UNIVERSAL IMAGE CODING USING MULTISCALE RECURRENT PATTERNS AND PREDICTION

Nuno M. M. Rodrigues[1,2], *Eduardo A. B. da Silva*[3], *Murilo B. de Carvalho*[4],
Sérgio M. M. de Faria[1,2], *Vítor M. M. da Silva*[1,5]

[1]Instituto de Telecomunicaçoes, Portugal; [2]ESTG, Instituto Politecnico Leiria, Portugal;
[3]PEE/COPPE/DEL/Poli, Univ. Fed. Rio de Janeiro, Brazil; [4]TET/CTC, Univ. Fed. Fluminense, Brazil;
[5]DEEC, Universidade de Coimbra, Portugal.

e-mails: nuno.rodrigues@co.it.pt, eduardo@lps.ufrj.br, murilo@telecom.uff.br, sergio.faria@co.it.pt, vitor.silva@co.it.pt

ABSTRACT

In this paper we present a new method for image coding that is able to achieve good results over a wide range of image types. This work is based on the Multidimensional Multiscale Parser (MMP) algorithm [1], allied with an intra frame image predictive coding scheme. MMP has been shown to have, for a large class of image data, including texts, graphics, mixed images and textures, a compression efficiency comparable (and, in several cases, well above) to the one of state-of-the-art encoders. However, for smooth grayscale images, its performance lags behind the one of wavelet-based encoders, as JPEG2000.

In this paper we propose a novel encoder using MMP with intra predictive coding, similar to the one used in the H.264/AVC video coding standard. Experimental results show that this method closes the performance gap to JPEG-2000 for smooth images, with PSNR gains of up to 1.5dB. Yet, it maintains the excellent performance level of the MMP for other types of image data, as text, graphics and compound images, lending it a useful universal character.

1. INTRODUCTION

Despite the well known intrinsic limitations of the transformed-based image encoding methods, this class of algorithms is generally considered as state-of-the-art, both in image and video compression. These methods assume that an image has a low-pass nature and expect most of the transform coefficients for the higher frequencies to be negligible or of little importance. This is then exploited by using coarse quantization or by simply ignoring these frequency components. When used to encode images with frequency distributions other then low-pass, like text and graphics, the efficiency of these methods deteriorates noticeably.

The method presented in this paper is built upon an algorithm that is not based on the transformation-quantization-encoding paradigm. It is referred to as Multidimensional Multiscale Parser (MMP) [1], because it uses an adaptive dictionary of vectors to approximate variable-length input vectors. These vectors result from recursively parsing an original input block of the image. Scaling transformations are used to resize each dictionary element to the dimension of the block segment that is being considered.

In this work we introduce a new development of MMP, called MMP-Intra, that combines MMP with predictive coding techniques, like those used in H.264/AVC Intra coding prediction. Experimental results show that this new method is able to achieve relevant coding gains over the original MMP, making it competitive with state-of-the-art encoders, like JPEG2000 [2] and H.264/AVC [3]. When text, graphics and texture images are considered, then MMP-Intra achieves significant gains over the other tested encoders, which indicates that it has a universal character.

Unlike previous extensions of the MMP [4], MMP-Intra does not make any assumptions about the nature of the input image, achieving good coding results both for smooth grayscale images as well as for text, graphics and combined text and grayscale images.

In the next section the MMP algorithm for image coding is briefly described. Section 3 presents the MMP-Intra method, discussing the joint use of Intra-like prediction schemes and MMP. In section 4 some experimental results are shown and section 5 presents the conclusions.

2. THE MMP ALGORITHM

Although the MMP algorithm was initially proposed as a generic lossy data compression method, it is easily extendable to work with n-dimensional data, and has been successfully applied to image coding. In this section we describe the most important aspects of the MMP algorithm applied to image coding. An exhaustive description of the method can be found in [1].

MMP is based on approximations of data segments (in

0-7803-9134-9/05/$20.00 ©2005 IEEE

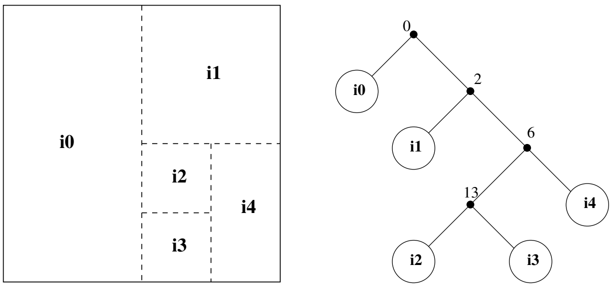

Fig. 1. Segmentation of a block and corresponding binary tree: the root corresponds to the original 4×4 block (level 4), while nodes i_2 and i_3 (1×1 blocks) belong to level 0.

this case image blocks), using words of an adaptive dictionary $\mathcal{D}$ at different scales. For each block X^l in the image, the algorithm first searches the dictionary for the element S_i^l that minimizes the Lagrangian cost function of the approximation. The superscript l means that the block X^l belongs to $level\ l$ of the segmentation tree. Square blocks, corresponding to even levels, are segmented into two vertical rectangles. Generally, a block of level l has dimensions $(2^{\lfloor \frac{l+1}{2} \rfloor} \times 2^{\lfloor \frac{l}{2} \rfloor})$.

