

Escolha da direção preferencial de processamento do algoritmo de extrapolação EI para codificação de imagens baseada em objeto

Flávia Magalhães Freitas e Abraham Alcaim

Resumo—Este trabalho contribui para o debate sobre qual das estratégias baseadas nas variâncias dos comprimentos dos segmentos do objeto - MILV ou MALV - é mais adequada à determinação da direção preferencial para o algoritmo de extrapolação EI - Extension Interpolation. A aplicação desse algoritmo na codificação orientada por objeto com DCT baseada em blocos tem mostrado excelentes resultados a taxas baixas e médias. Entretanto, a eficiência de codificação ainda pode ser melhorada com a escolha da estratégia mais adequada para determinar a direção preferencial de processamento na extrapolação dos blocos de contorno. Pretende-se também verificar a possibilidade de relacionar a eficiência das estratégias MILV e MALV ao indicador morfológico TNPO - Textura média dos blocos, de acordo com o Número de Pixels do Objeto.

Palavras-Chave— Codificação orientada por objeto, técnicas de extrapolação, DCT baseada em blocos, indicadores morfológicos.

Abstract— This paper contributes to the discussion on which of the strategies based on the variances of object segments lengths - MILV or MALV - is more adequate to determine the preferential direction for the EI - Extension Interpolation - padding technique. The application of this algorithm to object-based coding with block-based DCT has shown good results at low and medium rates. However, the coding efficiency can be improved by choosing the more adequate strategy for determining the preferential direction of the padding scheme. In this paper, we also intend to analyse the possibility of relating the efficiency of the strategies MILV and MALV to the morphological feature TNOP - Texture of blocks, according to the Number of Object Pixels.

Keywords— Content-based video coding, padding techniques, block-based DCT, morphological features.

I. INTRODUÇÃO

Nos últimos anos, um grande esforço tem sido focado na melhoria do desempenho de codificação das transformadas orientadas por objeto. Nesse campo, há basicamente duas abordagens diferentes para processar a imagem: a primeira usa a DCT adaptativa à forma (SA-DCT) [1]-[3],[7]-[8] e a segunda emprega esquemas de DCT baseada em blocos após processar os blocos de contorno usando-se algoritmos de extrapolação [4]-[6],[7]-[8].

Assim como na SA-DCT, que consiste de dois processamentos unidimensionais (aplicação de duas DCT-1D), também o algoritmo de extrapolação EI (Extension Interpolation) [4],[5]

processa os segmentos do objeto primeiro em uma direção e depois na outra. Esforços têm sido concentrados na concepção de estratégias para a determinação da direção a ser adotada como preferencial no processamento, com o intuito de melhorar a eficiência de codificação das transformadas.

Em [7], foi apresentado um critério ótimo no sentido R-D (taxa-distorção), que seleciona a direção preferencial que produz o menor produto $emq \times nbits$, onde emq é o erro médio quadrático medido após a recuperação do bloco e $nbits$ é o número total de bits usados na quantização dos seus coeficientes. O critério pode ser usado tanto na SA-DCT quanto no algoritmo de extrapolação EI e, obviamente, supera em desempenho todas as outras estratégias existentes para a escolha da direção preferencial de processamento, necessitando da transmissão de um bit por bloco como informação paralela. Contudo, apresenta a desvantagem de exigir o cálculo de um novo conjunto de coeficientes a cada vez que ocorrer mudança na taxa de transmissão.

Outras estratégias para a seleção da direção preferencial de processamento reportadas na literatura baseiam-se na forma do objeto segmentado [5],[7], o que não implica em nenhuma informação paralela adicional no caso das transformadas orientadas por objeto. Essas estratégias produzem coeficientes que não dependem da taxa de bits, permitindo codificações rápidas e fáceis para diferentes aplicações. A direção preferencial de processamento em um bloco de contorno é determinada pela variância dos comprimentos dos segmentos do objeto nesse bloco em cada uma das direções. Seguindo essa abordagem, há duas estratégias distintas: (1) a MILV (Mínimum Lengths Variance) [5], que calcula as variâncias dos comprimentos das linhas e das colunas do objeto e seleciona a direção associada à menor dessas variâncias e (2) a MALV (Máximum Lengths Variances) [7], que seleciona a direção associada à maior dessas variâncias.

