Progressive Contour Coding in the Wavelet Domain

Nicola Adami, Pietro Gallina, Riccardo Leonardi, and Alberto Signoroni

Università degli Studi di Brescia, Telecommunications Group
Dipartimento di Elettronica per l'Automazione,
via Branze, 38, I25123 Brescia, Italy
{firstname.lastname}@ing.unibs.it
http://www.ing.unibs.it/tlc

Abstract. This paper presents a new wavelet-based image contour coding technique, suitable for representing either shapes or generic contour maps. Starting from a contour map (e.g. a segmentation map or the result of an edge detector process), a unique one-dimensional signal is generated from the set of contour points. Coordinate jumps between contour extremities when under a tolerance threshold represent signal discontinuities but they can still be compactly coded in the wavelet domain. Exceeding threshold discontinuities are coded as side information. This side information and the amount of remaining discontinuity are minimized by an optimized contour segment sequencing. The obtained 1D signal is decomposed and coded in the wavelet domain by using a 1D extension of the SPIHT algorithm. The described technique can efficiently code any kind of 2D contour map, from one to many unconnected contour segments. It guarantees a fully embedded progressive coding, state-of-art coding performance, good approximation capabilities for both open and closed contours, and graceful visual degradation at low bit-rates.

1 Introduction

From a human observer point of view visual content can often be captured by just relying on image discontinuities, often referred to as edges, contours, shapes,... depending on the application context. An efficient and functional coding of contour information, inherent to or extracted from images or video frames, can be exploited in order to improve the content analysis and management capabilities of picture and video archiving or the effectiveness of visual communication systems. There is a wide literature regarding the so called "shape coding" case (i.e. the coding of one or more closed contour lines that typically correspond to the boundaries of segmented objects) because of its role in object based video coding approaches. However, as far as we know, there is no proposed method for effective lossy coding of generic contour maps.
Currently, the best performing techniques work in the data domain and can be subdivided into two families: line-based methods [2, 3] and bitmap-based methods [4, 5].

Transform coding of contours has already been introduced with the Fourier descriptors method [1], which is however suited only for connected and closed contour lines. Moreover, their coding performance are not competitive with respect to the state-of-art of shape coding. Instead of what has happened for image coding, the wavelet transform has not yet been considered in any kind of shape or contour map coding. In this work we set up the generic contour map coding problem as a one-dimensional (1D) signal coding problem. The 1D signal to code is directly obtained from the sequence of coordinates of the contour segments that define the contour map. The 1D signal generation is not unique because it depends on the contour segments concatenating order. We will show how to optimize the representation of this structural information so as to build a low complexity 1D signal and a limited amount of side information. The wavelet representation of the so obtained 1D signal guarantees an efficient coding of the contour map itself and allows to exploit the typical wavelet transform coding features and bit-stream properties (such as progressive quality reconstruction and, presumably, spatial scalability).
The paper is organized as follows: in Sec.2, after a more precise problem definition (2.1), we describe the contour encoding algorithm which consists in a proper 1D signal sequence creation (2.2) followed by the wavelet coding of such a signal and the required side information (2.3). In Sec.3 we show some experimental results and for both shape coding (3.1) and generic contour map coding (3.2) Sec.4 provides some concluding remarks.

2 Encoding Algorithm

2.1 Problem definition

We consider a *contour map* as a binary image where active pixels are in direct spatial relation with contour and/or shape information on the original image. A contour map can be defined as a set of non intersecting *contour tracts* or contour *segments*. A contour tract is a connected and non redundant (filiform), open or closed pixel sequence defined on the discrete spatial domain. Contour maps can be generated in various ways, for example by means of contour extraction operators, segmentation techniques, thinning or skeletonization algorithms.
In our framework the contour map can be interpreted as a single sequence of contour points and described by the corresponding sequence of cartesian coordinates. As already stated, this sequence is not at all unique because it depends on the scanning order of the single contour points and on the order the contour segments (if more than one) are concatenated. Supposing that the contour point scanning process on a single contour tract follows an adjacency criterion, the only signal discontinuities are generated by the coordinate jumps between subsequent contour segments (the coordinate differences between the end of a contour tract and the beginning of another one).
One can think wavelets are adequate to approximate discontinuities, however for accurate contour representation only moderate discontinuities can be tolerated

