MRC Compression of Compound Documents using H.264/AVC-I

Alexandre Zaghetto and Ricardo L. de Queiroz

Abstract— The Mixed Raster Content (MRC) ITU document compression standard (T.44) specifies a multi-layer multi-resolution representation of a compound document. It is expected that higher compression can be achieved if more efficient compression standards are used to compress each layer. In this paper we present an MRC compound document codec that uses the H.264/AVC operating in INTRA mode to encode background/foreground layers and JBIG2 to encode the binary Mask layer. The result is an unrivaled performance for compressing compound documents as demonstrated by our experiments.

Keywords— H.264/AVC, JPEG-2000, Image coding, Compound documents, Mixed Raster Content.

I. INTRODUCTION

The Mixed Raster Content (MRC) ITU document compression standard (T.44) [1]-[6] specifies a multi-layer multi-resolution representation of a compound document. In this paper we present a basic 3-layer MRC codec that uses the H.264/AVC [7] operating in INTRA mode to encode background/foreground layers and JBIG2 [18] to encode the binary Mask layer. The main objective is not to propose a new layer separation nor a data-filling algorithm, but to show that MRC coding based on H.264/AVC and JBIG2 can achieve better compression rates than schemes that use other state-of-the-art still image coders.

A. H.264/AVC-INTRA

The newest video coding standard, the H.264/AVC [7], has been well explained in the literature [8]-[13]. Many papers have illustrated its performance showing many comparative results against coders such as MPEG-2. All results point to at least a factor of two improvement over previous standards. The reasons why the AVC is so good are many small improvements over previous methods. Each improvement brings a small coding gain, adding up to significant gains. Apart from the factor-of-two improvement over other standards, there are a few unexpected advantages that come with the AVC package.

H.264/AVC is a video compression standard and it was not conceived to be applied as a still image compression tool. Nevertheless, the many coding advances brought into H.264/AVC, not only set a new benchmark for video compression, but they also make it a formidable compressor for still images [14], [15]. One of the components of these advances is the intra-frame macroblock prediction method,

which, combined with the context-adaptive binary arithmetic coding (CABAC), turns the H.264/AVC into a powerful still image compression engine. If we set our H.264/AVC implementation to work on a sole "INTRA" frame it will behave as a still image compressor. We refer to this coder as AVC-I. The big surprise is that it also outperforms previous state-of-art coders such as JPEG-2000 [16]. This is a surprise to many because it was not meant to be an image coder at all. However, results are consistent and unison. Gains of the AVC-I over JPEG-2000 are typically in the order of 0.25dB to 0.5dB in PSNR (peak signal-to-noise ratio) for pictorial images [14], [15], [17].

B. JBIG2

JBIG2 [18] is an international standard for lossy and lossless compression of bi-level images developed by the Joint Bi-level Image Experts Group. Published in 2000 as the ITU-T Recommendation T.88, and in 2001 as ISO/IEC Standard 14492, JBIG2 outperforms JBIG1 [19] and Group 4/MMR [20] by at least a factor of 2 and 3, respectively [21] . Typically, a JBIG2 coder decomposes a bi-level image into regions (text, halftone and generic) and encodes each region using a different method.

C. Compound Documents

Electronic documents are basically represented in two forms: vectorial or raster [3]. It is not much of a challenge to compress vectorized documents since each object can be compressed individually and the whole file can be compressed losslessly. The real challenge is to compress rasterized documents. Compound documents are assumed here as raster documents which contain a mix of text and pictorial contents. Compression algorithms are developed with a particular image type, characteristic and application in mind and no single algorithm is best across all types of images or applications. When compressing text, it is important to preserve the edges and shapes of characters accurately to facilitate reading. The human visual system, however, works differently for typical continuous-tone images, better masking high frequency errors [22]. Fig. 1 shows an example of a compound document.

D. Mixed Raster Content

Compound raster documents have typically been compressed as a single image. However, different compression algorithms may be applied to each of the regions of the document. That is the way multiple-coder based algorithms work [3]. The mixed raster content (MRC) imaging model [1], [2],

Fig. 1. "Compound1": example of a compound document, assumed here as a raster document which contain a mix of text and pictorial contents.

