

STRONG EDGE FEATURES FOR IMAGE CODING*

JOSEP R. CASAS and LUIS TORRES

Department of Signal Theory and Communications

Universitat Politècnica de Catalunya

Campus Nord UPC, edifici D5

Gran Capità s/n, 08034 Barcelona, Spain

E-mail: josep@gps.tsc.upc.es

Abstract. A two-component model is proposed for perceptual image coding. For the first component of the model, the watershed operator is used to detect strong edge features. Then, an efficient morphological interpolation algorithm reconstructs the smooth areas of the image from the extracted edge information, also known as *sketch data*. The residual component, containing fine textures, is separately coded by a subband coding scheme. The morphological operators involved in the coding of the primary component perform very efficiently compared to conventional techniques like the LGO operator, used for the edge extraction, or the diffusion filters, iteratively applied for the interpolation of smooth areas in previously reported sketch-based coding schemes.

Key words: morphological interpolation, perceptual models, feature-based image coding

1. Introduction

The coding of visual information cannot be performed efficiently without considering perceptual factors of the human vision [1]. Current standards for image compression [2] exploit some aspects of visual perception –for instance, in the design of quantization tables for DCT coefficients– but it is generally accepted that only the study of image models strongly related to the Human Visual System [3] will lead to the highest compression ratios needed for very low bit-rate applications. The so-called Second Generation models permit a graceful degradation of the perceived quality of reconstructed images, without the unnatural artifacts (blockiness, ringing and blurring) of waveform coding techniques at low bit-rates.

In a recent paper, Ran and Farvardin [4] stressed the role of strong edges in our perception of the images. They put forward a model for image compression based on a three component decomposition: namely a strong edge component, a smooth component and a texture component. The performance of such model was thoroughly investigated [5], proving its utility for most coding applications and showing a significant improvement over JPEG at low bit-rates. However, one of the drawbacks of their method, as well as of previously reported works in the same field [6, 7], is the large computation time spent in the edge extraction and interpolation processes applied to obtain the first and second components of the model. These are based on the solution of an energy minimization (variational) problem, governed by the heat or diffusion equation, that converges rather slowly to the final solution by means of an iterative space-variant filtering operation.

* This work was supported by the European RACE/Morpheco project 2053

Mathematical Morphology provides powerful operators to perform shape analysis. Morphological operators are very useful for the detection of edge features in ‘perceptually motivated’ Second Generation image coding applications, as has been shown in segmentation-based coding schemes [8]. Furthermore, a new morphological interpolation technique intended for the scattered data interpolation problem has been described in [9]. It has proven to be faster than linear diffusion techniques in order to generate the smooth component from the extracted edge features, with similar quality of the interpolation result.

The aim of this paper is to present a two-component model for perceptual image coding that strongly relies on morphological operators. For the first component, the watershed operator is used to detect strong edge features. Then, a morphological interpolation algorithm reconstructs the smooth textures from the edge information. This results in the *primary* component of the perceptual model, formed by strong edges and smooth textures. The residual component, or *texture* component, contains the fine textures and is separately coded by a subband coding scheme already presented in [10]. The morphological operators involved in the coding of the primary component perform very efficiently compared to previously reported techniques for edge extraction, like the LGO operator [6] or the diffusion filters iteratively applied for the interpolation of smooth areas in [6, 7, 4].

2. Strong Edge Extraction from the Morphological Laplacian

The morphological Laplacian, $L(f)$, is defined as the residue of the gradient by dilation, $g^+(\cdot)$, and the gradient by erosion, $g^-(\cdot)$, that is:

$$g^+(f) = \delta(f) - f \quad (1)$$

$$g^-(f) = f - \varepsilon(f) \quad (2)$$

$$L(f) = g^+(f) - g^-(f) \quad (3)$$

The morphological Laplacian is greater than zero at the lower edge of the transitions and smaller than zero at the upper edge. In flat surfaces or slanted planes without convexity changes, it cancels out. Indeed, it can be shown that the morphological Laplacian is an approximation of the signal second derivative.

The extrema of the second derivative locate the points with largest curvature values. These points occur at the upper and lower sides of the transitions bringing information about the transition width and the intensity change. The reader is kindly referred to the drawing in figure 1 (left) for an illustration of the one-dimensional case. In the two-dimensional case of figure 1 (right), the set of points where the morphological Laplacian reaches significant values mainly corresponds to the perceived image contours. The lines of largest curvature are called upper and lower *edge brims* by some authors [4]. Robinson [11] claims that these lines, which may be obtained as the ‘crest’ and ‘valley’ lines of a second derivative operator, are less noisy than Laplacian zero-crossings that follow the edge midpoints. Edge brims do not show so many random fluctuations because they do not represent a very rapid change in value with respect to position as edge midpoints do.

