

# JPEG-MATCHED DATA FILLING OF SPARSE IMAGES

*George Pavlidis, Sofia Tsekeridou and Christodoulos Chamzas*

Democritus University of Thrace, Vas. Sofias, 67100 Xanthi, Hellas  
{gpavlid;tsekerid;chamzas}@ee.duth.gr

## ABSTRACT

To efficiently compress rasterized compound documents, an encoder must be content-adaptive. Content adaptivity may be achieved by using a layered approach. In such an approach, a compound image is segmented into layers and then, the appropriate encoders are employed to compress these layers individually. A major factor in using these standard encoders efficiently is to match the layers' characteristics to those of the encoder by using data filling techniques to fill-in the initially sparse layers. In this work we propose a sub-optimal non-linear projections scheme that efficiently matches the baseline JPEG coder in compressing background layers, leading to smaller files with better image quality.

## 1. INTRODUCTION

It is long recognized that the compression of images is content dependent. Today, a wealth of compression schemes is available to effectively deal with different kinds of images according to the characteristics of their content: we have bi-level image compression, continuous-tone image compression, as well as artificial image (graphics) compression. The most difficult case to compress is that of compound images. Compound images are images comprising text, graphics and continuous-tone images. Since efficient compression is specific to each kind of image an effective way of dealing with such images is the layered approach. According to the layered approach, a compound image is segmented according to some features into three or more layers. These layers can then be efficiently compressed by standard encoders, such as JPEG [1] or JPEG2000 [2] for color layers and JBIG [3] or JBIG2 [4] for the masks. This way the overall image compression and quality are optimized in comparison with the one-scheme-overall-compression approach.

Many works ([9]-[16],[24]-[28]) have already pointed out how this approach results to significantly better results both in compression and final image quality. The three-layer representation of a compound RGB color image leads to an augmentation of the data bit depth from 24 bpp to 49 bpp, since the representation involves two 24 bpp images (foreground/background) and one 1 bpp image

(mask)<sup>1</sup>. Although the scheme seems extremely redundant, it is in this representation that lays the key to efficient compression. Though the model seems simple, it has proven to be extremely powerful, and, in minor variation, is already successfully employed in several commercial products, such as DjVu [5] and Digipaper [6]. In these cases the model, although based on the same idea, is not compliant to any standard. It is since the release of ITU MRC Rec. T.44 (Mixed Raster Content) [7], that the representation of a decomposed compound document image got a standard form. Based on this Rec., L. Sharpe and R. Buckley propose a layered imaging architecture for the JPEG2000 compression scheme [8] in 2000, D. Mukherjee et al proposed a fully MRC-compliant JPEG-matched compression scheme in 2001 [9] and a fully MRC-compliant JPEG2000-matched scheme in 2002 [10]. In [11], K. Jung and R. Seiler propose an integrated segmentation and JPEG2000 compression system for compound documents. In [14] R. de Queiroz et al based on prior work by R. de Queiroz [12],[13] get deeper into the problem of segmentation and data filling, in order to provide with a complete system for document image segmentation and compression with optimized characteristics, based on the MRC representation.

It has soon been realized that, since one of the consequences of the segmentation is that the produced layers are actually sparse matrices, there has to be an algorithm to fill-in the missing parts, so that existing compression schemes could actually operate efficiently on them. There are many techniques available for this task, although some of them did not initially appear as solutions to this problem. An important aspect is that the approach of filling-in the missing parts should be "compatible" with the characteristics of the encoder that is to be used on each layer. Thus, the foreground and background layers have to be filled with specific color information, usually deduced from existing parts of the corresponding images, so that some sort of "continuity" is ensured in all areas of the resulting images that will be compressed with standard lossy/lossless coding schemes such as JPEG or

---

<sup>1</sup> It should be noted that in our experimentations the foreground was considered to be monochrome for simplicity and it was discarded; this way the total data bit depth becomes 25 bpp.

