincident, which leads to the possibility of trivial solutions. Figure 6 provides an example of one such trivial solution for $T_{\max} = T_{\min} = 1$ *pixel*. This shows the approximating contour is only located between contour points $(83,117)$, $(84,118)$, $(83,118)$ and $(83,117)$, so though the decoded contour upholds the above definition, it actually only encodes a very small part of the contour. This situation can, to some extent, be resolved by using a *sliding window* (SW) [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998]. For example, the sample result in Figure 5(a) was obtained using a SW-length of 15 *pixels* and yet it produced erroneous distortions. A number of techniques to calculate the most appropriate SW-length are proposed in [Sohel, Dooley and Karmakar 2007; Sohel et al. 2006]. In addition, the TB can lead to increased quantisation errors as the approximating curve points must firstly be quantised to fit into the TB-grid, with the precision level of the TB being highly dependent on the TB-grid unit size. This said, from a performance perspective the only difference between DB and TB is that the former supports fixed admissible peak distortions, while the latter supports variable distortions, so without loss of generality, when $T_{\max} = T_{\min}$, DB and TB are identical.

Computational complexity: From the TB definition, every point on the candidate edge (curve) must be checked to see whether it belongs to the TB points set. The number of points compatible with the TB-grid on a candidate edge (curve) is $O(N_B)$, while the TB itself comprises a set of $O(N_B)$ points, so the full checking process for any candidate curve necessitates $O(N_B^2)$ time in the worst case for both polygon and BS-based encoding. To ensure high accuracy, the TB supports sub-pixel grids which increase the number of points, so the complexity is better expressed as $O(\psi \cdot N_B^2)$ for both approximations, where ψ is the maximum number of TB points associated with a contour point. This means with $T_{\max} = 1$ *pixel*, for grid-sizes of 1, $\frac{1}{2}$, $\frac{1}{3}$ and $\frac{1}{4}$ *pixels* the corresponding values of ψ are 5, 13, 29 and 49. An alternative interpretation is these ψ values represent $T_{\max} = 1, 2, 3$ and 4 *pixels* respectively on a grid-size of 1 *pixel*.

Despite their wide ORD adoption, both SAD and TB/DB techniques are unable to ensure the accuracy of the geometric distortion measurement process in certain shape scenarios. This provided the impetus for the development of the ADMSC algorithm which guarantees accurate distortion measurement.

3.3. Accurate distortion measurement for shape coding (ADMSC)

ADMSC was formulated within the SAD framework to address its main limitation, when the distance is measured from an extended part of the candidate-edge rather than the edge itself. From a distortion measurement perspective, any contour point can lie in one of three possible positions relative to a polygon edge. The contour point can firstly be perpendicularly connected by a line directly onto the polygon edge or onto either of its extended parts. Examples illustrating all three relative positions are given in Figure 7 for contour point *O* with respect to polygon edge *EF*. In Figure 7(a), since the perpendicular line from *O* directly intersects *EF* at *M*, *OM* is the minimum distance, while in Figure 7(b), the line intersects the extended *EF* at *M* and hence is closer to *E*, so *OE* is the minimum distance. Similar reasoning for Figure 7(c) gives *OF* as the minimum distance. SAD in contrast measures this shortest distance as *OM* in all three cases, so it can be concluded that the different relative positions of a point with respect to an edge will lead to three different distortion measurements. To resolve this inaccuracy, it is firstly necessary to check the relative positions before actually calculating the distortion, though this monitoring process



(a) (b) (c)

Figure 7: Relative positions of contour point *O* with respect to polygon edge *EF* – the perpendicular line from the point intersects (a) the line itself, (b) and (c) the extended line.

increases the overall computational overhead. ADMSC importantly provides a unified distortion measure that manages all three position scenarios without recourse to any checking and so guarantees an accurate measure of distortion. The ADMSC shortest distance *d* of point *O* from edge *EF* is formally expressed as:

$$d^2(E,F,O) = |OM|^2 + \tfrac{1}{4}\left(|EM| + |FM| - |EF|\right)^2 \qquad (7)$$

where $|\cdot|$ represents the absolute distance between two endpoints.

To demonstrate the performance of ADMSC, Figure 8 shows the approximated arbitrary shape in Figure 4(a) for $D_{\max} = 1$ *pixel*. It is visually apparent from the results in Figures 4(a) and 7 that ADMSC has accurately measured every distortion compared with SAD, which has failed to correctly measure the distortion at the four corners. This demonstrates the advantage of ADMSC in accurately measuring all distortions and so generates a shape reconstruction more congruent with perceptual distortion. It also implies the notion of *shape deformation* needs to be considered alongside geometric distortion (see Section 5).

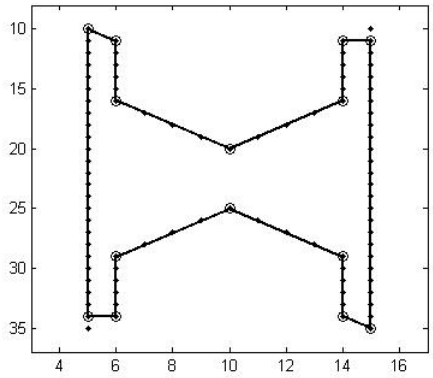
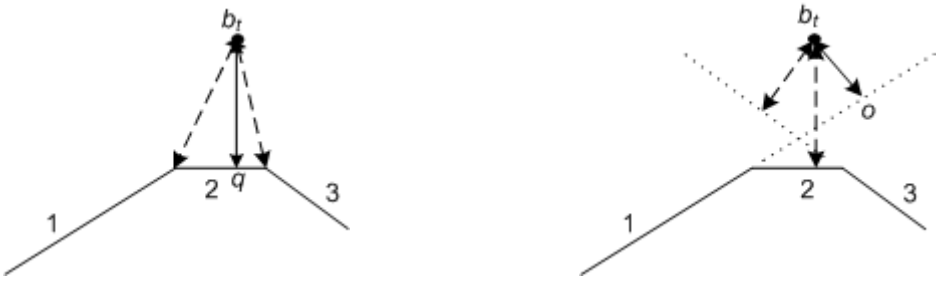


Figure 8: Approximating contour from the encoded data by vertex-based ORD optimal shape coding framework using ADMSC for $D_{\max} = 1$ *pixel.*

Computational complexity: The overall order of computational complexity for ADMSC is the same as SAD, namely $O(N_B)$ for polygon-based encoding and $O(N_B^2)$ for BS encoding, though as detailed in [Sohel, Dooley and Karmakar 2006], ADMSC does incur a slightly higher computational time due to the extra operations performed in solving the distortion measurement limitation in SAD-based calculations.