Fig. 2. Illustration of MRC imaging model. The original document is represented using 3 layers: Foreground (FG), Background (BG) and Mask.

(a)   (b)   (c)

Fig. 3. Layer decomposition of "compound1" before data-filling procedure: (a) Foreground; (b) Mask; (c) Background.

has been proposed as a multi-layer multi-resolution representation of a compound document. The basic 3-layer MRC model represents an image as two image layers (Foreground or FG and Background or BG) and a binary image layer (Mask), which determines if a pixel belongs to BG or FG. Fig. 2 illustrates the described model. Once the original single-resolution image is decomposed into layers, each layer can be processed and compressed using different algorithms. Background and Foreground processing operations can include a resolution change and a data-filling procedure. The compression algorithm used for a given layer would be matched to the layer's content, allowing for improved compression while reducing distortion visibility. The compressed layers are then packed and delivered to the decoder. At the decoder, each plane is retrieved, decompressed, processed and the image is composed using the MRC imaging model. MRC has been proposed and/or accepted for several standards [1], [23], [24], [25] as has been used in several products [26]-[29].

II. DATA-FILLING

The first step of MRC compression is the layer segmentation algorithm [3], [30]. In this paper, we consider a basic 3-layer MRC model, which uses a BG, FG and Mask representation.

Once the image is segmented there will be "don't care" regions on BG and FG layers. Pixels assigned to the BG will be marked as "don't care" on the FG, and vice-versa. These pixels can be replaced by anything to enhance compression [3], [31], [32]. There are many methods for the replacement (data-filling). This paper uses the iterative wavelet-based plane filling proposed by De Queiroz [33], which we describe next.

Let F and B represent the pixel positions where the Mask indicates FG or BG respectively. First, we compute averages as:
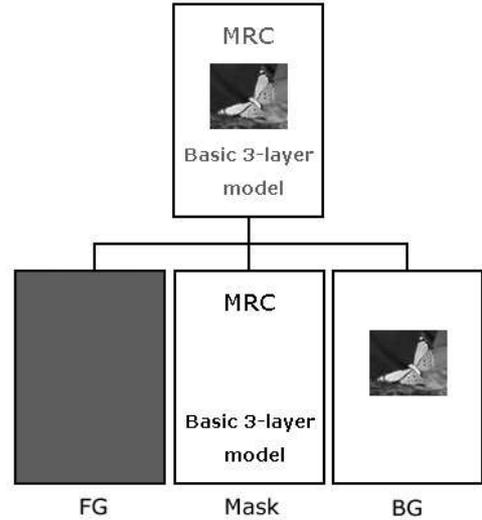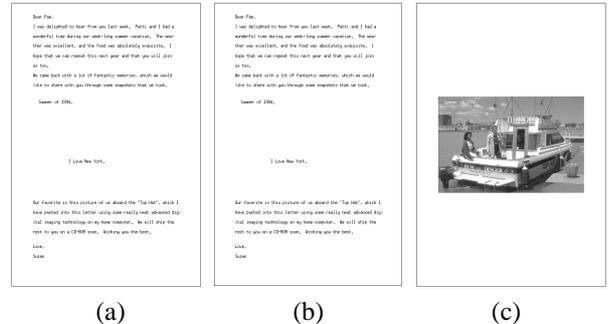
$$m_{BG} = mean(x(i,j)|(i,j) \in B)$$
$$m_{FG} = mean(x(i,j)|(i,j) \in F) \tag{1}$$

where $x(i,j)$ represents the original image.