JPEG2000. As expected, the filling-in of these layers is not a trivial task. The focus in this work is on manipulating the *background* layer. This layer has usually much less sparsity and can be filled-in using various methods. In most published works on this subject, two approaches can be identified:

- **Reconstruction-oriented**, in which the data filling is driven by the notion of reconstructing an ideal or fictional model image; there are techniques from *image restoration* and *reconstruction* (under the scope of filling-in or substituting lost and destroyed original image parts, such as in [15]), method-specific techniques for *removal of compression artifacts*, such as in [16], *image inpainting* techniques [17],[18] (that propagate information from existing parts to the missing ones, in an iterative manner), *disocclusion* methods [19] (that exploit level lines in a neighborhood around the missing parts and try to extend them into the missing areas), *texture synthesis* approaches [20] (where missing parts of images are totally replaced by specific existing parts according to their neighborhoods), *hybrid image-inpainting – texture-synthesis* methods [21],[22], and methods dealing with the reconstruction of missing parts that were lost during a noisy transmission [23]

- **Strictly compression-oriented**, which are driven by the notion of significantly improving compression performance, without any reconstruction notions in mind; there are methods like in [24] where J. Huang et al proposed a layered method for compressing bank check images, in [25] where H. Chen et al proposed an iterative Projections Onto Convex Sets (POCS) [26] algorithm to encode arbitrary shaped image segments, the work of L. Bottou and S. Pigeon in [27] (migrating from the DCT to the wavelet domain), in [28] where R. Stasinski and J. Konrad, propose a variation to [25] (the bandwidth limitation is implemented in the frequency domain on an oversampled grid), the works of R. de Queiroz et al in [14] as well as in [11] and [13] where they propose a data filling scheme that targets specifically the JPEG encoder.

## 2. DATA FILLING USING NON-LINEAR PROJECTIONS

Inspired by the works in compression-oriented schemes, we devised a system that actually implements a non-linear successive projections algorithm as described in [25] and briefly presented in [13]. Our purpose was to use a standard compression scheme (JPEG) and see if the extension of the idea of POCS with the introduction of non-linearity by a quantizer could actually lead to better compression results.

The baseline JPEG encoder uses DPCM coding to reduce the amplitude of the quantized DC coefficients and Huffman coding of the DPCM differences and the quantized AC coefficients. From this point of view, it seems

straightforward to assume that data filling should be a process of predicting the DC as well as several AC values for each block, in order to make the image less demanding of bits during the entropy coding stage. The overall scheme is as follows: given an image and a binary mask (that marks pixels for foreground), the system:

- performs an initial data filling (prefilling) so as to provide with a good starting point
- compresses the prefilled image using JPEG at a given quality, and measures compression rate
- decompresses the image, and measures quality of the originally existing pixels (where the mask is false)
- replaces background (initially existing – where the mask is false) pixels with their initial values
- updates entirely-empty-blocks' DC values (using the DC of the previous block)
- iterates until convergence; convergence in this algorithm is met when no further changes in compression ratio and in quality (PSNR) are measured.

The system can be described by using two sets:

(a) the set of images represented by using a group of transform coefficients quantized by the JPEG quantizer  $q$ :

$$Q = \{ \tilde{f} \mid \tilde{F}(i, j) = f(i, j)/q(i, j), \quad \forall i, j \}$$

(b) the set of images whose pixel values are equal to the initially existing ones (as defined by the mask):

$$P = \{ \tilde{f} \mid \tilde{f}(x, y) = f(x, y), \quad x, y \notin M \}$$

where  $f, \tilde{f}, q, M$  are the original sparse image, the filled image, the quantizer, and the mask respectively,  $i, j$  are coordinates in the transform domain and  $x, y$  are coordinates in the spatial domain.