Operational dissimilarity between ADMSC and SAD: Given the synergistic genesis of the ADMSC and SAD measurement strategies, Figure 9 provides a graphical illustration of the subtle differences between the two techniques for contour point b_t within a BS context, where the BS curve is represented as the concatenation of piecewise polygon-edges 1, 2 and 3. Figure 9(a) shows the distances from these edges using ADMSC, with $|b_t q|$ being the minimum (where $|\cdot|$ is the Euclidean distance). This represents the final distortion at b_t, while in Figure 9(b), SAD measures the distortions from either the edges or their extensions so the corresponding final (minimum) distortion is $|b_t o|$. In comparing these measurements, ADMSC has palpably calculated the accurate distortion whilst SAD has produced a much lower distortion than the actual value, which will in certain cases be erroneous. This measured distortion is compared with the admissible distortion to make certain the candidate BS-curve upholds this value.



(a) (b)

Figure 9: An illustration of the difference in distortion measures using – (a) ADMSC and (b) SAD for a BS-based framework.

While ADMSC successfully solves the problem of guaranteeing consistently accurate perceptual distortion measurement within the ORD shape coding framework, the complexity impost is still a significant overhead. From the aforementioned discussions, ADMSC incurs $O(N_B)$ time for polygon-

based shape coding, but as with SAD it takes $O\left(N_B^2\right)$ for BS-based coding. In contrast, TB/DB mandates $O\left(N_B^2\right)$ time for both polygon and BS-based encoding. With the distortion measurement being embedded within the kernel of all the various shape coding algorithms, it is vital to investigate faster measurement approaches to ameliorate the high computational complexity. It was in this context that the DMCLP strategy was designed, which incurs $O\left(N_B\right)$ time for both polygonal and BS encoding.

3.4. Fast distortion measurement technique using chord-length-parameterisation (DMCLP)

The philosophy behind this measurement algorithm is that if there is an associated approximation point for each contour point, the distortion can be measured as the Euclidean distance between these two points. To obtain the corresponding approximated point, a BS parametric representation is used. Since a BS curve Q_k for a *control point* (CP) set $\{s_{k-1}, s_k, s_{k+1}\}$ is defined by the control parameter *u*, for each value of *u*, an approximated point is generated. As every contour point has an associated *u* value, computing the distortion simply becomes a point-to-point distance calculation rather than finding the minimum of a number of edge-distortions as in SAD and ADMSC, which expedites the distortion calculation process. If the distortion between contour points and their corresponding curve points is less than or equal to the admissible distortion of the respective contour points, the curve upholds the distortion bound and is considered a candidate curve segment within the RD optimisation process. To determine this *u*, chord-length parameterisation, which is widely used in the development of parametric curve algorithms [Farin 1997], is applied to construct a smooth curve. For an arbitrary curve segment having start and end indices *i* and *j* respectively of the associated points in contour $B = \{b_0, b_1, \cdots, b_{N_B-1}\}$, the *u* values are determined from:

$$u_t = \begin{cases} 0, & \text{if } t = i \\ \dfrac{\left|b_i b_{i+1}\right| + \left|b_{i+1} b_{i+2}\right| + \cdots + \left|b_{t-1} b_t\right|}{\left|b_i b_{i+1}\right| + \left|b_{i+1} b_{i+2}\right| + \cdots + \left|b_{j-1} b_j\right|} & \text{otherwise} \end{cases} \qquad (8)$$

where u_t is associated with contour point b_t and $\left|b_{t-1} b_t\right|$ is the Euclidean distance. Once u_t is obtained, a BS point corresponding to b_t can be located from:

$$Q_k(s_{k-1}, s_k, s_{k+1}, u_t) = \begin{bmatrix} {u_t}^2 & u_t & 1 \end{bmatrix} \cdot \begin{bmatrix} 0.5 & -1.0 & 0.5 \\ -1.0 & 1.0 & 0.0 \\ 0.5 & 0.5 & 0.0 \end{bmatrix} \cdot \begin{bmatrix} s_{k-1} \\ s_k \\ s_{k+1} \end{bmatrix} \qquad (9)$$

When a BS curve point is generated using u_t, the peak distortion at b_t is the Euclidean distance between these two points, so it is then only required to ensure the contour point upholds the admissible distortion. Similarly for polygon-based encoding, a corresponding approximating point for each contour point can be obtained from:

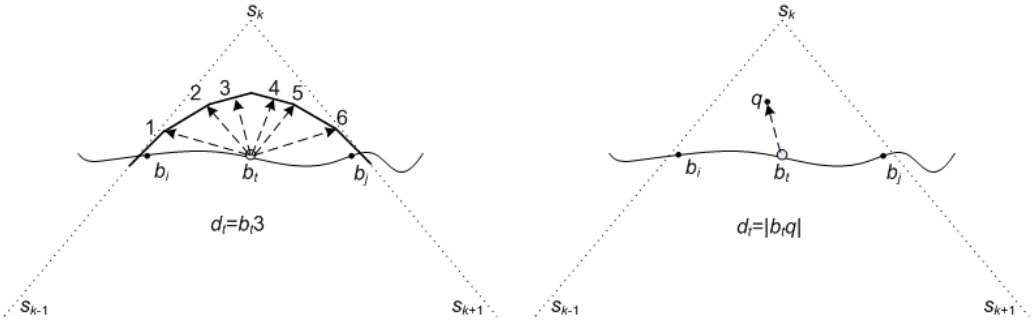
$$q = u_t \cdot s_k + (1 - u_t) \cdot s_{k-1} \qquad (10)$$

which is subsequently used in the distortion calculation.

Computationally DMCLP is faster than both SAD and ADMSC for BS-based encoding, while its time complexity is equivalent to SAD and ADMSC for polygon-based coding [Sohel, Karmakar and Dooley 2007]. It however provides a more relaxed measure of the true distortion by virtue of relying on only the distance between an approximated point and its associated contour point. This can lead to over-estimation of the actual distortion with extra bits being incurred to encode a contour and an ensuing impact on RD

performance. DMCLP's accuracy is still superior to both SAD and DB/TB since both these strategies ignore certain distortion components (see Sections 3.1 and 3.2) and so do not always correctly uphold the peak distortion. Conversely, DMCLP is bounded within the admissible distortion limit.

Operational dissimilarity between ADMSC and DMCLP: ADMSC measures the distortion for a particular contour point using edge-distortion criteria both in polygon and BS-based coding. For polygon-based coding this is obvious while for BS encoding, it measures the edge-distortion from the piecewise edges that form the BS curve. Conversely for both polygon and BS cases, DMCLP employs point-to-point Euclidean distance calculations, with an approximating point corresponding to each contour point generated, with the distance between these points considered as the distortion.



(a) (b)

Figure 10: Example illustrating the key difference between – (a) ADMSC and (b) DMCLP distortion metrics for a BS-based framework.

