

A Modified Distortion Measurement Algorithm for Shape Coding

Ferdous Ahmed Sohel, Laurence S Dooley and Gour C Karmakar

Gippsland School of Computing and Information and Technology

Monash University, Churchill, Victoria – 3842, Australia.

E-mail: Ferdous.Sohel@infotech.monash.edu.au

Abstract: Efficient encoding of object boundaries has become increasingly prominent in areas such as content-based storage and retrieval, studio and television post-production facilities, mobile communications and other real-time multimedia applications. The way distortion between the actual and approximated shapes is measured however, has a major impact upon the quality of the shape coding algorithms. In existing shape coding methods, the distortion measure do not generate an actual distortion value, so this paper proposes a new distortion measure, called a modified distortion measure for shape coding (DMSC) which incorporates an actual perceptual distance. The performance of the Operational Rate Distortion optimal algorithm [1] incorporating DMSC has been empirically evaluated upon a number of different natural and synthetic arbitrary shapes. Both qualitative and quantitative results confirm the superior results in comparison with the ORD algorithm for all test shapes, without any increase in computational complexity.

Keywords – Distortion measure, shortest absolute distance, perceptual distance, shape coding, object based coding.

1. Introduction

Object-oriented video coding based upon shape information is a very challenging topic in the literature currently as it facilitates retrieval, interactive editing and manipulation of both natural and synthetic videos. Due to the inherent bandwidth limitations of existing communication technology, video applications such as video over the Internet, video on demand, wired and mobile video transmission for hand-held devices will all be benefited immensely from fast and efficient shape coding techniques.

Within the object-oriented framework, a video sequence is represented by the evolution of video object planes (VOPs), with each frame comprised of one or more VOPs, which in turn are described in terms of shape, texture and motion information. In MPEG-4 [7] for example, shape or boundary encoding is widely used.

In a classical vertex-based shape coding system, a shape is encoded to an optimal number of bits within some prescribed distortion value. The total number of bits required to encode a particular shape is known as

the *rate* (cost), and the distortion is usually measured as the maximum distance between the original and approximated shape. Distance measurement is an important issue in rate-distortion algorithms. For instance, the shortest absolute distance is used in [1-6], where the vertical distance from a point on a shape to a line or extended line segment of the approximated shape is defined as the distortion for that point. This measure however does not produce the actual distortion for those points where the vertical distance is measured from the extended line segment. To address this issue, a distortion measure algorithm, namely *a modified distortion measurement algorithm for shape coding (DMSC)* is proposed on the bedrock of Operational Rare Distortion (ORD) optimal algorithms in [1-6]. The performance, in terms of subjective assessment of the decoded shape and the actual distortion with respect to the constrained maximum distortion, of the DMSC has been tested for a number of different natural and synthetic arbitrary shapes, and will be shown that it can make guarantee on the maximum distortion value for the decoding.

The remainder of this paper is organized as follows: Section 2 provides a brief review of existing shape measurement techniques, while Section 3 presents the new proposed distortion measure algorithm. Section 4 gives a description of the ORD optimal algorithms in [1-6], while Section 5 provides the experimental results, analysis and performance comparison of the various ORD techniques. Finally, some concluding remarks are made in Section 6.

2. Existing Shape Distortion Measurement Algorithms

A shape is generally approximated as a polygon or a higher order parametric curve. Usually parametric curves are comprised of a sequence of discrete points that ultimately form a polygon.

Amongst the distance measurement algorithms, the Euclidian distance is extensively used to determine the distance between two points. The Euclidian distance $d(p_1, p_2)$, between two points p_1 and p_2 in a plane, having Cartesian coordinates (x_1, y_1) and (x_2, y_2) respectively, is defined as:-

$$d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

If Euclidian distance is used as the distortion measure, the distance of each point on the original shape from all points on the approximated polygon are needed to be calculated to find the minimum distance, as shown in Figure 1. In Figure 1, the shortest distance of q from line $ad = qc$ while from p it is pa .

