

Melhorando a Compensação de Movimentos da Crominância para Codificadores de Vídeo Baseados na Recomendação H.263

Marcos Moecke e Rui Seara

Resumo—O procedimento de compensação de movimento (CM) para os componentes de crominância da recomendação H.263 é analisado e uma proposta para a sua melhoria é apresentada. Mostramos que é possível aprimorar a qualidade da CM do codificador H.263-*baseline*, utilizando a informação de movimento já presente no fluxo de *bits*, através do aumento da precisão dos vetores de movimento e redução do tamanho de blocos, usados na CM da crominância. Desse modo, melhoramos a qualidade da CM sem acarretar aumento na taxa de *bits* usada para a CM. Resultados experimentais em seqüências de vídeo padrão confirmam a eficiência da abordagem proposta.

Palavras-Chave—Recomendação H.263, compensação de movimento, codificação de vídeo colorido, vetores de movimento.

Abstract—This paper investigates the H.263 chrominance motion compensation (MC) procedure and presents an enhanced approach. We show that the quality of the H.263-*baseline* MC can be improved by increasing the accuracy of the chrominance MC vector, using the motion information already existing in the bit stream, and also reducing the block size used for the chrominance MC. In this way, we attain an additional improvement of the MC quality without increasing the MC bit rate. Experimental results using standard video sequences ratify the effectiveness of the proposed approach.

Index Terms— H.263 standard, motion compensation, color video coding, motion vectors.

I. INTRODUÇÃO

A codificação de vídeo, na maioria dos casos, é feita através de sistemas híbridos, que utilizam a técnica da predição por estimação/compensação de movimentos (EM/CM) em blocos para a exploração da redundância temporal e a transformada cosseno discreta (DCT) ou transformada wavelet discreta (DWT) para a codificação do erro de predição [1]. O tamanho dos blocos pode variar entre

Marcos Moecke e Rui Seara, LINSE – Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, SC, E-mails: moecke@ieee.org, seara@linse.ufsc.br.

Marcos Moecke, (também) Coordenadoria de Telecomunicações, Unidade de São José, Centro Federal de Educação Tecnológica de Santa Catarina.

Este trabalho foi parcialmente financiado pelo CNPq.

4×4 e 32×32 *pixels*. Na codificação de vídeo colorido, o espaço de cores usado é geralmente o YCbCr [2], que possui um componente de luminância (Y) e dois componentes de crominância (Cb e Cr). A EM é normalmente realizada levando-se em conta apenas as informações de luminância, sendo que os mesmos vetores de movimento obtidos para o componente Y são utilizados para a CM de Y, Cb e Cr. Em função da menor resolução espacial do sistema visual humano (HVS) à crominância [3], sua resolução é normalmente a metade da usada pela luminância, correspondendo ao formato de crominância 4:2:0. Os padrões de vídeo H.261, H.263, MPEG-1 e MPEG-2 utilizam tal abordagem para a codificação de sinais de vídeo. Para acomodar a diferença de resolução dos componentes de luminância e crominância no processo de CM, diferentes procedimentos podem ser adotados. Neste artigo, mostramos que a abordagem prevista pela recomendação H.263 [4] pode ser melhorada, pois a informação de movimento existente no fluxo de *bits* não é utilizada plenamente na compensação de movimento dos componentes de crominância. Apresentamos uma proposta para a melhoria de qualidade da compensação dos componentes de crominância, sem qualquer aumento adicional na quantidade de *bits* utilizados, empregando uma maior precisão nos vetores de movimento e também usando blocos de menor tamanho no processo de CM da crominância. Com tal abordagem, obtivemos ganhos médios de qualidade superiores a 1 dB em relação à H.263. Ganhos idênticos também podem ser obtidos quando a mesma proposta é usada para o padrão MPEG-2 com formatos de crominância 4:2:0 ou 4:2:2.

II. COMPENSAÇÃO DE MOVIMENTO DA CROMINÂNCIA NA RECOMENDAÇÃO H.263

Na recomendação H.263, os quadros da seqüência de imagens são particionados em macroblocos de luminância de 16×16 *pixels* e de crominância de 8×8 *pixels*. Além disso, cada macrobloco de luminância (16×16 *pixels*) é subdividido em 4 blocos de 8×8 *pixels*.

A. Um Vetor de Movimento por Macrobloco

No procedimento de EM/CM do modo básico (*baseline*) da recomendação H.263, um vetor de movimento (VM) é

considerado para todos os *pixels* do macrobloco de luminância. Esse mesmo vetor, escalado por um fator de 2 em ambas as direções (horizontal e vertical), é usado para os blocos de crominância Cb e Cr, conforme ilustrado na Fig. 1. Como a precisão utilizada para o VM de luminância (VM_Y) é de $1/2$ *pixel*, obtém-se para os respectivos VMs de crominância (VM_C) precisão de $1/4$ *pixel*, a qual é convertida para $1/2$ *pixel* segundo a Tabela 1.

