IMPROVING MULTISCALE RECURRENT PATTERN IMAGE CODING WITH LEAST-SQUARES PREDICTION MODE

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

[1]Instituto de Telecomunicações, Portugal;
[2]ESTG, Instituto Politécnico de Leiria, Portugal;
[3]PEE/COPPE/DEL/EE, Univ. Fed. do Rio de Janeiro, Brazil.
[4]TET, Univ. Fed. Fluminense, Brazil.

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

ABSTRACT

The Multidimensional Multiscale Parser-based (MMP) image coding algorithm, when combined with flexible partitioning and predictive coding techniques (MMP-FP), provides state-of-the-art performance. In this paper we investigate the use of adaptive least-squares prediction in MMP.

The linear prediction coefficients implicitly embed the local texture characteristics, and are computed based on a block's causal neighborhood (composed of already reconstructed data). Thus, the intra prediction mode is adaptively adjusted according to the local context and no extra overhead is needed for signaling the coefficients. We add this new context-adaptive linear prediction mode to the other MMP prediction modes, that are based on the ones used in H.264/AVC; the best mode is chosen through rate-distortion optimization.

Simulation results show that least-squares prediction is able to significantly increase MMP-FPs rate-distortion performance for smooth images, leading to better results than the ones of state-of-the-art, transform-based methods. Yet with the addition of least-squares prediction MMP-FP presents no performance loss when used for encoding non-smooth images, such as text and graphics.

***Index Terms*—** Image Coding, Recurrent Pattern Matching, Block Intra Prediction, Least Square Minimization

1. INTRODUCTION

Transform-based image encoders are the state-of-the-art in image compression algorithms. Their powerful compression capability is based on the fact that images are smooth and have low energy high frequency components. New coding tools have been developed to further enhance their compression efficiency. A good example is the intra coding of the H.264/AVC video coding standard [1], where the introduction of directional prediction modes has greatly improved its coding efficiency; H.264/AVC-Intra even outperforms wavelet-based encoders such as JPEG2000 [2].

Although very useful in most practical cases, the smoothness assumption does not hold for all kind of images, namely text and compound images (e.g. scanned documents) or computer graphics. Previous works led to the development of an algorithm based on recurrent pattern matching in multiple scales, referred to as MMP (Multidimensional Multiscale Parser) algorithm [3]. MMP presents consistently better results than most state-of-the-art encoders for non-smooth images such as text and graphics, although a rate-distortion performance gap was still present for smooth images.

With the introduction of an image prediction step, as well as flexible dictionary adaptation and image segmentation, MMP-based encoders [4] [5] were able to consistently outperform image encoders such as H.264/AVC [1] and JPEG2000 [6] also for smooth images. The prediction modes used in MMP-FP are based on the intra-prediction used in the H.264/AVC standard [1].

Nevertheless, the prediction techniques used in these compression algorithms assume that the pixels inside a block are similar along the prediction's direction, and are not so efficient for predicting less homogenous images. Many attempts have been made to improve general intra prediction schemes. A good example is the work in [7], that modifies the traditional prediction modes by reusing the predicted pixels.

In this paper we propose the use of least-squares prediction (LSP) as an additional prediction mode. LSP has been successfully applied both for image [8] and video coding [9], and its idea was also extended to block-based prediction in [10]. The coefficients are calculated from the decoded data using a causal training window. Therefore, the use of an LSP-based prediction mode in MMP requires no side information. Nevertheless, since MMP uses different block sizes for segmenting and encoding an image, we have to adapt LSP for block prediction. The implementation details are explained in the rest of the paper.

This paper is organized as follows. Section 2 briefly reviews the MMP algorithm. The Least-Square Prediction mode, as well as the derivation of LSP for block prediction, are presented in section 3. Results are given in section 4 and section 5 concludes the paper.

2. THE MMP ALGORITHM

MMP divides the image into non-overlapping blocks, and uses patterns at different scales to approximate the image's blocks. If no pattern from the dictionary satisfactory matches the new block, it is iteratively split, until appropriate matches are found for all the remaining blocks. Concatenation of patterns used for encoding the blocks are then added to the dictionary at multiple scales, therefore adapting it to the image's characteristics.

In order to find the best block segmentation and index encoding, a rate-distortion optimized algorithm based on Lagrangian cost is used. The best segmentation is determined by comparing the node cost in the segmentation tree. The output bitstream is then formed by arithmetic encoding the segmentation flags and the dictionary indexes. More information can be found in [3].

