

Image Segmentation Techniques for Object-Based Coding

Junaid Ahmed, Joseph Bosworth, and Scott T. Acton

The Oklahoma Imaging Laboratory
School of Electrical and Computer Engineering
Oklahoma State University
{ajunaid,bosworj,sacton}@okstate.edu
<http://OIL.okstate.edu>

Abstract

Two image segmentation methods are presented and compared in terms of rate-distortion within an object-based coding scheme. The LOMO segmentation exploits the relationship between mathematical morphology and local monotonicity in producing a multiscale segmentation. The process is a morphological analogy to the Laplacian of Gaussian. The level set approach uses area morphology to generate segmented regions having a specified minimum area. Segments are optimally chosen from the connected components of the image level sets. A simple object-based coding scheme using the discrete cosine transform is used to avoid the artifacts produced by conventional block-based coding at segment boundaries. Results of each segmentation method are given and compared to one another and to conventional JPEG coding by rate-distortion and the presence of boundary artifacts.

1. Introduction

A dilemma within image segmentation research is the search for an effective measure of segmentation quality. Different methods of segmentation exist, utilizing different image characteristics, *e.g.* shape, texture, motion, etc. These methods perform differently depending on the application and are often compared only subjectively. One important application is clear, however.

This work was supported by the National Aeronautics and Space Administration under EPSCOR grant NCC5-171.

With the emergence of new standards such as MPEG-4, object-based coding is a rapidly developing field, allowing greater functionality than previous encoding methods. We therefore contend that segmentation quality can be measured, in this context, by the performance in such coding schemes, *i.e.* by rate-distortion.

Although no standard segmentation technique is widely accepted for object-based coding, the coding results depend critically on the segmentation. To complicate matters, different algorithms have been developed for the object-based coding itself. Here we present one such method, using a combination of block-based classification, JPEG compression and the method of successive projection onto convex sets as given in [1].

The remainder of the paper is organized as follows. In Section 2 we discuss the object-based coding method technique. In Sections 3 and 4, two novel grayscale segmentation methods are given: locally monotonic (LOMO) segmentation and level set-based segmentation. Finally, in Section 5, sample results are given, along with the corresponding rate-distortion measures. The results compare favorably with those of JPEG in terms of boundary quality and object-based functionality.

2. Object-Based Coding Theory

In the object-based coding method used in this investigation, rectangular blocks are used to provide boundaries for each segmented object. These sub-blocks are then divided into 8x8 pixel sub-blocks for further processing. Three types of sub-blocks are possible. First, sub-blocks that are not part of the segment are ignored (not coded). Second, we have blocks in which all constituent pixels are part of the object; these are coded in a manner similar to the standard JPEG, using the discrete cosine transform (DCT) coefficients and quantization. The third class includes sub-blocks containing segment boundary pixels, requiring special processing to ensure the preservation of the segment boundary.

The objective of processing sub-blocks of the third type is the reduction of the number of coefficients needed to represent the object and to preserve the edges. We use the successive projection onto convex sets approach. The first convex set can be defined using the energy compactness property of the DCT and the fact that the effect of the boundary will be mostly within the high frequency components of the transform. We set the high frequency coefficients to zero and retain only the low frequency coefficients.

The second convex set can be obtained from the boundary of the object in that particular sub-block. In an iterative manner, pixels that are part of the object are restored to their original values, while those outside of the object are given by the reconstruction defined by the DCT coefficients. The sub-block is then successively projected onto the convex sets in this manner, reducing the high frequency components inherent in the boundary. The coefficients are then encoded using the resultant DCT coefficients.

Finally, the boundaries of each object are encoded using a lossless chain-code. Thus, boundary information is encoded separately from intra-segment DCT coefficients.

3. LOMO Segmentation

Here, we present the first of two segmentations for use in the object-based coding scheme. This segmentation is derived by a linear combination of morphological filters, and is designed to generate signals with the property of local monotonicity. We now summarize the motivation behind this segmentation method and give an outline of the algorithmic steps of its implementation.

In the one-dimensional (1-D) case, a signal is locally monotonic (LOMO) of a given degree or scale n if and only if it is either non-increasing or non-decreasing within every window of length n . This definition can be restated in terms of morphology by the statement that a discrete 1-D signal is locally monotonic of a given degree n if and only if it is a root signal of both the open (\circ) and close (\bullet) filters simultaneously, assuming a zero-valued symmetric structuring element k of length $n-1$. In [2], this definition is generalized from 1-D to higher dimensions by requiring that a LOMO- n signal f be a root signal of the filter:

$$f \leftarrow \frac{f \circ k + f \bullet k}{2} . \quad (1)$$

Such signals may be derived from general inputs by iterative application of (1) until convergence. In practice, LOMO signals can be obtained by the alternative filter:

$$f \leftarrow \frac{(f \circ k) \bullet k + (f \bullet k) \circ k}{2} , \quad (2)$$

which converges more rapidly to a root signal.