The algorithm then segments the original block into two blocks, X_1^{l-1} and X_2^{l-1}, with half the pixels of the original block, and searches the dictionary of level $(l-1)$ for the elements $S_{i_1}^{l-1}$ and $S_{i_2}^{l-1}$ that minimize the cost functions for each of the sub-blocks.

After evaluating the R-D results of each of the previous steps, the algorithm decides whether to segment the original block or not. Each non-segmented block is approximated by one word of the dictionary (S_i^l). If a block is segmented, then the same procedure applied to the original block is recursively applied to each segment.

The resulting binary segmentation tree is encoded using two binary flags: flag '0' represents the tree nodes, or block segmentations and flag '1' represents the tree leafs (sub-blocks that are not segmented). These flags are not used for blocks of level 0, that can't be further segmented.

The binary tree is encoded using a preorder approach: for each node, the sub-tree that corresponds to the left branch is first encoded, followed by the right branch sub-tree. In the final bit-stream, each leaf flag is followed by an index, that identifies the word of the dictionary that should be used to approximate the corresponding sub-block. These items are encoded using an adaptive arithmetic encoder.

Figure 1 represents the segmentation of an example block and the segmentation tree that MMP uses to encode it. In this example, $i_0 \ldots i_4$ are the indexes that were chosen to encode each of the sub-blocks, and so this block would be encoded using the following string of symbols:

$$0 \quad 1 \quad i_0 \quad 0 \quad 1 \quad i_1 \quad 0 \quad 0 \quad i_2 \quad i_3 \quad 1 \quad i_4.$$

The R-D optimization of the segmentation tree, $\mathcal{T}$, that is used to encode each block, is performed evaluating the Lagrangian cost for every segmentation decision, given by $J(\mathcal{T}) = D(\mathcal{T}) + \lambda R(\mathcal{T})$, where $D(\mathcal{T})$ is the distortion obtained when using $\mathcal{T}$ and $R(\mathcal{T})$ is the corresponding rate.

Unlike conventional vector quantization (VQ) algorithms, MMP uses *approximate block matching with scales* and an adaptive dictionary.

Block matching with scales is an extension of the ordinary pattern matching, in the sense that it allows the matching of vectors of different lengths. In order to do this, MMP uses a separable scale transformation T_N^M to adjust the vectors' sizes before trying to match them. For example, in order to approximate an original block X^l using one block S^k of the dictionary, MMP has to first determine $S^l = T_k^l[S]$. Detailed information about the use of scale transformations in MMP is presented in [1].

MMP uses an adaptive dictionary that is updated while the data is encoded. Every time a block is approximated by the concatenation of two dictionary blocks, of any given level, the resulting block is used to update the dictionary, becoming available to encode future blocks of the image, independently of their size. This updating procedure for the dictionary uses only information that can be inferred by the decoder exclusively from the encoded segmentation flags and dictionary indexes.

3. THE MMP-INTRA ALGORITHM

In this new algorithm we use predictive methods, based on intra-frame prediction techniques used in the H.264/AVC standard [3]. The use of predictive coding modifies the image patterns encoded by MMP. When prediction is successful, the residual blocks tend to be more homogeneous with lower (and similar) energy. This favours the adaptation of the dictionary, thus improving approximation of the encoded blocks, resulting in a more efficient method.

In MMP-Intra, the predictive coding of one image block refers to neighbouring samples of previously-coded blocks, which are to the left and/or above the block to be predicted. For each of the available prediction modes, MMP-Intra first determines the prediction block and the residue values. The block with the residue pixels is then encoded using MMP.

Intra prediction is used hierarchically for blocks of dimensions 16×16, 16×8, 8×8, 8×4 and 4×4 (corresponding to levels 8 to 4 of the segmentation tree). This hierarchical prediction scheme, allied to the use of Lagrangian R-D cost functions, allows the encoder to determine the best trade-off between the prediction accuracy and the additional overhead introduced by the prediction data.

The size of the blocks used for the prediction can be different from the size MMP uses to encode the prediction error block. Therefore, in order for the decoder to be able to determine the prediction block that was used by the encoder, the size of the prediction blocks, as well as the used prediction mode, must be sent to the decoder. This information is encoded together with the original MMP flags and indexes, using a set of flags that represent the prediction modes and prediction error segmentation. These flags are encoded using an adaptive arithmetic coder.

