

# Casamento Aproximado de Padrões Multiescala com Segmentação Flexível e Treino do Dicionário

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

**Resumo**—Neste artigo apresentamos um novo esquema de segmentação para o algoritmo de compressão de imagens MMP (Multidimensional Multiscale Parser), aliado a um procedimento de treino do dicionário. Estas técnicas permitiram ao MMP ultrapassar os melhores algoritmos baseados em transformadas, como o JPEG2000 e o H.264/AVC, para vários tipos de imagens e taxas de compressão. Os esquemas propostos neste artigo permitem ao MMP explorar com mais eficiência a estrutura da imagem, com ganhos consideráveis na codificação de imagens suaves. Adicionalmente, as novas técnicas não comprometem os resultados apresentados pelo MMP em imagens não suaves, que superavam consideravelmente os dos algoritmos baseados em transformadas.

**Palavras-Chave**—Codificação de imagens, casamento de padrões, segmentação de imagens, processamento de imagens.

**Abstract**—In this paper we present a new segmentation method for the MMP (Multidimensional Multiscale Parser) algorithm, allied with a dictionary training procedure. These improvements allowed MMP to outperform state-of-the-art transform-based encoders like JPEG2000 and H.264/AVC, for both smooth and non-smooth images, at low to medium compression ratios. The new approach allowed MMP to exploit the images' structure in a more effective way, increasing significantly its performance for smooth images. Additionally, the new techniques do not compromise the results achieved for non-smooth images, where MMP already outperformed transform-based algorithms.

**Keywords**—Image coding, pattern matching, image segmentation, image processing

## I. INTRODUÇÃO

Os algoritmos baseados em transformadas, que são utilizados na compressão de imagens, fazem depender muito a sua eficiência da natureza passa-baixo destas imagens. Esta consideração leva-os a desprezar os coeficientes correspondentes às alta-frequências ou a tratá-los com um passo de quantização elevado. Esta abordagem proporciona ótimos resultados quando o pressuposto é verdadeiro, mas torna-os especialmente limitados para codificar imagens que não cumprem com estas especificações.

<sup>1</sup>Instituto de Telecomunicações, Portugal;

<sup>2</sup>ESTG, Instituto Politécnico Leiria, Portugal;

<sup>3</sup>PEE/COPPE/DEL/Poli, Univ. Fed. Rio de Janeiro, Brazil;

<sup>4</sup>TET/CTC, Univ. Fed. Fluminense, Brazil;

<sup>5</sup>DEEC, Univ. de Coimbra, Portugal.

<sup>6</sup>DE/CITAB, Univ. de Trás-os-Montes e Alto Douro, Portugal.

e-mails: [nelson.carreira@co.it.pt](mailto:nelson.carreira@co.it.pt), [nuno.rodrigues@co.it.pt](mailto:nuno.rodrigues@co.it.pt), [eduardo@lps.ufjr.br](mailto:eduardo@lps.ufjr.br), [murilo@telecom.uff.br](mailto:murilo@telecom.uff.br), [sergio.faria@co.it.pt](mailto:sergio.faria@co.it.pt), [vitor.silva@co.it.pt](mailto:vitor.silva@co.it.pt), [mcabral@utad.pt](mailto:mcabral@utad.pt).

O algoritmo de compressão apresentado neste artigo não se baseia no paradigma transformada-quantização-codificação. O algoritmo é designado por Multidimensional Multiscale Parser (MMP) [1], nome que deriva do fato de se basear no casamento aproximado de padrões, de escala variável. O algoritmo utiliza vetores de dimensões variáveis, para aproximar blocos da imagem de entrada. Estes vetores são armazenados num dicionário adaptativo e resultam da concatenação de padrões da imagem previamente processados. A imagem é dividida em blocos que são individualmente processados. Para cada um destes blocos, o algoritmo adota um esquema de segmentação aplicado recursivamente, até se atingir a sua representação ótima, segundo uma critério taxa-distorção. A utilização de transformações de escala permite ao MMP ajustar as dimensões dos vetores existentes no dicionário às dos blocos resultantes de cada estágio de segmentação.

A segmentação no MMP é realizada segundo um esquema rígido, sendo imposta uma direção fixa para cada dimensão de bloco. Testes experimentais demonstraram, no entanto, que a direção segundo a qual os blocos são segmentados, tem uma grande importância no desempenho de codificação do MMP. A diferença de desempenho chega aos 0.4dB na relação sinal-ruído de pico (PSNR) em algumas imagens, mostrando como este esquema de segmentação rígido pode limitar nitidamente o desempenho do algoritmo. Por outro lado, a utilização de esquemas de segmentação flexível provaram melhorar consideravelmente o desempenho de outros codificadores [2].