The LOMO segmentation method relies upon the generation of a multi-scale representation, or scale-space, of such locally monotonic images. Beginning with the original image, each scale-space layer is derived by the iterative application of (2) to the image of the previous scale. As described in [2], this filtering procedure possesses many desirable scaling properties, as well as edge-localization and robustness to noise.

At each given scale, edge detection is performed. The edge-detection employs an unbiased (self-dual) combination of morphological erode (\ominus) and dilate (\oplus) filters, referred to as the morphological Laplacian:

$$d = \frac{f \oplus k + f \ominus k}{2} - f , \quad (3)$$

where the structuring element k is circular of diameter $2n-3$ (guaranteeing the equivalence to a discrete second derivative approximation in the 1-D case [2]). Zero-crossings of this difference function d represent edges, and (3) can be viewed as an analogy to the linear Laplacian of Gaussian (LoG) edge detection: a morphological approximation of a Laplacian operator acts on an appropriately chosen layer of a morphological scale-space.

Insignificant edges are removed by a threshold on the morphological gradient (structuring element diameter $n-1$). Edges detected at a given scale n are then dilated by a circular structuring element of diameter $n-1$ and then thinned. This process alleviates a difficulty of zero-crossing detection by closing edge gaps occurring at junctions between regions. With each single-scale segmentation complete, segments are then linked through scale from coarse to fine. A given segment is associated with the next-higher layer segment possessing the greatest spatial intersection. In this manner, the boundaries of an initial coarse-scale segmentation are refined to finer and finer detail, until the final segmentation is reached.

4. Area Morphology and Level Set-Based Segmentation

A disadvantage of segmentation approaches such as anisotropic diffusion or the watershed algorithm is the inability to prescribe the minimum region *area* in the segmentation. These segmentation techniques may be sensitive to small regions of high contrast. Using an area morphology approach, we can extract segments from the image that exceed a minimum area.

Area morphology is based on the area of connected components (regions) in the *level sets* of an image [3]. In a threshold decomposition of the image \mathbf{I} , an associated level set $L(\mathbf{I}, t)$ is a set of pixels that meet a given threshold t : $(x, y) \in L(\mathbf{I}, t)$ if $I(x, y) \geq t$. For a

discrete range of K intensities,

$I(x, y) = \sum_{i=0}^{K-1} 1_{(x,y) \in L(\mathbf{I}, t)}$ where $1_{\{\cdot\}}$ is the set indicator function.

The area open operator, denoted by $\circ^s(\mathbf{I})$, removes all connected components within the level set $L(\mathbf{I}, t)$ that do not have a minimum area of s . An area close operator,

denoted by $\bullet^s(\mathbf{I})$, removes connected components in the complemented level set $L^c(\mathbf{I}, t)$ that do not possess the minimum area s – area close removes small regions of zeros. Applied to each level set in the image (for each intensity in the range of possible intensities), area open removes small bright objects, while area close removes small dark objects.

The area open-close (AOC) operator, denoted by $\bullet^s \circ^s(\mathbf{I})$, is defined as the concatenation of the area open and close operators. We use the AOC operator to construct a scale-space for level set-based image segmentation. Processing can be implemented independently on each image level set; the grayscale image is reconstructed by a standard stacking operation after processing the level sets.

Given the AOC-scaled image, a segmentation may be derived. In such a segmentation, only the connected components within image level sets (or complements of level sets) can define a segment. The scale/area parameter of the AOC scale-space representation defines the minimum area of a segment. The segmentation algorithm finds the optimal combination of segments that minimize the summed internal intensity variance of each segment. So, the goal of the level set approach is to provide a segmentation that maximizes region homogeneity. The approach guarantees closed regions and boundary preservation, which are important features in object-based coding.

5. Results and Conclusions

For the original 256x256 pixel 8-bit grayscale images of Figures 1, 4, and 7, we show example results for each segmentation method. Figures 2, 5, and 8 show a LOMO segmentation and Figures 3, 6, and 9 show the result of level set segmentation method. Figure 10 shows the rate-distortion curves corresponding to the ‘Old Central’ image (Figure 1) for the two segmentation methods and JPEG compression. Similarly, Figures 11 and 12 show the rate-distortion curves for the ‘cameraman’ and ‘swan’ images.

These rate-distortion curves show that level set method gives better compression than the LOMO segmentation method. This difference is more evident in the curve for the ‘swan’ image and less distinct in the other two. Both object-based coded methods show poorer performance in the rate-distortion curves when compared to JPEG, but this is due to the fact that JPEG does not have the extra overhead of the segmentation. (Note that the compression here is due solely to quantization of DCT coefficients – no Huffman or run-length coding is used.)

