Transactions Letters

Optimal Bit Allocation for Joint Texture-Aware Contour-Based Shape Coding and Shape-Adaptive Texture Coding

Saurav K. Bandyopadhyay and Lisimachos P. Kondi

***Abstract*—In this paper, an operationally optimal joint shape and texture coding algorithm is proposed that uses shape-adaptive texture coding as well as texture-aware shape coding. The solution is optimal in the operational rate distortion sense, i.e, given the coding setup, the solution will guarantee the smallest possible distortion for a given rate. The shape is approximated using polygons or higher order curves. We also consider biasing the cost function to favor horizontal and vertical edges for the case of polygon approximation (biased polygon approximation). The texture is encoded using shape-adaptive discrete cosine transform or shape-adaptive discrete wavelet transform of the MPEG-4 video codec. A comparison is drawn between the two techniques. Both a fixed-width and a variable-width tolerance band for shape coding are considered. The variable width of the tolerance band is a function of the texture profile, i.e, the width is inversely proportional to the magnitude of the image gradient. Experimental results are presented and conclusions are drawn.**

***Index Terms*—Optimization, shape-adaptive discrete cosine transform (SA-DCT), shape-adaptive discrete wavelet transform (SA-DWT), shape coding, texture coding.**

I. Introduction

THE second generation video coding techniques represent an image by the shape, motion and texture information of its constituent objects. However, in order to build an efficient video encoder, optimal bit allocation between shape and texture is necessary. Again, the coding of the shape is not independent of the coding of the texture of the object, in other words, a joint coding scheme needs to be developed.

Shape coding is a problem with relatively long history [1]–[19]. In [20] and [21], a vertex based shape coding method is proposed that takes into consideration the texture information. It utilizes the texture information to create a variable-width tolerance band. However, no algorithm is proposed for the Rate-Distortion optimal bit allocation between shape and texture coding. In [22], a joint shape and texture rate control algorithm for MPEG-4 encoders is proposed. However, no rate-distortion optimal solution is provided. In [23], bit allocation between shape and texture for MPEG-4 codec is provided. However, only bitmap-based shape coding is considered and no shape-adaptive texture coding techniques are used.

In this paper, for the first time, an operationally optimal joint shape and texture coding algorithm is proposed that uses shape adaptive texture coding as well as texture-aware shape coding. The initial results of this work are reported in [24] and [25]. The algorithm is based on the use of polygons or B-splines to encode the shape and shape-adaptive discrete cosine transform (SA-DCT) or shape-adaptive discrete wavelet transform (SA-DWT) of the MPEG-4 codec to encode the texture. Contour based shape coding techniques are known to perform better than the bitmap-based shape encoding method adopted by MPEG-4 [26]. Hence, our algorithm becomes a natural choice. The solution is optimal in the operational rate distortion sense. For the polygon approximation we also consider biasing the cost function to favor horizontal and vertical edges (biased polygon approximation).

The rest of the paper is organized as follows. In Section II, we discuss the contour based shape coding and in Section III, we discuss the SA-DCT and SA-DWT based texture coding techniques used. In Section IV, the problem formulation and the optimal solution are presented. Section V provides our experimental results. In Section VI, we draw conclusions.

II. Shape Coding

The goal of shape coding is to encode the shape information of a video object to enable applications requiring content-based video access. We have used contour-based shape coding for our purpose. Rate-Distortion optimal shape coding techniques have been shown to outperform bitmap based techniques adopted by MPEG-4 [26]. Hence, the shape is approximated using a polygon or B-splines for lossy shape coding. In all cases, the problem reduces to finding the shortest path in a directed acyclic graph (DAG). Both a fixed-width and a variable-width tolerance band are considered. The reader interested in the details of contour based shape coding is referred to [26].