Neste artigo será apresentado um novo esquema de segmentação flexível, que confere ao MMP a capacidade de optar em cada nó, pela direção que minimiza o custo de representação do bloco, de uma forma adaptativa.

Foi também desenvolvido um esquema de treino do dicionário para o MMP. O MMP utiliza um dicionário inicial bastante simples, sendo atualizado ao longo da codificação com concatenações de blocos utilizados na aproximação da imagem. Apesar da capacidade de adaptação revelada pelo algoritmo, a utilização de um dicionário inicial pobre compromete o desempenho do codificador [3].

O estudo da utilização do MMP na codificação de sinais de vídeo [4] também fornece alguns indícios que fundamentam a necessidade de crescimento do dicionário. De fato, os resultados demonstraram que para este tipo de sinais onde o dicionário tem tempo de aumentar a sua diversidade de padrões, o desempenho do MMP se destaca claramente do H.264/AVC, o que não acontecia para imagens estáticas.

Estas constatações motivaram um esquema de treino do dicionário, onde através da codificação sucessiva de um conjunto

de imagens de teste, se gera um conjunto representativo de blocos que são disponibilizados ao codificador e decodificador, para complementar o dicionário inicial gerado.

Na próxima seção, será descrito genericamente o funcionamento do MMP. Na seção III, apresentamos o novo esquema de segmentação e na seção IV o procedimento de treino do dicionário. Os resultados dos testes experimentais serão apresentados na seção V e, por último, serão apresentadas algumas conclusões acerca do trabalho desenvolvido.

## II. O ALGORITMO MMP

O MMP é um algoritmo genérico de compressão de dados, que foi aplicado com sucesso, entre outras áreas, à compressão de imagens. Nesta seção, será apresentada uma descrição genérica do MMP.

### A. O MMP-RD

O algoritmo MMP foi apresentado pela primeira vez em [1]. O seu funcionamento baseia-se na aproximação de segmentos de dados (neste caso de blocos da imagem), usando elementos de um dicionário adaptativo  $\mathcal{D}^l$ . Para cada segmento ou bloco  $X^l$  da imagem, o algoritmo começa por procurar no dicionário o elemento  $S_i^l$  que minimiza a função de custo Lagrangeana, definida por  $J(T) = D(X^l, S_i^l) + \lambda R(S_i^l)$ , onde  $D()$  é o erro quadrático entre o bloco original e o bloco candidato à aproximação, e  $R()$  é o número de bits necessários para representar este último. O sobre-escrito  $l$  indica que o bloco pertence à escala  $l$ , cuja dimensão é dada por  $(2^{\lfloor \frac{l+1}{2} \rfloor} \times 2^{\lfloor \frac{l}{2} \rfloor})$ .

O bloco original é posteriormente dividido em dois sub-blocos,  $X_1^{l-1}$  e  $X_2^{l-1}$ , sendo que a dimensão de cada um é metade da do bloco original. O procedimento de seleção dos elementos do dicionário  $S_{i_1}^{l-1}$  e  $S_{i_2}^{l-1}$  da escala  $l-1$ , que minimizam o custo Lagrangeano para cada um dos sub-blocos resultantes, é aplicado recursivamente a cada uma das metades. O procedimento de segmentação prossegue até se atingir a escala 0 do dicionário, ou seja, até se chegar a blocos constituídos por uma única amostra.

Obtém-se assim uma árvore de segmentação, onde o bloco original surge dividido sucessivamente em metades. Numa segunda fase, o MMP analisa esta árvore de segmentação a partir da sua base, a fim de determinar a sua configuração ótima. Para um dado bloco de escala  $l$ , é comparado o custo Lagrangeano associado à sua representação pelo melhor elemento disponível no dicionário, com a soma dos custos associados aos dois sub-blocos que originou. Se este for menor, a melhor opção é não segmentar o bloco, e os nós da árvore correspondentes aos dois sub-blocos são podados. Caso contrário, estes permanecem na árvore.

Após esta iteração, obtém-se uma árvore de segmentação binária que otimiza a representação do bloco em causa. A árvore irá ser directamente utilizada para gerar o fluxo binário final do arquivo. A sua representação é feita com recurso a duas *flags*: uma *flag* '0' que representa os nós da árvore que correspondem a segmentações, e uma *flag* '1' que representa os nós terminais ou blocos não segmentados.