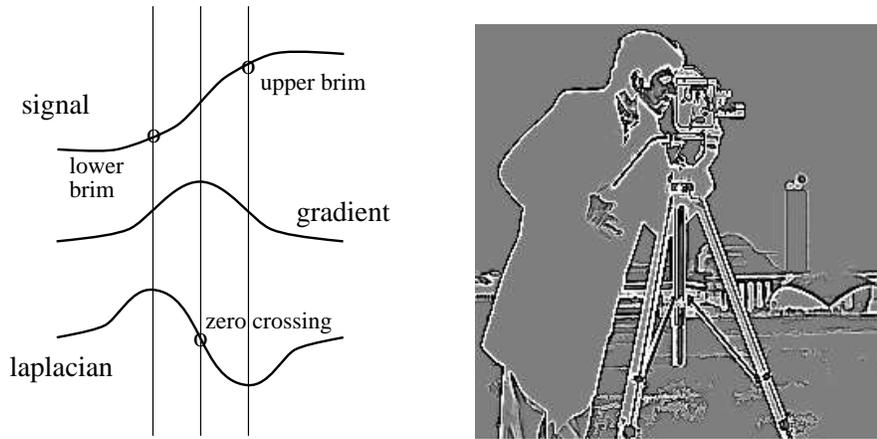


Fig. 1. Laplacian extrema locates edge brims: left, upper and lower brims; right, morphological Laplacian of the cameraman image (Note: mid grey corresponds to zero level)

A smooth approximation of an image may be obtained by interpolation from the set of pixels with large curvature values. The following experiment has been carried out in order to investigate its possible application to image compression. In the left image of figure 2, a set of pixels having large absolute values of the morphological Laplacian is shown. If we attempt to interpolate the remaining pixels of the smooth areas in between, the result will be the one presented in the right image –the interpolation technique will be briefly explained in the following section. About one tenth of the pixels of the image have been used as initial points for the interpolation algorithm. The peak to noise ratio of the interpolated image of figure 2 is only 23 dB but its subjective quality is not bad, because our attention is primarily drawn to the strong transitions which have been correctly placed and reproduced. Edge brims do look promising for a primitive-based characterization of visual information from a perceptual point of view. This experiment proves that it is possible to obtain a fair approximation of the original image from the amplitudes and positions of some pixels in the neighborhood of the strong edges. Moreover, the morphological Laplacian performs as an effective enhancement operator for the detection of such set of initial pixels.

Obviously, the application of this idea to image coding relies on the selection of a proper set of initial pixels for the interpolation process. The *initial set* should lead to a compact representation and, at the same time, allow a good approximation of the original image by interpolation. A good strategy could be to select the crest and valley lines of the Laplacian so that a chain code technique may be applied for the coding of the pixels' positions at low cost. The amplitudes of the initial pixels should also may be coded with a few number of bits. Given that intensity values along the edge brims should keep rather constant, some approximation may be employed to code these values within each brim line. In the current work, polynomial approximations will be used.

In the left image of figure 3, the white and black lines correspond, respectively, to



Fig. 2. Morphological interpolation from pixels with large Laplacian values: left, initial image (about 10% pixels); right, interpolation result

the crest and valley lines of the Laplacian or, likewise, to the positions of the lower and upper edge brims of the initial image. Edge brims may be detected by computing the watershed of the Laplacian and of its dual with an appropriate set of markers. In order to obtain the lower brims (crest lines of the Laplacian), the set of markers is formed by the union of two sets: the flat areas of the original image larger than a given size and some connected components with negative Laplacian values indicating the presence of valleys. For the upper brims (valley lines of the Laplacian), the second set is formed by connected components with positive Laplacian values indicating peaks and ridges. Please notice that some pieces of contour have been removed from the watershed result, either because the Laplacian was not significant enough at these positions or because the lines were too short. The necessary thresholds have been chosen on an empirical basis.

If the initial set is composed of the pixels at the positions indicated by the watershed lines shown in figure 3 with the approximated intensity values, the interpolation results in the right image of the same figure.

3. Reconstruction of the Strong Edge Component by Morphological Interpolation

The interpolation technique used in the previous examples has been already presented in [9] but, for completeness purposes, it will be briefly described in the present section.