Figure 10 illustrates the central difference between the DMCLP and ADMSC measuring techniques in a BS-based framework for the CP set $\{s_{k-1}, s_k, s_{k+1}\}$. In Figure 10(a), the BS curve generates a series of piecewise edges (1, 2, 3, 4, 5, 6), with the shortest distance of these polygons from the contour point b_t being determined and the minimum amongst these ($b_t 3$ in the example) designated as the final distortion. In contrast for DMCLP, each associated contour point has its own *u* value, so there is no need to generate the complete BS curve (only the corresponding BS point *q*), so the distortion for this measurement technique is $|b_t q|$ (see Figure 10(b)).

DMCLP employs chord-length parameterisation to determine a *u* value for every contour point, which is subsequently used to obtain the approximating BS (or polygon) point for that contour point. Crucially the location of these approximating points depends not only on *u*, but also on the distance between consecutive CP, with the concentration of these points being greater in the area close to the smaller control polygon-edge than the larger edge. In Figure 11 for instance, the density of BS points is higher in the area close to edge $s_k s_{k+1}$ than in the vicinity of edge $s_k s_{k-1}$. The 17th point of both the contour and approximating BS-curve has been encircled in Figure 11, from which it is evident the distortion at this particular contour point with respect to the entire approximating BS curve, is much lower than the point-to-point distance determined by DMCLP. This demonstrates the intrinsically relaxed nature of the DMCLP measure in the sense that it does not calculate the minimum distortion, so despite producing a smaller distortion this particular CP set will be rejected leading to a higher bit-rate requirement. To address this restriction, the next section presents a recently developed distortion metric that uses *arc-length parameterisation* (ALP).

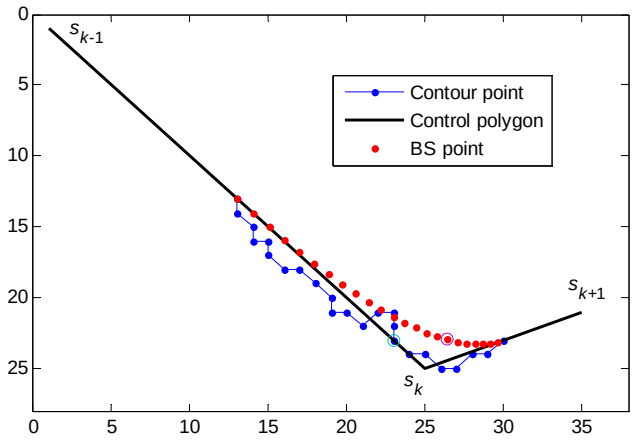


Figure 11: Example showing the measurement problem with the DMCLP.

3.5. Distortion measurement technique using arc-length-parameterisation (DMALP)

The rationale of ALP is to define a set of weights for the parametric curve coefficients so the generated curve points obtain *unit speed*, that is, the distances between two consecutive curve points are equal [Farouki 1997]. If the number of BS points, i.e., the number of u steps equals the number of contour points, then a BS point lies within the vicinity of the corresponding contour point. The ALP process involves two key steps. In the first, the value of u is determined using a modified CLP. The approach formalised in (8) assumes u monotonically increases with t though in many cases, while the contour point index is incremented, there is no commensurate advancement of the position of the contour point with respect to the *baseline* (the line joining the two end CPs). The corollary is that while the contour point has not moved the corresponding BS point has, leading to an over-estimation of the true distortion. This scenario is evinced in Figure 12, where the 3rd contour point and its corresponding BS point appear very close. For the next few contour points there is little advancement in the direction of baseline, while the BS points have moved away so that for instance, for the 6th contour point and its corresponding BS point (both encircled), the distance between them has now become significant.

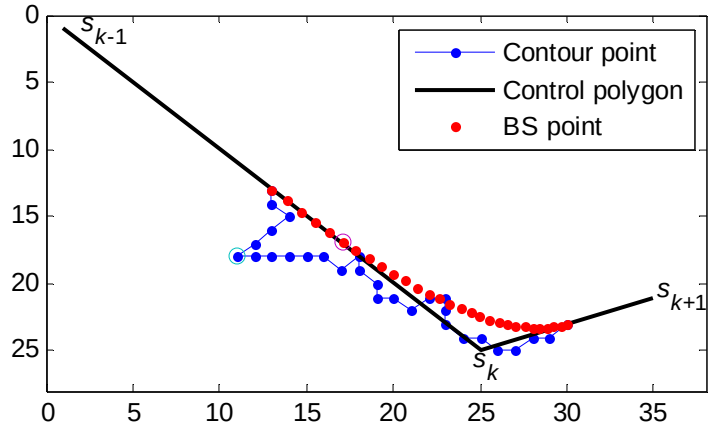


Figure 12: Illustration of the monotonic incremental problem associated with DMCLP.

To resolve this inadequacy it is essential to consider the projection of the distance between b_{t-1}, b_t on the baseline, as an alternative to the geometric distance. If β is the angle between the baseline and edge b_{t-1}, b_t, then this projection will be $distance(b_{t-1}b_t) \times \cos\beta$. Since a 1-*pixel* neighbourhood (see Figure 13) is employed for the contour chain, if the angle between the baseline and the global coordinate is translated to the same pixel grid, then depending on β, the value of $|b_{t-1}b_t|$ will be either 0,1 or -1. As illustrated in Figure 13, if $\beta = \pm \pi/4$ then the distance is $\sqrt{2}$ and $|b_{t-1}b_t| = 1$. If the angle is either $\pi/2$ or $3\pi/2$ then $|b_{t-1}b_t| = 0$, while for either $\beta = 135^0$ or 225^0, $|b_{t-1}b_t| = -1$.

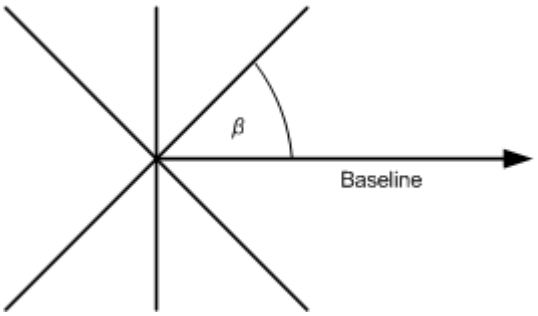


Figure 13: Illustration of the value of $|b_{t-1}b_t|$.

In the second module, ALP is performed to obtain the weights w so that an equal distance between the CP is achieved. As detailed in [Farouki 1997], for a quadratic curve, ALP can be obtained as follows:

$$Q_1 = \left(|\Delta s_{k-1}|^2 + 6 \cdot |\Delta s_k|^2 + 3 \cdot |\Delta s_{k-1}| \cdot |\Delta s_k|\right) \text{ and}$$
$$-Q_0 = \left(6 \cdot |\Delta s_{k-1}|^2 + |\Delta s_k|^2 + 3 \cdot |\Delta s_{k-1}| \cdot |\Delta s_k|\right) \quad (11)$$

where $|\Delta s_k| = s_{k+1} - s_k$.

$$\text{So } \alpha_{\min} = \left(1 + \sqrt{-Q_1 / Q_0}\right)^{-1} \quad (12)$$

Hence, the optimal parameterisation coefficients are

$$w_j = (1 - \alpha_{\min})^j \cdot (\alpha_{\min})^{2-j} \text{ for } j = 0,1,2. \quad (13)$$

The combination of these weights and respective value of *u*, is employed to calculate the approximating point corresponding to a contour point. The distortion is then calculated using the point-to-point distance between them.