2.1. MMP with Predictive Coding and Flexible Partitioning

In an MMP framework, prediction may be used as a primary step, in order to determine a prediction residue block for MMP coding. When prediction is successfully applied, the residual blocks tend to be more homogeneous with lower (and similar) energy, favoring the adaptation of the dictionary and improving the encoding efficiency.

Like the H.264/AVC standard [1], MMP-Intra uses adaptive block size and several prediction modes. Each prediction mode is optimized together with the prediction block size, that may vary from 16×16 down to 4×4 blocks. The prediction step is optimized based on a Lagrangian RD cost function. For each prediction mode and available block size, the residue is encoded using MMP; the MMP coding cost for the residue block is added to the cost for sending the prediction data; by repeating this process for all available modes and block sizes, the encoder is able to determine the best trade-off between prediction accuracy and the associated overhead.

Other improvements were introduced in [4], related to: reducing the entropy of the dictionary indexes' symbols, limiting the insertion of new dictionary elements in every scale, improving the dictionary adaptation process and increasing the efficiency of the arithmetic encoding using adaptive contexts.

In [5] a new flexible partitioning scheme was proposed (MMP-FP). The rigid dyadic block partitioning scheme used in MMP-Intra was relaxed. The block can be segmented into vertical or horizontal direction, according to the best RD compromise. An additional flag is sent to indicate the direction of the segmentation.

MMP-FP was able to improve considerably MMP's coding efficiency, outperforming state-of-the-art transform-based algorithms for medium to low compression rates, even for smooth images. Furthermore, the MMPs performance gains for non-smooth images (like text and compound images) were further expanded. Details and results for MMP-FP can be found in [5].

3. MMP WITH LEAST-SQUARES PREDICTION

As was previously mentioned, the MMP-FP algorithm uses a set of prediction modes inspired by the ones defined in the H.264/AVC standard. Prediction generates residue signals that have a distribution function concentrated around a reduced set of low luminance values, generating more regular patterns and increasing the probability of using patterns already present in the dictionary.

H.264/AVC's prediction uses a set of nine possible prediction directions, where the pixels inside the block are predicted using the filtered neighborhood pixel values according to a chosen direction. Although quite effective, it is based on the supposition that the pixels do not vary along the chosen direction. This is generally not the case for blocks with complex textures and few homogeneous regions. In order to overcome the limitations of these prediction modes, the use of a new, non-directional prediction mode was investigated.

3.1. Least-Squares Prediction

The Least-Squares Prediction (LSP) method determines each prediction pixel by adaptively filtering a selected neighborhood. The set of filter coefficients are determined based on a training window using reconstructed data [8]. The coefficients are selected according to an N*th* order Markovian model, and often the nearest pixels are used. The prediction can be described with the following equation

$$\hat{X}(\mathbf{n}_0) = \sum_{i=1}^{N} a_i X(\mathbf{n}_i) \tag{1}$$

The coefficients used in LSP prediction are locally optimized using a causal training window in a least-square sense. A convenient choice of a training window is the double-rectangular window that contains $M = 2T(T+1)$ elements (see Figure 1 (a)). The training sequence can be arranged in a $M \times 1$ column vector $\mathbf{y} = [X(n-1)\ldots X(n-M)]^T$ (See Figure 1 (b)). With the prediction neighbors we form an $M \times N$ matrix

$$C = \begin{bmatrix} X(n-1-1) & \ldots & X(n-1-N) \\ \vdots & & \vdots \\ X(n-M-1) & \ldots & X(n-M-N) \end{bmatrix}$$

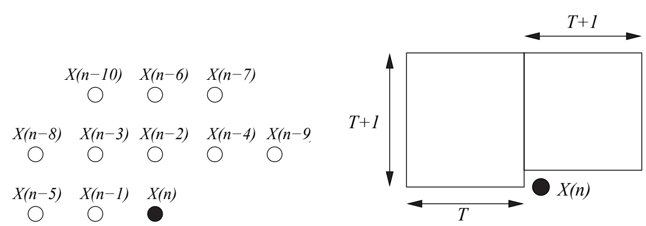
The coefficients can be determined by LS optimization, finding the solution for $\min(||\mathbf{y} - C\mathbf{a}||^2)$. A well-known closed form solution for this problem is

$$\mathbf{a} = (C^T C)^{-1}(C^T \mathbf{y}) \tag{2}$$

LSP is well suited for predicting arbitrarily oriented edges, due to its edge-directed property [8], where the edge pixels play a dominant role in the LS optimization process. If the training window has enough edge pixels, LS prediction will successfully identify the edge direction and also correctly predict the new pixel.