A. Notations

Let $B = \{b_0, \ldots, b_{N_B-1}\}$ denote a connected boundary which is an ordered set, where b_j is the jth point of B and N_B is the total number of points in B. Let $P = \{p_0, \ldots, p_{N_{p-1}}\}$ denote the set of control points of the polygon approximation, which is also an ordered set with p_k the kth vertex of P and N_p the total number of vertices in P. The kth edge starts at p_{k-1} and ends at p_k. In case of a closed boundary, $b_0 = b_{N_B-1}$. A polygon edge is defined by two control points, its vertices.

Manuscript received March 30, 2006; revised October 6, 2006 and February 9, 2007. This paper was recommended by Associate Editor S. Li.

S. K. Bandyopadhyay is with W&W Comm./DSP Research, Sunnyvale, CA 94086 USA (e-mail: sauravb@wwcoms.com).

L. P. Kondi is with the Department of Computer Science, University of Ioannina, Ioannina 45110, Greece (e-mail: lkon@cs.uoi.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2008.918784

B-splines can be used to approximate the contour instead of a polygon. A B-spline curve segment is defined by three control points.

B. Tolerance Band

A fixed-width tolerance band has a width $D_{\max}$ along the boundary B. The approximating contour must lie within the tolerance band.

A variable-width tolerance band requires a $D_{\max}$ for every boundary point. We denote this by $D_{\max}[i], i = 0, \ldots, N_B - 1$ where N_B is the number of boundary points. In order to construct the tolerance band, we draw circles from each boundary point b_i with diameter $D_{\max}[i]$. The tolerance band consists of the set of points inside the circles. In order to find $D_{\max}[i]$, the gradient is computed for an image $f(x, y)$ as

$$\nabla f(x,y) = \left[\frac{\partial f}{\partial x}\ \frac{\partial f}{\partial y}\right]^T = [f_x\ f_y]^T. \tag{1}$$

In practice, the Sobel edge detector masks are used to calculate the gradient. The magnitude of the gradient is then computed, that is

$$|\nabla f(x,y)| = \sqrt{f_x^2(x,y) + f_y^2(x,y)}. \tag{2}$$

Let us denote by $\mathrm{gradmin}$ and $\mathrm{gradmax}$, respectively, the minimum and maximum of the magnitude of the image gradient for the whole image. Let us also denote the desired minimum and maximum values of $D_{\max}[i]$ as $T_{\min}$ and $T_{\max}$, respectively. Then, a linear mapping is performed between the gradient value of each boundary point and the width of the tolerance band. If the magnitude of the gradient at the boundary point b_i is $\mathrm{grad}[i]$, then the width of the tolerance band at this point is given by

$$D_{\max}[i] = T_{\min} + \xi\,(\mathrm{grad}[i] - \mathrm{gradmax}) \tag{3}$$

where

$$\xi = \frac{T_{\max} - T_{\min}}{\mathrm{gradmin} - \mathrm{gradmax}}. \tag{4}$$

In practice, we need to define a threshold (Th) for the gradient magnitude. The boundary points whose gradient magnitude exceeds the threshold should have the minimum possible $D_{\max}[i]$. Therefore,

$$D_{\max}[i] = T_{\min},\ \text{if}\ \mathrm{grad}[i] \geq \mathrm{Th}. \tag{5}$$

C. Shape Coding Problem Formulation

In this paper, we approximate boundaries using B-splines, polygons and the biased polygon approximation. A polygon edge is defined by two control points, its vertices. The segment rates and the segment distortion depend on two points and are given by $r(p_{k-1}, p_k)$ and $d(p_{k-1}, p_k)$, respectively. The bit rate $R(p_0, \ldots, p_{N_{p-1}})$ for the entire curve is given by

$$R(p_0, \ldots, p_{N_{p-1}}) = \sum_{k=0}^{N_{p-1}} r(p_{k-1}, p_k). \tag{6}$$

The segment distortion for the polygon case is given as

$$d(p_{k-1}, p_k) = \begin{cases} 0: & \text{all points of } G_k(p_{k-1}, p_k) \\ & \text{are inside the tolerance band} \\ \infty: & \text{any point of } G_k(p_{k-1}, p_k) \\ & \text{is outside the tolerance band} \end{cases} \tag{7}$$