Computational complexity: The overall computational complexity for DMALP in the worst case is $O(N_B)$, which is exactly the same as DMCLP, and one degree lower than all other existing distortion measurement techniques for BS-based encoding. The calculation of the reparameterisation coefficients does mean however, the time complexity of DMALP is slightly higher than DMCLP.

4. Results and analysis

All experiments were implemented in Matlab on a 2.8 GHz (Giga Hertz) Pentium-4 processor, with 512 Megabytes of random access memory under a Windows XP operating system and applied to a number of natural and synthetically generated shapes and standard video test sequences having various spatial and temporal resolutions. The specifications of the different standard test sequences used in this paper are summarised in Table 1.

Table 1: Test sequence specifications

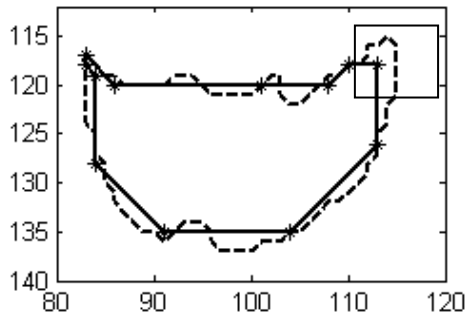
Video sequence	Format	Spatial resolution (*pixels)*	Number of frames
MissAmerica.qcif	QCIF	176 × 144	100
Akiyo.qcif	QCIF	176 × 144	300
Bream.qcif	QCIF	176 × 144	300
Kids.sif	SIF	352 × 240	100
Stefan.sif	SIF	352 × 240	450
Kids.sdtv	SDTV	720 × 486	300
Stefan.sdtv	SDTV	720 × 480	300

To assess the comparative performance of the different geometric distortion measurement techniques, the vertex-based ORD optimal shape coding framework [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] was employed as the test bed. The reasons for this were twofold: firstly it has been proven ORD is optimal and secondly, it provides an ideal application sphere for geometric distortion. For presentational clarity, this section will concentrate on analysing the results for the *Neck* region of the 31st frame of the *MissAmerica*.qcif sequence, with results for the other sequences being summarised in tabular form. To clarify the nomenclature adopted, the following two-parameter notation is used:

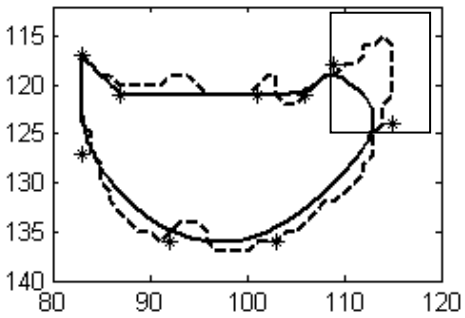
Approximation type–Distortion measurement type

Approximation type refers to either polygon or quadratic BS-based approximation while *Distortion measurement type* refers to the choice of SAD, TB (or DB), ADMSC, DMCLP and DMALP, so for instance, *Polygon–ADMSC* means that the algorithm is based on a polygon approximation with ADMSC being the distortion measurement technique.

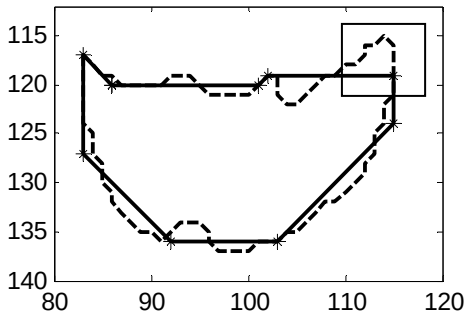
The first set of experiments concentrated on peak distortion measurement for different *approximation-measurement* pairings, for a prescribed set of admissible values. The various results for the admissible distortion setting $T_{\max} = 2$, $T_{\min} = 2$ *pixels* are displayed in Figures 14(a)-(j), while the corresponding numerical results also with other selected distortion pairings are summarised in Table 2. Figure 14 reveals that in general, all algorithms produced similar perceptual shapes with the notable exceptions of *Polygon–SAD*, *B-spline–SAD*, *Polygon–TB* and *B-spline–TB*, where as highlighted by the rectangular boxes, the distortion was greater than the prescribed peak value. The selection of $T_{\max} = T_{\min}$ was deliberate so all contour points would have identical admissible distortions, as well as to highlight the measurement problems inherent in SAD and TB. Furthermore, while both SAD and TB failed to sustain the peak admissible distortion, in contrast DMCLP, ADMSC, DMALP and ADMA all consistently maintained a bounded peak distortion.



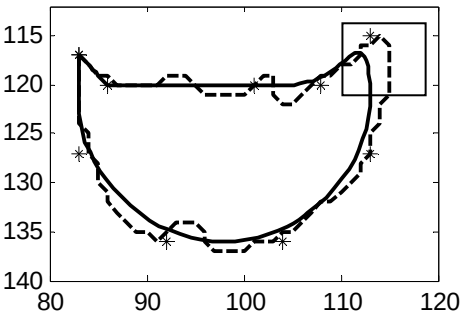
(a) *Polygon–SAD*



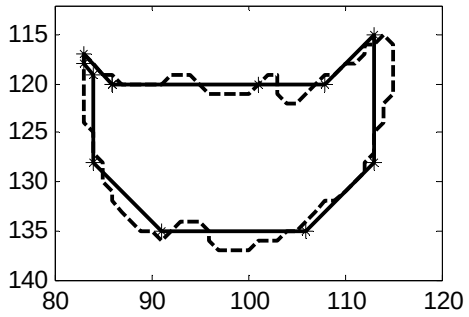
(b) *B-spline–SAD*



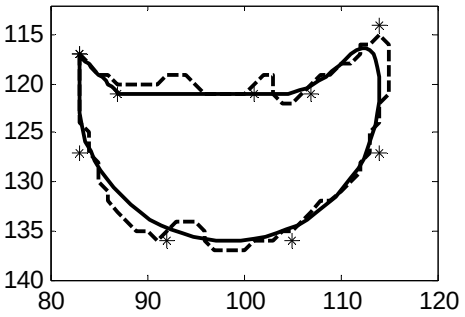
(c) *Polygon–TB* (or *DB*)



