MULTISCALE RECURRENT PATTERN IMAGE CODING WITH A FLEXIBLE PARTITION SCHEME

Nelson C. Francisco[1,6], *Nuno M. M. Rodrigues*[1,2], *Eduardo A. B. da Silva*[3], *Murilo B. de Carvalho*[4], *Sérgio M. M. de Faria*[1,2], *Vitor M. M. da Silva*[1,5], *Manuel J. C. S. Reis*[6]

[1]Instituto de Telecomunicações, Portugal; [2]ESTG, Instituto Politécnico Leiria, Portugal; [3]PEE/COPPE/DEL/Poli, Univ. Fed. Rio de Janeiro, Brazil; [4]TET/CTC, Univ. Fed. Fluminense, Brazil; [5]DEEC, Univ. de Coimbra, Portugal. [6]DE/CITAB, Univ. de Trás-os-Montes e Alto Douro, Portugal.

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

ABSTRACT

In this paper we present a new segmentation method for the Multidimensional Multiscale Parser (MMP) algorithm. In previous works we have shown that, for text and compound images, MMP has better compression efficiency than state-of-the-art transform-based encoders like JPEG2000 and H.264/AVC; however, it is still inferior to them for smooth images.

In this paper we improve the performance of MMP for smooth images by employing a more flexible block segmentation scheme than the one defined in the original algorithm. The new partition scheme allows MMP to exploit the image's structure in a much more adaptive and effective way. Experimental tests have shown consistent performance gains, mainly for smooth images. When employing the new block segmentation scheme, MMP outperforms the state-of-the-art JPEG2000 and H.264/AVC Intra-frame image coding algorithms for both smooth and non-smooth images, at low to medium compression ratios.

Index Terms— Image Coding, Pattern Matching, Image Segmentation, Image Processing

1. INTRODUCTION

Transform-based image encoding methods assume that an image has a low-pass nature, and expect most of the transform coefficients at higher frequencies to be of little importance. This is exploited by using coarse quantization or by simply ignoring these frequency components. This usually results in a very good performance for smooth images. Nevertheless, when used to encode images that are not of a low-pass nature, as 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 to approximate variable-length input vectors. These vectors result from recursively parsing an original input block of the image. A dyadic segmentation is used, such that each block is split in two equal-sized blocks along a pre-established direction. Scaling transformations are used to resize each dictionary element to the dimension of the block segment that is being considered.

Experimental results demonstrate that the direction in which blocks are partitioned in the original algorithm may have a significant influence on the compression results. Differences of up to 0.4dB in peak signal-to-noise ratio (PSNR) were observed for some images. These observations motivated the development of a new segmentation scheme for MMP, that adaptively selects the partition direction for the input block, based on a local criterion. This new method is able to more efficiently exploit the image's structure, and achieves good coding performances for a wide range of image types. Flexible segmentation schemes have been known to improve the performance of other block based image encoders [2]. In the next section the state-of-the-art of MMP-based algorithms for image coding is briefly presented. Section 3 introduces the new adaptive segmentation scheme, discussing its most significant advantages. In section 4 experimental results are shown and section 5 presents the conclusions of this work.

2. MMP IMAGE CODING

The MMP algorithm is a generic lossy data compression method that has been successfully applied to image coding. In this section we present the state-of-the-art of MMP image coding algorithms.

2.1. The original MMP

The MMP algorithm was first presented in [1]. It is based on approximations of data segments (in this case image blocks)

Authorized licensed use limited to: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Downloaded on March 3, 2009 at 15:24 from IEEE Xplore. Restrictions apply.

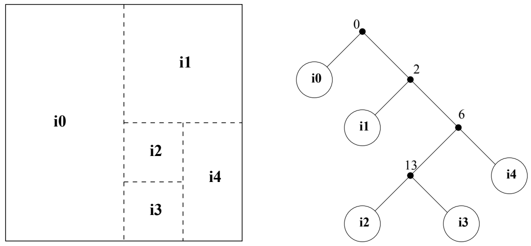

Fig. 1. Segmentation of a 4×4 block (scale 4)(a) and the corresponding 5 scale binary tree (b).

of scale l, using words from an adaptive dictionary $\mathcal{D}^l$. 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 $J(\mathcal{T}) = D(X^l, S_i^l) + \lambda R(S_i^l)$, where $D()$ is the sum of square differences (SSD) function and $R()$ is the rate needed to encode the approximation. The superscript l means that the block X^l belongs to $scale\ l$, that corresponds to a block size of $(2^{\lfloor \frac{l+1}{2} \rfloor} \times 2^{\lfloor \frac{l}{2} \rfloor})$. The algorithm then segments the original block into two blocks, X_1^{l-1} and X_2^{l-1}, each with half the pixels of the original block, and searches the dictionary of scale $(l-1)$ 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. The costs for each of the previous steps are then evaluated and the algorithm decides whether to segment or not the original block. Each non-segmented block of scale l is approximated by one word S_i^l of the dictionary $\mathcal{D}^l$. If a block is segmented, then the same procedure is recursively applied to each segment.