Cada *flag* indicativa de um nó terminal é seguida de um índice que identifica o elemento do dicionário selecionado

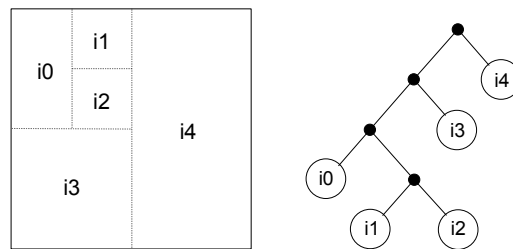


Fig. 1. Segmentação de um bloco 4x4 (escala 4) (a) e árvore de segmentação binária correspondente com 5 níveis (b).

para aproximar o sub-bloco em causa. As *flags* e os índices são codificados com recurso a um codificador aritmético adaptativo, com contexto dependente da escala  $l$  do bloco.

A Figura 1 representa a segmentação de um bloco e a respectiva árvore binária resultante, onde  $i_0, \dots, i_4$  representam os índices escolhidos para codificar cada sub-bloco. A resultado é a seguinte sequência de símbolos:

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

A utilização de um dicionário adaptativo é uma característica importante do MMP. Cada segmentação de um bloco de escala  $l$  origina um novo padrão composto pela concatenação dos dois blocos de escala  $l-1$ . Este novo bloco é utilizado para atualizar o dicionário, passando a estar disponível em todas as suas escala. Para permitir que um elemento gerado na escala  $l$  possa estar disponível na escala  $s$  do dicionário, é utilizada uma transformação de escala  $T_l^s$ , de modo a ajustar as dimensões dos vetores. É ainda importante salientar que este procedimento de atualização não requer o envio de nenhuma informação adicional, visto que o decodificador mantém uma cópia do dicionário, sincronizada apenas com a informação contida na árvore de segmentação.

### B. O MMP com esquema preditivo: MMP-I

O MMP-I [5] é uma combinação de sucesso do algoritmo MMP com esquemas de predição intra-frame, semelhantes aos utilizados pela norma H.264/AVC [6].

Para cada bloco original,  $X^l$ , é determinado um bloco de predição  $P_m^l$  e o seu respectivo bloco de resíduo  $Q_m^l$ , sendo este último codificado com o algoritmo MMP. O MMP-I utiliza um conjunto de modos de predição baseados nos utilizados pelo H.264/AVC, apresentados em detalhe em [5].

A predição Intra é utilizada hierarquicamente para os blocos de dimensões  $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 8$ ,  $8 \times 4$  e  $4 \times 4$  (que correspondem aos níveis 8 a 4 da árvore de segmentação). O esquema de predição hierárquica, aliado à utilização da função de custo Lagrangeano R-D, possibilitam ao codificador determinar a melhor relação entre a qualidade da predição e o número de bits associados à sua representação.

Sempre que o bloco de predição é segmentado, é imposta a mesma segmentação ao bloco de resíduo. Os blocos de resíduo podem continuar a ser segmentados para além da segmentação da predição, no sentido de minimizar o seu custo de representação. O decodificador reconstrói o bloco de predição utilizando a informação relativa à sua dimensão (imposta pela sua escala  $l$ ) e o índice do modo de predição utilizado. Esta informação adicional é codificada tal como as

*flags* e índices recorrendo ao codificador aritmético adaptativo [5], com contexto na escala do bloco.

Esta nova abordagem aplicada ao MMP possibilitou um aumento muito significativo de desempenho na compressão de imagens naturais. O MMP passou a apresentar para este tipo de imagens resultados próximos dos atingidos pelos algoritmos que constituem o estado da arte da codificação de imagens, mantendo resultados muito acima destes para imagens não suaves.

### C. Gestão eficiente do dicionário: MMP-II

O MMP-I utiliza o mesmo procedimento de atualização do dicionário utilizado pelo MMP original.

Testes experimentais demonstraram que o processo utilizado é claramente sub-ótimo do ponto de vista da minimização da entropia dos índices utilizados, motivando o estudo de novas técnicas que permitem otimizar estas rotinas e consequentemente aumentar o desempenho global do codificador. A implementação destas técnicas resultou num novo algoritmo referido como MMP-II [7], [8].

Nesta seção, serão descritas as principais inovações introduzidas pelo MMP-II (para obter mais detalhes, consultar [7], [8]).