Equation (7) takes a curve segment $G_k(p_{k-1}, p_k)$ defined by two control points p_{k-1} and p_k and checks if it lies with the tolerance band. The curve distortion can be expressed in terms of the segment distortion as

$$D\left(p_0, \ldots, p_{N_p-1}\right) = \max_{k \in [1, \ldots, N_{p-1}]} d(p_{k-1}, p_k). \tag{8}$$

Thus, the optimization problem we are solving for the shape coding is

$$\min\ R\left(p_0, \ldots, p_{N_p-1}\right),\ \text{subject to}: D\left(p_0, \ldots, p_{N_p-1}\right) = 0. \tag{9}$$

This guarantees that the entire boundary approximation will lie within the tolerance band. The problem reduces to finding the shortest path in a DAG [26]. The formal proof of the DAG-shortest path algorithm can be found in [27].

The shape-adaptive texture coding using SA-DCT or SA-DWT is expected to be more efficient if the edges of the object are horizontal or vertical because the spatial correlation between neighboring pels will be better maintained. We allow a $\mathrm{bias} < 1$ multiplicative factor for the rates of segments defined by control points p_{k-1} and p_k, which correspond to horizontal or vertical edges [20]. Thus

$$r'(p_{k-1}, p_k) = \mathrm{bias} \times [r(p_{k-1}, pq_k)] \tag{10}$$

if p_{k-1} and p_k define the horizontal or vertical edge. Thus, the boundary coding algorithm will favor horizontal and vertical edges at the expense of increased bit rate for shape coding.

B-splines can also be used for shape coding instead of a polygon. The motivation in using B-splines is better coding efficiency for objects in natural images. Such objects tend to have fewer straight lines and narrow corners. The problem formulation and solution are similar to the polygon case.

III. Texture Coding

Texture is encoded using SA-DCT or SA-DWT using an MPEG-4 compliant codec.

SA-DCT [28] provides a way of encoding blocks using a number of coefficients that is equal to the number of object pels in the block. This is accomplished by shifting the object pels towards the origin of the block and then taking one dimensional DCTs row-wise and then column-wise. The length of these one dimensional DCTs can be less than eight. The SA-DCT always shifts the samples in an arbitrarily shaped block towards a certain edge of the rectangular bounding box before performing 1-D DCT along the row or the column. Hence, some spatial correlation may be lost. Again, intuitively, it is not efficient to perform column DCT on a set of coefficients that are from different frequency bands after the row DCT transforms.

To overcome the above difficulties we consider the SA-DWT [29] for coding arbitrarily shaped still texture. The SA-DWT transforms the samples in the arbitrary shaped region into the

TABLE I
PARAMETERS CONTAINED IN S

Shape coding techniques	Parameters in S
Polygons using fixed-width of the tol. band	D_{max}
Polygons using var.-width of the tol. band	Th, T_{min}, T_{max}
B-splines using fixed-width of the tol. band	D_{max}
B-splines using var.-width of the tol. band	Th, T_{min}, T_{max}
Biased polygon using fixed-width of the tol. band	$D_{max}, bias$
Biased polygon using var.-width of the tol. band	$Th, bias, T_{min}, T_{max}$

same number of coefficients in the subband domain while keeping the spatial correlation, locality property and self-similarity across subbands. Encoding and decoding the SA-DWT coefficients are the same as encoding and decoding the regular wavelet coefficients except for keeping track of the locations of the wavelet coefficients according to the shape information. A 2-D SA-DWT is applied to the texture object, producing the dc component (low frequency subband) and a number of ac (high-frequency) subbands. The dc subband is quantized, predictively encoded using a form of differential pulse code modulation (DPCM) and entropy coded using an arithmetic encoder. The ac subbands are quantized and reordered ("scanned"), zerotree encoded and entropy coded.