Obviously set (b) is a convex set, while (a) is not. Since this holds, the system cannot be described by the POCS theory and, therefore, cannot be guaranteed to converge to a global minimum. Despite the difficulties in theory we implemented the system, and found that the system always converges. Convergence speed as well as finding the global optimum depends a lot on the initial prefilling method used. As we are bound to sub-optimality, it is reasonable to assume that the starting point (or the prefilling method) is a key point to improve the approximation of the overall optimum. To verify this, we have tested thirteen prefilling methods in order to evaluate the effect of each algorithm on the efficiency of the overall system: one method that sets all masked pixels to zero (method 0) and twelve methods that fill-in according to an algorithm. As expected, method 0 was the worst case, leading to bigger image files with the lowest quality. Methods based on global features failed to capture local characteristics and therefore to match the encoder. Methods based on inpainting gave impressive results in terms of reconstruction but the compression efficiency was not as good as with the "adaptive" methods, where prefilling actually takes the given

encoder into account. The method proposed in this work gave the best results on the average (when considering both compression and quality gain). We named this method *DC-Huffman* since it fills-in with DC values, having the baseline JPEG's Huffman coder in mind. In FIG. 1 we show the overall system performance evaluation with respect to the prefilling methods for a typical test image and JPEG quality factor fixed at 50. In the figure, methods are sorted from left to right according to their descending order of efficiency. Improvement is expressed in terms of comparison with the case of using prefill method 0.

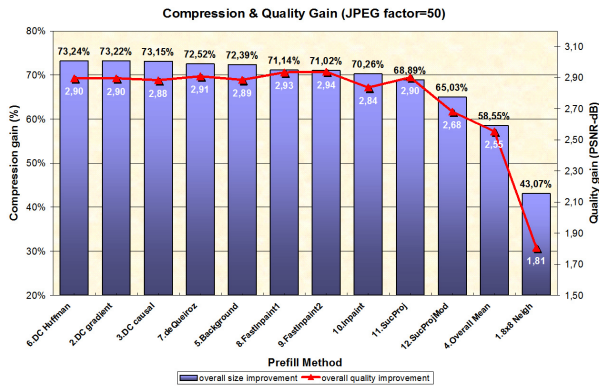


FIG. 1. Compression and quality gain of each prefilling method prior to projections compared with no-prefilling prior to projections

All quality measurement are expressed in PSNR terms (in dB); specifically, since color images are involved, a PSNR metric proposed in [29] is being used for simplicity, according to which:

$$PSNR = \frac{(4 \cdot PSNR_Y + PSNR_U + PSNR_V)}{6}$$

where  $PSNR_Y$ ,  $PSNR_U$ , and  $PSNR_V$  are the corresponding PSNR values of the Y, U, and V image components.

In FIG. 2 we show a typical example of a mixed image, a mask and the corresponding foreground and background images to be processed and compressed. In FIG. 3 we show a zoomed portion of the resulting images, after projections convergence, for the case of prefilling method 0 and method DC-Huffman. JPEG quality factor is fixed to 50. If we examine the images we can see that when using zero-prefilling the final image is not filled with data matching the encoder; since sharp edges are preserved, ringing effects are significant in the areas around the masked pixels, lowering the overall image quality. On the other hand, DC-Huffman prefilling leads to a smooth image without discontinuities and without any ringing artifacts. The objective and subjective quality of the non-masked background pixels is significantly improved.

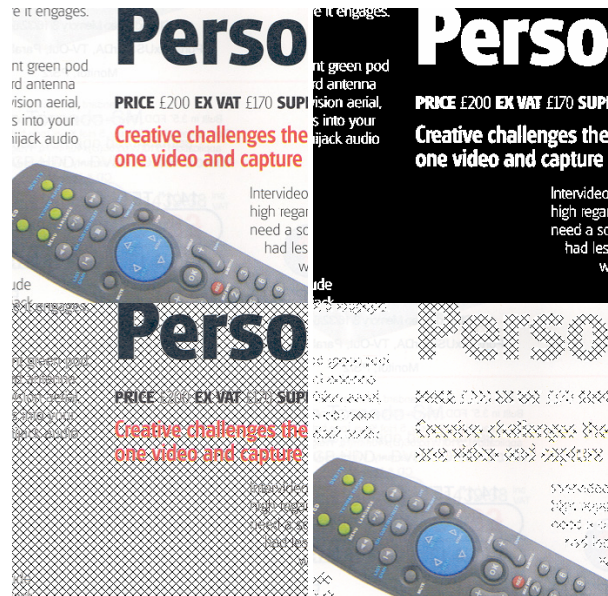


FIG. 2. Typical example of a mixed image, a mask and the corresponding foreground and background images (the crosshatched areas represent empty pixel positions)