Ainda não há na literatura um consenso sobre qual das estratégias - MILV ou MALV - proporciona os melhores resultados quando se emprega o algoritmo EI em codificação de imagens baseada em objeto. Em [5], foi proposto o emprego da estratégia MILV na escolha da direção preferencial de processamento do algoritmo de extrapolação EI. Contudo, essa escolha foi criticada em [7], onde a estratégia MALV foi apontada como a mais eficiente tanto para a SA-DCT, quanto para o algoritmo de extrapolação EI. Neste artigo, busca-se avaliar essas conclusões contraditórias através de uma análise criteriosa dessas estratégias, com base em indicadores que relacionem características morfológicas e de textura. A Seção

II apresenta uma breve descrição do algoritmo EI, resultados preliminares relativos aos desempenhos das estratégias MALV e MILV e o indicador a ser utilizado na análise dessas estratégias. Na Seção III são descritas as condições dos testes experimentais e a Seção IV contém resultados e discussão. Finalmente, a Seção V é dedicada às conclusões do trabalho.

II. O ALGORITMO EI, RESULTADOS PRELIMINARES E INDICADORES MORFOLÓGICOS

O algoritmo EI [4],[5] é um esquema de extrapolação para codificação de imagens orientada a objeto que utiliza DCT baseada em blocos de dimensão $N \times N$ (em oposição à SA-DCT [1], que utiliza dimensões variáveis nas duas direções). Esse algoritmo aplica a DCT-1D em uma direção — horizontal ou vertical — aos M pixels do objeto (na linha ou coluna correspondente), gerando M pontos no domínio da transformada. Em seguida, são inseridos $(N - M)$ zeros nesse domínio e aplicada a DCT-1D inversa aos N pontos resultantes. O resultado representa N pontos interpolados no domínio espacial. Considerando que o número de segmentos não nulos na direção preferencial é S , resultarão S segmentos com N pixels. O mesmo procedimento é realizado na outra direção, onde a DCT-1D é aplicada sobre N segmentos de tamanho S .

Como mencionado na Introdução, em [5] a estratégia MILV foi empregada para a escolha da primeira direção de processamento (direção preferencial) no algoritmo EI. Por outro lado, em [7] a estratégia MALV foi apresentada como a melhor opção para a escolha da direção preferencial. Entretanto, nossos resultados experimentais preliminares indicaram que para alguns objetos, normalmente aqueles de baixa textura, a estratégia MALV é superior à MILV e para outros, normalmente os de alta textura, o contrário se verifica. Portanto, é incorreto afirmar que MALV é melhor que MILV ou *vice-versa*. Na verdade, para cada bloco de contorno específico, uma das estratégias pode ser preferível à outra, produzindo um *emq* menor.

A partir de resultados preliminares nós observamos, também, que normalmente a escolha da direção mais adequada em um determinado bloco depende de características morfológicas e também da textura. Um possível indicador nesse sentido pode ser obtido dividindo-se os blocos de contorno em grupos que contêm números de pixels do objeto semelhantes — considerando desde grupos mais vazios a grupos mais cheios — e determinando as texturas médias em cada grupo. Esse indicador, denominado TNPO (distribuição das texturas de acordo com o número de pixels do objeto em cada grupo) foi apresentado em [8], por Freitas e Alcaim, como uma ferramenta potencial para a seleção do esquema de codificação mais adequado a uma dada taxa de bits. O que se pretende verificar é se pode ser feita a opção por MILV ou MALV em um bloco específico, com boa precisão, com base no indicador TNPO, ou seja, observando-se quão cheio ou vazio de pixels do objeto ele é, bem como a textura média nos blocos com um número de pixels equivalente.

III. CONDIÇÕES EXPERIMENTAIS

Os experimentos foram realizados sobre um plano de objeto de vídeo (VOP) de três seqüências com 11 quadros: WEATHER (144x176 pixels), CHILDREN (144x176 pixels) e FISH AND LOGO (240x352 pixels). Foram ainda codificados um objeto dos quadros AKYIO (320x384 pixels), LENA (512x512 pixels) e PEPPER (512x512 pixels). Utilizaram-se blocos de imagem de tamanho 8×8 , os blocos de contorno foram extrapolados usando o algoritmo EI associado às estratégias MALV e MILV e finalmente, todos os blocos foram codificados usando a DCT-2D. Como medida de desempenho, foi utilizado o ganho em dB da razão pico/ruído (*RPR*) versus a taxa de bits por pixel do objeto (*bppo*).

Os blocos de contorno foram divididos em 10 grupos, G1 a G10. Em G1 ficaram os blocos contendo de 1 a 7 pixels do objeto, em G2 blocos com 8 a 13 pixels, em G3 blocos com 14 a 19 pixels, em G4 blocos com 20 a 25 pixels, em G5 blocos com 26 a 32 pixels, em G6 blocos com 33 a 38 pixels, em G7 blocos com 39 a 44 pixels, em G8 blocos com 45 a 50 pixels, em G9 blocos com 51 a 56 pixels e em G10, blocos com 57 a 63 pixels do objeto. O desvio padrão das luminâncias dos pixels em cada bloco foi calculado e a média dos desvios padrões dos blocos pertencentes a um determinado grupo irá representar a textura média desse grupo, sendo armazenada no indicador TNPO.