The optimal block partitioning is represented by a binary segmentation tree, that is encoded using two binary flags: flag '0' represents the tree nodes or block segmentations, and flag '1' represents the tree leaves (sub-blocks that are not segmented). 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 all encoded using an adaptive arithmetic encoder, with a different context for each tree level, (that corresponds to a block scale).

Figure 1 represents the segmentation of a block and its corresponding segmentation tree. In this example, $i_0,\ldots,i_4$ are the indexes chosen to encode each sub-block. In this example, the resulting string of symbols is as follows:

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

The use of an *adaptive dictionary* is one important feature of MMP. Every segmentation of a block of scale l originates a new pattern formed by the concatenation of two dictionary blocks of scale $l-1$. This new block is used to update the dictionaries in every scale. In order to update the dictionary of scale s using an element of scale l, one uses a separable scale transformation T_l^s to adjust the vector' scale. It is important to note that this dictionary adaptation process requires no extra overhead. The decoder uses the segmentation flags and dictionary indexes information to keep a synchronized copy of the dictionary, allowing for a correct reconstruction of the image.

2.2. MMP with predictive coding: MMP-I Algorithm

MMP-I [3] is a successful combination of the original MMP algorithm and intra-frame prediction techniques, like those used in the H.264/AVC standard [4]. For each original block X^l, MMP-intra first determines the prediction block P_m^l and the corresponding residue block Q_m^l, that is encoded by the MMP algorithm. MMP-I uses essentially the same prediction modes of H.264/AVC; details can be found in [3].

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 segmentation of the residue block can be different from the one applied to the MMP compression of the residue. The decoder uses information about the block size and the used prediction mode to reconstruct the prediction block. This additional data is encoded together with the original MMP flags and indexes using an adaptive arithmetic coder [3]. With this information, the decoder is able to reconstruct the image blocks by calculating the corresponding prediction block and adding it to the decoded residue block.

2.3. Efficient Dictionary Adaptation - MMP-II Algorithm

MMP-I uses the same dictionary updating procedure as the original MMP. However, experimental tests revealed some inefficiencies in this procedure. These tests motivated the investigation of several dictionary adaptation techniques that improved the performance of MMP-I. We refer to the improvement of MMP-I with these dictionary adaptation techniques as MMP-II [5]. MMP-II is able to consistently improve on the coding performance of MMP-I for all bit-rates and image types. In what follows we describe in general terms the dictionary design methods used by MMP-II (the reader is referred to [5, 6] for a complete description).

An increase in the arithmetic encoding's performance was achieved by improved context modeling. The dictionary elements are organized into partitions and each dictionary element is identified using a partition index followed by its index inside that partition. The scale at which the block was originally created by concatenation of two dictionary words is used as a context, exploiting the fact that blocks generated at different levels have different matching probabilities.

The use of an efficient redundancy control scheme for dictionary elements limits the insertion of a new block if its distance relatively to one block that is already available in the dictionary is inferior to a given threshold d. This prevents the

Authorized licensed use limited to: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Downloaded on March 3, 2009 at 15:24 from IEEE Xplore. Restrictions apply.

creation of a new index (and thus, an increase in the overall rate) for blocks that bring very little distortion gains.

The use of extra patterns to update the dictionary, in order to improve its approximation power, was also proven to be effective. The concatenated block is transformed in order to create new code-vectors. Several methods were defined for this, namely the use of geometric transformations and translations of the concatenated block.

A norm-equalization procedure is also used, in order to adapt the new code-vector patterns to the residue signal's statistical distribution.