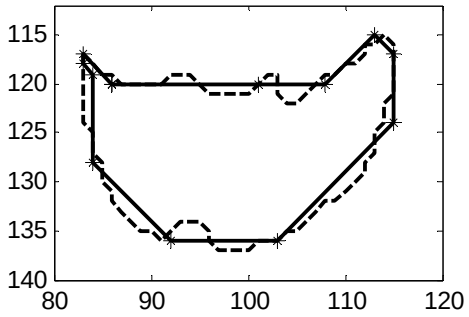
(d) *B-spline–TB* (or *DB*)



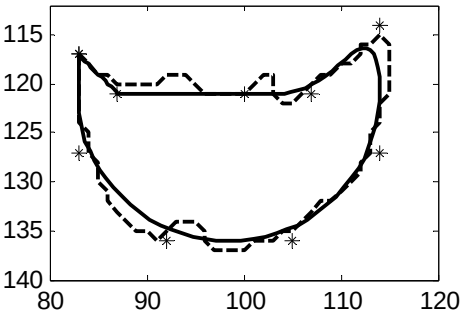
(e) *Polygon–ADMSC*



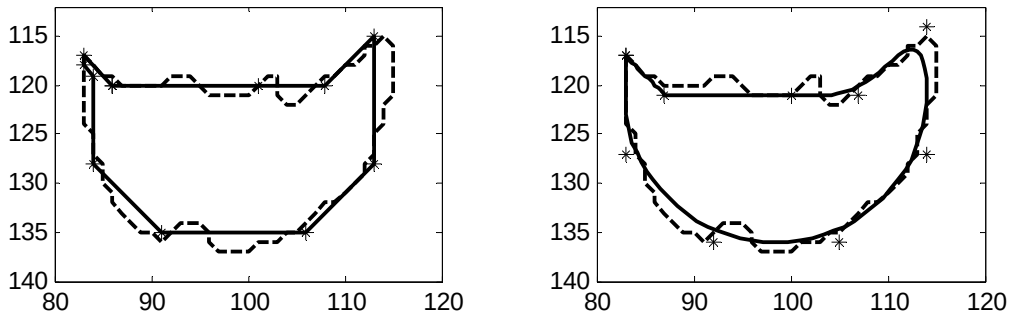
(f) *B-spline–ADMSC*



(g) *Polygon–DMCLP*



(h) *B-spline–DMCLP*



(i) *Polygon–DMALP* (j) *B-spline–DMALP*

Figure 14: Results for the *Neck* region of the 31st frame of the *MissAmerica*.qcif sequence with $T_{\max} = 2$ and $T_{\min} = 2$ *pixels* (**Legend** – *solid line*: approximating contour; *dashed line*: original contour; *asterisk*: CP).

Table 2: Bit-rate (*bits*) requirements (with obtained distortion in parenthesis whenever it is different from the admissible peak distortion) for the *Neck* region of the 31st frame of *MissAmerica*.qcif with different admissible distortion pairs ($T_{\max}$ & $T_{\min}$ *pixels*) using various ORD optimal shape-coding algorithms.

Admissible distortion →	$T_{\max} = 1$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 2$	$T_{\max} = 3$, $T_{\min} = 1$	$T_{\max} = 3$, $T_{\min} = 2$
Algorithms ↓	**Bit-rate**	**Bit-rate**	**Bit-rate**	**Bit-rate**	**Bit-rate**
Polygon–SAD	116 (1.42)	100 (2.24)	87 (2.23)	92	85
Polygon–TB	115 (2.24)	95 (2.24)	79 (4.0)	71 (5.0)	70 (5.0)
Polygon–ADMSC	138	109	86	86	86
Polygon–DMCLP	146	112	93	92	88
Polygon–DMALP	140	109	87	87	86
B-spline–SAD	125 (2.0)	83 (2.45)	80 (5.65)	80 (7.0)	80 (8.0)
B-spline–TB	133 (2.0)	87 (3.6)	78 (2.8)	76 (6.0)	75 (6.0)
B-spline–ADMSC	127	100	78	78	78
B-spline–DMCLP	132	102	78	80 (2.83)	78
B-spline–DMALP	128	100	78	78	78

A cursory review of the SAD and TB results in Table 2, i.e., $T_{\max} = 2, T_{\min} = 1$ *pixels*, fallaciously reveals they consistently mandated fewer bits for encoding, until cognisance is taken that these algorithms did not always uphold the admissible distortion constraint in the *B-spline–SAD* and *B-spline–TB* cases, which respectively generated maximum distortions of 2.45 and 3.6 *pixels* despite being supposedly bound to a peak of 2 *pixels*. This implies SAD and TB ignored certain parts of the shape leading to a lower bit requirement than ADMSC, DMCLP and DMALP-based algorithms, which in contrast, all guarantee the peak admissible distortion. Interestingly for $T_{\max} = 3, T_{\min} = 1$ *pixels*, *B-spline–DMCLP* produced a maximum distortion of only 2.83 *pixels* and so did not fully exploit the admissible distortion limit of 3 *pixels*, reflecting the earlier comment over DMCLP being a more relaxed measure. From a bit-rate perspective, this relaxation is manifest by a negligibly small increase in the number of bits incurred for DMCLP-based algorithms, so with $T_{\max} = 3, T_{\min} = 1$ *pixels* the respective bit-rate requirements for *B-*

spline–ADMSC, B-spline–DMCLP and *B-spline–DMALP* are 78, 80 and 78 *bits* respectively. The results also confirm the lowest bit-rate is achieved by ADMSC which guarantees to measure the absolute minimum distortion.

Table 3 summarises the numerical results for different combinations upon various test sequences for the setting $T_{\max} = 2$, $T_{\min} = 1$ *pixels*. These corroborate the same observations, namely ADMSC, DMCLP and DMALP always maintain the admissible distortion which is not the case for either SAD or TB, while their respective bit-rates are all analogous.

Table 3: Average bit-rate (*bits per frame*) requirements (with the obtained distortion in parenthesis whenever it is different from the admissible peak) for the various test sequences with $T_{\max} = 2$, $T_{\min} = 1$ *pixels* using various combinations of polygon-based algorithms.