in order to avoid the introduction of annoying artifacts. In fact, wavelet approximation smooths the signal discontinuity and can generate ringing artifacts near the discontinuity boundaries. The above effects cause false contour segment connections (due to signal smoothing) and a "crumbling" of the contour extremities (due to the "ringing") on the reconstructed contour map. These effects, other than be linked in their entity to the coding rate, are more or less visible depending on the value of the single coordinate jumps. The above observations are illustrated in Fig.1 where a detail view of a low bit-rate (1 bit/contour point) sequence coding is shown: the larger (rightmost) discontinuity corresponds to strong oscillations in the reconstructed sequence, while the smaller (leftmost) discontinuity is completely smoothed and the corresponding contour tracts turn out to be joined. These impairing effects should be kept under control and there-

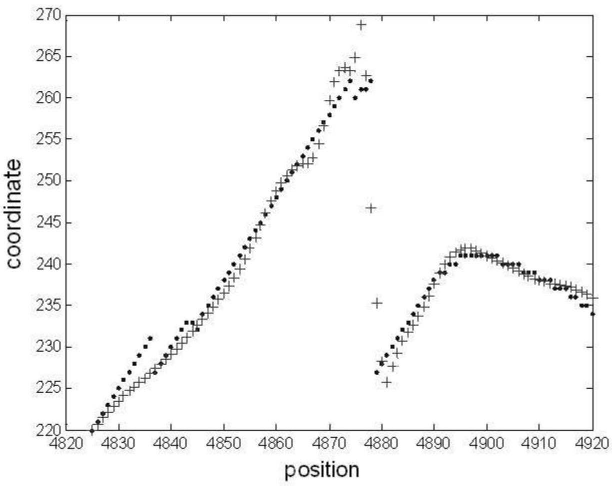

Fig. 1. Particular of a sequence of coordinates: with symbol "·" the original 1D signal sequence, with "+" a 1*bit/contour point* coded one.

fore the discontinuity amplitudes on the 1D signal sequence must be controlled and possibly minimized. To do this we adopt specific solutions for handling and sequencing disconnected contour segments in order to improve both objective and visual coding performances.

2.2 Creation of the One-dimensional Sequence

Discontinuity handling has been achieved by the following two steps:

a. *global discontinuity minimization*: it consists in creating the 1D signal and finding an optimal or sub-optimal sequencing of individual contour segments,

b. *local discontinuity control*: it consists in treating the coordinate discontinuity between consecutive segments as side information when their amplitude exceeds a certain threshold.

The aim of the *global discontinuity minimization* (step a.) consists in minimizing the "total discontinuity path" which is covered by the coordinate signal. This mean finding a contour segment sequencing mechanism which minimizes the global amount of coordinate jumps. This problem can be set as a typical "Traveling Salesman Problem", with the inherent unmanageable complexity in finding the optimal solution. Sub-optimal suitable heuristics, which lead to the generation of contour segment linking paths, has been found and tested. Here we briefly describe the adopted solution which has been experimentally selected to be the most effective among other similar ones:

1. The first considered contour point is selected as the leftmost-upper one, and its position is stored.
2. according to a minimum path length criterion[1], the coordinates of the point that is nearest to the last added one are concatenated iteratively and generate two integer valued sequences $x(n)$ and $y(n)$ corresponding to the various abscissa and ordinates respectively; the total path T is computed as the sum of the euclidian distances (hereafter referred as "jumps") among adjacent contour points
3. The extremities $(n, n+1)$ of the greatest not yet considered jump are found, which correspond to two coordinates pair $(C(n), C(n+1))$, where a coordinate pair is defined by $C(n) = (x(n), y(n))$.
4. The right-hand sub-sequence $(C(n+1), ..., C(N_{right}))$ is found with $N_{right} > n+1$ such that $C(N_{right})$ is the contour point closest (in terms of euclidian distance) to $C(n)$; then the above sub-sequence is reversed (flipped) and the resulting new total path T_I is computed; in Fig. 2 this process is shown where $C(n) = A$, $C(n+1) = B$ and $C(N_{right}) = D$.
5. A dual processing with respect to 4. is performed for a left-hand sequence $(C(N_{left}), ..., C(n))$; in this case the total path T_{II} is computed.
6. $T_m = min(T, T_I, T_{II})$ is selected along with the corresponding coordinate sequences; one can observe that by maintaining the synchronization of the inversions on the abscissa and ordinate sequences no side information is needed when T_I or T_{II} are selected.
7. The process iterates on i from step 3., until, a stop criterion is achieved, for example, for a selected ϵ, $T_m(i-1) - T_m(i) < \epsilon$.