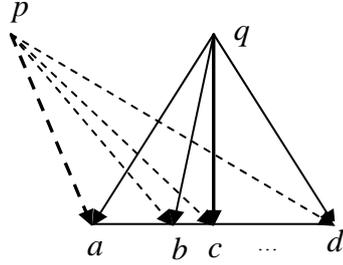


Figure 1: Shortest distance using Euclidian Distance.

This requires a computational complexity of $O(n)$, where n is the number of points in the approximated shape (polygon), to measure the distortion for each point of the original shape.

The shortest absolute distance had been introduced to conquer the high computational complexity, where the perpendicular distance is measured from each point on the original shape to the polygon. Let $d(a,d,m)$ be the shortest absolute (perpendicular) distance from a point m on a line passing through points a and d . This is formulated as:-

$$d(a,d,m) = \frac{|(m_x - a_x)(b_y - a_y) - (m_y - a_y)(b_x - a_x)|}{\sqrt{(b_x - a_x)^2 + (b_y - a_y)^2}} \quad (2)$$

where subscripts x and y represent the x and y coordinate values respectively. This is clearly computationally more efficient, as it takes only $O(1)$ to calculate the distance from a point to a line segment; however, the perceived shape distortion is not properly reflected in (2). To understand why, consider the example in Figure 2:

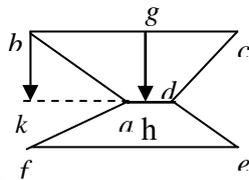


Figure 2: The minimum distance example using (2).

Eq. (2) measures the shortest geometric distance between a point and a line segment. The minimum distance point of the line segment may either be on the line or its extension in Figure 2. The minimum distance point is on the line segment (gh is the distance of point g from ad), the minimum distance point is on the extended line segment (bk is the distance of point b from ad , which lies on the extension ak). This leads to an inaccuracy in the distortion measure from a perceptual point of view, since the actual perceptual distance of b from line

ad is ba , while according to (2) it is bk . This algorithm therefore does not take account of human visual perception with respect to shape and essentially introduces further distortion in sharp edges, especially in concave or convex areas of a shape. This was the primary motivation behind the development of the new shape distortion measure that is described in Section 3.

3. Proposed Distortion Measurement Algorithm for Shape Coding

This Section introduces a novel distortion measurement algorithm, namely a *modified distortion measurement algorithm for shape coding* (DMSC) which considers the minimum distance of a shape point from a respective line segment connecting two approximate polygon points. Note that, the extended portion of the line segment (ad in Figure 2) is not taken into account because of the need to measure the true shape distortion.

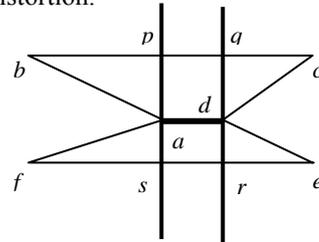


Figure 3: Vertex positioning strategy.

Let the upper part of the shape $abcdefa$, $abcd$ be approximated by the polygon segment ad . The distance of every point of shape $abcd$ from segment ad is then calculated. Two perpendicular lines are drawn at a and d (Figure 3) and it is assumed that they intersect the original boundary at points p, q respectively. When the shape points lie within or on the perpendicular lines, the distortion is measured using (2). For all other shape points, the distances are measured either from a or d depending on the candidate shape point position with respect to the perpendicular lines. In Figure 3 for example, if the candidate shape point lies before the vertical line at a or after the vertical line at d , the distortion is measured from a or d respectively.