For the object-based methods, some compression is necessarily compromised for the added object-based functionality and removal of visually undesirable artifacts at segment boundaries. Figure 16 shows a close-up of a region within the original image of Figure 13. Within the decoded JPEG image of Figure 17, blocking artifacts and sacrifice of edge localization are noticeable. These artifacts are avoided in the object-based results shown in Figures 14 and 15.

In conclusion, the level set segmentation method tends to outperform the LOMO segmentation in terms of rate-distortion within an object-based coding scheme. While the object-based method fails to give better overall compression than the block-based JPEG method, it provides superior boundary preservation and increased functionality for object-based searches.

6. References

- [1] H.H. Chen, M.R. Civanlar, and B.G. Haskell, “A block transform coder for arbitrarily shaped image segments”, *Proc. IEEE Int. Conf. on Image Processing*, Austin, Texas, November 13 - 16, 1994.
- [2] J. Bosworth and S.T. Acton, "Morphological Image Segmentation by Local Monotonicity," *Proc. of the Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, October 24-27, 1999.
- [3] D.P. Mukherjee and S.T. Acton, "Document page segmentation using multiscale clustering," *Proc. IEEE Int. Conf. on Image Processing*, Kobe, Japan, Oct. 25-29, 1999.

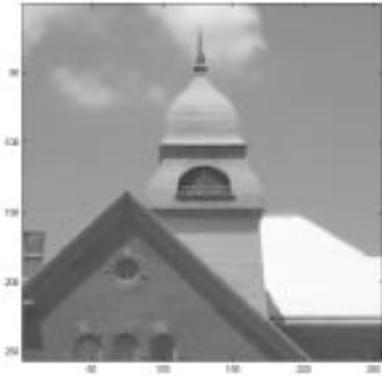


Figure 1: 'Old Central'



Figure 2: LOMO Segmentation

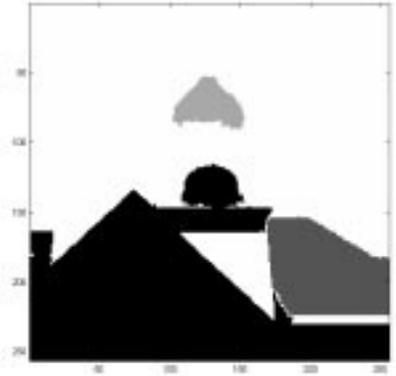


Figure 3: Level set Segmentation



Figure 4: 'Cameraman'



Figure 5: LOMO Segmentation



Figure 6: Level set Segmentation

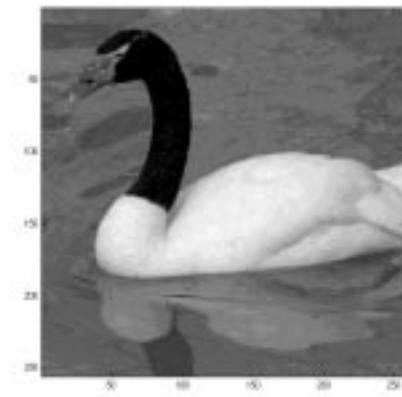


Figure 7: 'Swan'



Figure 8: LOMO Segmentation

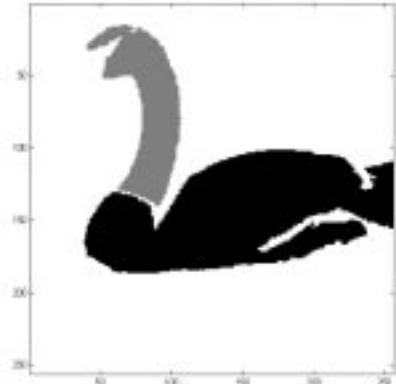


Figure 9: Level set Segmentation

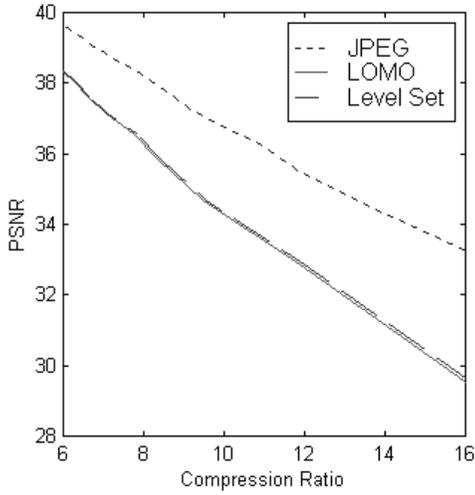


Figure 10: Rate-Distortion Curve for the Image 'Old Central'

