Robust Image Coding with Perceptual-Based Scalability
Marcia G. Ramos, Sheila S. Hemami
School of Electrical Engineering, Cornell University, Ithaca, NY 14853

We present a multiresolution-based image coding technique that achieves high visual quality through perceptual based scalability and robustness to transmission errors. To achieve perceptual coding, the image is first segmented at a block level (16 × 16) into smooth, edge, and highly-detailed regions, using the Hölder regularity property of the wavelet coefficients as well as their distributions. The activity classifications are used when coding the high-frequency wavelet coefficients.

The image is compressed by first performing a 3-level hierarchical decomposition, yielding 10 subbands which are coded independently. The LL band is coded using reconstruction-optimized lapped orthogonal transforms, followed by quantization, run-length encoding, and Huffman coding [1]. The high-frequency coefficients corresponding to the smooth regions are quantized to zero. The high-frequency coefficients corresponding to the edge regions are uniformly quantized, to maintain Hölder regularity and sharpness of the edges, while those corresponding to the highly-detailed regions are quantized with a modified uniform quantizer with a dead zone [2]. Bits are allocated based on the scale and orientation selectivity of each high-frequency subband as well as the activity regions inside each band corresponding to the edge and highly-detailed regions of the image. The quantized high-frequency bands are then run-length encoded.

The resulting compressed data is organized into a single bit stream from which images at different spatial resolutions as well as bit rates can be extracted in a straightforward manner. The LL band is given first, followed by all high-frequency information for the edge regions with increasing resolution, and finally all high-frequency information for the highly-detailed regions with increasing resolution. As such, extracted streams of progressively higher bit rates provide increasing visual quality. After all edge information is included, the predominant errors are in the highly detailed regions, such as patterns and textures, and are therefore less noticeable. Furthermore, the LL band coding technique facilitates reconstruction of this visually critical band in the event of data loss.

A comparison of the coding technique with JPEG was performed both by comparing the PSNR obtained for a desired bit rate for the full-resolution images and subjectively. While the PSNRs for the proposed technique were generally lower, the perceived quality was higher. In a subjective test involving 28 (non-image-processing) people and 10 images, the perceived visual quality of the images coded using the proposed technique surpassed that of the JPEG-encoded images for 74% of the comparisons at low bit rates. The proposed technique produces ringing artifacts at low bit rates, compared with severe blocking artifacts for JPEG encoded images. At high bit rates, JPEG and the proposed technique achieve similar results; however, the proposed technique allows scalability and provides robustness to loss of data at all bit rates.