If the equation of a straight line is known, using elementary geometry, the position of any point can be easily determined whether it lies on the straight line or to which side of it. For a given straight line ($\alpha x + \beta y + \delta = 0$), the position determining function is $f(x, y) = \alpha x + \beta y + \delta$. The position of a point (c, d) with respect to the straight line $\alpha x + \beta y + \delta = 0$ is then determined using the following functions:

$$f(c, d) = \begin{cases} < 0, & \text{for below (before) the line} \\ > 0, & \text{for above (after) the line} \\ = 0, & \text{on the line} \end{cases} \quad (3)$$

The perpendicular lines on ad at a and d can respectively be defined as,

$$\begin{aligned} x(a_x - d_x) + y(a_y - d_y) + (a_x \times d_x + a_y \times d_y - a_x^2 - a_y^2) &= 0 \quad \text{and} \\ x(a_x - d_x) + y(a_y - d_y) + (d_x^2 + d_y^2 - a_x \times d_x - a_y \times d_y) &= 0 \end{aligned} \quad (4)$$

The complete DMSC algorithm is summarized in the following Algorithm:

Algorithm 1: Modified distortion measurement algorithm for shape coding:

Input: A shape point m and a line segment bounded by points a and d inclusive.

Output: $d(a, d, m)$, the minimum distortion between the shape point m and the line segment ad .

Step 1: Construct two perpendicular lines on the current candidate segment endpoints.

Step 2: Locate the position of any shape point m within ad with respect to the relative orientation of the perpendicular lines derived in Step 1 using (3) and (4).

Step 3: IF position of m is within perpendicular line inclusive, THEN calculate distortion using (2).

Step 4: ELSE IF position of m is outside of the perpendicular line at a , on the further side of perpendicular line at d , THEN the distortion is :

$$d(a, d, m) = \sqrt{(a_x - m_x)^2 + (a_y - y_m)^2}$$

Step 5: ELSE position of m is outside of the perpendicular line at d , on the farther side of the perpendicular line at a , the distortion is :

$$d(a, d, m) = \sqrt{(d_x - m_x)^2 + (d_y - y_m)^2}$$

Step 6: Stop.

The computational complexity of this algorithm is same as the shortest absolute distance measure algorithm, which is $O(1)$, once is only required to be.

This algorithm DMSC is mathematically modelled and the potentiality of DMSC has tested using the ORD optimal algorithms described in [1-6], which had introduced shape coding as ORD optimal, for different types of shapes with different size and different orientation. The following Section will provide a brief description of the ORD algorithm.

4. The ORD Optimal Technique Incorporating the Proposed Distortion Measurement Algorithm for Shape Coding

The ORD optimal techniques [1-6] approximate a boundary of a given shape by a set of significant points and encode them instead of all the actual shape points. Let $B = \{b_1, b_2, \dots, b_{N_B}\}$ be an ordered set of boundary points, where N_B is the total boundary points. Note $b_1 = b_{N_B}$ represents a closed boundary. Let $S = \{s_1, s_2, \dots, s_{N_S}\}$ be an ordered set of significant points used to approximate the boundary (B), where N_S is the total number of points. The k^{th} edge starts from s_{k-1} and ends at s_k . Since S is an ordered set, it uniquely represents the approximated polygon and is differentially encoded. If $r(s_{k-1}, s_k)$ is the requisite bit rate for differentially encoding of significant point s_k , given that s_{k-1} has already encoded, the bit rate $R(S)$ required for set S is then:-

$$R(S) = \sum_{k=1}^{N_S} r(s_{k-1}, s_k) \quad (5)$$

where $r(s_0, s_1)$ is the number of bits required to encode the absolute position of the first point s_1 . For a closed boundary $r(s_{N_S-1}, s_{N_S}) = 0$ as the last point does not need to be encoded. The k^{th} edge which connects two consecutive significant points, s_{k-1} and s_k is an approximation to a portion of shape $\{b_j = s_{k-1}, b_{j+1}, \dots, b_{j+l} = s_k\}$ consisting of $l+1$ shape points. $d(s_{k-1}, s_k)$ denotes the edge distortion between the shape and its approximated boundary defined using s_{k-1} and s_k . The distortion measure is then the maximum of all distortion measures.