O coeficiente DC foi uniformemente quantizado com 8 bits e os coeficientes AC foram quantizados não uniformemente, com um número de bits dependente das respectivas variâncias, considerando uma distribuição de probabilidade laplaciana. A alocação de bits depende dos coeficientes finais da transformada e é expressa por

$$B_i = B_0 + \frac{1}{2} \log_2 \frac{\sigma_i^2}{(\prod_{j=1}^{64} \sigma_j^2)^{\frac{1}{64}}} \quad i = 2, \dots, 64 \quad (1)$$

onde B_i é o i -ésimo coeficiente AC, σ_i^2 é a sua variância e B_0 é um parâmetro que permite variar a taxa de bits média. Nas simulações a seguir, B_0 foi variado de 0,2 a 8 bppo (bits por pixel do objeto), em 17 valores distintos.

Para uma dada taxa B_0 requerida para a codificação do objeto como um todo, foram computados os desempenhos de todos os blocos de contorno e também os resultados tomando-se apenas os blocos de um determinado grupo. Cada uma das curvas apresentadas na Seção III terão 17 pontos, que serão os resultados obtidos para cada B_0 . Obviamente, para um dado B_0 , os blocos pouco preenchidos apresentarão taxas de bppo mais altas, ao passo que os blocos muito preenchidos apresentarão baixas taxas de bppo.

IV. RESULTADOS E DISCUSSÃO

A Figura 1 mostra o indicador morfológico TNPO para um objeto de AKYIO, CHILDREN, FISH AND LOGO, WEATHER, LENA e PEPPER. Esse indicador será relacionado à eficiência das estratégias MALV e MILV para a determinação da direção preferencial do algoritmo de extrapolação EI, nos blocos pertencentes a cada um dos 10 grupos.

A partir do gráfico da Figura 1, constata-se que os objetos AKYIO e PEPPER apresentam baixas texturas para todos os grupos, ao passo que CHILDREN e FISH AND LOGO normalmente apresentam texturas mais elevadas. LENA e WEATHER apresentam, em geral, texturas de médias a altas, dependendo do grupo.

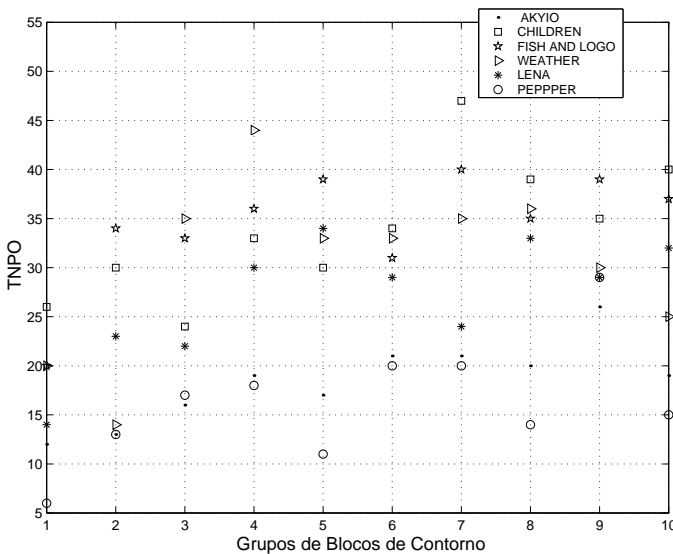


Fig. 1. Indicador Morfológico TNPO - Textura de acordo com o Número de Pixels do Objeto

Nas Figuras 2 a 7, são apresentados os resultados de desempenho de codificação dos objetos ao se usar as estratégias MALV e MILV associadas ao algoritmo de extrapolação EI empregado na DCT baseada em blocos. Nessas figuras, observa-se uma tendência: a baixas taxas de bits, aproximadamente até 2 bppo, a estratégia MILV geralmente é mais eficiente para todos os objetos. A partir dessa taxa, normalmente os objetos em que todos os grupos têm textura baixa, como AKYIO e PEPPER, são melhor codificados se a direção preferencial para a extrapolação dos blocos de contorno for indicada pela estratégia MALV, ao passo que os objetos que apresentam textura elevada em praticamente todos os grupos, como CHILDREN e FISH AND LOGO, são mais eficientemente codificados se os blocos de contorno forem processados primeiro pela direção apontada pela estratégia MILV.