Starting from the set of initial pixels the, *morphological interpolation* technique approximates the amplitudes of the unknown pixels of the image by an efficient algorithm that consists of two steps, namely a *geodesic propagation step* followed by a *smoothing step*. This two steps are successively iterated.



Fig. 3. Interpolation from lower and upper edge brims: left, brims' positions; right, interpolation result at 0.18 bpp

3.1. GEODESIC PROPAGATION STEP

In the geodesic propagation step, the amplitude values of the known pixels are propagated by geodesic dilation to fill the empty areas of the image. This is performed by a fast algorithm, using FIFO queues. Some intermediate images corresponding to the geodesic propagation step are shown in figure 4 for a synthetic initial image consisting of two small geometric figures.

3.2. PROGRESSIVE SMOOTHING STEP

At the positions where two or more propagation fronts originated from initial pixels of different amplitudes meet, the process stops and a false transition is created. The false transitions appearing outside the set of initial pixels will be smoothed in the second step. The morphological Laplacian is used as a transition detector in order to obtain these false transitions. Pixels on both sides of the false transitions compose the set of *secondary pixels*. A grey level value equal to the average of the intensity values on both sides of the transition is assigned to each secondary pixel. This is the smoothing step. Secondary pixels will be used in the next iteration of the algorithm in order to smooth out these transitions.

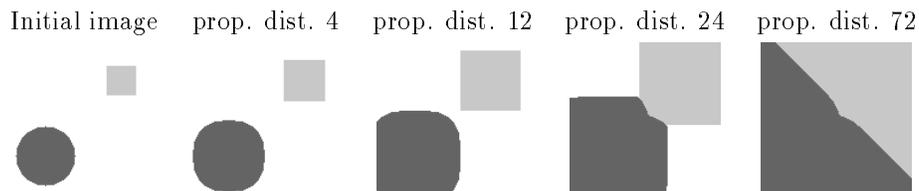


Fig. 4. Geodesic propagation step: initial pixels and some intermediate images

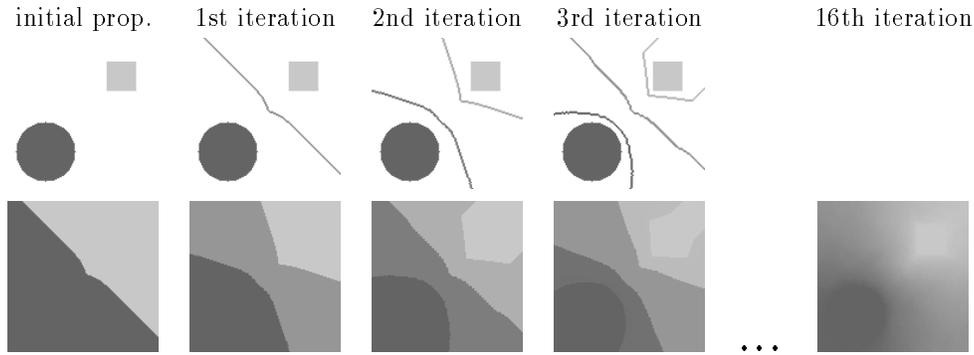


Fig. 5. Smoothing step: first row, initial and secondary pixels; second row: propagation

3.3. ITERATION

Next, a second iteration is performed: the propagation step propagates the grey level values from the sets of initial as well as secondary pixels. The propagation creates new false transitions which define a new set of secondary pixels where grey level values are smoothed. Note that this new set of secondary pixels generally does not include the first set of secondary pixels. This process of 1) propagation of values from the initial and secondary pixels, and 2) smoothing of the grey levels of the new secondary pixels, is iterated until idempotence. Figure 5 illustrates several iterations of the algorithm. After a few number of iterations, the algorithm quickly converges to the final interpolated image. For instance, only twelve iterations were necessary for the morphological interpolation result presented in the example of figure 3. The CPU time was of 9 seconds in a *SUN SparcClassic* workstation. Applying iterated space variant linear diffusion filters to solve the problem of interpolation from sparse sets, and even using multigrid techniques to speed up convergence, execution times about two orders of magnitude higher are reported [4].

4. Two-component Image Coding

The interpolation result of figure 3 (right) corresponds to the first component of our perceptual model, formed by strong edges and smooth areas generated from the initial pixels' positions and amplitudes. A derivative chain code is used to code the pixels' positions, whereas the amplitude values have been coded by polynomial approximation. More precisely, the network of brim lines is broken at each triple point (points with more than two branches). Then, the amplitudes of the pixels located under the resulting curves are approximated by a first order polynomial. The two coefficients defining each polynomial are quantized, entropy coded and transmitted. The overall bit-rate for this coded image is 0.18 bits per pixel. About 16% of this rate is spent in the coding of amplitudes, 70% for the chain-code information and the remaining 9% for the initial positions of each brim line.