Let $d(s_{k-1}, s_k, q)$ be the shortest absolute distance of q from the line connecting points s_{k-1} and s_k . This distance is measured in the ORD algorithm by (2). Note that this will also be measured using the new algorithm described in Algorithm 1 and will be compared and contrasted with (2) in order to test its potential. The maximum absolute distance between the portion of a shape $\{b_j = s_{k-1}, b_{j+1}, \dots, b_{j+l} = s_k\}$ and the edge of the approximated polygon (s_{k-1}, s_k) is then given by:

$$d(s_k, s_{k-1}) = \max_{q \in \{b_j = s_{k-1}, b_{j+1}, \dots, b_{j+l} = s_k\}} d(s_{k-1}, s_k, q) \quad (6)$$

Hence, the maximum polygon distortion is :-

$$D(S) = \max_{k \in \{1, \dots, N_S\}} d(s_{k-1}, s_k) \quad (7)$$

where $d(s_0, s_1) = 0$.

Hence the problem becomes a solution to the following constrained optimization problem:

$$\min_{s_1, \dots, s_{N_S}} D(S), \text{ subject to } R(S) \leq R_{\max} \quad (8)$$

where R_{\max} is the maximum allowable rate.

And also its complementary problem,

$$\min_{s_1, \dots, s_{N_S}} R(S), \text{ subject to: } D(S) \leq D_{\max} \quad (9)$$

D_{\max} is the maximum allowable distortion.

This can be recursively defined as a dynamic problem formulation:

$$R_k(s_k) = R_{k-1}(s_{k-1}) + w(s_{k-1}, s_k) \quad (10)$$

where

$$w(s_{k-1}, s_k) = \begin{cases} \infty & : d(s_{k-1}, s_k) > D_{\max} \\ r(s_{k-1}, s_k) & : d(s_{k-1}, s_k) \leq D_{\max} \end{cases} \quad (11)$$

A detailed description of rate measure $r(s_{k-1}, s_k)$ is provided in [3] which uses both linear and logarithmic schemes for a combined chain code and run-length coding along with orientation encoding scheme for all significant points. The recursion needs to be initialized by setting $R_0(s_0)$ equal to the number of bits required to code the first point; so $R_{N_S}(s_{N_S}) = R(S)$, the rate for the entire set of significant points for the complete boundary.

The Directed Acyclic Graph (DAG) [8] algorithm has the least time complexity of currently known shortest path algorithms for a graph. The selection of a starting point plays an important role in such algorithms and [1] and [3] have presented a clear definition on how to select this point. Thus a weighted DAG is formed with edge list $E = \{(b_i, b_j) \in B^2, i < j\}$ and the weighting value is determined by (11) to find the shortest path. A DAG formation example for an arbitrary shape point 9 is shown in Figure 4.

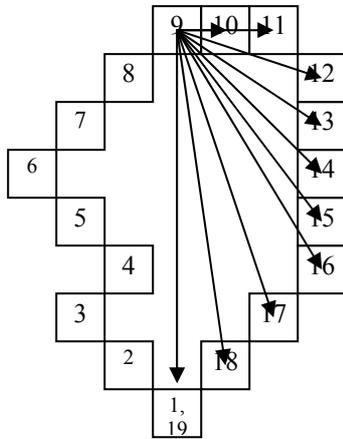


Figure 4: Formation of DAG for point 9, i.e., $(b_9, b_j) \in B^2, j > 9$

Algorithm 2: The ORD algorithm incorporating DMSC.