Algorithms→	***Polygon–SAD***	***Polygon–TB***	***Polygon–ADMSC***	***Polygon–DMCLP***	***Polygon–DMALP***
Video sequence↓	**Bit-rate**	**Bit-rate**	**Bit-rate**	**Bit-rate**	**Bit-rate**
MissAmerica.qcif	343 (3.0)	338 (3.0)	348	355	350
Akiyo.qcif	312 (2.8)	310 (3.0)	313	320	314
Bream.qcif	421 (3.0)	415 (3.0)	421	430	422
Kids.sif	1592	1593	1593	1600	1595
Stefan.sif	580	582	585	589	587
Kids.sdtv	4500	4499	4507	4512	4510
Stefan.sdtv	1080	1080	1085	1092	1090

Table 4: CPU time (*seconds*) required for the *Neck* region of the 31st frame of the *MissAmerica*.qcif by different ORD optimal shape coding algorithms for various admissible distortion pairs ($T_{\max}$, $T_{\min}$ in *pixels*) (the distortions produced by SAD/TB are in parentheses)

Admissible distortion →	$T_{\max} = 1$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 2$	$T_{\max} = 3$, $T_{\min} = 1$	$T_{\max} = 3$, $T_{\min} = 2$
Algorithms ↓	**Time**	**Time**	**Time**	**Time**	**Time**
Polygon–SAD	1.59 (1.42)	1.80 (2.24)	1.90 (2.23)	2.0	2.0
Polygon–TB	4.26 (2.24)	6.03(2.24)	7.73 (4.0)	11.35 (5.0)	12.66 (5.0)
Polygon–ADMSC	1.63	1.89	2.01	2.15	2.25
Polygon–DMCLP	1.61	1.83	1.92	1.97	2.02
Polygon–DMALP	1.62	1.84	1.95	1.99	2.08
B-spline–SAD	120 (2.0)	550 (2.45)	560 (5.65)	565 (7.0)	570 (8.0)
B-spline–TB	90.60 (2.0)	510.50 (3.6)	545.50 (2.8)	620.30 (6.0)	680.40 (6.0)
B-spline–ADMSC	554.20	575.00	582.10	587.80	591.60
B-spline–DMCLP	270.20	290.30	297.00	312.50	314.30
B-spline–DMALP	271.00	291.05	298.20	313.90	315.20

Since the computational complexity of the distortion measurement process has a direct impact on the shape coding framework, the next series of experiments were conducted to compare the time requirements incurred by the different algorithm combinations for various admissible distortion pairs. Table 4 summarises the total *central processing unit* (CPU) times for these assorted implementations. For BS

encoding, it is readily apparent that the algorithms which employ either DMCLP or DMALP as their distortion metric are computationally faster than their SAD, TB/DB, and ADMSC counterparts. For instance, with $T_{\max} = 3, T_{\min} = 1$ *pixels*, *B-spline–SAD*, *B-spline–TB*, *B-spline–ADMSC*, *B-spline–DMCLP* and *B-spline–DMALP* required 565, 620.3, 578.8, 312.5 and 313.9 *secs* respectively, so vindicating the earlier analysis that while SAD, TB and ADMSC incur $O\left(N_B^2\right)$, the cost for both DMCLP and DMALP is only $O\left(N_B\right)$. For polygon-based encoding, DMCLP algorithms incur less time than ADMSC-based techniques, while as anticipated, TB algorithms consistently required higher measurement times reflecting its overall $O\left(N_B^2\right)$ overhead. The recently proposed DMALP technique has a slightly higher time overhead than DMCLP, though it proved faster than either ADMSC or TB/DB.

To establish confidence bounds for the results in Table 4, cognisance of the time variations in the different distortion measurement techniques needs to be made so a statistical *t-test* has been applied to the CPU data obtained from ten separate runs of each algorithm at 95% confidence intervals. The endpoints of the confidence intervals namely, the *confidence limits* (CL) and *probability* (*p*) of observing a value as either extreme or more extreme of the test data are summarised in Table 5, for the same shape coding algorithms and admissible distortion pair combinations given in Table 4. It is noteworthy to mention that the result of the test was always $H = 1$, indicating a rejection of the null hypothesis at the 5% significance level. The negligible *p* values reveal the very low probabilities of data going to the extreme of the test statistics. It also shows that the time results in Table 4 always fall inside the limits of the corresponding statistical results in Table 5. The results also further corroborate the superior computational speeds of both DMCLP and DMALP compared with the other distortion measurement techniques.

Table 5: CL (*secs*) and *p* values of a *t-test* at confidence intervals of 95% for CPU times to encode the *Neck* region of the 31st frame of *MissAmerica*.qcif, for the same shape coding algorithms and admissible distortion pair ($T_{\max}$, $T_{\min}$ in *pixels*) combinations used in Table 4.

Admissible distortion →	$T_{\max} = 1$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 1$	$T_{\max} = 2$, $T_{\min} = 2$	$T_{\max} = 3$, $T_{\min} = 1$	$T_{\max} = 3$, $T_{\min} = 2$
Algorithms ↓			***t-test* results**		
Polygon–SAD	CL: [1.57 1.63]	[1.77 1.83]	[1.88 1.93]	[1.98 2.07]	[1.99 2.08]
	p: 1.00e-015	1.16e-015	2.89e-016	5.18e-015	7.60e-015
Polygon–TB	[4.21 4.30]	[6.02 6.09]	[7.69 7.81]	[11.32 11.40]	[12.64 12.73]
	4.97e-018	1.77e-020	2.49e-019	4.58e-022	5.85e-022
Polygon–ADMSC	[1.61 1.67]	[1.84 1.92]	[1.98 2.07]	[2.11 2.19]	[2.23 2.30]
	6.63e-016	2.44e-015	8.03e-015	1.66e-015	5.76e-016
Polygon–DMCLP	[1.59 1.65]	[1.83 1.89]	[1.88 1.96]	[1.95 2.04]	[1.99 2.07]
	2.75e-016	1.23e-016	1.28e-014	3.65e-015	2.07e-015
Polygon–DMALP	[1.61 1.66]	[1.83 1.89]	[1.93 2.02]	[1.97 2.07]	[2.02 2.12]
	1.31e-016	1.33e-016	1.13e-014	1.12e-014	1.11e-014
B-spline–SAD	[118.8 125.4]	[547.8 555.2]	[558.7 565.5]	[564.2 569.0]	[568.6 573.5]
	4.07e-013	9.69e-020	3.54e-020	1.57e-023	1.61e-023
B-spline–TB	[88.7 93.5]	[508.2 513.0]	[543.8 548.9]	[618.9 623.5]	[678.7 683.1]
	1.86e-014	3.86e-021	3.30e-021	4.58e-022	1.11e-024
B-spline–ADMSC	[553.3 558.1]	[573.9 578.5]	[580.7 584.6]	[586.0 589.4]	[590.6 594.5]
	1.61e-021	7.67e-022	1.74e-022	4.23e-025	1.49e-022
B-spline–DMCLP	[269.1 273.6]	[289.1 293.5]	[296.0 300.0]	[311.7 315.6]	[313.0 317.9]
	5.81e-019	2.76e-019	6.55e-020	3.64e-020	3.34e-019
B-spline–DMALP	[269.6 273.9]	[289.6 293.9]	[296.2 300.1]	[312.8 316.5]	[313.7 319.5]
	3.85e-019	2.23e-019	7.13e-020	3.36e-020	1.50e-018