3.2. Using LSP for block prediction

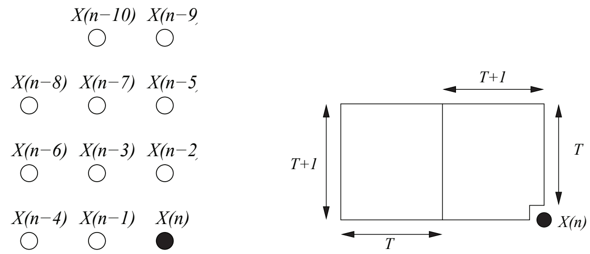
In [10], a block prediction formulation using LSP was proposed. In its lossy approach only a set of previously decoded pixels, on the top and to the left of the current block, are used as the neighborhood for predicting all the pixels inside the block. This is so even for pixels that are far away from these prediction borders, what decreases prediction effectiveness.



(a) Pixel Neighborhood (b) Training Pixels

Fig. 1. Implementation of block prediction using LSP

The above restriction on the neighborhood is used due to the fact that, since the encoding is block-based, when encoding all the pixels in a block only pixels from the previous block are available for use in the predictor. In our approach, we decided to waive such a restriction. The closest neighbors in a pixel-by-pixel basis are always chosen, even for positions that have not been previously decoded. In such cases, we replace the unavailable reconstructed pixel value by its predicted value, previously determined with LSP. Figure 1 (a) shows the neighborhood chosen for prediction, and Figure 1 (b) illustrates the pixels that will be selected for training.



(a) Causal Pixel Neighborhood (b) Causal Training Pixels

Fig. 2. Modifications for block prediction

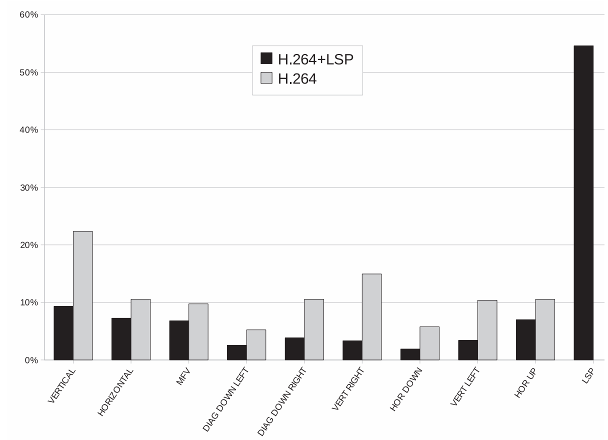


Fig. 3. Percentages of choice for each prediction mode for image Barbara.

Another limiting factor of the method in [10] is that it performs the training using pixels that are to the right of the predicted position, as shown in Figures 1 (a) and (b). However, for block prediction, these pixels may not be available for training, since some of them may belong to a block that still needs to be encoded. Therefore, for such elements we replace the training window and the available neighborhood by one that has only causal elements. Figures 2 (a) and (b) show the modified training window and the neighborhood used in such cases; notice that elements on the right are not used for prediction. It is important to note that in both cases (Figures 1 and 2), we still need to use the predicted pixels instead of the decoded pixels in order to predict pixels inside a block.

On the image's edges, where there are no available pixels for LSP to perform the training (leftmost and uppermost blocks), the LSP mode is deactivated, and only the H.264/AVC based prediction modes can be used.

As in most pattern-matching based encoders, such as VQ, MMP's complexity is mainly affected by the search of each optimum index. Therefore, the impact of prediction process on the complexity of MMP-FP is negligible. However, when using LSP prediction a matrix inversion operation needs to be performed for each pixel, which has a non-negligible impact on the computational complexity. Fast implementations of LSP have been proposed [8]; however, since the main focus of this paper is the algorithm's compression performance, its implementation will be the focus of a future work.