1. $R^*(s_1) = r(s_0, s_1)$;
2. for $i = 2, \dots, N_B$;
3. $R^*(b_i) = \infty$;
4. for $i = 1, \dots, N_B - 1$;
5. for $j = i + 1, \dots, N_B$;
6. calculate edge distortion $d(b_i, b_j)$
using algorithm (1);
7. look up edge rate $r(b_i, b_j)$;
8. assign $w(b_i, b_j)$;
9. if $(R^*(b_i) + w(b_i, b_j) < R^*(b_j))$;
10. $R^*(b_j) = (R^*(b_i) + w(b_i, b_j))$;
11. $\text{prev}(b_j) = b_i$;

In Algorithm 2, $R^*(b_i)$ represents the minimum rate to reach the shape point b_i from the source $p_1 = b_i$ via a polygon approximation. $R^*(b_{N_B})$ is the solution to (8) or (9). Prev described in step 11 of algorithm 2 will contain all the significant points in a recursive fashion.

5. Results and Analysis

Both the original ORD optimal algorithm [1-5] and the ORD algorithm incorporating DMSC were implemented using Matlab 6.1 (The Mathworks Inc.) for a number of different natural and synthetic arbitrary shapes, some of which were selected by collecting the objects from IMSI¹. All object shapes were manually segmented and the results were derived using the coded shape information generated by each respective algorithm.

The experimental results produced by the original ORD and ORD with DMSC algorithms for the butterfly object in Figure 5(a), are shown in Figures 5(b)-(d), with a maximum distortion value $D_{\max} = 10$ pel. Note the two front antenna of the butterfly object were not separated by manual segmentation, because it was difficult to intuitively separate. The results in Table 1 show that the decoded shape (Figure 5 (c)) encoded using ORD with DMSC has a maximum distortion of 10 pel i.e. equal to D_{\max} , while the corresponding distortion value for Figure 5(b) using the original ORD algorithm is 27 pel. This improvement is due to the fact that the original ORD algorithm does not consider distance in its distortion

¹ IMSI's Master Photo Collection, 1895 Francisco Blvd. East, San Rafael, CA 94901-5506, USA.

measure. To further illustrate the superiority of ORD incorporating the new DMSC over the original, two results are superimposed in different colours in Figure 5(d). If the results in Figures 5(b)-(d) are compared with the original shape shown in Figure 5(a), it is visibly apparent that ORD with DMSC produces a much more accurate shape representation i.e. has a lower shape distortion.

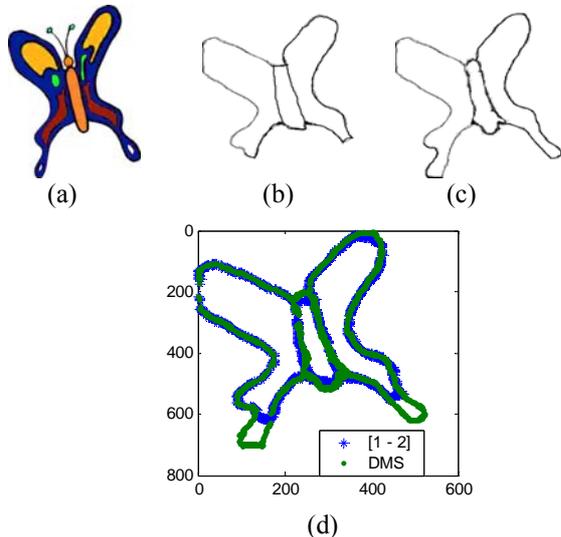


Figure 5: a) Original butterfly object, b) The decoded shape using the original ORD algorithm, c) The decoded shape using ORD with DMSC, d) A comparative plot showing the shapes produced by the original ORD and ORD incorporating DMSC.

A second series of experiments were conducted using the 31st frame of the *Miss America* video sequence (Figure 6) which was extensively used in [1-6] and the fish image (Figure 8(a)). The results for both are shown in Figures 7 and 8(b)-(d) respectively. The corresponding distortion values for the original ORD and ORD incorporating DMSC for prescribed maximum distortion values are given in Table 1.

