Recent Advances in Object-Based Image Compression

Mark S. Schmalz

Center for Computer Vision and Visualization
University of Florida, Gainesville, FL 32611-6120
em: mssz@cise.ufl.edu, *ph*: 352-392-6831, *fx*: 352-392-1220

Image compression has progressed from pixel-level entropy coding, to transform coding applied to fixed regions (e.g., JPEG-2K with rectangular blocks). Of interest here are substitutions applied to arbitrarily-shaped regions. For example, object-based compression (OBC) segments disjoint, contiguous image regions, and compactly approximates each region's boundary and content (e.g., represented by a codebook exemplar). When region boundaries closely approximate scene object boundaries, OBC resembles object recognition. Disadvantages include cost of segmentation and region boundary representation, and pixel- or object-level distortion in the decompressed image.

Given a source image with values in $\mathbf{Z}_m$, let segmentation produce n contiguous regions, each having N_P boundary pixels enclosing N_A pixels. Let the source have $log(m)$ bpp, and let boundary pixels in the j^{th} segmented region be represented by $N_S(j)$ symbols, at an average of k bits per symbol. If each region is labelled with an N_i-bit tag (e.g., texture, color, intensity, and variance descriptors), and N_x bits point to a landmark of the region and describe rotation of the texture pattern, then the compression ratio is given by

$$CR = \frac{log(m) \cdot \sum_{j=1}^{n} N_A(j) + N_P(j)}{n(N_i + N_x) + k \cdot \sum_{j=1}^{n} N_S(j)} \quad , \qquad (1)$$

where quantities on the left- (resp. right-)hand side of the addition operations describe effects of region contents (resp. boundaries). Bit rate is given by $(CR \,/\, log(m))^{-1}$.

The author has developed efficient boundary coding algorithms reducing N_S to $\mathbf{O}(\log(N_P))$, thus increasing CR up to 2,500:1 for selected outdoor scenes. Fourier descriptors can also yield low N_S, and MPEG-4 and -7 support efficient boundary representation. Segmentation algorithms, while of primary interest to image understanding researchers, are key to the practical success of OBC. We analyze the complexity of Lee's and Munoz' segmentation methods [2-4] and relate these to Equation (1) in a time-space complexity metric. Additional metrics include semantic similarity between the source and OBC-decompressed image, which are especially useful for CR > 500:1. This would eventually support a time-space-distortion metric, for OBC transform optimization.

[1] Schmalz, M.S. "Object-based image compression", *Proc. SPIE* **4793**:13-23 (2002).

[2] Schmalz, M.S. and G.X. Ritter. "Region segmentation techniques for object-based image compression", *Proc. SPIE* **5561**:62-75 (2004).

[3] Muñoz, X., J. Freixenet, X. Cufi, and J. Marti. "Strategies for image segmentation combining region and boundary information", *Pat. Rec. Lett.* **24**:375-392 (2003).

[4] Lee, K.L. and L.H. Chen. "A new method for extracting primitives of regular textures based on wavelet transform", *International Journal of Pattern Recognition and Artificial Intelligence* **16**(1):1-25 (2002).

Proceedings of the 2005 Data Compression Conference (DCC'05)
1068-0314/05 $ 20.00 **IEEE**

IEEE COMPUTER SOCIETY