IV. PROBLEM FORMULATION

The shape of the object is encoded using polygons, B-splines, or the biased polygon approximation. Both the fixed-width and the variable-width of the tolerance band are considered for each of these three boundary approximations giving a total of six different shape coding techniques as follows:

1) polygons using fixed-width of the tolerance band;
2) polygons using variable-width of the tolerance band;
3) B-splines using fixed-width of the tolerance band;
4) B-splines using variable-width of the tolerance band;
5) biased polygons using fixed-width of the tolerance band;
6) biased polygons using variable-width of the tolerance band.

In the case of fixed-width of the tolerance band, the parameter of interest is the band width $(D_{\max})$. Considering variable-width, the band is determined by the threshold (Th), minimum $(T_{\min})$ and maximum $(T_{\max})$ width of the tolerance band. In the biased polygon approximation case, the parameter of interest is the bias for the horizontal and vertical edges. Let S be a set that contains the shape coding parameters. It thus specifies one of the six shape coding methods along with the parameters associated with it. For example, for the biased polygon approximation case with fixed width of the tolerance band, S contains the width of the tolerance band $(D_{\max})$ and the bias for the horizontal and vertical edges. The parameters in S to be determined for each of the six shape coding methods are shown in Table I.

Let T denote the texture coding parameters (QP for SA-DCT based texture coding and decoded bitrate for texture coding using SA-DWT), as shown in Table II.

Hence, the optimization problem can be written as follows:

$$S^*, T^* = \arg\min_{S,T} R_{\text{total}}(S,T) \tag{11}$$

subject to

$$D^{\text{YUV}}_{\text{texture}}(S,T) \le D^{\text{YUV}}_{\text{budget}} \tag{12}$$

where, $R_{\text{total}}(S,T) = R_{shape}(S) + R_{\text{texture}}(S,T)$.

TABLE II
PARAMETERS CONTAINED IN T

Texture coding methods	Parameters in T
SA-DCT	QP
SA-DWT	$Rate$

$D^{\text{YUV}}_{\text{texture}}(S,T)$ is the texture distortion of the image in the region of the reconstructed shape. The $D^{\text{YUV}}_{\text{budget}}$ is the maximum allowable texture distortion. The distortion in all of our simulation results is calculated based on the Mean Square Error (MSE) between the original and the reconstructed image in the region of the reconstructed shape. The distortion can be mathematically written as:

$$D^{\text{YUV}}_{\text{texture}} = \frac{1}{N}\left\{ \sum_{(x,y)\in C_Y} \delta_y^2(x,y) + \sum_{(x,y)\in C_{U,V}} \left(\delta_u^2(x,y) + \delta_v^2(x,y)\right) \right\} \tag{13}$$

where $\delta_y(x,y)$ is the Y component differential intensity value at pixel position (x,y), $\delta_u(x,y)$ is the U component differential intensity value at pixel position (x,y), $\delta_v(x,y)$ is the V component differential intensity value at pixel position (x,y). The input video sequence is in YUV 4:2:0 color space. C_Y is the reconstructed shape region for the Y component and $C_{U,V}$ is the reconstructed shape region for the U and V components. For YUV 4:2:0, $C_{U,V}$ is subsampled in both the horizontal and vertical dimensions by a factor of 2. N is the total number of pels in the region of the reconstructed shape for Y, U, and V components. The peak signal-to-noise ratio (PSNR) can be calculated using

$$\text{PSNR}^{\text{YUV}}_{\text{texture}} = 10\log_{10}\left(\frac{255^2}{D^{\text{YUV}}_{\text{texture}}}\right). \tag{14}$$