3. FLEXIBLE PARTITIONING SCHEME

All MMP-based algorithms described above use a rigid dyadic block partitioning scheme. Even scales (square) blocks are always divided in the vertical direction, while odd scale blocks are segmented horizontally. This limits MMP's capability to adapt freely to the image structure, compromising its efficiency. This paper proposes the use of an alternative MMP segmentation scheme. In it, any block may be segmented along either the horizontal or the vertical direction, based on a local R-D criterion. Prior to its encoding, each image block X^l is segmented in both the vertical and horizontal directions. This procedure is applied recursively for each child node, until the level 0 of the segmentation tree is reached. This means that a large set of new block dimensions become available for the compression. For example, in the case of a 16×16 block size, the original MMP block segmentation procedure only uses square and rectangular blocks with 2:1 proportions, generating a total of 9 different scales (16×16, 16×8, 8×8,..., 1×1). The new block segmentation scheme may generate blocks of 25 possible scales (16×16, 16×8, 8×16 (new), 16×4 (new), 4×16 (new),..., 1×1), with no restriction on the aspect ratio, meaning that all blocks with dimensions $2^m \times 2^n$ are available, for $m, n = 0, \ldots, 4$. Figure 2 illustrates the residue segmentation for both schemes (white lines) and the additional partitions used to encode the predicted residue patterns (represented by the black lines).

From the bottom of the tree up, the value of the Lagrangian cost function for each segmentation option is evaluated, and the option with lower cost is chosen. If the decision to segment the block using one direction is taken, the child nodes generated by the other direction of the segmentation tree are pruned. If the lowest Lagrangian cost corresponds to a non-segmentation decision (*i.e.* the block corresponds to a tree leaf), all child nodes are pruned. An additional segmentation flag is used to explicitly indicate the segmentation direction. As a direct consequence of this new segmentation scheme, the block partition dimensions become very flexible and the method is able to adapt much more efficiently to the input signal's features. The new flexible segmentation scheme is used by MMP-II both for the compression of the predicted residue and for the prediction step. This results in an much more accurate prediction process, creating a predicted residue with lower energy, that is more efficiently compressed by MMP. The new partitioning method also uses block sizes that favor the prediction process - a good illustration for this is given by the very thin blocks (*e.g.* 16×1) - when one has a thin vertical detail, these blocks tend to generate a more accurate prediction signal than, for example, 4×4 prediction blocks (that have the same number of pixels).

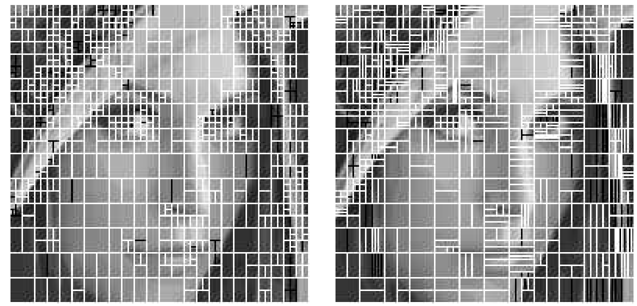

Fig. 2. Residue block segmentation for the encoding of the LENA image with the conventional (left) and the new (right) segmentation scheme.

4. EXPERIMENTAL RESULTS

The new flexible partitioning scheme (FPS) was implemented in MMP-II, and some experimental tests were performed. The algorithm was used to encode grayscale images, initially divided into blocks of dimension 16×16, with prediction segmentation defined only for blocks with more than 16 pixels. This is a compromise between compression efficiency and computational complexity. The encoding performance is compared against that of two state-of-the-art image compression algorithms: H.264/AVC Intra and JPEG2000.

The performance of MMP for text and graphic images is considerably above that of the transform-based algorithms. MMP-II achieves gains of up to 5dB and 6dB, when compared with H.264/AVC [4] and JPEG2000 [7], respectively. Because of this, the focus of our tests was on the performance of the proposed method for smooth images. Experimental results for a representative set of well known natural scene test images are presented in Figures 3 to 5 [8].

These figures clearly show a consistent improvement in the MMP-II encoding performance for smooth images. This allows MMP-II to outperform the state-of-the-art image compression algorithms for almost all compression ratios. In addition, other experiments not reported here due to lack of space, show that, like the previous versions of MMP, the proposed method outperforms transform-based algorithms also for text and compound images.

Authorized licensed use limited to: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Downloaded on March 3, 2009 at 15:24 from IEEE Xplore. Restrictions apply.