4. EXPERIMENTAL RESULTS

The main motivation for using LSP was to improve the efficiency of the MMP encoder for smooth, textured images, without degrading its performance for non-smooth images. In this section we present some experimental results that justify the use of LSP as an additional prediction mode in MMP-FP. Results of two state-of-the-art image encoding algorithms, JPEG2000 [6] and H.264/AVC *High* profile intra encoder [1], are also presented.

In order to fully exploit the prediction capability of the LS mode, a high order model was chosen. In our experiments, a fixed neighborhood of 10 elements is used as depicted in Figures 1 and 2. The training window size is set to $T = 7$. Empirical studies conducted in [8] suggested that windows larger than 7 do not further improve the prediction performance.

Figure 3 shows the usage of the available prediction modes, for the original MMP-FP and MMP-FP with LSP. One may observe the high adoption of the LSP mode in comparison with the other modes (more than 50%). The LSP mode was chosen mainly in areas with strong edges, specially in the woman's clothes. In such areas, the presence of strong edges in the training window contribute for the correct prediction of the pixel during the LS optimization process.

The RD curves for different test images are shown in Figures 4 to 7. One may notice that the gain achieved by LSP prediction depends on the input image. For smooth images with complex texture, like Barbara, gains up to 1.2 dB are achieved, while for other pictures, like Lena, they amount to about 0.25dB. We can also notice that the performance increase is higher for middle to high rates. Since LSP is strongly dependent on the training accuracy, it performs better for those rates, due to the fact that the pixels used in the optimization process are more accurate. In all cases we may observe the PSNR advantage of MMP over the transform-based state-of-the-art methods, JPEG2000 and H.264/AVC.

For text and compound images, like images PP1205 and PP1209[1], the edges have very steep variations and occur very often, making it difficult for LSP (or any other prediction method) to "learn" an edge from the training window data. Nevertheless, adding the extra prediction mode did not affect the algorithm's rate-distortion performance, as we can see in Figures 6 and 7.

5. CONCLUSION

In this paper we propose an image encoding algorithm where least-squares prediction is used in a multiscale recurrent pattern image encoding framework. The addition of a LS prediction mode can successfully estimate local texture features using linear prediction coefficients, derived on the fly. Due to this fact, LSP prediction mode tends to be predominant over the other ones. This yields in some cases gains of more than 1 dB in the range of middle to high rates.

Besides increasing the encoding performance for smooth images, MMP with LSP prediction presented no rate-distortion performance losses for text and compound images. The proposed method outperforms state-of-the-art, transform-based compression algorithms for all image types, from smooth to text and compound images.

LSP was thus shown to be able to significantly increase MMP's RD performance. In future works, alternative methods for fast coefficient estimation such as Recursive Least Square (RLS) methods will be investigated.

[1] These and other text and compound test images are available for download at http://www.lps.ufrj.br/profs/eduardo/MMP