Dessa forma, a afirmação em [7] de que MALV é melhor que MILV, ao se apresentar o resultado da codificação de apenas um quadro da seqüência WEATHER, não é verdadeira. Na Figura 2, apresentam-se os resultados de codificação para os 11 primeiros quadros da seqüência WEATHER, onde se observa que MALV foi superior a MILV somente a taxas de médias a altas. Para taxas baixas, embora ambas as estratégias proporcionem desempenhos praticamente idênticos, ocorreu uma ligeira supremacia de MILV em relação a MALV.

Dos resultados experimentais nós verificamos, ainda, que as eficiências das estratégias MILV e MALV apresentam uma variação não apenas de objeto para objeto, como também, em

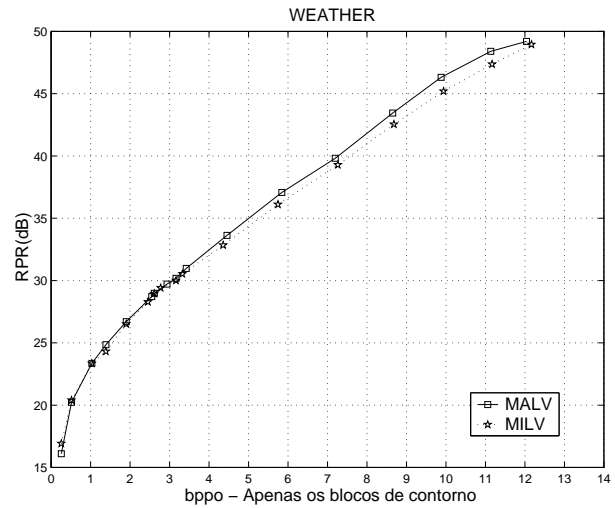


Fig. 2. Resultados de codificação para 11 quadros de um objeto da seqüência WEATHER

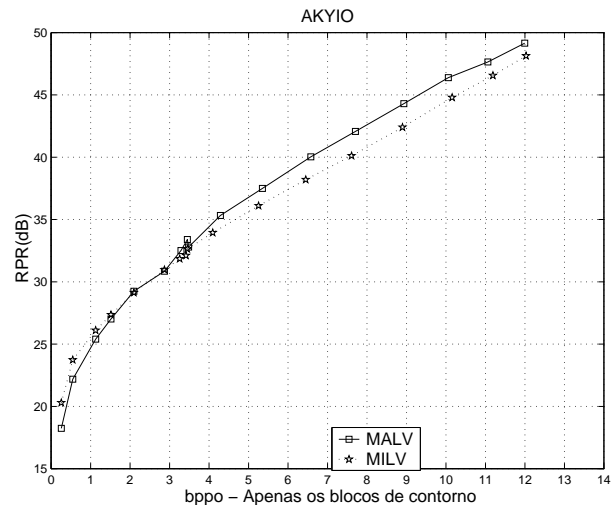


Fig. 3. Resultados de codificação para um objeto do quadro AKYIO

um mesmo objeto, de grupo para grupo. Para grupos com texturas muito baixas ou muito altas, a avaliação do indicador TNPO fornece uma aproximação razoável para a escolha da estratégia mais adequada. Nos grupos com texturas médias, onde pode haver um número razoável de blocos com alta textura e também outros com baixa textura, a análise com base unicamente no indicador TNPO é dificultada. Observamos também que o limiar das texturas classificadas como altas ou baixas variam de grupo para grupo. Nos blocos medianamente preenchidos por pixels do objeto, torna-se mais difícil a escolha desses limiares para a determinação da estratégia mais eficiente.

Na estratégia MALV, a direção preferencial será aquela onde a maioria dos pixels do objeto está localizada em segmentos cheios. Portanto, em blocos de baixa textura, onde todos os segmentos tendem a estar mais suavizados mesmo antes da extrapolação, é melhor usar a estratégia MALV, que aplica

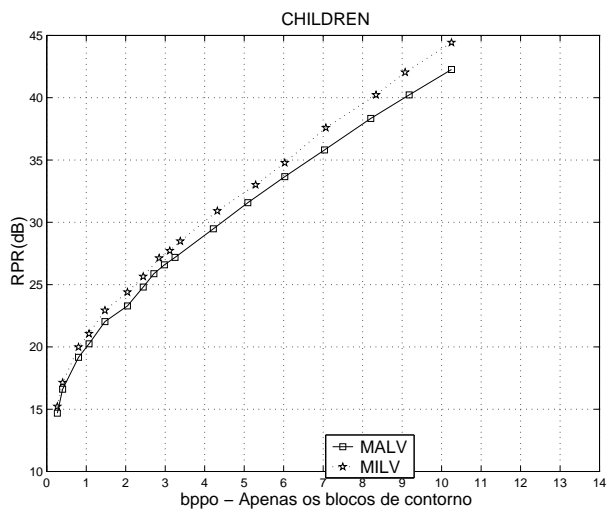


Fig. 4. Resultados de codificação para 11 quadros de um objeto da seqüência CHILDREN