In the problem formulation of (11) and (12), the constraint is on the texture distortion instead of a weighted sum of texture and shape distortions. The reason for this is that, in the latter case, the results would be highly dependent on the (arbitrary) weighting of the two distortions. In our case, we can guarantee an acceptable shape distortion by limiting the range of parameters $D_{\max}$ and $T_{\max}$. The constrained minimization problem stated in (11) and (12) is converted to an unconstrained problem by using the Lagrangian multiplier method as

$$J_\lambda(S,T) = R_{\text{total}}(S,T) + \lambda \cdot D^{\text{YUV}}_{\text{texture}}(S,T) \tag{15}$$

where λ is the Lagrangian multiplier. Now, according to (15), if there is a λ^* for which $S^*, T^* = \arg\min_{S,T} J_\lambda(S,T)$ and which satisfies $D^{\text{YUV}}_{\text{texture}}(S^*,T^*) = D^{\text{YUV}}_{\text{budget}}$ then S^*, T^* is the optimal solution to the constrained problem of (11) and (12).

TABLE III
OPTIMAL RESULTS OF EXPERIMENT 1 USING FRAME 0 OF THE "CHILDREN" SEQUENCE

R_{total} (bits)	Case	Th/D_{max}	Bias	$D^{YUV}_{texture}$
32604	6	600	0.5	2.39
25588	6	600	0.5	5.50
21356	6	600	0.5	10.14
16012	6	600	0.5	23.62
14378	5	2.5	0.5	31.75
10914	5	2.5	0.5	66.13
4362	5	2.5	0.5	382.12
4304	5	3.0	0.5	395.82

TABLE IV
OPTIMAL RESULTS OF EXPERIMENT 2 USING FRAME 0 OF THE "CHILDREN" SEQUENCE

R_{total} (bits)	Case	Th/D_{max}	Bias	$D^{YUV}_{texture}$
27824	6	600	0.5	0.83
21848	6	600	0.5	2.63
14904	5	3.0	0.5	18.38
9920	5	3.0	0.5	68.18
4952	5	3.0	0.7	315.92
4896	6	600	0.5	322.40

The optimal solution involves finding the λ that satisfies the target distortion. In this work, we use the bisection algorithm to determine λ^*.

V. EXPERIMENTAL RESULTS

In this section, we present experimental results for the optimization problem in (11) and (12). A number of experiments were conducted using several video sequences, some of which are reported below. In Experiment 1, polygons, B-splines and the biased polygon approximation are used for shape coding while SA-DCT of MPEG-4 codec is used for texture coding.

In our discrete optimization problem, the various parameters are allowed to take the following values. The fixed width $(D_{\max})$ of the tolerance band is varied from 0.8 to 3.0 in steps of 0.1. In case of variable width of the tolerance band, the threshold (Th) is varied from 100 to 600 is steps of 50, $T_{\min} = 0.8$ and $T_{\max} = 3.0$. In case of the biased polygon approximation, the bias is varied from 0.1 to 0.9 in steps of 0.1. QP is varied from 2 (fine quantization) to 31 (coarse quantization) in steps of 1. Table III shows the results of Experiment 1 using frame 0 of the "Children" sequence. Cases 5 and 6 are chosen as the optimal solution for the reported results. The third column in the Table III shows the threshold (Th) or the tolerance band width $(D_{\max})$ for the variable and the fixed width tolerance band cases, respectively.

In Experiment 2, the six shape coding techniques are used along with SA-DWT based texture coding of MPEG-4. The encoded bit stream is decoded at different rates. Table IV shows the results of Experiment 2 using, respectively the frame 0 of the "Children" sequence. Similar observations can be made.

Fig. 1 shows a rate distortion comparison between Experiments 1 and 2 for frame 0 of the "Children" sequence. In Fig. 2, a subjective comparison for the same frame is drawn. It is clearly observed that the joint shape and texture coding using SA-DWT based texture coding outperforms that using SA-DCT. The experimental results show that the biased polygon approximation is chosen as the optimal solution among the six different shape coding techniques for a specific texture coding technique.