At this point (step b.), a *local control of the residual discontinuities* is needed. If a discontinuity is higher than a properly defined threshold it is convenient to remove it from the sequences of the coordinates. The positions of the discontinuities and the offsets are lossless entropy coded, memorized in a header of

[1] The nearest contour point which abscissa and ordinate are concatenated to the respective sequences is found by using a minimum 2D euclidian distance criterion. In case of two or more contour points at the same euclidian distance the contour point which better preserve the current contour direction is seleted.

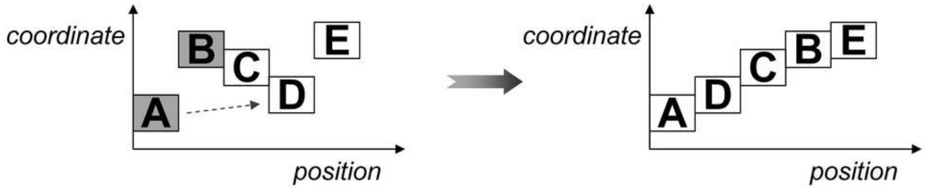

Fig. 2. Original sequence (on the left) and new generated sequence.

the bit-stream, and sent as side information to the decoder. In addition, the first point immediately subsequent to a discontinuity is removed from the sequence of contour point because it is redundant. The offset is then subtracted from the coordinates of all points which follow the same discontinuity. The signal sequences of abscissas and of ordinates are also concatenated in order to obtain a single 1D signal. The threshold that is used for the local control of the residual discontinuities should be calculated to be optimal or near optimal in a rate-distortion (R-D) sense with respect to a proper distortion metrics. Because of its "structural" nature and of its direct influence on the visual results, an experimental threshold (e.g. found to be optimal with respect to a set of contour maps) could be used, in a reliable way, as an "invariant" of the method.

2.3 Wavelet Encoding with 1D-I-SPIHT

The obtained 1D signal and its side information perfectly represent the contour map. The 1D signal is ready to be encoded in the wavelet domain. Because of the peculiarity of the produced signals we experimentally tested various wavelet filter basis. The 16/4 biorthogonal spline basis demonstrated slightly better coding performance with respect to others and therefore it has been used for the results presented in Sec.3.

Thanks to the fact that wavelet transform does not modify the signal support, the proposed contour coding actually preserves the original number of contour points on all decoded contour maps. This turns to be a nice property of our scheme and facilitates the evaluation of the quality metrics considered hereafter.

To encode the wavelet coefficients we used a one-dimensional and improved version of the SPIHT [6] algorithm. We adopted the algorithm called I-SPIHT which has already been tested for 2D and 3D data [7]. In I-SPIHT some redundant bits (whose value can be deducted unambiguously) are removed and the arithmetic coding part has been improved. The interested reader can refer to [7] for a detailed explanation of the solutions adopted in I-SPIHT. The adaptation of the SPIHT or I-SPIHT algorithm to a one-dimensional transformed domain is straightforward and so details are omitted.

Heading and side information is entropy coded by using a simple Huffman encoder. The compressed header contains the total number of contour points, the 1D-I-SPIHT setup information and the x and y coordinates and amplitudes of the residual discontinuities. The size of the coded header on the total bitstream length depends on the amount and position of the contour tracts of the contour map to code. At 2 bit/contour point we experimentally determined a 1% ratio in case of shape coding while this ratio is about 20% in case of complex image contour map coding. The decoder first decode the header and then it progressively reconstruct an approximate version of the original shape or contour map.

3 Experimental Results

We tested the proposed technique on several contour-map data. Representative results are described here for shape coding and generic contour-map coding.