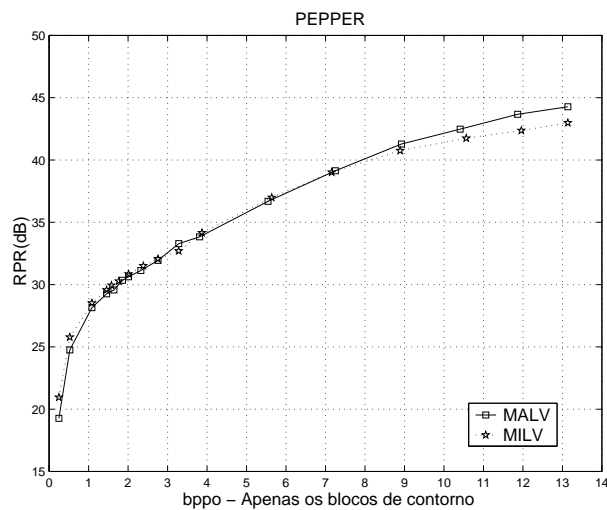


Fig. 7. Resultados de codificação para um objeto do quadro PEPPER

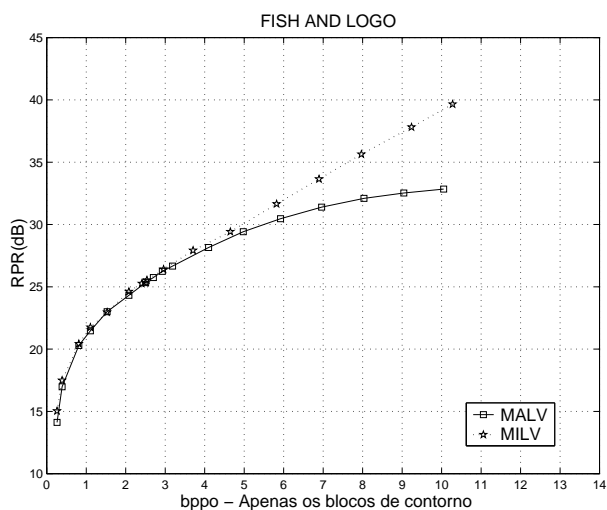


Fig. 5. Resultados de codificação para 11 quadros de um objeto da seqüência FISH AND LOGO

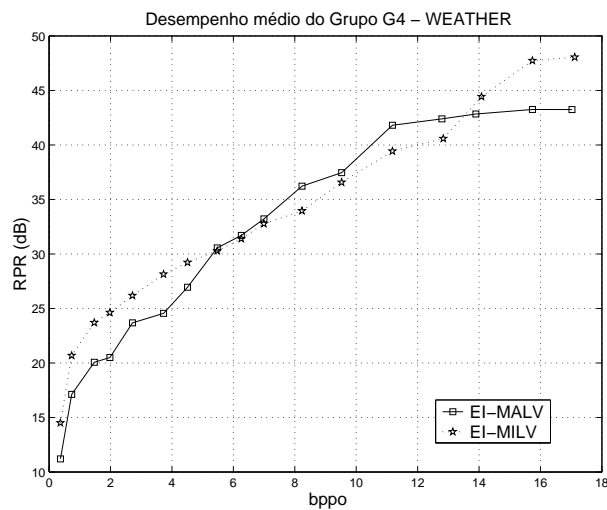


Fig. 8. Resultados de codificação dos blocos do grupo G4 de WEATHER (alta textura)