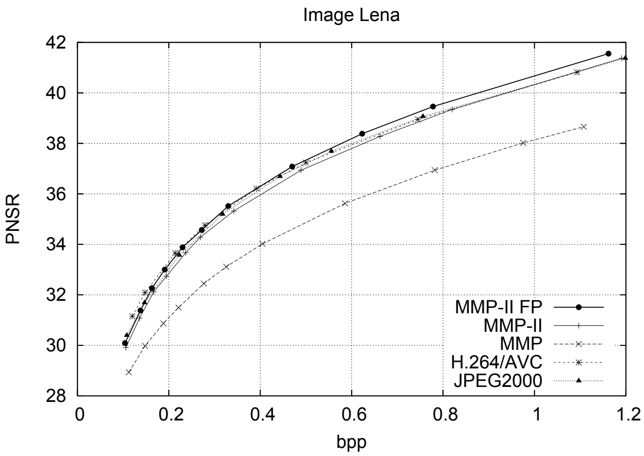

Fig. 3. Experimental results for image LENA 512×512.

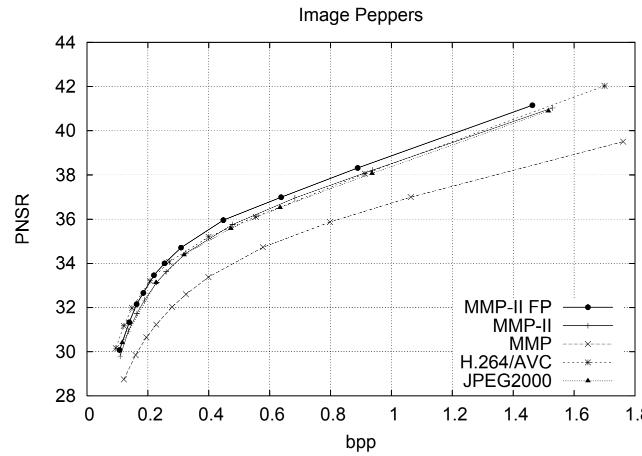

Fig. 5. Experimental results for image PEPPERS 512×512.

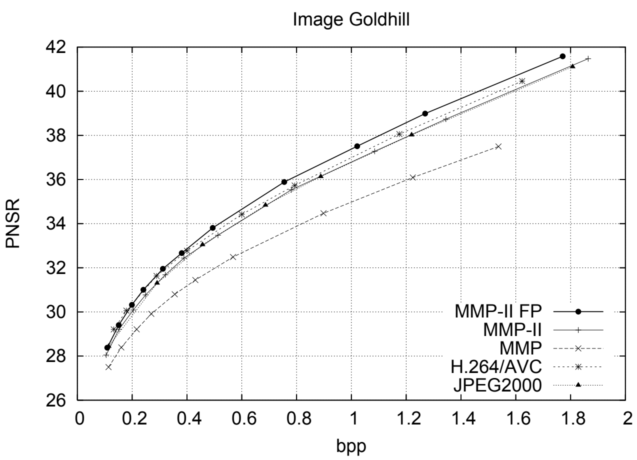

Fig. 4. Experimental results for image GOLDHILL 512×512.

5. CONCLUSIONS

In this paper we present a new flexible segmentation scheme for multiscale pattern matching image coding, that is able to achieve considerable coding gains for smooth images, while maintaining the excellent efficiency for other image types. This new method allows MMP-based algorithms to outperform state-of-the-art DWT and DCT-based encoders for a broad class of images and a wide range of compression ratios. This demonstrates that MMP-based algorithms are an alternative to the dominant transform-coding paradigm that is worth investigating, due to their very useful universal character and top compression performances.

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] Yan Huang, I. Pollak, M.N. Do, and C.A. Bouman, "Fast search for best representations in multitree dictionaries," *Image Processing, IEEE Transactions on*, vol. 15 no. 7, pp. 1779–1793, July 2006.

[3] N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, and Vitor M. M. Silva, "Universal image coding using multiscale recurrent patterns and prediction," *IEEE International Conference on Image Processing*, Genova, Italy, September 2005.

[4] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), *Draft of Version 4 of H.264/AVC (ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 part 10) Advanced Video Coding)*, March 2005.

[5] N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, Vitor M. M. Silva, and Frederico Pinagé, "Efficient dictionary design for multiscale recurrent patterns image coding," *ISCAS 2006 IEEE International Symposium on Circuits and Systems*, Island of Kos, Greece, May 2006.

[6] E. B. L. Filho, E. A. B. da Silva, M. B. de Carvalho, and F. S. Pinagé, "Universal image compression using multiscale recurrent patterns with adaptive probability model," *Accepted for publication in IEEE Transactions on Image Processing*.

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

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

Authorized licensed use limited to: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Downloaded on March 3, 2009 at 15:24 from IEEE Xplore. Restrictions apply.