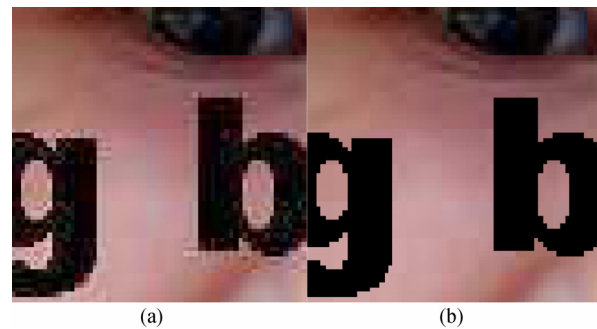


FIG. 3. Comparison between using no prefilling and prefilling with DC-Huffman: zoomed portion of image “fay” at (a) 24967 bytes (32,85 dB PSNR) and (b) 19479 bytes (33,58 dB PSNR)

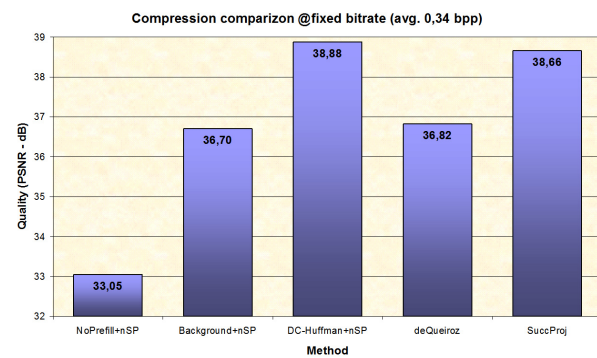


FIG. 4. Average quality for fixed bitrate

The algorithm was tested on a set of standard document images, both gray-level and color, and turned out to give similar results. The background was filled using five

methods and compressed with baseline JPEG. In FIG. 4 we show the average quality obtained for fixed bitrate (at 0.34 bpp). Results show that the proposed method achieves a better overall average image quality.

### 3. CONCLUSIONS

In the fast growing world of digital compound or documents images, efficient compression can be achieved by using layered schemes. These layers, defined by following the MRC model, can be coded with standard encoders and are expected to be of small amount of bytes and improved quality. An important aspect of the whole process is the correct data filling of the sparse layers, as this would ensure improvement both in compression and quality. In this work, inspired by the work on POCS, we proposed a non-linear projections scheme for data filling that matches the baseline JPEG coder and produces good compression results and improved image quality.