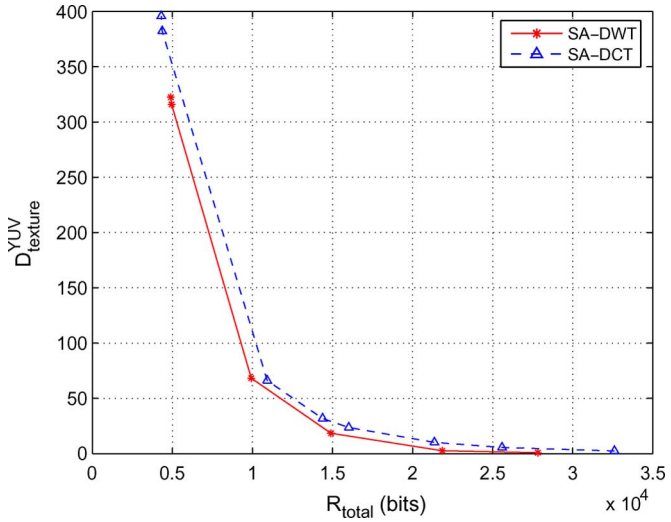
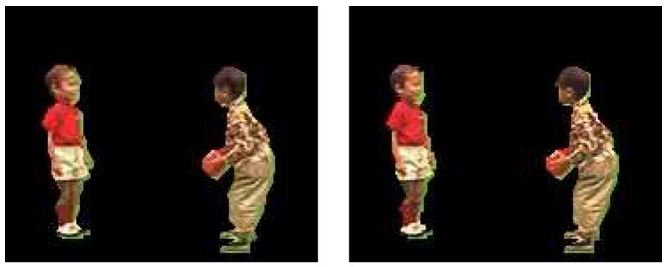


Fig. 1. Comparison of joint shape and texture coding with SA-DCT- or SA-DWT-based texture coding of the MPEG-4 codec. The results are shown for frame 0 of the "Children" sequence.



(a) (b)

Fig. 2. Frame no. 0 ("Children" sequence). (a) SA-DCT based texture coding, $\mathrm{Th} = 600, \mathrm{Bias} = 0.5, QP = 4, R_{\mathrm{total}} = 21356$ bits, $D^{\mathrm{YUV}}_{\mathrm{texture}} = 10.14$. (b) SA-DWT based texture coding, $\mathrm{Th} = 600, \mathrm{Bias} = 0.5, \mathrm{Decoding\ rate} = 21848$ bits, $D^{\mathrm{YUV}}_{\mathrm{texture}} = 2.63$.

In Experiment 3, a comparison is made between the results from Experiments 1 and 2 and the optimal points obtained by using context arithmetic encoding (CAE) along with SA-DCT or SA-DWT based texture coding of MPEG-4. Two fields, $alpha\_threshold\,(\alpha_{th})$ and $change\_CR\_disable\,(\beta)$, are used for controlling lossless or lossy coding using bitmap based shape coding method of MPEG-4 standard.

If only shape coding is concerned and the objective is to encode the shape using the minimum number of bits given a shape distortion constraint, the B-spline method performs best, followed by the polygon and the biased polygon methods. This is because B-splines can follow a natural contour better than a polygon. The polygon approximation is the optimal approximation using polygons (if only shape coding is considered). Thus, biasing the cost function (biased polygon approximation) will lead to a less efficient polygon approximation. However, when shape and texture coding are jointly considered, as in this paper, the interactions between shape coding and shape-adaptive texture coding become evident. Shape adaptive texture coding is

