Application of Single-Pass Adaptive VQ to Bilevel Images

Cornel Constantinescu and James A. Storer
Department of Computer Science, Brandeis University

Abstract: In past work (Constantinescu and Storer [CS94a, CS94b]), we introduced a new single pass adaptive vector quantization algorithm that maintains a constantly changing dictionary of variable sized rectangles by "learning" larger rectangles from smaller ones as an image is processed. For lossy compression of gray scale images, this algorithm with no advance information or training typically at least equals and often exceeds the compression obtained by the JPEG standard for a given quality. All of our past work with this approach has been with lossy compression of images where pixels are 8 or more bits.

Here we provide experimental evidence that our generic single pass adaptive VQ algorithm is highly effective for bilevel images. We examine not only lossless compression, but also very high quality lossy compression as well as mixtures of lossless and lossy compression applied to scanned images that contain text, gray scale images, and line drawings. New distortion measures are introduced for high quality lossy compressed bilevel images.

Although our generic algorithm does not always beat the best available methods on a specific type of data, it does well over diverse data sets, including documents with mixed types of data. In addition, our algorithm has the advantage that it can gracefully go between lossy, near lossless, and lossless compression. For example, on the standard CCITT image test set, we are slightly worse than JBIG (one of the best currently available lossless bilevel image compression standards) for pure lossless compression, but get about twice the compression of JBIG for near lossless compression (where visual quality of the decompressed image is nearly identical to the original). We have also experimented with an image that is a mixture of text and gray scale imagery; the following table compares the compression of our algorithm using high quality lossy on images and lossless on text to using JPEG on images (at the same SNR as our algorithm) and several different lossless compression standards:

<table>
<thead>
<tr>
<th>OUR (lossless on text, lossy on images)</th>
<th>JPEG on images, different lossless algorithms on text</th>
</tr>
</thead>
<tbody>
<tr>
<td>JBIG</td>
<td>G4</td>
</tr>
<tr>
<td>9.0</td>
<td>8.4</td>
</tr>
</tbody>
</table>

References


1068-0314/95$4.00 © 1995 IEEE