As novas técnicas de otimização do dicionário trouxeram alterações em toda a estrutura de gerenciamento, manutenção e representação dos elementos do dicionário, visando essencialmente:

- O aumento do desempenho do codificador aritmético. Esta melhoria foi conseguida com o uso de um condicionamento estatístico otimizado. Os elementos do dicionário passaram a estar organizados em partições ou contextos distintos no interior de cada escala  $l$ , passando a ser identificados através do índice da partição seguido do índice que identifica o elemento no interior dessa mesma partição. A escala na qual o bloco foi originalmente criado define o contexto no qual o elemento é inserido na escala  $l$ , possibilitando assim explorar a diferença de probabilidade de utilização de um bloco de acordo com a sua escala original, o que permite reduzir a entropia média dos índices.
- O controle de redundância. Esta alteração adveio da constatação de que, novos blocos muito semelhantes a blocos já existentes eram frequentemente inseridos, sendo que o ganho associado à melhor representação do bloco atual não era suficiente para compensar o aumento da entropia dos restantes índices. O esquema de controle de redundância adotado visa impedir a inserção de novos elementos sempre que a sua distância relativamente a outro bloco existente seja inferior a um dado limiar  $d$ , função da taxa de compressão pretendida.
- A super-atualização do dicionário. Com o intuito de acelerar o crescimento do dicionário, a atualização passou a ser feita com o bloco original e com novos blocos originados pelas transformações geométricas a ele aplicadas. As transformações geométricas contempladas incluem simetrias, rotações de  $\pm 90^\circ$  e translações de  $1/2$  e  $1/4$  de bloco. Todos os blocos serão sujeitos ao controle de

redundância, de forma a garantir que o processo de super-atualização não prejudica o dicionário, ao forçar a sua atualização com blocos demasiado semelhantes.

- A restrição das transformações de escala. Constatou-se que os blocos originalmente criados numa escala do dicionário distante da escala onde foram inseridos, têm uma probabilidade muito baixa de virem a ser utilizados na codificação da imagem, sendo esta probabilidade maior quanto maior for a proximidade entre escalas. A limitação de escalas resulta na inserção de um bloco recém criado apenas nas escalas vizinhas da sua escala original, e não em todas as escalas como acontecia anteriormente.
- A equalização da norma. Os blocos de resíduo apresentam uma distribuição estatística estreita, que pode ser modelada utilizando uma função Gaussiana Generalizada. A equalização de norma visa adaptar a norma dos blocos sujeitos a transformações de escala à distribuição estatística determinada [7].

O MMP-II apresenta um desempenho superior ao do MMP-I para todos os tipos de imagens e taxas de compressão.

### III. ESQUEMA DE SEGMENTAÇÃO FLEXÍVEL: MMP-FP

Todos os algoritmos anteriormente apresentados utilizam um esquema de segmentação rígido, onde os blocos apenas podem ser segmentados segundo uma direção previamente imposta. Por exemplo, os blocos das escalas pares (blocos quadrados), são sempre divididos verticalmente, enquanto que os blocos de escalas ímpares são segmentados horizontalmente, ou vice-versa.

Testes experimentais demonstraram que a primeira abordagem tende a gerar melhores resultados para a maioria das imagens, existindo entretanto um conjunto de imagens onde tal não se verifica. Por outro lado, é importante observar que o fato de uma dada segmentação ser em média melhor na globalidade da imagem, não garante que isso seja verdade para todos os blocos que a compõem.

A partir destas observações, conclui-se que o esquema de segmentação rígido dificulta a adaptação do MMP à estrutura presente na imagem. Por esse motivo, foi desenvolvido um novo esquema de segmentação que permite que em cada ponto, o MMP possa optar pela melhor direção segundo a qual o bloco será segmentado, baseando-se num critério local e não numa imposição prévia.

Segundo o novo esquema implementado, cada bloco  $X^l$  da imagem a codificar pode ser segmentado em qualquer direção, aplicando-se este procedimento recursivamente a cada um dos sub-blocos resultantes até atingir o nível 0 da árvore de segmentação. Isto implica a existência de novas escalas para o dicionário, de modo a contemplar as novas dimensões de blocos que este esquema impõe. Por exemplo, o MMP original apenas utilizava blocos quadrados e blocos retangulares que cumpriam com a relação 2:1, resultando num total de 9 escalas distintas ( $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 8$ , ...,  $1 \times 1$ ). O novo esquema de segmentação eleva o número de escalas para 25 ( $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 16$  (novo),  $16 \times 4$  (novo),  $4 \times 16$  (novo), ...,  $1 \times 1$ ), sem restrições do ponto de vista da relação das suas dimensões, o que significa que todos os blocos

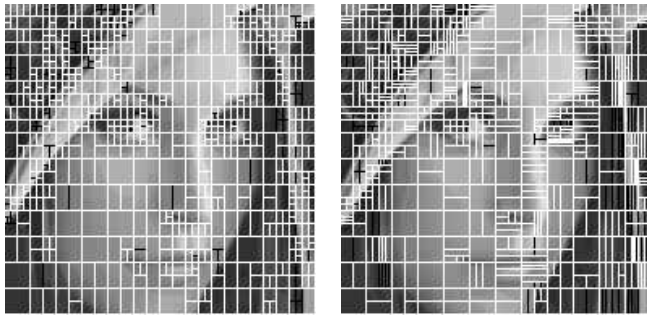


Fig. 2. Segmentação na codificação da imagem LENA com esquema de segmentação convencional (esquerda) e com novo esquema (direita).