3.1 Shape Coding Results

For a performance evaluation on the shape coding case, we compare our results to those obtained with state-of-art techniques: the line-based methods "baseline-based" [2] and "vertex-based" [3] and the bitmap-based methods MMR (Modified Modified Read) [4] and CAE (Context-Based Arithmetic Encoding) [5]. In particular CAE is the solution adopted by the MPEG-4 standard. To evaluate the distortion on the reconstructed shape, we evaluate the D_n and D_p quality measure. These measures have been defined in MPEG-4 and used for shape coding performance evaluations [8]. D_n represents the number of erroneously represented pels of the coded shape divided by the total number of pels belonging to the original shape. D_p is the peak deviation, measured in pixel, where the deviation is calculated as the euclidian distance between the center of mass of a reconstructed pixel and the center of mass of the nearest original pixel. As test data we used the first frame of the test sequence "Kids", in SIF format (352x240 pels). Fig. 3 shows the total bits required for coding the shape as function of D_n. We also limited our analysis to the condition $D_p \leq 3$ which determines the right end point of the various curves of Fig. 3. In fact, it has been evaluated that a $D_p > 3$ produces decoded shapes which are not suitable for video coding purpose [8]. For near-lossless rates the proposed technique loses efficiency. This problem is actually common to the considered contour-based techniques (and in general to most lossy contour coding algorithms). When a near lossless condition is required, the algorithms usually start to employ a chain-code. This solution could also be adopted for the described technique (or another solution explored). For lower bit-rates the proposed technique is more efficient than the 2 bitmap-based methods and also than the vertex-based technique; it approaches the performance of the baseline-based method that is currently the most efficient method reported in the literature. Moreover, the proposed technique reaches lower-rates and higher D_n without violating the condition $D_p \leq 3$ pels. This can be interpreted as a good artifact control and actually corresponds to good perceptual quality performance at low bit-rates. Fig. 4 shows in grey line the decoded shapes at the minimal allowed bitrate (702 bits, $D_p = 3$) on top of the original ones (in black line). Another important aspect is that the proposed technique is the only one that produces a progressively decodable bit-stream.

On this basis and other objective and visual tests we can conclude that, for the case of shape coding, our method presents performance comparable to state-of-art techniques and even at very high compression ratio (well under 1 bit/contour point) it doesn't introduce heavy or annoying artifacts while it preserves most details.

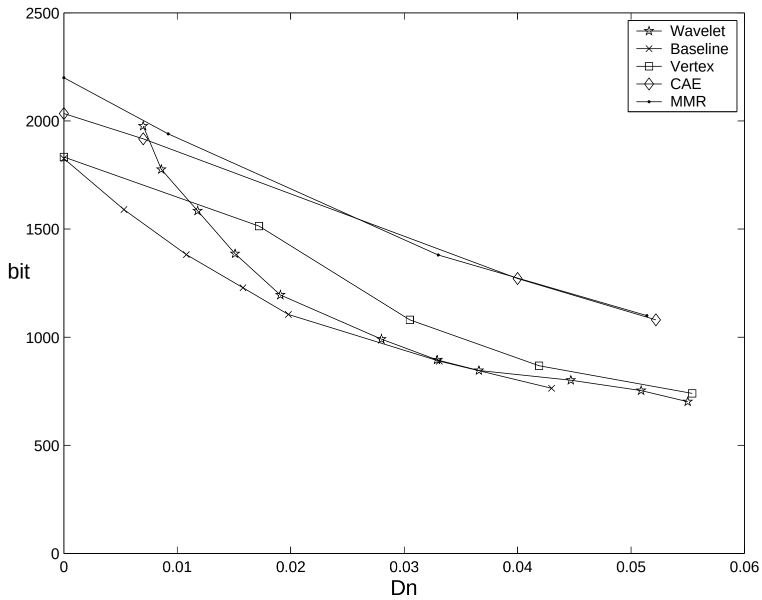

Fig. 3. Shape coding bits in function of D_n for the first frame of the sequence "Kids"; the proposed technique is labeled "Wavelet".