Let I_0 be the starting FG plane with "don't care" pixels replaced by m_{FG}. Also, let $\tilde{I}_n$ be the compressed and decompressed version of I_n using a given coder at a target bit rate. If we plan to use a wavelet coder, $\tilde{I}_n$ can be approximated as:

$$\tilde{I}_n = W^{-1}(round(W(I_n)/Q) * Q) \tag{2}$$

where W denotes the Wavelet Transform of 5 or 6 levels, $round(.)$ is a rounding operator, and Q is a step size to quantize the wavelet coefficients. It is expected to use quite large Q numbers, like those that would yield very high compression ratios. Then, for $n = 0$ until $n = \nu$, where ν limits the number of cycles to the maximum of let us say 3 or 4, we compute

$$I_{n+1}(i,j) = \begin{cases} I_n(i,j) & : & (i,j) \in F \\ \tilde{I}_n(i,j) & : & (i,j) \in B. \end{cases} \tag{3}$$

We stop the loop either after ν cycles or when

$$mean(|I_n(i,j) - I_{n-1}(i,j)|) < \xi, \quad (i,j) \in B, \tag{4}$$
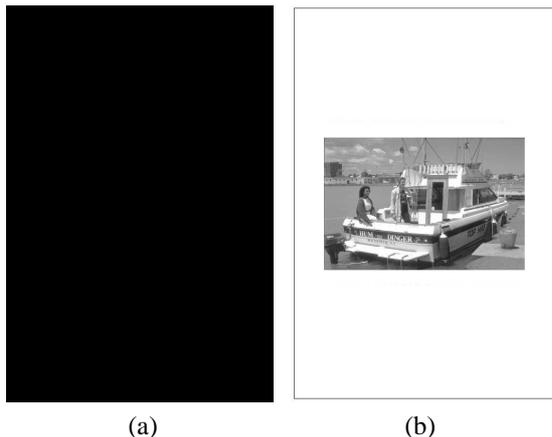
Fig. 4.   (a) Foreground and (b) Background of "compound1" after data-filling procedure.

Fig. 5.   Zoomed part of the Background of "compound1": (a) Original document; (b) After data-filling.

where ξ is some tolerance number, i.e. it stops when the filling in the "don't care" region converges. The same process applies to the BG plane, replacing foreground by background notation and vice-versa. Fig. 4 shows BG and FG of "compound1" processed using the above method.

Fig. 5 shows a zoomed part of the BG, where the effect of the data-filling algorithm can be observed more clearly.

III. RESULTS

The Imàge "Compound1" was encoded using AVC-I, JPEG-2000 and using the proposed MRC model. In MRC compression, the Mask was encoded using JBIG2. BG and FG were
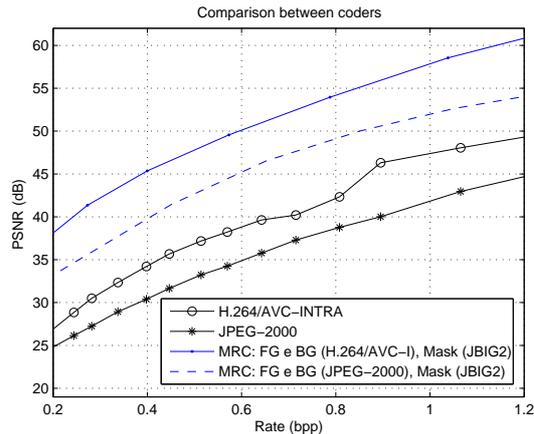

Fig. 6.   PSNR plots for "coumpound1", comparing: (a) AVC-I; (b) JPEG-2000; (c) MRC: FG and BG encoded using AVC-I and Mask encoded using JBIG2; (d) MRC: FG and BG encode using JPEG-2000 and Mask encoded using JBIG2. PSNR plots show that the MRC model based on H.264/AVC-INTRA outperforms the MRC model based on the state-of-the-art still image coder JPEG-2000 by more than 5 dB at 1 bit/pixel.
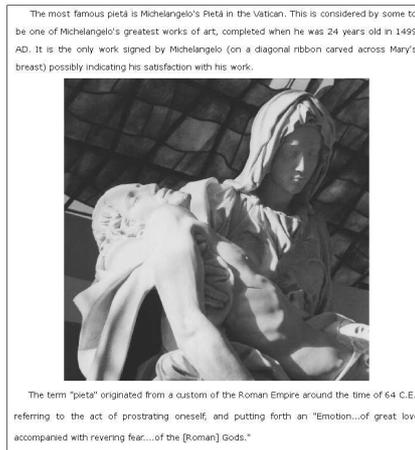

Fig. 7.   "Pietà": example of compound document.

both encoded using JPEG-2000 and AVC-I. PSNR plots are shown in Fig. 6.