com dimensões  $2^m \times 2^n$  passam a estar disponíveis, para  $m, n = 0, \dots, 4$ . A Figura 2 ilustra a segmentação do bloco de predição para ambos os esquemas (linhas brancas) e as partições adicionais utilizadas para codificar o resíduo (linhas pretas). É fácil visualizar na figura o aparecimento de blocos com configurações mais variadas do que acontecia na versão anterior.

O passo de otimização da árvore de segmentação passa a avaliar o custo inerente a três opções, segmentar na vertical, horizontal ou não segmentar, de modo a obter o custo mínimo de representação do bloco. Quando a decisão é segmentar o bloco numa dada direção, os nós da árvore originados pela segmentação na direção alternativa são eliminados. Se o menor custo estiver associado à não segmentação do bloco, então todos os nós filhos associados às segmentações são eliminados, e o nó transforma-se num nó terminal, tal como acontecia anteriormente.

É imposta a utilização de uma *flag* adicional na representação da árvore de segmentação a somar às já existentes, de modo a especificar a direção segundo a qual a partição do bloco foi efectuada. Como consequência directa deste novo esquema de segmentação, a partição dos blocos torna-se muito mais flexível, permitindo que o MMP se adapte de uma forma muito mais eficiente às características do sinal a codificar.

Tendo em conta que o novo esquema de segmentação passa a ser utilizado tanto no passo de predição como na codificação do resíduo, torna-se possível gerar blocos de predição com uma qualidade bastante superior ao conseguido pelo esquema anterior. Tal tem como consequência a diminuição da energia presente nos blocos de resíduo, permitindo a sua codificação de uma forma bastante mais eficiente. Um exemplo é o caso dos blocos estreitos (*e.g.*  $16 \times 1$ ), onde a predição tende a ser bastante melhor do que aquela que se verifica para um bloco quadrado, dada a proximidade com os pixels de referência.

#### IV. TREINO DO DICIONÁRIO

O dicionário inicial utilizado pelo MMP é bastante simples, contendo apenas para cada escala, blocos homogêneos distribuídos ao longo da faixa dinâmica do sinal de entrada. O aumento do seu poder de aproximação depende da diversificação dos padrões disponíveis, conseguida pela inserção de novos padrões resultantes de concatenações de blocos da parte da imagem já processada [1].

Assim, um crescimento rápido e bem estruturado do dicionário é de extrema importância, de forma a reduzir a percentagem de blocos que necessita ser codificada até o dicionário começar a apresentar uma diversidade de padrões que permita reduzir o número médio de segmentações impostas a cada bloco. O maior número de segmentações nos primeiros blocos processados terá um impacto significativo, que deverá ser minimizado com vista a um aumento do desempenho global do algoritmo [3].

O processo de treino que apresentamos tem como objetivo gerar um conjunto de blocos, que sendo disponibilizados ao codificador e ao decodificador, servirão como complemento ao dicionário inicial do MMP. Para que o seu efeito seja efetivo, o conjunto de padrões deverá ser o mais diversificado possível, dentro da distribuição probabilística dos blocos de resíduo a codificar.

O procedimento adotado passou por criar dois grupos de quatro imagens de teste. As imagens de cada um dos grupos foram codificadas sequencialmente, todas com o mesmo valor de  $\lambda$ , carregando para o dicionário inicial de cada codificação, todos os blocos previamente utilizados nas anteriores. Os blocos foram inseridos no contexto e escala onde foram originalmente criados, mas a frequência de ocorrência dos blocos não foi mantida, de modo a evitar condicionar o codificador a usar blocos anteriormente usados. A única informação tida em conta refere-se ao fato de o bloco já ter sido ou não usado ao longo do processo de treino. Após terminar a codificação de um grupo de quatro imagens com um dado valor de  $\lambda$ , adotou-se um novo valor para este parâmetro, e o procedimento continuou. Ao alterar assim a taxa de compressão, pretendeu-se diversificar as características dos blocos criados. Valores de  $\lambda$  maiores, impõem menos segmentações e por isso inserção de blocos em escalas superiores, enquanto que  $\lambda$  pequenos resultam na inserção de mais blocos nas escalas pequenas. Pretendia-se assim que o dicionário de referência gerado pudesse ser o mais independente possível da taxa de compressão. Foram utilizados três valores distintos para  $\lambda$ , distribuídos ao longo da gama de taxas de compressão pretendidas. A sequência de codificação foi realizada por ordem crescente e decrescente de valores de  $\lambda$ , tendo-se verificado que tal não tem influência nos resultados obtidos.