3.2 Generic contour map coding results

In this experiment we start from the contour map generated by the Canny method [9] on an image of "Audrey Hepburn". Coding results are shown in Fig. 5 in terms of E_m, i.e. the mean error (deviation) measured (in pixel/contour point). A *6 pixel distance* among the discontinuity coordinate extremities has been experimentally found a suitable threshold value for local discontinuity control (see Sec.2.2). In fact, it guarantees good visual quality performance and in most cases it minimizes (in a R-D sense) the mean error deviation E_m. Lossless rates are also reported as a reference for two lossless techniques: an 8-connectivity chain-code technique (line-based) [10] and the bitmap-based JBIG standard (bitmap-based) [11]. In this case multiple open contours must be represented and this generates a certain amount of side information. Visual results are very encouraging as shown in Fig. 6. Even at bit-rates of about 1 bit/contour point no annoying artifacts have been observed.

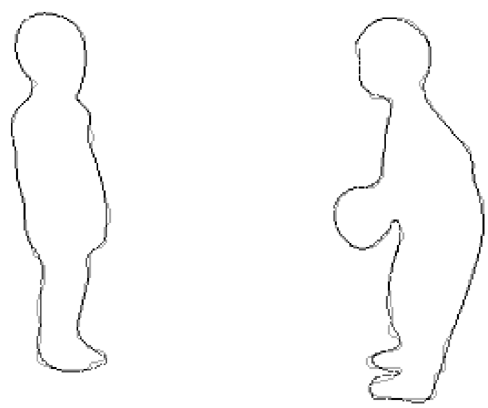

Fig. 4. Original shapes (black lines) and shapes encoded with 702 bits.

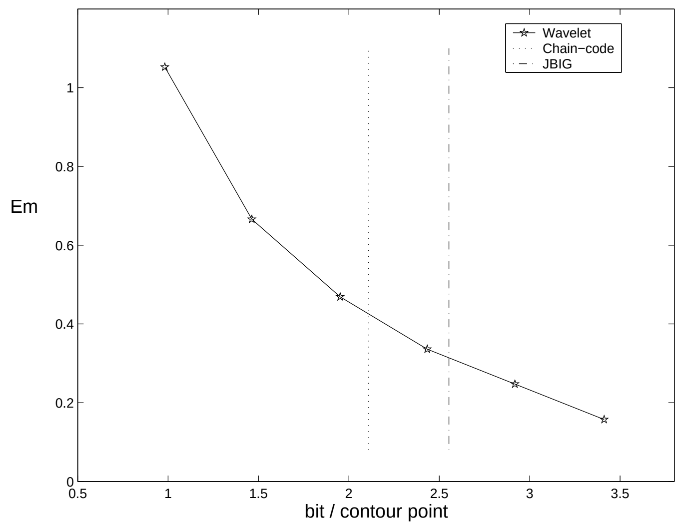

Fig. 5. Coding results in terms of Em for the "A.E." contour map. Lossless coding rates are indicated for the chain-code and JBIG techniques.