Results for the highlighted neck region in the 31st frame of the *Miss America* sequence are summarized in Figure 7, coded with a distortion of 1 pel. The maximum distortion for the decoded shape using ORD [1] is 2 pel (indicated by the ellipse in Figure 7(a)), while DMSC maintained a $D_{max} = 1$ pel (Figure 7(b)), so confirming the superiority of this new approach.



Figure 6: The 31st frame of the *Miss America* sequence with the 'neck' region highlighted in white.

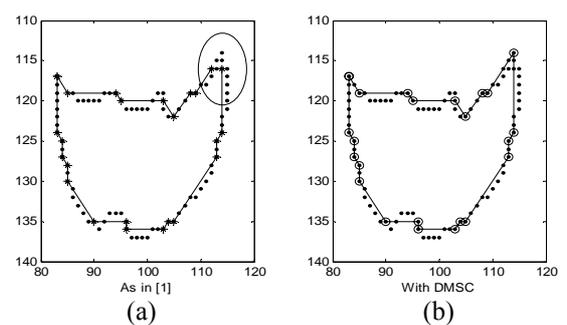


Figure 7: Decoded shapes using shortest absolute distance measure with $D_{max} = 1$ pel (a) Original ORD [1-6], note highlighted area has a distortion of 2 pel, b) shape coding with DMSC has $D_{max} = 1$ pel.

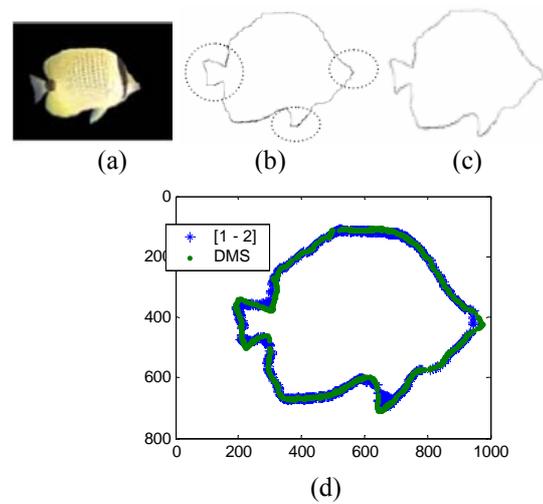


Figure 8: a) The original fish object, b) The decoded shape using the original ORD algorithm, c) The decoded shape using ORD with DMSC, d) A comparative plot of the original ORD and ORD incorporating DMSC.

The algorithm in [1-6] produces more distortion than the stipulated maximum distortion at the sharp corners as indicated with dotted ellipses, whereas the new algorithm shows the minimum rate distortion and obeys the tolerable distortion. Because, DMSC is considering all perceptual distances.

A final series of experiments were performed upon the synthetic shape in Figure 9(a) to confirm the superiority of the new DMSC algorithm. The results are presented in Figures 9(b-c) and Figure 10 for $D_{max} = 1$ and $D_{max} = 2$ pel respectively. In Figure 9(b), when the top solid line is considered to approximate the upper boundary, the shortest absolute distance will have a distortion of 1 pel, but in reality $\sqrt{3^2 + 1} = 3.1$ pel will be perceived by a human. The same thing occurs in Figure 10(a).

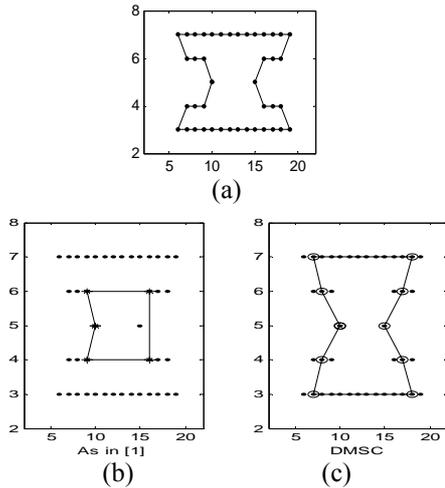


Figure 9: a) An artificial object, b) results of original ORD algorithm in [1-6] and c) results of ORD incorporating DMSC for $D_{\max} = 1$ pel. Distortion is greater than 1 pel (actually 3.1 pel) at any corner of the decoded object for algorithms in [1-6].