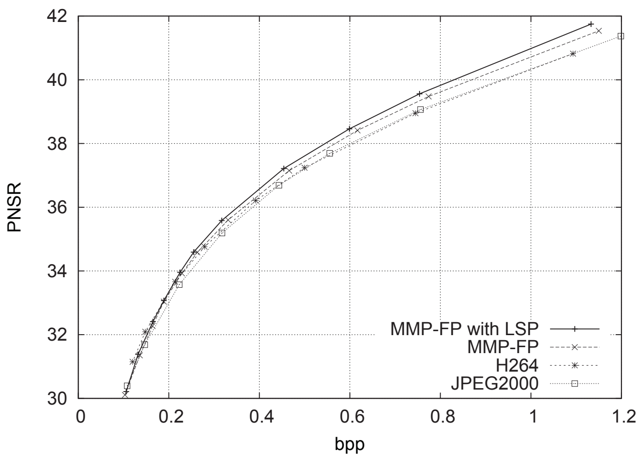


Fig. 4. PSNR result for Lena

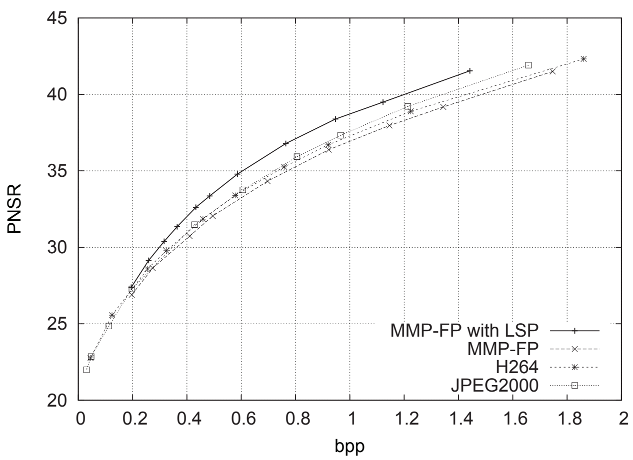


Fig. 5. PSNR result for Barbara

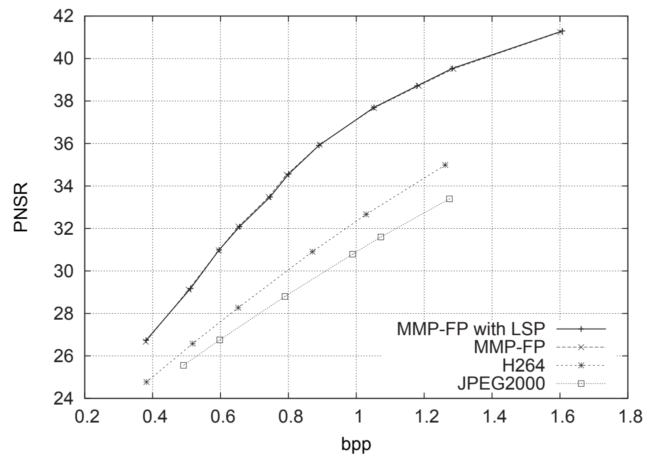


Fig. 6. PSNR result for PP1205

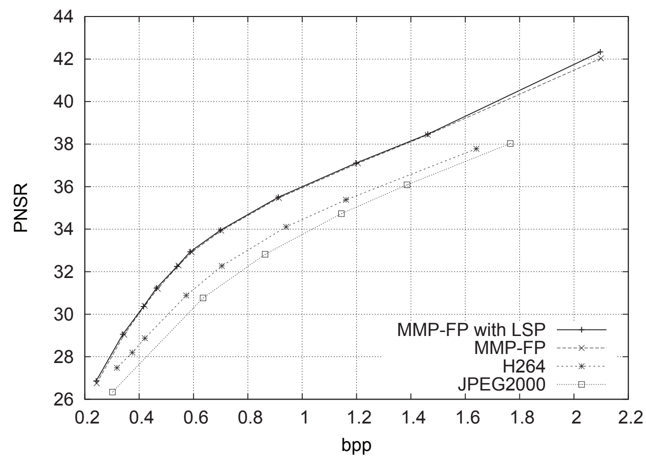


Fig. 7. PSNR result for PP1209

6. REFERENCES

[1] 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.

[2] D. Marpe, T. Wiegand, and S. Gordon, "H.264/MPEG4-AVC fidelity range extensions: Tools, profiles, performance, and application areas," *IEEE International Conference on Image Processing 2005*, vol. 1, pp. 593–596, September 2005.

[3] M. de Carvalho, E. da Silva, and W. Finamore, "Multidimensional signal compression using multiscale recurrent patterns," *Elsevier Signal Processing*, vol. 82, pp. 1559–1580, November 2002.

[4] N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, and Vitor M. M. Silva, "On dictionary adaptation for recurrent pattern image coding," *Image Processing, IEEE Transactions on*, vol. 17, no. 9, pp. 1640–1653, September 2008.

[5] Nelson C. Francisco, N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, Vitor M. M. Silva, and Manuel J. C. S. Reis, "Multiscale recurrent pattern image coding with a flexible partition scheme," *IEEE International Conference on Image Processing*, October 2008.

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

[7] Zhang Xiong-Wei Huang Hui, Cao Tie-Yong, "The enhanced intra prediction algorithm for H.264," *IEEE International Conference on Image Processing*, October 2008.

[8] Xin Li and Michael T. Orchard, "Edge-directed prediction for lossless compression of natural images," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 813–817, 2001.

[9] Xin Li, "Least-square prediction for backward adaptive video coding," *EURASIP J. Appl. Signal Process.*, vol. 2006, no. 1, pp. 126–126, 2006.

[10] Limin Liu, Yuxin (Zoe) Liu, and Edward J. Delp, "Enhanced intra prediction using context-adaptive linear prediction," *PCS2007 - Picture Coding Symposium*, November 2007.