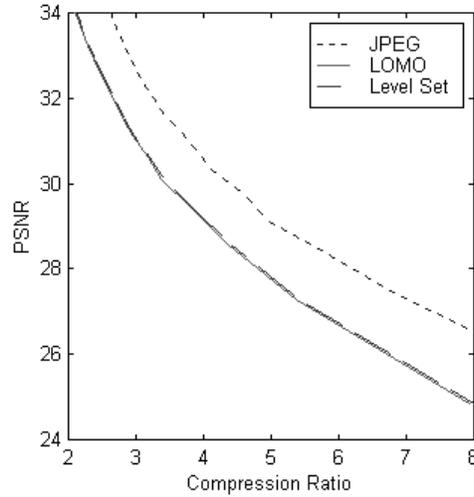


Figure 11: Rate-Distortion Curve for the Image 'Cameraman'

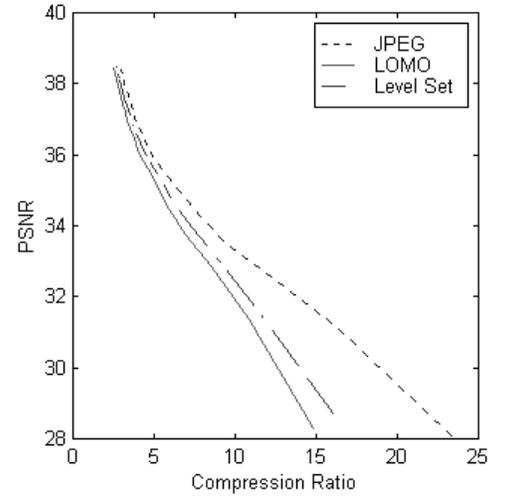


Figure 12: Rate-Distortion Curve for the Image 'Swan'

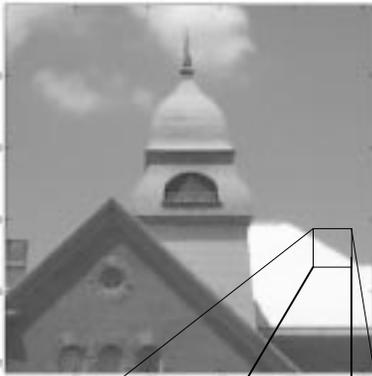


Figure 13: Original Image

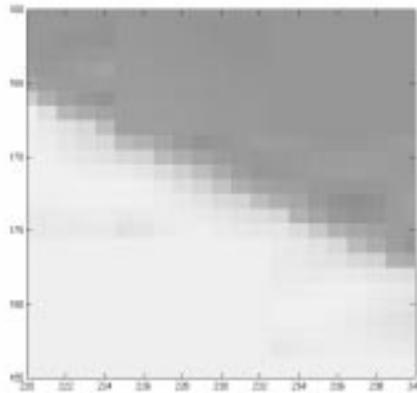


Figure 14: Close-up of decoded image using LOMO segmentation

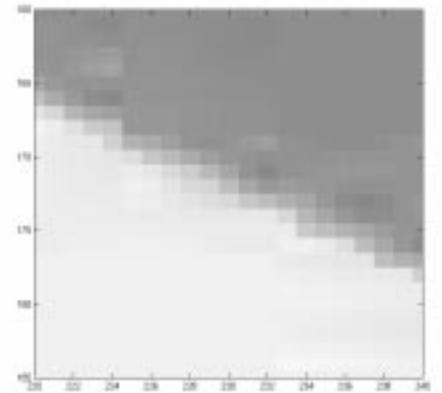


Figure 15: Close-up of decoded image using level set segmentation

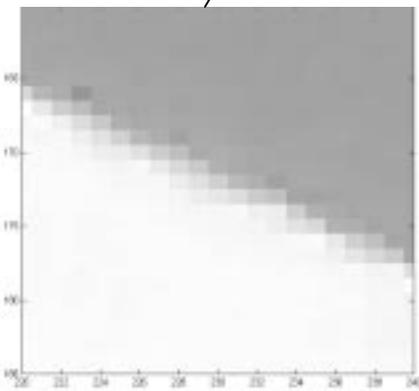


Figure 16: Close-up of Original

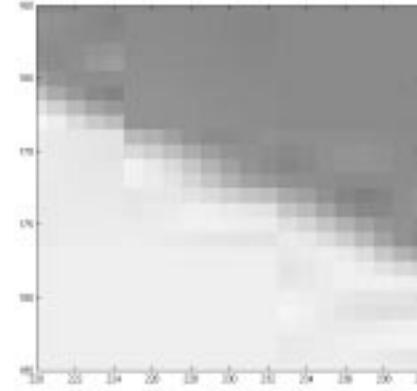


Figure 17: Close-up of decoded JPEG