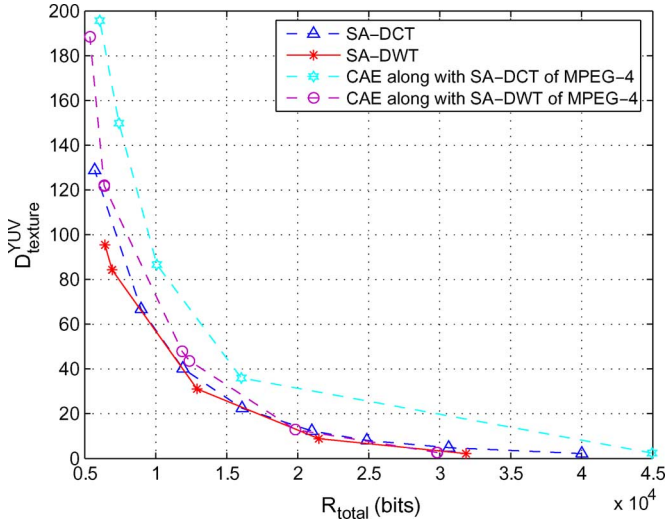


Fig. 3. Comparison of optimal solution from Experiments 1 and 2 with joint shape and texture using CAE based shape coding of MPEG-4 along with SA-DCT or SA-DWT based texture coding of the MPEG-4 codec. The results are shown for frame 0 of the "Bream" sequence.

more efficient when the shape edges are horizontal or vertical. Thus, the biased polygon approximation, although it provides less efficient shape coding, it leads to more efficient shape-adaptive texture coding and is thus selected by our joint shape and texture coding optimization algorithm.

VI. Conclusion

We presented an operational rate-distortion optimal bit allocation scheme between texture and shape for the encoding of the object based video. Experimental results show that the SA-DWT based texture coding performs better than SA-DCT based texture coding in the joint encoding of shape and texture. The biased polygon approximation is the least efficient amongst the considered cases for shape coding. However, we used it for efficient texture coding. As expected the optimal solution selected the techniques that uses biased polygon approximation for shape coding. The solution is determined using the Lagrangian multiplier method.