Após terminado o processo de treino, o dicionário de referência gerado pelo primeiro grupo de imagens foi usado para codificar duas imagens representativas do segundo, e vice-versa.

O teste de desempenho de codificação utilizando o dicionário de referência foi realizado segundo duas abordagens. Na primeira, os elementos do dicionário de referência foram inseridos no contexto a que pertenciam originalmente, tal como foi feito ao longo do processo de treino. Tentou-se com isto preservar alguma da informação de entropia decorrente das diferenças de probabilidades entre contextos. A segunda abordagem visou a criação de um novo contexto para cada escala, especialmente destinado aos blocos provenientes do treino. Por outras palavras, o dicionário continuou a apresentar a mesma configuração e inicialização, mas foi adicionado um novo contexto onde todos os blocos provenientes do treino são inseridos. Com isto, pretendeu-se garantir que mesmo na

hipótese menos favorável de nenhum dos blocos de treino vir a ser utilizado, o codificador aritmético adaptativo pudesse rapidamente minimizar o impacto da inserção do novo contexto na taxa, não prejudicando o desempenho global do codificador.

Observou-se que a primeira abordagem tende a apresentar melhores resultados a taxas de compressão elevadas do que a segunda. Todavia, para taxas de compressão baixas onde a probabilidade de utilização dos blocos provenientes do dicionário de treino é baixa, a entropia adicional criada por estes faz com que os resultados tendam a ser inferiores aos obtidos sem recurso ao dicionário de treino. O fato da segunda abordagem contemplar a introdução destes elementos num contexto próprio, permite que o impacto na taxa advindo da introdução do contexto correspondente ao treino se anule rapidamente, pelo que esta abordagem possibilita que os resultados da versão com treino nunca sejam piores do que os da versão sem treino, mesmo para baixas taxas de compressão. Porém, para taxas de compressão elevadas, os resultados são marginalmente inferiores aos conseguidos pela primeira abordagem. Adotou-se então a segunda abordagem, por considerar-se que apresenta resultados mais consistentes ao longo da gama de taxas de compressão pretendidas.

## V. RESULTADOS EXPERIMENTAIS

O novo esquema de segmentação flexível foi implementado no MMP-II, tendo sido realizados alguns testes experimentais. O novo algoritmo foi utilizado para codificar imagens suaves em escala de cinza, e com tamanho inicial de bloco de  $16 \times 16$  pixels. A predição foi definida para todas escalas cujos blocos contemplam 16 ou mais pixels. Esta limitação visou atingir um compromisso entre o desempenho e a complexidade computacional do algoritmo. As imagens utilizadas encontram-se disponíveis para download em [9].

Verificou-se para este tipo de imagens que o novo esquema de segmentação possibilitou atingir ganhos em relação ao MMP-II, que se situam, em média, nos 0.4dB, para o PSNR das imagens testadas. Este acréscimo de desempenho permitiu ao MMP superar o H.264/AVC [6] e o JPEG2000 [10], na codificação de imagens suaves. O MMP-FP passou a apresentar resultados consistentemente superiores aos destes algoritmos, acima dos 0.3bpp (ver Tabela I).

A versão com treino do dicionário foi testada com o MMP-FP, de modo a observar qual o efeito combinado destas duas novas implementações na codificação de imagens suaves. Estes testes experimentais realizados visaram igualmente testar o desempenho comparativo das duas opções distintas de inserção dos blocos provenientes do treino. A primeira observação relativa aos resultados é que, o treino do dicionário proporciona um aumento mais significativo do desempenho para taxas de compressão elevadas do que para taxas de compressão baixas. Esta particularidade era esperada, uma vez que para taxas de compressão elevadas ocorrem poucas segmentações de blocos, resultando num crescimento muito mais limitado do dicionário. Desta forma, qualquer auxílio ao seu crescimento torna-se bastante mais significativo do que em casos onde este atinge dimensões elevadas.

Na Tabela I, são apresentados os resultados detalhados obtidos com o H.264/AVC, o JPEG2000, o MMP-II, o MMP

com segmentação flexível, e com cada um dos procedimentos de codificação com treino apresentados, para a imagem LENA.

