TAP Notes Report

IP_Number	**A20310**
Submitters	James E Beach; Krockenberger, Martin
IP Title	A SEGMENTATION ALGORITHM FOR MIXED RASTER CONTENT COMPRESSION SCHEMES
TAP Panel	20I
TAP Members	Marshall Bern, Doug Curry, Chris Dance, Leigh Klotz, Kris Popat, Ruth Rosenholtz, Doron Kletter, Dan Swinehart
TAP Date	04/05/2002
Rating	3(h)
TAP Comment	RATING: 3, with high potential business value.

IP/A20310 addresses the problem of rapidly generating a highly compressed image from an electronic source image, such as may be produced by an office application such as Word or PowerPoint. The most relevant application area in mind is print drivers, especially those that will be written to emerging compressed-image standards, which are being promoted by Xerox. The method applies best to images that contain little or no noise and often have quite sharp edges, although it works fine on all kinds of images. Most existing MRC algorithms binarize the image and then attempt monochrome segmentation, searching for features or connected components. By contrast, the proposed algorithm performs clustering in color space on small portions of the image, extending and modifying a similar Xerox CEF segmentation algorithm. The proposed algorithm has two major stages. First, each nxn region of the b-bit color input image is "downsampled" to m b-bit color pixels and (m-1) 1-bit, nxn pixel selectors (Advantageously, n=4, m=2, and b=24, but other choices are possible.) The color pixels and selectors are chosen to approximate the original image as closely as possible when combined. Optionally, at the end of the first stage, very high-frequency components, assumed to be half-tones, can be smeared out. Second, color pixels are exchanged between layers in a manner that optimizes the color layers for the best possible JPEG compression. As the exchanges are done, the selectors are modified so as to keep the image unchanged. Any color pixels that are completely masked off by their corresponding selector pixels do not affect the output image and so are set to optimize the JPEG algorithm.

The panel believes this is an effective method for achieving high compression for a large percentage of the documents that will be applied to the MRC-based printer drivers. Panelists concurred with the inventor that this invention is a useful combination of known techniques and a few improvements, particularly in the treatment of black and white as special points in the clustering and in the detection of invisible pixels and their treatment as don't-cares. Panelists cited as related work an ISIP2000 conference paper dealing with filling in holes in the background; the same material is believed to be covered in one of the cited patents. The panel determined that detection of other implementations of the exact technique would be possible, although other compression methods are clearly possible. In particular, JPEG2000 may emerge as a strong competitor. Perhaps the proposed method can be applied in that context as well.

Status Comment	None

Monday, April 08, 2002