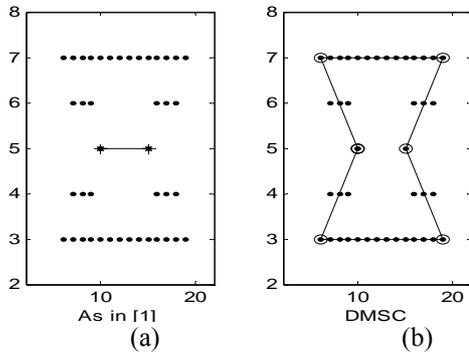


Figure 10: Comparison of (a) Original ORD algorithm in [1-6] and (b) ORD algorithms incorporating DMSC for $D_{\max} = 2$ pel. DMSC guarantees a bounded maximum distortion, however in Figure (a) distortion is higher than 2 pel (4.5 pel) at any corner of the decoded object.

Table 1: Actual distortions in pel.

Image	D_{\max}	DMSC	[1-6]
Butterfly 1	10	10	27
Miss America (Neck)	1	1	2
Fish	5	5	8
Artificial Object	1	1	3.2
Artificial Object	2	2	4.5

It can be concluded from the subjective assessment of the decoded shape and actual distortion with respect to the stipulated limit, that the shortest absolute distance measure does not maintain the constraint of maximum distortion. Table 1 confirms this because the perceptual distortion exceeds the maximum allowable

distortion in all experiments, while the new DMSC algorithm takes cognisance of human visual perception by considering every distortion is always bounded.

6. Conclusions

Much research has already been conducted in respect of developing optimal shape coding algorithms, though none address the difference between geometric and perceptual distance in the shape coding technique. The original ORD algorithm is unable to ensure a bounded shape distortion after decoding, so to address this issue, a new distortion measurement algorithm called a *modified distortion measurement algorithm for shape coding (DMSC)* has been presented in this paper. Experimental results using a number of different natural and synthetic arbitrary shapes confirmed the improved performance of the DMSC algorithm, and proved it overcomes the limitation of the original ORD algorithm, without any increase in computational complexity.

7. 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, pp. 1126-1154, June 1998.
- [2] G. M. Schuster, G. Melnikov, and A. K. Katsaggelos, "Operationally Optimal Point-Based Shape Coding," IEEE Signal Processing Mag., vol. 15, pp. 28-40, Nov. 1998.
- [3] G. M. Schuster and A. K. Katsaggelos, Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding, Kluwer Academic Publishers, 1997.
- [4] G. M. Schuster and A. K. Katsaggelos, "An optimal boundary encoding scheme in the rate distortion sense," IEEE Trans. Image Processing, vol. 7, pp. 13-26, Jan. 1998.
- [5] H. Wang, G. M. Schuster, A. K. Katsaggelos, and T. N. Pappas, "An Efficient Rate-Distortion Optimal Shape Coding Approach Utilizing a Skeleton-Based Decomposition," IEEE Trans. Image Processing, vol. 12, pp. 1181-1193, Oct. 2003.
- [6] L. P. Kondi, G. Melnikov, and A. K. Katsaggelos, "Joint Optimal Object Shape Estimation and Encoding," IEEE Trans. Circuits Syst. Video Technol., vol. 14, pp. 528-533, Apr. 2004.
- [7] R. Koenen, "MPEG-4 multimedia for our time," IEEE Spectrum, vol. 36, pp. 26-33, Feb. 1999.
- [8] T. H. Cormen, C. h. Leiserson, R. L. Rivest, C. Stelin, Introduction to Algorithms, 2nd edition, MIT Press, 2001.