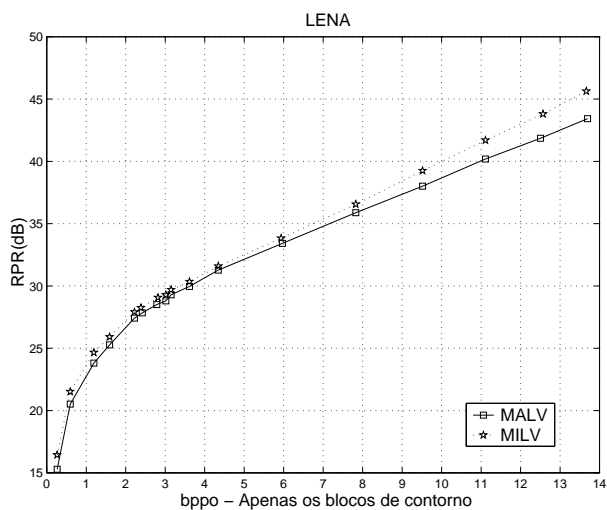


Fig. 6. Resultados de codificação para um objeto do quadro LENA

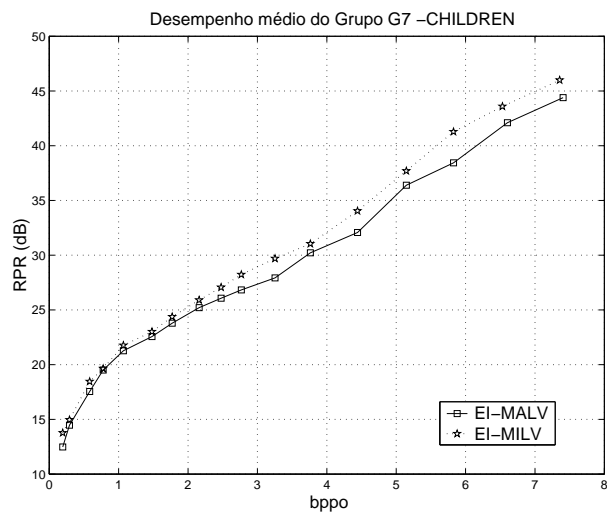


Fig. 9. Resultados de codificação dos blocos do grupo G7 de CHILDREN (alta textura)

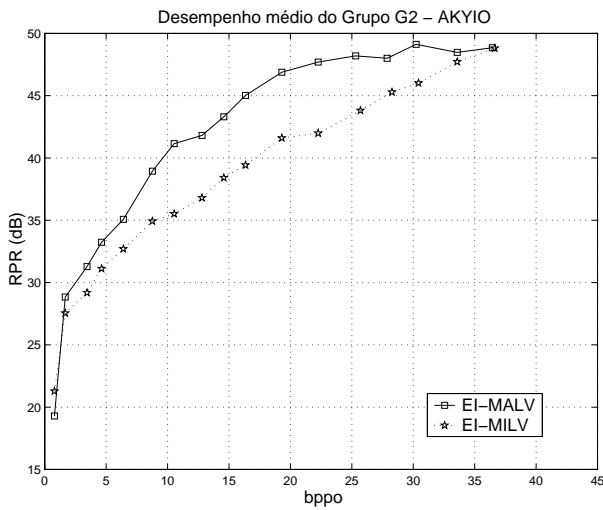


Fig. 10. Resultados de codificação dos blocos do grupo G2 de AKYIO

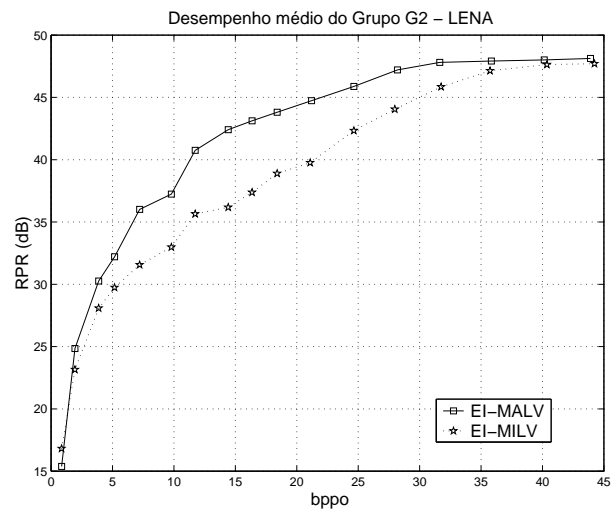


Fig. 11. Resultados de codificação dos blocos do grupo G2 de LENA

a primeira DCT-1D sobre segmentos maiores. Na estratégia MILV, a direção preferencial será aquela onde os pixels estão mais uniformemente distribuídos por segmentos de tamanhos mais uniformes. Portanto, em blocos de alta textura, após a primeira DCT-1D esses segmentos serão mais suavizados que os segmentos longos de MALV, proporcionando a compactação de energia em um menor número de coeficientes. Como exemplo são mostradas as Figuras 8 e 9, que ilustram casos onde o indicador TNPO apresenta valores altos (veja Figura 1).

Nos grupos G1 e G10, as estratégias MALV e MILV mostraram desempenhos razoavelmente equivalentes para todos os objetos. As diferenças maiores ocorreram nos blocos medianamente preenchidos. No grupo G2, por exemplo, observou-se uma forte dependência entre o indicador TNPO e a eficiência das estratégias MILV e MALV. As Figuras 10 a 12 apresentam o desempenho de codificação dos blocos do grupo G2 para os objetos examinados. Em AKYIO (Figura 10), LENA (Figura 11) e PEPPER (Figura 12), onde os blocos do grupo G2 apresentam texturas de valores baixos, a estratégia MALV foi superior à estratégia MILV. Por outro lado, em CHILDREN (Figura 13) e FISH AND LOGO (Figura 14), onde os blocos do grupo G2 apresentam texturas mais altas, a estratégia MILV foi superior. Entretanto, para os blocos do grupo G2 de WEATHER, que também tem textura baixa, os desempenhos de MILV e MALV se alternam, como mostra a Figura 15.