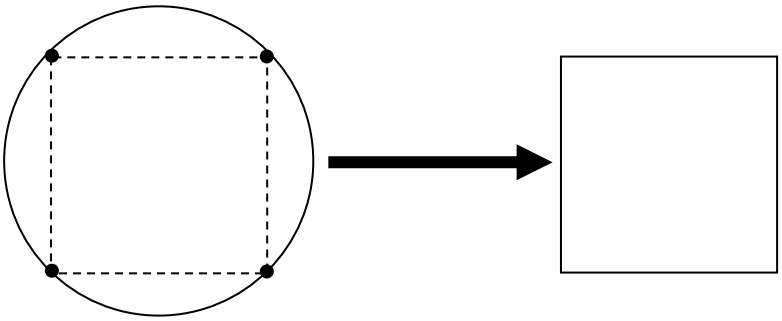
Table 6: Comparative chart for the various distortion measurement techniques

Quality criteria	**SAD**	**DB/TB**	**ADMSC**	**DMCLP**	**DMALP**
Always reflects accurate distortion	NO	NO	YES	NO	NO
Guaranteed admissible distortion	NO	NO	YES	YES	YES
Computational time complexity for a polygon-based framework	$O(N_B)$	$O(N_B^2)$	$O(N_B)$	$O(N_B)$	$O(N_B)$
Computational complexity for a BS-based framework	$O(N_B^2)$	$O(N_B^2)$	$O(N_B^2)$	$O(N_B)$	$O(N_B)$
Distortion type: peak or MS	Both	Peak only	Both	Both	Both

To conclude the experimental analysis, Table 6 presents a comparative summary of the key performance features and characteristics of the different distortion measurement algorithms examined. It confirms ADMSC always accurately measures the distortion and maintains the peak admissible distortion. DMCLP and DMALP over-estimate the distortion resulting in a higher than actual value, and though this does not adversely affect the preservation of the admissible distortion, it can compromise RD performance, though their respective bit-rate requirements are very similar. Although well-established within an ORD context, SAD and TB guarantee neither an accurate distortion measure nor uphold the admissible distortion. From a computational speed standpoint, DMCLP and DMALP are the most efficient mandating $O(N_B)$ time for both polygon and BS-based encoding, while DB/TB is the least efficient requiring $O(N_B^2)$. In addition, while DB/TB have been designed just for peak admissible distortion, SAD, ADMSC, DMCLP and DMALP can all successfully operate in both a peak (4) and MS (5) distortion measuring paradigm.

5. Current Research Challenges

This paper has concentrated so far upon reviewing assorted distortion metrics and geometric distortion measurement techniques. In this section, the attention moves towards more contemporary challenges, one of which is the importance of the subjective appearance of the reconstructed shape and the impact of *structural deformation*. The dynamic vertex-based shape coding framework focuses on RD optimisation, with the sole aim of reducing the required bit-rate for a prescribed admissible distortion and vice versa. It does not consider what the shape actually looks like when compressed, i.e., no account is taken of the perceptual structural deformation a shape may undergo during encoding. As a result, a shape may become deformed, such that it loses its actual structure and causes visual discrepancies, as well as recognition and matching problems in for example, content-based searching and retrieval applications, especially when shape data are highly compressed. This is evidenced in both Figure 4 and the example in Figure 15, where a circular shape (Figure 15(a)) is optimally encoded for a prescribed distortion shown by dotted lines, using just four CP. Figure 15(b) shows the reconstructed shape from the encoded information, where the circular shape has now become a square with significant changes of curvature/cornerity around the entire contour, especially at the four CP. The introduction of the term *shape deformation* and appropriate strategies to effectively quantify and integrate it within shape coding frameworks, directly attempt to address the challenge of how best to retain a shape's visual structure within some prescribed perceptual threshold. Despite various approaches including MS distortion minimisation [Schuster and Katsaggelos 1997], this objective has proven elusive and has not been satisfactorily resolved.



(a) (b)

Figure 15: Shape deformation – (a) a circular object shape to be encoded and (b) the reconstructed shape becomes a square, while still upholding the relevant admissible distortion criterion.

A unified framework can be formulated to manage both shape distortion and deformation minimisation. One of the proposed future directions is to redefine the cost function for distortion in [Katsaggelos, Kondi, Meier, Ostermann and Schuster 1998] be integrating deformation information. If $R_k(s_k)$, $r(s_{k-1}, s_k)$, $d(s_{k-1}, s_k)$, and $w(s_{k-1}, s_k)$ are the number of bits required to encode up to CP s_k, the bit-requirement to differentially encode s_k with respect to s_{k-1}, the distortion between the CPs s_{k-1} and s_k, and the cost to differentially encode point s_k with respect to s_{k-1} respectively, then the overall cost function can be recursively expressed as:

$$R_k(s_k) = R_{k-1}(s_{k-1}) + w(s_{k-1}, s_k) \tag{14}$$

$$\text{where } w(s_{k-1}, s_k) = \begin{cases} \infty & : d(s_{k-1}, s_k) > D_{\max} \\ r(s_{k-1}, s_k) & : d(s_{k-1}, s_k) \le D_{\max} \end{cases}.$$

Now considering shape deformation, the cost function $w(s_{k-1}, s_k)$ can be redefined as:

$$w(s_{k-1}, s_k) = \begin{cases} \infty & : d(s_{k-1}, s_k) > D_{\max} \\ r(s_{k-1}, s_k) + \alpha \cdot SD_{k-1,k} & : d(s_{k-1}, s_k) \le D_{\max} \end{cases} \tag{15}$$

where $SD_{k-1,k}$ is the shape deformation caused by edge $s_{k-1}s_k$ and α is a positive real number which weights the importance of the deformation. Clearly when α =0, the normal ORD result is obtained since shape deformation is not considered. If ΔR represents the smallest possible difference between the rates for the optimal polygon and the polygon when shape deformation is considered, then the largest feasible shape deformation is $SD_{\max}$, and α can accordingly be selected from:

$$\frac{\Delta R}{SD_{\max}} > \alpha > 0. \tag{16}$$

It needs to be emphasised the amount of extra bandwidth (ΔR) available, governs the level of structure that can be preserved, so as ΔR becomes larger, the corresponding shape deformation will be lower and vice versa, with this trade-off being directly controlled by parameter α in (16).

6. Conclusion

This paper has presented a contemporary review of both distortion metrics and geometric distortion measurement techniques. It has proven that among the gamut of different strategies developed, ADMSC always provides the most accurate distortion measurement, while DMCLP is computationally more efficient under all shape conditions with a higher-order curve based approximation, though the recently proposed DMALP algorithm exhibits slightly superior *rate-distortion* performance for comparable DMCLP processing speed.

7. References

AGHITO, S.M. AND FORCHHAMMER, S. 2004. Context based coding of binary shapes by object boundary straightness analysis. In *Data Compression Conference (DCC)*, Snowbird, Utah, USA, 399-408.