The document shown in Fig. 7 was also compressed using the same encoders. PSNR plots are shown in Fig. 8.

AVC-I seems to have an extra capacity of adapting itself to heterodox content [17]. For the compound documents shown in Fig. 1 and Fig. 7, the PSNR gains are substantial, even surpassing the mark of 4 dB improvement over JPEG-2000 at 1 bit/pixel. In spite of this extra capacity of AVC-I, the multiple-coder MRC model proposed here offers results that outperform the AVC-I single coder approach, surpassing the mark of 4 dB improvement at 1 bit/pixel. PSNR plots shown if Figs. 6 and 8 also demonstrate that the MRC model based on AVC-I outperforms the MRC model based on the state-of-the-art still image coder JPEG-2000.

IV. CONCLUSIONS

In this paper, we presented a basic 3-layer MRC codec that uses the H.264/AVC operating in INTRA mode to encode
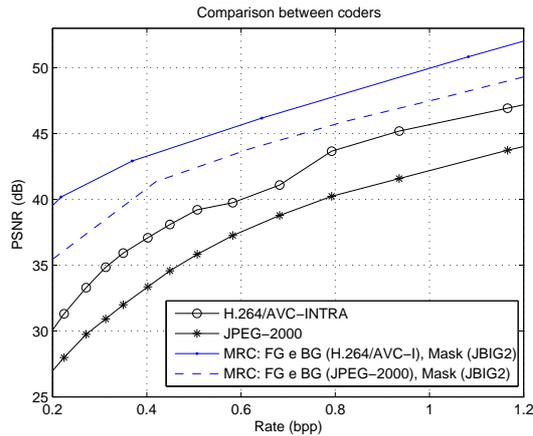
Fig. 8. PSNR plots for "pietà", comparing (a) AVC-I; (b) JPEG-2000; (c) MRC: FG and BG encoded using AVC-I and Mask encoded using JBIG2; (d) MRC: FG and BG encode with JPEG-2000 and Mask encoded with JBIG2. PSNR plots show that the MRC model based on H.264/AVC-INTRA outperforms JPEG-2000 by more than 2 dB at 1 bit/pixel.

BG/FG layers and JBIG2 to encode the binary Mask layer. Results show that with the MRC model it is possible to achieve better performance than single coder approaches, such as JPEG-2000 and AVC-I. Furthermore, using AVC-I to compress BG and FG yields better results than schemes based on JPEG-2000. Without a doubt MRC schemes based on AVC-I sets a new level of performance that is unrivaled by other standards. We just carried tests for computer generated images. Further developments and tests need to be done for scanned documents.

REFERENCES

[1] "Mixed Raster Content (MRC)", *ITU-T Recommendation T.44*, 1999.

[2] R. L. de Queiroz, R. Buckley, M. Xu, "Mixed raster content (MRC) model for compound image compression", *Proc. SPIE, Visual Communications and Image Processing*, Vol. 3653, pp. 1106-1117, Jan 1999.

[3] R. L. de Queiroz, "Compressing Compound Documents", in *The Document and Image Compression Handbook*, edited by M. Barni, Marcel-Dekker, 2005.

[4] D. Mukherjee, N. Memon, A. Said, , "JPEG-matched MRC Compression of Compound Documents", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Vol. 3, Thessaloniki, Greece, pp. 434-437, Oct. 2001.

[5] D. Mukherjee, C. Chrysafis, A. Said, "JPEG2000-matched MRC compression of compound documents", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Vol. 3, pp. 73-76, 2002.

[6] G. Feng and C.A. Bouman, "High-quality MRC document coding", *IEEE Trans. on Image Processing*, Vol. 15, No. 10, pp. 3152-3169, Oct. 2006.

[7] "Final Draft International Standard of Joint Video Specification", *ITU-T Recommendation H.264, ISO/IEC 14496-10 AVC*, March 2003.

[8] I. E. G. Richardson, *H.264 and MPEG-4 Video Compression*, Wiley, 2003.

[9] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp. 560-576, July 2003.