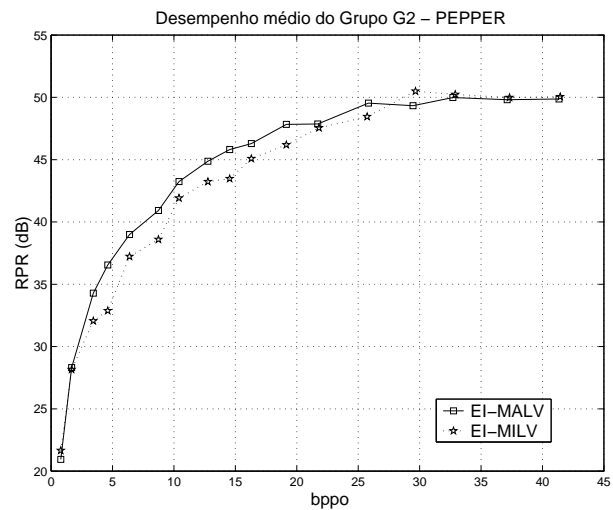


Fig. 12. Resultados de codificação dos blocos do grupo G2 de PEPPER

Os resultados mostraram que apesar da tendência ser de em blocos com textura pequena a estratégia MALV fornecer melhores resultados, infelizmente nem sempre isso pode ser garantido. Apesar disso, é razoável argumentar que uma escolha individual por grupo de blocos da estratégia mais adequada para a determinação da direção pode aumentar a eficiência de codificação da DCT baseada em blocos usando-se o algoritmo de extrapolação EI.

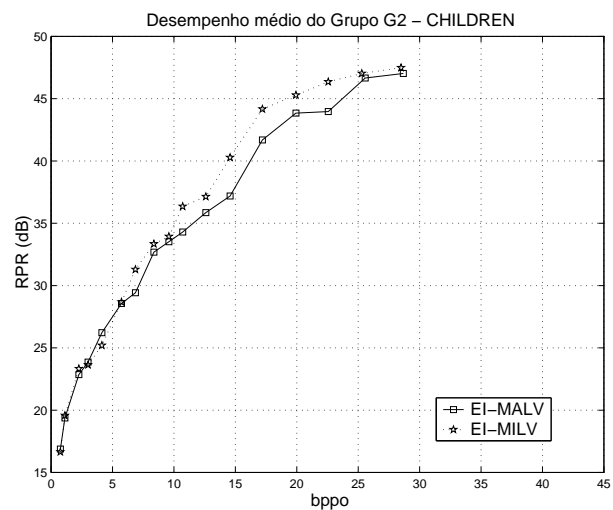


Fig. 13. Resultados de codificação dos blocos do grupo G2 de CHILDREN

Obviamente, além da textura do grupo, outras variáveis, como os comprimentos e os deslocamentos dos segmentos de objeto, o número de segmentos do objeto em cada direção e a textura do bloco específico interferem bastante