As Figuras 3 a 5 apresentam os resultados finais obtidos com a aplicação das duas técnicas apresentadas neste artigo. Para imagens suaves foi possível obter ganhos que se situam por volta dos 0.6dB em relação ao MMP-II, ultrapassando os resultados dos algoritmos baseados em transformadas para este tipo de imagens. Os resultados do MMP passaram a superar os destes algoritmos, para taxas a partir dos 0.15bpp, chegando a estar 0.4dB acima do melhor deles. A Figura 6 apresenta os resultados obtidos na codificação de uma imagem de texto (PP1205). Pode-se verificar pelo gráfico que o fato do treino ter sido efetuado apenas com imagens suaves, não só não prejudicou o desempenho do algoritmo para uma imagem não suave como ainda permitiu obter alguns ganhos marginais. A segmentação flexível permitiu também um acréscimo de desempenho para estas imagens, embora inferior ao conseguido para imagens suaves. O fato da predição não apresentar um desempenho tão bom para este tipo de imagens significa que a presença da informação adicional inerente à predição prejudica o ganho conseguido na codificação mais eficiente do resíduo. Para esta imagem de texto a diferença de desempenho entre o MMP e os algoritmos baseados em transformadas situa-se nos 4.5dB em relação ao H.264/AVC e nos 6dB em relação ao JPEG2000.

Estes resultados demonstram a universalidade do algoritmo, colocando o desempenho do MMP acima do apresentado pelos melhores algoritmos de compressão baseados em transformadas, para todas as imagens testadas.

TABELA I  
RESULTADOS PARA IMAGEM LENA EM DB

Taxa	0.15	0.30	0.45	0.6	0.75	0.90
H.264/AVC	<b>32.14</b>	35.03	36.75	37.93	38.97	39.77
JPEG2000	31.77	34.89	36.74	37.99	39.01	39.80
MMP-II	31.55	34.72	36.50	37.80	38.87	39.78
MMP-FP	31.81	35.03	36.85	38.18	39.26	40.12
MMP-FP Treino 1	32.09	<b>35.27</b>	<b>37.06</b>	38.29	39.23	40.02
MMP-FP Treino 2	32.05	35.20	37.03	<b>38.29</b>	<b>39.31</b>	<b>40.96</b>

## VI. CONCLUSÕES

Neste artigo, foram apresentadas duas novas técnicas aplicadas ao algoritmo MMP, que permitiram obter ganhos na compressão de imagens suaves, sem comprometer o seu excelente desempenho em imagens de texto. De fato, estas duas novas técnicas colocaram o desempenho do MMP acima do apresentado pelos algoritmos baseados nas transformadas DCT e DWT, aceites como os melhores algoritmos para compressão de imagens naturais.

O esquema de treino aqui apresentado, apesar de simples, permitiu validar o conceito de treino do dicionário ao apresentar ganhos significativos. Este fato deixa alguns indícios de que esquemas que envolvam a implementação de técnicas de treinamento de dicionário devam ser mais investigados.

Os resultados atingidos evidenciam a universalidade do MMP, sendo que este algoritmo supera agora o H.264/AVC e

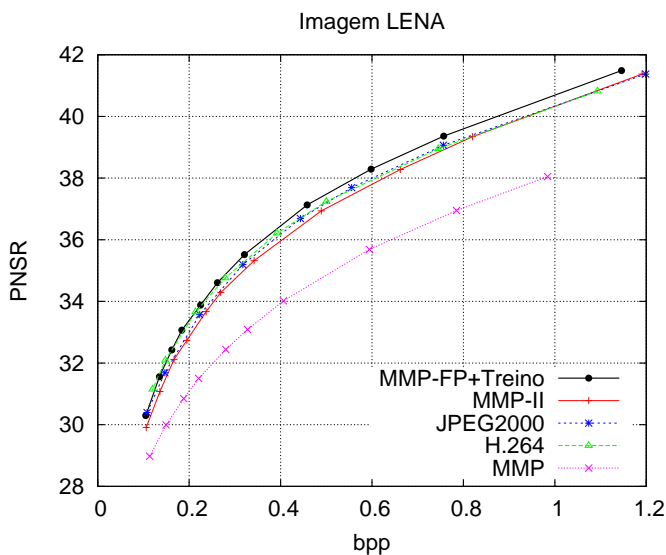


Fig. 3. Resultados experimentais para a imagem LENA 512x512.