[10] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp. 657-673, July 2003.

[11] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding stan-

dards", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp. 688-703, July 2003.

[12] G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions", *Proc. of SPIE Conference on Applications of Digital Image Processing XXVII*, Special Session on Advances in the New Emerging Standard: H.264/AVC, August, 2004.

[13] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity", *IEEE Circuits and Systems Magazine*, Vol. 4, No. 1, pp. 7-28, March 2004.

[14] D. Marpe, V. George, and T. Wiegand, "Performance comparison of intra-only H.264/AVC and JPEG2000 for a set of monochrome ISO/IEC test images", *Contribution JVT ISO/IEC MPEG and ITU-T VCEG, JVT M-014*, Oct. 2004.

[15] D. Marpe, V. George, H. L. Cycon, and K. U. Barthel, "Performance evaluation of Motion-JPEG2000 in comparison with H.264/AVC operated in pure intra coding mode", Wavelet Applications in Industrial Processing, in *Proc. of SPIE,* Vol. 5266, pp. 129-137, 2004.

[16] D. S. Taubman and M. W. Marcellin, JPEG2000: *Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic, 2002.

[17] R. L. de Queiroz, R. S. Ortis, A. Zaghetto, and T. A. Fonseca, "Fringe benefits of the H.264/AVC", *Proc. of Intl. Telecom. Symp.*, Fortaleza, Brazil, pp. 208-212, Sep. 2006.

[18] "Information technology - Coded representation of picture and audio information - Lossy/lossless coding of bi-level images", *ITU-T Recommendation T.88*, March 2000.

[19] "Information technology - Coded representation of picture and audio information - Progressive bi-level image compression", *ITU-T Recommendation T.82*, March 1995.

[20] "Facsimile coding schemes and coding control functions for Group 4 facsimile apparatus", *ITU-T Recommendation T.6*, Nov. 1988.

[21] "JBIG Maui Meeting Press Release: New document compression standard quadruples compression of today's fax standards and runs at unprecedented speeds", *ISO/IEC JTC1/SC29/WG1*, Dec. 1999.

[22] D. Marr, *Vision*, San Francisco, CA: Freeman, 1982.

[23] R. Buckley, D. Venable and L. McIntyre, "New developments in color facsimile and internet fax", *Proc. of IS&T's Fifth Color Imaging Conference*, pp. 296-300, Scottsdale, AZ, Nov. 1997.

[24] IETF RFC 2301. File Format for Internet Fax. March 1998. ftp://ftp.isi.edu/in-notes/rfc2301.txt.

[25] ISO/IEC JTC1/SC29 WG1, JPEG 2000 Committee, Working Draft 2.0, June 25, 1999.

[26] D. Huttenlocher and W. Rucklidge, "DigiPaper: a versatile color document image representation", *Proc. IEEE Intl. Conf. Image Proc.*, 25PS1.3, Kobe, Japan, Oct. 1999.

[27] 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.

[28] M. Thierschmann, K. Bartel, S. McPartlin, U. Martin, "New technology for raster document image compression", *EI'2000, Document Recognition and Retrieval VII*, Proc. SPIE, Vol. 3967, Jan. 2000.

[29] J. Huang, Y. Wang and E. Wong, "Check image compression using a layered coding method", *Journal of Electronic Imaging*, 7(3), pp. 426-442, July 1998.

[30] R. L. de Queiroz, Zhigang Fan, T. D. Tran, "Optimizing block-thresholding segmentation for multilayer compression of compound images", *IEEE Trans. on Image Processing*, Vol. 9, No. 9, pp. 1461-1471, Sep. 2000.

[31] R. L. de Queiroz, "On data-flling algorithms for MRC layers", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Vancouver, Canada, Vol. II, pp. 586-589, Sep. 2000.

[32] G. Lakhani, R. Subedi, "Optimal filling of FG/BG layers of compound document images", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Atlanta, USA, pp. 2273-2276, , Oct. 2006.

[33] R. L. de Queiroz, "Pre-processing for MRC layers of scanned images", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Atlanta, USA, pp. 3093-3096, , Oct. 2006.