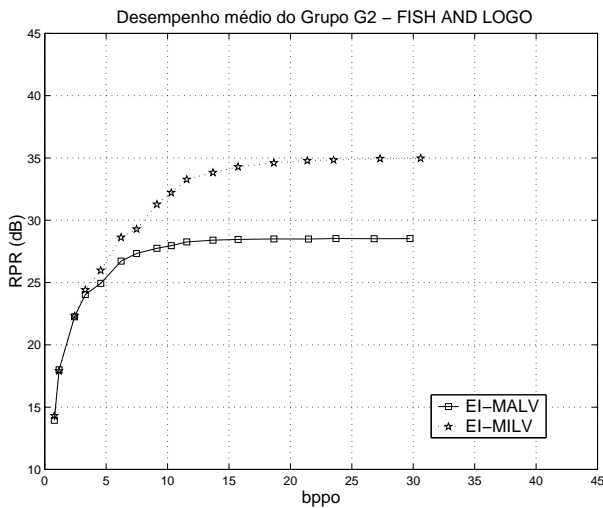


Fig. 14. Resultados de codificação dos blocos do grupo G2 de FISH AND LOGO

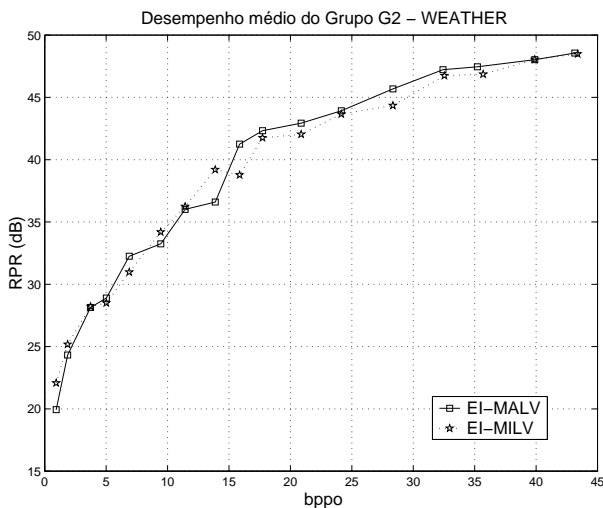


Fig. 15. Resultados de codificação dos blocos do grupo G2 de WEATHER

na determinação da direção preferencial de processamento mais eficiente. Contudo, há de se convir que uma avaliação que levasse em consideração todos esses parâmetros seria excessivamente complexa, dificultando a sua aplicação.

V. CONCLUSÕES

Este artigo examinou estratégias de escolha da direção preferencial de processamento do algoritmo EI, para codificação de imagens baseada em objeto. Foi mostrado que nem sempre a eficiência de MALV é melhor que a de MILV, como havia sido afirmado em [7]. Na verdade, para alguns objetos, normalmente os de alta textura, MILV chega a ser bem superior a MALV. Observou-se ainda que um mesmo objeto pode apresentar variações nos comportamentos médios dos desempenhos de MILV e MALV para grupos de blocos distintos, apontando para a possibilidade da escolha em separado da estratégia mais adequada a cada grupo.

Foi mostrado também que para grupos com textura alta, a estratégia MILV normalmente é preferível à estratégia MALV.

Para grupos de baixa textura, normalmente a estratégia MALV proporciona melhores ganhos. Contudo, para grupos de textura média, a correlação entre o indicador TNPO e a eficiência das estratégias MILV e MALV é reduzida.

Os resultados obtidos sinalizam para a importância de se pensar uma estratégia para a determinação da direção preferencial de processamento para os casos em que a correlação entre as eficiências das estratégias MILV e MALV e o indicador TNPO seja baixa. Obviamente, para que sua eficiência possa ser melhor relacionada ao indicador de textura do objeto, essa estratégia deverá ser baseada não apenas na forma do objeto, mas também nos coeficientes obtidos para o bloco ao se usar como preferencial ora a direção horizontal, ora a vertical. Mesmo que isso implique na transmissão de um bit extra para cada bloco, esse inconveniente poderá ser compensado pela melhoria da eficiência de codificação. A vantagem dessa nova estratégia sobre a estratégia ótima no sentido R-D [7] seria a utilização dos mesmos coeficientes no domínio das frequências espaciais, qualquer que seja a taxa de transmissão.

REFERÊNCIAS

- [1] T. Sikora and B. Makai, *Shape-Adaptive DCT for Generic Coding of Video*, IEEE Transactions on Circuit and Systems for Video Technology, Vol. 5, No. 1, pp. 59-62, February 1995.
- [2] M. Bi et al., *Comment on Shape-Adaptive DCT for Generic Coding of Video*, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 6, No. 6, pp. 686-688, December 1996.
- [3] E. C. Acocella and A. Alcaim, *Alignment by Phase of Vertical Coefficients in SA-DCT*, IEEE Signal Processing Letters, Vol. 8, No. 12, pp. 42-44, February 2001.
- [4] S. J. Cho et al., *Arbitrarily-shaped image segment coding using extension-interpolation*, J. Korean Institute of Communication Science 20 (9), pp. 2453-2463, 1995.
- [5] J. W. Yi et al., *A new coding algorithm for arbitrarily shaped image segments*, Signal Processing Image Communication, Vol. 12, pp. 231-242, 1998.
- [6] T. Xie et al., *A New Approach to Arbitrary Shaped DCT*, PCS'1997, Berlin.
- [7] G. Shen, B. Zeng and M.L. Liou, *Achieving Optimal Rate-Distortion Performance in Arbitrarily-Shaped Transform Coding*, ISCAS 2000 - IEEE International Symposium on Circuits and Systems, May 28-31, 2000, Geneva, Switzerland, pp II-1 to II-4.
- [8] F.M. Freitas and A. Alcaim, *Performance of the SA-DCT and Block-Based DCT Algorithms for Content-Based Video Coding in Terms of Morphological Features*, ISCIT 2002 - International Symposium on Communications and Information Technology, 2002, Pattaya, Thailandia.