Using this information, the decoder determines each resulting decoded image block by finding the corresponding prediction block and adding it to the decoded residue block.

The prediction modes used by MMP-Intra are the same ones used by H.264/AVC intra coded blocks, with one important exception: the DC mode was replaced by a new prediction mode, that uses the most frequent value (MFV) among the pixels used for the prediction, instead of the average value of those pixels. Because MMP-Intra encodes the prediction error using blocks from the dictionary, the DC value of the prediction error is not as good a choice for it as for the case of H.264/AVC, that uses transform coding. Experiments showed that for text and graphics images, characterised by regions of white background with a few very dark pixels as foreground, the use of DC prediction generates error blocks with very different gray values, depending on the prediction pixels. This has the highly adverse result of creating a set of scattered words in the dictionary, with different luma values, which limited the adaptation process of the dictionary. In these cases, the use of the most frequent value for the prediction has the advantage of creating prediction error blocks consistently centered around zero, narrowing the range of new blocks that are used in the dictionary update procedure and enhancing the overall coding efficiency. In addition, for smooth images, the use of the MFV instead of DC prediction has no effect on the performance.

4. EXPERIMENTAL RESULTS

MMP and MMP-Intra were implemented and some experimental tests were performed. Both algorithms were used to encode grayscale images, initially divided in blocks of dimensions 16×16.

When used to encode prediction error blocks, MMP uses an initial dictionary in the scale 1×1 (level 0) with values in the range from -255 to 255. The initial dictionaries for the other levels are obtained from this one by scale transformation. The scale transformation and dictionary update procedures are the same as the ones described in [1].

MMP-Intra used four prediction modes in the prediction of 16×16 blocks (level 8 of the segmentation tree): the H.264/AVC $Intra\_16 \times 16$ *Vertical, Horizontal* and *Plane* prediction and the MFV prediction described in section 3. For all other block sizes nine prediction modes were used, corresponding to the nine modes defined by H.264/AVC for $Intra\_4\times4$ prediction, with the DC mode replaced by MFV.

The R-D results of MMP-Intra and MMP where compared with JPEG2000 encoder [2] and with H.264/AVC (JM9.2) [3], used as a still image encoder. In spite the fact that H.264/AVC was developed for video coding, its performance in still image coding, using its intra-frame coding tools, is well known to match (and sometimes exceed) that of top image encoders. This, together with the fact that H.264/AVC uses essentially the same prediction of MMP-Intra, but in the context of the transform-quantization-encoding paradigm, led us to include the performance of the H.264/AVC in the evaluation of MMP-Intra. The used profile was the recently developed H.264/AVC's FRExt high profile. It was chosen because of its ability to also use 8×8 blocks for intra prediction and transform coding, besides the 16×16 and 4×4 blocks defined in the baseline and main profiles. This gives a closer approximation to the hierarchical prediction performed by MMP-Intra.

Figures 2 to 5 present the experimental results achieved for four grayscale test images: LENA (downloaded from [5]), pp1205, pp1209 and D108. The images pp1205 and pp1029 are a text image and a compound (text and grayscale). They were scanned, respectively, from pages 1205 and 1209 of the *IEEE Transactions on Image Processing*, volume 9, number 7, July 2000 and are available for download at [6]. Test image D108 is a texture image from the Brodatz Database, downloaded from [7].

The presented plots are limited to rates where the visual distortion of the compressed images doesn't make them unusable, namely for the images with text. These plots clearly show that the use of predictive schemes associated with MMP result in an better objective quality for the encoded image.

Its clear that MMP-Intra presents a considerable quality gain for smooth images when compared with MMP, specially for smaller compression rations. For image LENA, for example, MMP-Intra provided a gain of more than 1.5dB over MMP. This reduced the gap between JPEG2000 or H.264/AVC and image encoders based on multiscale recurrent patterns, to about 0.5dB.

For text and compound images, as well as for texture images, MMP and MMP-Intra achieved a significant performance gain over JPEG2000 and H.264/AVC high profile. For the mixed grayscale and text image, pp1209, the gains over JPEG2000 exceed 1.5dB and MMP-Intra consistently tops H.264/AVC and the original MMP. For the text and equations image, pp1205, both MMP and MMP-Intra outperform JPEG2000 by up to 5dB and H.264/AVC by up to about 3dB. For the texture image, in spite of the minor performance losses relative to the original MMP, MMP-Intra achieves a gain up to 1dB over H.264/AVC and more than 2dB over JPEG2000. In this case, the losses of MMP-Intra for higher compression rations result from the additional