Fig. 6. Texture component: left, coding residue; right, subband coded texture at 0.15 bpp

The coding residue of the first component –computed as the difference between the original image and the interpolation result– mainly consists of fine textures. This second component of the model is shown in the left image of figure 6. It lacks of significant transitions and may be approximated by a waveform coding technique. A coded reconstruction at low bit-rate (0.15 bpp) is shown in the right image of the same figure. It has been obtained by the application of the linear subband coding scheme presented in [10]. Information about the edge structure –available from the first component– is used for the texture coding of the second component, so that the masking effect of strong transitions may be considered. Amplitude errors in the neighborhood of these areas are less noticeable for the human eye than in other parts of the image [12]. Therefore, the quantization process is allowed to introduce large errors near the transitions by employing adaptive quantizers and bit allocations over arbitrarily shaped sub-edge regions in order to reduce the total number of bits.

5. Results and Conclusion

The compression ratio achieved with the above strategy is equal to 24 (0.33 bpp) for the addition of the strong edge and the fine-texture components of figures 3 and 6. The result is shown in the right image of figure 7. For comparison, the application of the JPEG standard [2] at the same bit-rate is shown in the left image of the same figure. The subjective quality of the described technique is significantly better because of the good rendition of the strong edges. The PSNR value (25.5 dB) is also larger than for the JPEG reconstruction (24 dB). The artifacts produced by the block-based DCT coding –blockiness in smooth areas and ringing in the neighborhood of strong transitions– are not present in the result of the two-component coding scheme. However, a different kind of visual artifacts may be observed. A certain smoothing effect is visible in some parts of the image and there are some missing objects, for instance the neck of the shirt has been almost removed.

The separate coding of strong edges permits the adaptation of the coding scheme to the visual perception of the images, avoiding unnatural degradations produced



Fig. 7. Results: left, JPEG at 0.35 bpp; right, described technique at 0.33 bpp

by waveform coding techniques at high compression ratios. A number of different artifacts are introduced by the new method at low bit-rates. It is hoped that such effects are more naturally perceived than those of waveform coding techniques by the subjective judgement of the observer. Furthermore, the application of morphological techniques for the extraction and reconstruction of the first component – watershed and morphological interpolation– perform very efficiently compared to previously reported techniques for edge extraction and scattered data interpolation.

References

1. N. Jayant, J. Johnston, and R. Safranek. Signal compression based on models of human perception. *Proceedings of the IEEE*, 81(10):1383–1421, October 1993.
2. G. K. Wallace. The JPEG still picture compression standard. *Communications of the ACM*, 34(4):30–44, April 1991.
3. M. Kunt, A. Ikonomopoulos, and M. Kocher. Second generation image coding techniques. *Proceedings of the IEEE*, 73(4):549–575, April 1985.
4. X. Ran and N. Farvardin. A perceptually motivated three-component image model. Part I: Description of the model. *IEEE Transactions on Image Processing*, 4(4):401–415, April 1995.
5. X. Ran and N. Farvardin. A perceptually motivated three-component image model. Part II: Application to image compression. *IEEE Transactions on Image Processing*, 4(4):430–447, April 1995.
6. S. Carlsson. Sketch based coding of grey level images. *EURASIP, Signal Processing*, 15(1):57–83, July 1988.
7. P. Grattoni and A. Guiducci. Contour coding for image description. *Pattern Recognition Letters*, 11:95–105, February 1990.
8. P. Salembier, L. Torres, F. Meyer, and C. Gu. Region-based video coding using mathematical morphology. *Proceedings of IEEE (Invited Paper)*, 83(6):843–857, June 1995.
9. J.R. Casas, P. Salembier, and L. Torres. Morphological interpolation for texture coding. In *IEEE International Conference on Image Processing*, volume I, pages 526–529, Washington DC, USA, October 1995.
10. J.R. Casas and L. Torres. A feature-based subband coding scheme. To be presented at ICASSP'96, May 1996.
11. John A. Robinson. Image coding with ridge and valley primitives. *IEEE Transactions on Communications*, 43(6):2095–2102, June 1995.
12. T.N. Cornsweet. *Visual Perception*. Academic Press, New York, 1970.