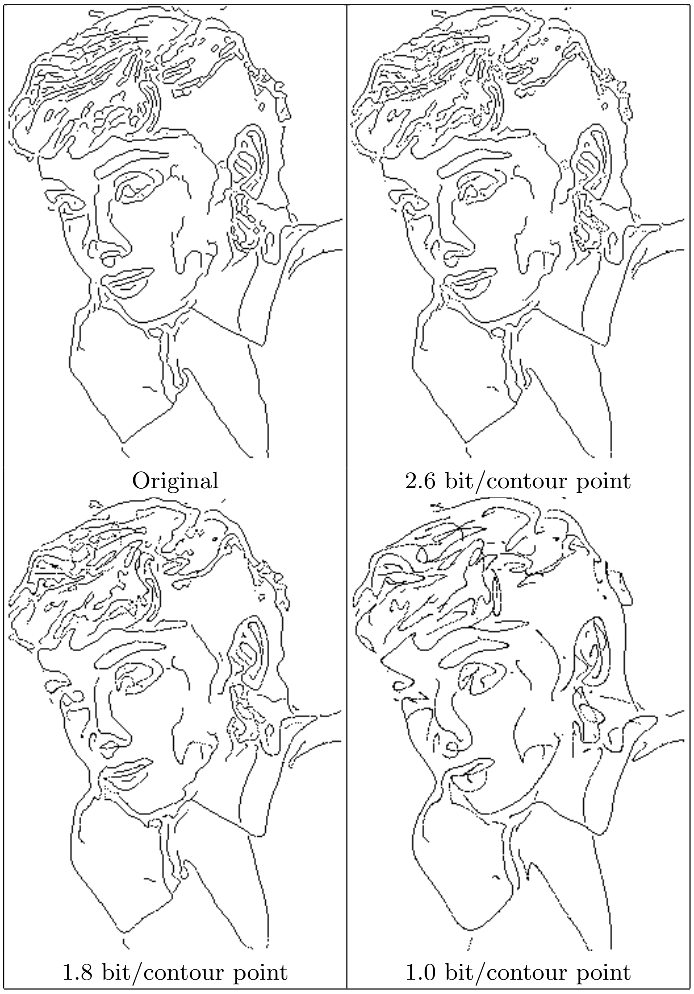

Fig. 6. Visual coding results for "A.E." contour map.

4 Conclusions

In this paper we proposed a new progressive shape coding technique based on the wavelet transform, suitable for any kind of contour images. The use of the wavelet transform, the unconstrained contour coding capabilities, the progressive structure of the coded bit-stream are the most original aspects of the proposed method, while the obtained coding performance makes this technique suitable for different image and video analysis and representation applications. The possibility of progressively decoding simple shapes or entire contour-maps can be useful or even required in many concrete situations (e.g. preview generation), it enables a fine-grain quality scalability and makes it easy to implement an unequal error protection. Progressive decoding is not possible in all the other considered techniques. In the case of shape coding, simulation results show that the proposed technique is as efficient as the best techniques reported in literature, while for application scenarios of lossy contour map coding perceptual results are already encouraging. Moreover, further optimizations of the entropy coding part are possible while other useful bit-stream properties, such as spatial scalability, seem to be achievable and are under consideration for future work.

References

1. Zahn, C. T., Roskies, R. Z.: Fourier Descriptors for Plane Closed Curves. IEEE Trans. Computers 21(**3**) (1972) 269–281
2. Lee, S., Cho, D., Cho, Y., Son, S., Jang, E., Shin, J., Seo,Y.: Binary Shape Coding Using Baseline-Based Method. IEEE Trans. Circ. and Sys for Video Technol. 9(**1**) (1999) 44–58
3. O'Connell, K.J.: Object-adaptive vertex-based shape coding method. IEEE Trans. Circ. and Sys for Video Technol. 7(**1**) (1997) 251–255
4. Yamaguchi, N., Ida, T., Watanabe, T.: A binary shape coding method using modified MMR. In: Proc. of Int. Conf. on Image Proc. (1997) 504–508
5. Brady, N., Bossen, F., Murphy, N.: Context-based arithmetic encoding of 2D shape sequences. In: Proc. of Int. Conf. on Image Proc. (1997) 29–32
6. Said, A., Pearlman, W. A.: A new, fast, and efficient image codec based on Set Partitioning in Hierarchical Trees. IEEE Trans. Circ. and Sys for Video Technol. 6(**3**) (1996) 243–250
7. Signoroni, A., Arrigoni, M., Lazzaroni, F., Leonardi, R.: Improving SPIHT-based compression of volumetric medical data. In: Proc. of Picture Coding Symp. (2001) 187–190
8. Katsaggelos, A. K., Kondi, L. P., Meier, F. W., Ostermann, J., Schuster, G. M.: MPEG-4 and rate-distortion-based shape-coding techniques. Proceedings of the IEEE 86(**6**) (1998) 1126–1154
9. Canny, J.: A computational approach to edge detection. IEEE Trans. on Pattern Anal. and Machine Intell. 8(**6**) (1986) 679–698
10. Eden, M., Kocher, M.: On the performance of a contour coding algorithm in the context of image coding - Part I: Contour segment coding. Signal Processing **8** (1985) 381–386
11. Kuhn, M.: JBIG-KIT lossless image compression library. http://www.cl.cam.ac.uk/∼mgk25/jbigkit/