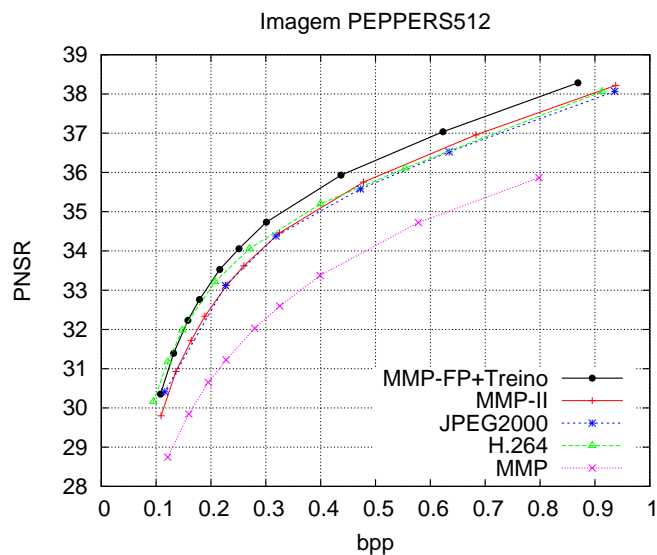


Fig. 5. Resultados experimentais para a imagem PEPPERS 512x512.

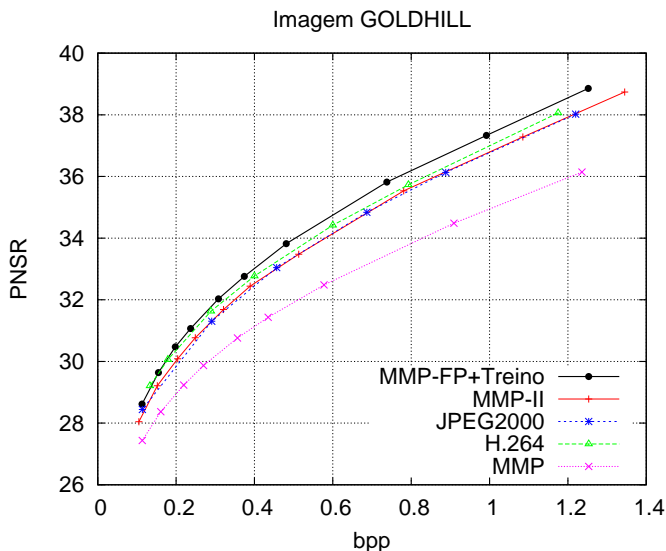


Fig. 4. Resultados experimentais para a imagem GOLDHILL 512x512.

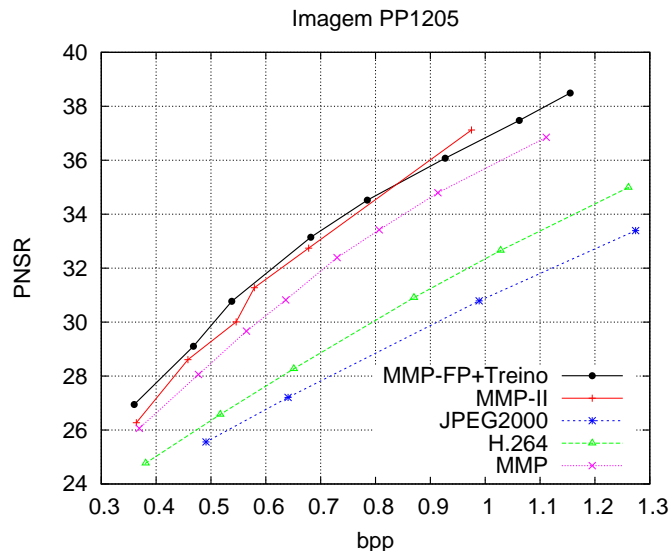


Fig. 6. Resultados experimentais para a imagem PP1205 512x512.

o JPEG2000 para imagens suaves, aumentando a vantagem de 4 e 6dB, respectivamente, sobre estes algoritmos quando aplicada a imagens de texto. O MMP afirma-se assim como uma excelente alternativa ao paradigma transformada-quantização-codificação vigente.

REFERÊNCIAS

[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, Novembro 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, Julho 2006.  
 [3] José F. L. Oliveira, E. A. B. da Silva, and Murilo B. Carvalho, "Recorrência de padrões multiescalas com dicionários treinados," *XXV Simpósio Brasileiro de Telecomunicações*.  
 [4] N. M. M. Rodrigues, E. A. B. da Silva, M. B. de Carvalho, S. M. M. de Faria, and Vitor M. M. Silva, "Improving h.264/avc inter compression with multiscale recurrent patterns," *IEEE International Conference on Image Processing*, Outubro 2006.

[5] 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*, Génova, Itália, Setembro 2005.  
 [6] 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)*, Março 2005.  
 [7] 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," *Aceito para publicação no IEEE Transactions on Image Processing*.  
 [8] 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*, Ilha de Kos, Grécia, Maio 2006.  
 [9] <http://www.estg.ipleiria.pt/~nuno/MMP/>.  
 [10] D. S. Taubman and M.W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, 2001.