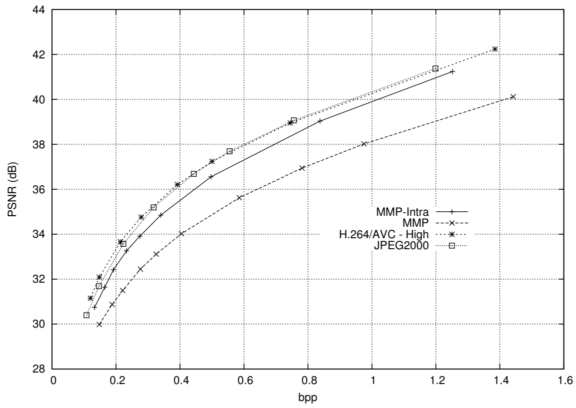

Fig. 2. Experimental results for image LENA.

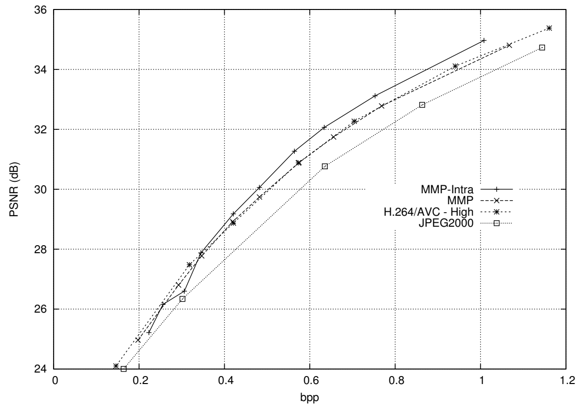

Fig. 4. Experimental results for image pp1209.

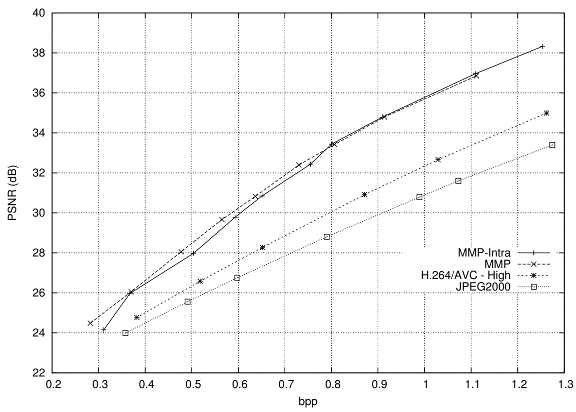

Fig. 3. Experimental results for image pp1205.

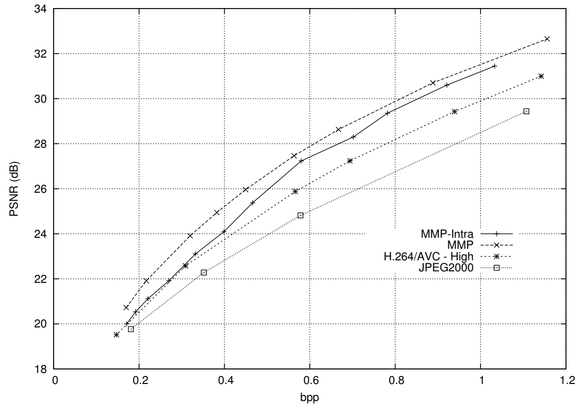

Fig. 5. Experimental results for image D108.

overhead needed to transmit the prediction information.

5. CONCLUSIONS

In this paper we presented a new image encoding method, that uses hierarchical predictive coding combined with the Multiscale Multidimensional Parser (MMP) algorithm. This new method is able to achieve coding gains relative to the MMP algorithm of about 1.5dB for smooth images while maintaining its excellent performance for text and mixed images, in spite of the increased overhead associated with the intra prediction information.

This shows that MMP-Intra has excellent performance for a wide range of image types, making it a good alternative for state-of-the-art DWT and DCT based encoders when encoders with a universal character are needed.

6. REFERENCES

[1] M. de Carvalho, E. da Silva, and W. Finamore, “Multidimensional signal compression using multiscale recurrent patterns,” *Elsevier Signal Processing*, no. 82, pp. 1559–1580, November 2002.

[2] D. S. Taubman and M.W. Marcelin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2001.

[3] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, “Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T rec. H.264/ISO/IEC 14 496-10 AVC,” JVT-G050 2003.

[4] E. Filho, M. de Carvalho, and E. da Silva, “Multidimensional signal compression using multi-scale recurrent patterns with smooth side-match criterion,” *IEEE International Conference on Image Processing*, October 2004.

[5] http://sipi.usc.edu.

[6] http://www.estg.ipleiria.pt/∼nuno/MMP/.

[7] http://www.ux.his.no/∼tranden.