AGHITO, S.M. AND FORCHHAMMER, S. 2006. Context-based coding of bilevel images enhanced by digital straight line analysis. *IEEE Transactions on Image Processing 15*, 2120-2130.

BANDYOPADHYAY, S.K. AND KONDI, L.P. 2005. Optimal bit allocation for joint contour-based shape coding and shape adaptive texture coding. In *International Conference on Image Processing (ICIP)*, Genoa, Italy, 589-592.

BRADY, N. 1999. MPEG-4 standardized methods for the compression of arbitrarily shaped video objects. *IEEE Transactions on Circuits and Systems for Video Technology 9*, 1170-1189.

BRADY, N., BOSSEN, F. AND MURPHY, N. 1997. Context-based arithmetic encoding of 2D shape sequences. In *International Conference on Image Processing (ICIP)*, Washington, DC, USA, 29-32.

CHEN, Z. AND NGAN, K.N. 2004. Linear rate-distortion models for MPEG-4 shape coding. *IEEE Transactions on Circuits and Systems for Video Technology 14*, 869-873.

FARIN, G.E. 1997. *Curves and surfaces for computer-aided geometric design: a practical guide*. Academic Press.

FAROUKI, R.T. 1997. Optimal parameterizations. *Computer Aided Geometric Design 14*, 153-168.

FREEMAN, H. 1961. On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronic Computers EC-10*, 260-268.

HÖTTER, M. 1990. Object-oriented analysis-synthesis coding based on moving two-dimensional objects. *Signal Processing 2*, 409-428.

HÖTTER, M. 1994. Optimization and efficiency of an object-oriented analysis-synthesis coder. *IEEE Transactions on Circuits and Systems for Video Technology 4*, 181-194.

KATSAGGELOS, A.K., KONDI, L.P., MEIER, F.W., OSTERMANN, J. AND SCHUSTER, G.M. 1998. MPEG-4 and rate-distortion-based shape-coding techniques. *Proceedings of the IEEE 86*, 1126-1154.

KONDI, L.P., MEIER, F.W., SCHUSTER, G.M. AND KATSAGGELOS, A.K. 1998. Joint optimal object shape estimation and encoding. In *SPIE Visual Communication and Image Processing*, San Jose, California, USA, 14-25.

KONDI, L.P., MELNIKOV, G. AND KATSAGGELOS, A.K. 2001. Jointly optimal coding of texture and shape. In *International Conference on Image Processing (ICIP)*, Thessaloniki, Greece, 94-97.

KONDI, L.P., MELNIKOV, G. AND KATSAGGELOS, A.K. 2004. Joint optimal object shape estimation and encoding. *IEEE Transactions on Circuits and Systems for Video Technology 14*, 528-533.

KOPLOWITZ, J. 1981. On the performance of chain codes for quantization of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-3*, 180-185.

MATSUBARAA, T., FUKUOKAB, D., YAGIC, N., HARAC, T., FUJITAC, H., INENAGAD, Y., KASAID, S., KANOD, A., ENDOE, T. AND IWASEF, T. 2005. Detection method for architectural distortion based on analysis of structure of mammary gland on mammograms *International Congress Series 1281*, 1036-1040.

MEIER, F.W., SCHUSTER, G.M. AND KATSAGGELOS, A.K. 2000. A mathematical model for shape coding with B-splines. *Signal Processing: Image Communication 15*, 685-701.

MELNIKOV, G., SCHUSTER, G.M. AND KATSAGGELOS, A.K. 2000. Shape coding using temporal correlation and joint VLC optimization. *IEEE Transactions on Circuits and Systems for Video Technology 10*, 744-754.

O'CONNELL, K.J. 1997. Object-adaptive vertex-based shape coding method. *IEEE Transactions on Circuits and Systems for Video Technology 7*, 251-255.

QIAOA, P., LUC, K., LESTARID, W. AND WANGE, J. 2007. Curvature mode shape-based damage detection in composite laminated plates. *Composite Structures 80*, 409-428.

RICHARDSON, I.E. 2003. *H.264 and MPEG-4 video compression: video coding for next generation multimedia*. John Wiley & Sons.

SCHUSTER, G.M. AND KATSAGGELOS, A.K. 1997. *Rate-distortion based video compression: optimal video frame compression and object boundary encoding*. Kluwer Academic Publishers.

SCHUSTER, G.M., LI, X. AND KATSAGGELOS, A.K. 2004. Shape error concealment using Hermite splines. *IEEE Transactions on Image Processing 13*, 808-820.

SCHUSTER, G.M., MELNIKOV, G. AND KATSAGGELOS, A.K. 1998. Operationally optimal vertex-based shape coding. *IEEE Signal Processing Magazine 15*, 91-108.

SOARES, L.D. AND PEREIRA, F. 2004. Spatial shape error concealment for object-based image and video coding. *IEEE Transactions on Image Processing 13*, 586-599.

SOHEL, F.A. AND BENNAMOUN, M. 2008. Distortion measurement using arc-length-parameterisation within vertex-based shape coding framework. In *International Conference on Computers and Information technology*, Khulna, Bangladesh, 329-334.

SOHEL, F.A., DOOLEY, L.S. AND KARMAKAR, G.C. 2006. Accurate distortion measurement for generic shape coding. *Pattern Recognition Letters 27*, 133-142.

SOHEL, F.A., DOOLEY, L.S. AND KARMAKAR, G.C. 2007. New dynamic enhancements to the vertex-based rate-distortion optimal shape coding framework. *IEEE Transactions on Circuits and Systems for Video Technology 17*, 1408-1413.

SOHEL, F.A., KARMAKAR, G.C. AND DOOLEY, L.S. 2006. Dynamic sliding window width selection strategies for rate-distortion optimal vertex-based shape coding algorithms. In *International Conference on Signal Processing (ICSP)*, Guilin, China.

SOHEL, F.A., KARMAKAR, G.C. AND DOOLEY, L.S. 2007. Fast distortion measurement using chord-length parameterisation within the vertex-based rate-distortion optimal shape coding framework. *IEEE Signal Processing Letters 14*, 121-124.

TOPIWALA, P.N. 1998. *Wavelet image and video compression*. Kluwer Academic Publishers.

VQEG 1999. Final report from the video quality experts group on the validation of objective models of video quality assessment.

WANG, H., SCHUSTER, G.M. AND KATSAGGELOS, A.K. 2005. Rate-distortion optimal bit allocation for object-based video coding. *IEEE Transactions on Circuits and Systems for Video Technology 15*, 1113-1123.

WANG, H., SCHUSTER, G.M., KATSAGGELOS, A.K. AND PAPPAS, T.N. 2003. An efficient rate-distortion optimal shape coding approach utilizing a skeleton-based decomposition. *IEEE Transactions on Image Processing 12*, 1181-1193.

WU, H.R. AND RAO, K.R. 2006. *Digital video image quality and perceptual coding*. CRC press, Taylor and Francis.