### 5. REFERENCES

- [1] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Compression Standard*, Van Nostrand Reinhold, 1993.
- [2] D. S. Taubman, M. W. Marcellin, "JPEG2000, Image Compression, Fundamentals, Standards and Practice", Kluwer Academic Publishers, 2002.
- [3] ISO/IEC International Standard 11544, ITU Recommendation T.82, JBIG, Progressive bi-level image compression, 1993.
- [4] ISO/CEI International Standard 14492, ITU-T Recommendation T.88, JBIG2, 2000 (E).
- [5] L. Bottou, P. Haffner, P. Howard, P. Simard, Y. Bengio and Y. LeCun, High quality document image compression using DjVu, *Journal of Electronic Imaging*, 7(3), pp. 410-425, July 1998.
- [6] D. Huttenlocher, P. Felzenszwalb and W. Rucklidge, Digi-paper: A versatile color document image representation, in *Proc. IEEE ICIP 1999*, Kobe, Japan, 1999.
- [7] ITU-T Recommendation T.44, Mixed Raster Content, 1999.
- [8] L. H. Sharpe II and R. Buckley, JPEG 2000 .jpm file format: a layered imaging architecture for document imaging and basic animation on the web, *Proceedings of SPIE*, vol. 4115, pp. 464-475, 2000.
- [9] D. Mukherjee, N. Memon, A. Said, JPEG-matched MRC compression of compound documents, in *Proc. IEEE ICIP 2001*, Thessalonica, Greece, pp. 434-437, 2001.
- [10] D. Mukherjee, C. Chrysafis, A. Said, JPEG2000-matched MRC compression of compound documents, in *Proc. IEEE ICIP2002*, Rochester, NY, USA, Sept. 22-25, 2002.
- [11] K. Jung and R. Seiler, Segmentation and Compression of Documents with JPEG2000, *First International JPEG2000 Workshop*, Lugano, Switzerland, July 8, 2003.
- [12] R. de Queiroz, Compression of compound documents, in *Proc. IEEE ICIP 1999*, Kobe, Japan, 1999.
- [13] R. de Queiroz, On data filling algorithms for MRC layers, in *Proc. IEEE ICIP 2000*, Vancouver, Canada, Sept. 10-13, 2000.
- [14] R. de Queiroz, Z. Fan and T. Tran, Optimizing block-thresholding segmentation for multilayer compression of compound images, *IEEE Trans. On Image Processing*, Vol. 9, No. 9, pp.1461-1471, Sept. 2000.
- [15] Y. Yang, N. Galatsanos and A. Katsaggelos, Projection-based spatially adaptive reconstruction of block-transform compressed images, *IEEE Trans. On Image Processing*, vol. 4, no. 7, pp. 896-908, Jul. 1995.
- [16] Y. Yang and N. Galatsanos, Removal of compression artifacts using projections onto convex sets and line process modeling, *IEEE Trans. On Image Processing*, vol. 6, no. 10, pp. 1345-1357, Oct. 1997.
- [17] M. Bertalmio, G. Sapiro, C. Ballester and V. Caselles, Image inpainting, in *Proc. Computer Graphics, SIGGRAPH 2000*, New Orleans, USA, pp. 417-424, Jul. 2000.
- [18] M. Oliveira, B. Bowen, R. McKenna and Y. Chang, Fast digital image inpainting, *Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001)*, pp. 261-266, Marbella, Spain, Sept. 3-5, 2001.
- [19] S. Masnou and J. Morel, Level-lines based occlusion, in *Proc. IEEE ICIP 98*, Chicago, IL., USA, Oct. 1998.
- [20] H. Igehy and L. Pereira, Image replacement through texture synthesis, in *Proc. Of IEEE ICIP 1997*, Santa Barbara, CA, USA, Oct. 26-29, 1997.
- [21] M. Bertalmio, L. Vese, G. Sapiro and S. Osher, Simultaneous structure and texture image inpainting, *IEEE Trans. On Image Processing*, vol. 12, no. 8, Aug. 2003.
- [22] S. Rane, G. Sapiro and M. Bertalmio, Structure and texture filling-in of missing image blocks in wireless transmission and compression applications, in *Proc. IEEE ICIP2002*, Rochester, USA, Sept. 22-25, 2002.
- [23] S. Shirani, F. Kossentini and R. Ward, Reconstruction of baseline JPEG coded images in error prone environments, *IEEE Trans. On Image Processing*, vol. 9, pp. 1292-1299, Jul. 2000.
- [24] J. Huang, Y. Wang and E. Wong, Check image compression using a layered coding method, *Journal of Electronic Imaging* 7(3), pp. 426-442, 1998.
- [25] H. Chen, M. Civanlar and B. Haskell, A block transform coder for arbitrary shaped image segments, in *Proc. IEEE ICIP 1994*, vol.1, pp. 85-89, 1994.
- [26] H. Bauschke and J. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review*, 38, pp. 367-426, 1996.
- [27] L. Bottou and S. Pigeon, Lossy compression of partially masked images, in *Proc. IEEE Data Compression Conference*, Snowbird, UT, 1998.
- [28] R. Stasinski, J. Konrad, POCS-based image reconstruction from irregularly-spaced samples, in *Proc. IEEE ICIP 2000*, Vancouver, Canada, Sept. 10-13, 2000.
- [29] L. W. Kang and J. J. Leou, A new error resilient coding scheme for JPEG image transmission based on data embedding and vector quantization, in *Proc. IEEE Int. Symposium on Circuits and Systems (ISCAS2003)*, Bangkok, vol. 2, pp. 532-535, Thailand, May 2003.
- [30] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62--66, 1979.