References

[1] H. Freeman, "On the encoding of arbitrary geometric configurations," *IRE Trans. Electron. Comput.*, vol. EC-10, pp. 260–268, Jun. 1961.
[2] M. Eden and M. Kocher, "On the performance of a contour coding algorithm in the context of image coding. Part I: Contour segment coding," *Signal Process.*, vol. 8, no. 4, pp. 381–386, Jul. 1985.
[3] P. Gerken, "Object-based analysis-synthesis coding of image sequences at very low bit rate," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 3, pp. 228–235, Jun. 1994.
[4] M. Hotter, "Object-oriented analysis-synthesis coding based on moving two-dimensional objects," *Signal Process.: Image Commun.*, vol. 2, pp. 409–428, 1990.
[5] G. M. Schuster and A. K. Katsaggelos, "A optimal segmentation encoding scheme in the rate distortion sense," in *Proc. Int. Conf. Circuits Syst.*, Atlanta, Georgia, May 1996, vol. 2, pp. 640–643.
[6] G. M. Schuster and A. K. Katsaggelos, "An optimal boundary encoding scheme in the rate distortion sense," *IEEE Trans. Image Process.*, pp. 13–26, Jan. 1998.
[7] H. Wang, G. M. Schuster, and A. K. Katsaggelos, "Rate-distortion optimal bit allocation scheme for object-based video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 9, pp. 1113–1123, Sep. 2005.
[8] K. J. O'Connell, "Object-adaptive vertex-based shape coding methid," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 251–255, Feb. 1997.
[9] C. T. Zahn and R. Z. Roskies, "Fourier Desciptors for plane closed curves," *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 269–281, Mar. 1972.
[10] P. Salember, F. Marques, and A. Gasull, "Coding of partition sequences," in *Video Coding: The Second Generation Approach*, L. Torres and M. Kunt, Eds. Norwell, MA: Kluwer.
[11] A. J. Saghri and H. Freeman, "Analysis of the precision of generalized chain codes for the representation of planar curves," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 5, pp. 533–540, May 1981.
[12] J. Koplowitz, "On the performance of chain codes for the quantization of line drawings," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 3, pp. 180–185, Mar. 1981.
[13] D. L. Neuhoff and K. G. Castor, "A rate and distortion analysis of chain codes for line drawings," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 1, pp. 53–68, Jan. 1985.
[14] T. Kaneko and M. Okudaira, "Encoding of arbitrary curves based on the chain code representation," *IEEE Trans. Commun.*, vol. COMM-33, pp. 697–707, Jul. 1985.
[15] R. Prasad, J. W. Vieveen, J. H. Bons, and J. C. Arnbak, "Relative vector probabilities in differential chain coded line-drawings," in *Proc. IEEE Pacific Rim Conf. Commun., Comput. Signal Process.*, Victoria, BC, Canada, Jun. 1989, pp. 138–142.
[16] J.-Y. Suh, K. J. Kim, and M. G. Kang, "Inter-frame vertex selection scheme for object-based image sequence coder," *IEEE Trans. Consumer Electron.*, vol. 45, no. 3, pp. 835–841, Aug. 1999.
[17] K. J. Kim, J.-Y. Suh, and M. G. Kang, "Generalized interframe vertex-based shape encoding," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1667–1676, Oct. 2000.
[18] N. Brady, "MPEG-4 standardized methods for the compression of arbitrary shaped video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 8, pp. 1170–1189, Dec. 1999.
[19] H. Luo, "Image-dependent shape coding and representation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 3, pp. 345–354, Mar. 2005.
[20] L. P. Kondi, G. Melnikov, and A. K. Katsaggelos, "Joint optimal object shape estimation and encoding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 4, pp. 528–533, Apr. 2004.
[21] L. P. Kondi, F. W. Meier, G. M. Schuster, and A. K. Katsaggelos, "Joint optimal object shape estimation and encoding," in *Proc. Conf. Vis. Image Process.*, San Jose, CA, Jan. 1998, pp. 14–25.
[22] A. Vetro, H. Sun, and Y. Wang, "Joint shape and texture rate control for MPEG-4 encoders," in *Proc. IEEE Int. Conf. Circuits Systems*, Jun. 1998, pp. 285–288.
[23] H. Wang, G. M. Schuster, and A. K. Katsaggelos, "Object-based video compression scheme with optimal bit allocation among shape, motion and texture," in *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, Sep. 2003, vol. III, pp. 785–788.
[24] S. K. Bandyopadhyay and L. P. Kondi, "Joint rate distortion optimal shape and texture coding," in *Proc. IEEE Can. Conf. Electrical Comput. Eng.*, Niagara Falls, Canada, May 2004, vol. II, pp. 841–844.
[25] S. K. Bandyopadhyay and L. P. Kondi, "Optimal bit allocation for joint contour-based shape coding and shape adaptive texture coding," in *Proc. IEEE Int. Conf. Image Process.*, Genoa, Italy, Sep. 2005, vol. I, pp. 589–592.
[26] A. K. Katsaggelos, L. P. Kondi, F. W. Meier, J. Ostermann, and G. M. Schuster, "MPEG-4 and rate-distortion based shape coding techniques," *Proc. IEEE*, vol. 86, no. 6, pp. 1126–1154, Jun. 1998.
[27] T. Cormen, C. Leiserson, and R. Rivest, *Introduction to Algorithms*. New York: McGraw-Hill, 1991.
[28] T. Sikora and B. Makai, "Shape-adaptive DCT for generic coding of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 1, pp. 59–62, Feb. 1995.
[29] S. Li and W. Li, "Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 5, pp. 725–743, Aug. 2000.
[30] G. M. Schuster and A. K. Katsaggelos, *Rate Distortion Based Video Compression, Optimal Video Frame Compression, and Object Boundary Encoding*. Norwell, MA: Kluwer, 1997.