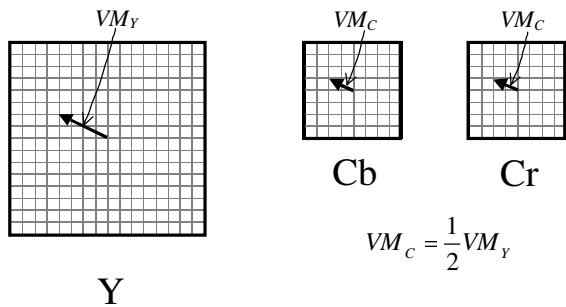


Fig. 1. Uso da informação dos vetores de movimento conforme a recomendação H.263 (*baseline*).

TABELA 1
CONVERSÃO DE PRECISÃO DOS VETORES DE MOVIMENTO
DE ACORDO COM A RECOMENDAÇÃO H.263 (*BASELINE*)

posição de $1/4$ <i>pixel</i>	0	$1/4 \dots 3/4$	1
posição de $1/2$ <i>pixel</i>	0	$1/2$	1

B. Quatro Vetores de Movimento por Macrobloco

Quando o modo de predição avançado (Anexo F da recomendação H.263 – *advanced prediction mode*) ou o filtro de desbloqueamento (Anexo J da recomendação H.263 – *deblocking filter mode*) é usado, um vetor de movimento (VM_{Yi}) é determinado para cada um dos quatro blocos 8×8 *pixels* do macrobloco Y. Nesses casos, um único vetor de movimento para a CM do bloco de crominância 8×8 *pixels* (VM_C) é determinado pela média dos vetores dos 4 blocos de luminância dividida por 2, conforme ilustrado na Fig. 2. O valor final do vetor de movimento tem precisão de $1/16$ *pixel* que é convertida para $1/2$ *pixel* segundo a Tabela 2.

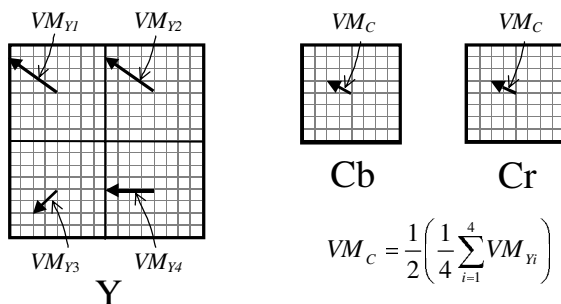


Fig. 2. Uso da informação dos vetores de movimento conforme a recomendação H.263 (Anexo F).

TABELA 2
CONVERSÃO DE PRECISÃO DOS VETORES DE MOVIMENTO
CONFORME A RECOMENDAÇÃO H.263 (ANEXO F)

posição de $1/16$ <i>pixel</i>	$0 \dots 2/16$	$3/16 \dots 13/16$	$14/16 \dots 1$
posição de $1/2$ <i>pixel</i>	0	$1/2$	1

III. PROPOSTA DE MODIFICAÇÃO DA COMPENSAÇÃO DE MOVIMENTO DOS COMPONENTES DE CROMINÂNCIA

A análise da abordagem utilizada para a CM da crominância na recomendação H.263 indica que existem dois pontos frágeis nesta abordagem que podem resultar em perda de resolução e, conseqüentemente, perda de qualidade da CM de crominância:

- redução da precisão dos VM_C de $1/4$ *pixel* para $1/2$ *pixel*, mantendo a precisão da CM do componente de crominância idêntica ao da CM do componente de luminância, apesar da diferença de resolução entre esses componentes;
- uso de um vetor médio no lugar dos 4 vetores de movimento disponíveis no fluxo de *bits*.

Neste artigo, propomos a realização de duas modificações no procedimento de CM dos componentes de crominância, visando superar esses pontos frágeis da recomendação H.263.

A. Aumento da Precisão dos Vetores de Movimento

No modo básico (*baseline*) de EM/CM da H.263, pode-se evitar a redução da precisão dos VMs, utilizando uma interpolação bilinear na CM da crominância, através do seguinte procedimento:

- Em cada macrobloco o vetor de movimento (VM_Y) com precisão de $1/2$ *pixel*, obtido para o bloco de 16×16 *pixels* de Y, é dividido por 2 para a obtenção de um vetor de movimento (VM_C) com precisão de $1/4$ *pixel* a ser usado nos blocos 8×8 *pixels* de Cb e Cr.
- A compensação de movimentos dos blocos de crominância é realizada utilizando o vetor obtido, mantendo a precisão de $1/4$ *pixel*.
- Para obter os valores das posições com precisão de $1/4$ *pixel* nos componentes Cb e Cr, utiliza-se a interpolação bilinear, dada pela seguinte expressão:

$$a_{ij} = \left\langle (1/16)[(5-i)(5-j)A_1 + (5-i)(j-1)A_2 + (i-1)(5-j)A_3 + (i-1)(j-1)A_4 + 8] \right\rangle, \quad (1)$$

onde a_{ij} representa o valor do *pixel* na posição interpolada, A_i corresponde aos valores dos *pixels* das posições inteiras conforme representado na Fig. 3, e $\langle \cdot \rangle$ denota o operador de extração da parte inteira do argumento.

$$\begin{array}{cccc}
\odot A_1 & \times a_{12} & \times a_{13} & \times a_{14} & \odot A_2 \\
& \times a_{21} & \times a_{22} & \times a_{23} & \times a_{24} \\
& \times a_{31} & \times a_{32} & \times a_{33} & \times a_{34} \\
& \times a_{41} & \times a_{42} & \times a_{43} & \times a_{44} \\
\odot A_3 & & & & \odot A_4 \\
\odot A_i - \text{posição inteira de pixel} \\
& \times a_{ij} - \text{posição interpolada de } \frac{1}{4} \text{ pixel}
\end{array}$$

Fig. 3. Interpolação das posições com precisão de $\frac{1}{4}$ pixel.

B. Redução do Tamanho do Bloco

No caso do uso do modo de predição avançado ou do filtro de desbloqueamento, quando estão disponíveis quatro VMs em cada macrobloco, utilizamos também quatro vetores na CM da crominância no lugar de um vetor médio. Em virtude da menor resolução dos componentes de crominância, é necessário reduzir o tamanho do bloco utilizado na CM da crominância para 4×4 pixels. A seguir, é descrita a abordagem empregada:

- i) Em cada macrobloco os VMs obtidos para os blocos 8×8 pixels de Y (VM_{Y1} , VM_{Y2} , VM_{Y3} , VM_{Y4}), com precisão de $\frac{1}{2}$ pixel, são divididos por 2 para a obtenção de quatro VMs com precisão de $\frac{1}{4}$ pixel (VM_{C1} , VM_{C2} , VM_{C3} , VM_{C4}), que serão usados na CM dos componentes de crominância.
- ii) Os blocos de crominância com tamanho 8×8 pixels são subdivididos em 4 sub-blocos de 4×4 pixels. Cada vetor de movimento VM_{Ci} é utilizado na CM do seu sub-bloco correspondente, conforme ilustrado na Fig. 4.
- iii) A compensação de movimentos dos blocos de crominância é realizada mantendo a precisão de $\frac{1}{4}$ pixel do vetor.
- iv) Para a obtenção dos valores das posições com precisão de $\frac{1}{4}$ pixel nos componentes Cb e Cr, utiliza-se interpolação bilinear.

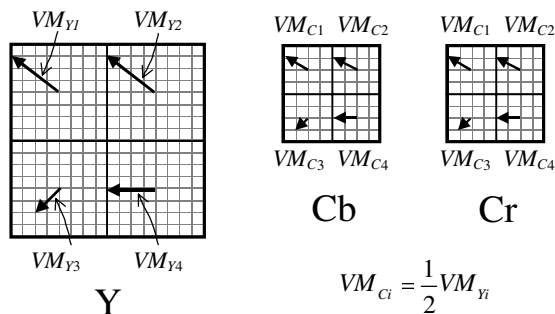


Fig. 4. Uso da informação dos vetores de movimento de acordo com a abordagem proposta.

IV. RESULTADOS EXPERIMENTAIS

A avaliação da qualidade foi realizada considerando-se seqüências de vídeo no formato QCIF (Y: 176×144 pixels, e Cb e Cr: 88×72 pixels). Utilizamos os primeiros 100 quadros das seqüências *Suzie*, *Miss America*, *Mother and Daughter*, *Salesman*, *Foreman* e *Carphone*. Para mostrar o ganho obtido na CM, utilizamos como quadro de referência na EM sempre o quadro anterior da seqüência original, fazendo com que a escolha do tipo de codificação do erro não influencie nos resultados. Nesse tipo de avaliação, a taxa de bits é considerada suficientemente grande, resultando em uma compressão sem perdas. A qualidade das imagens foi comparada quantitativa e subjetivamente após a compensação de movimento, sendo os VMs estimados através do algoritmo de busca exaustiva apresentado em [5]. Confrontamos a abordagem descrita na recomendação H.263 com aquela proposta neste trabalho. O parâmetro de comparação de qualidade objetiva usado foi a razão sinal-ruído de pico (PSNR).

A. Efeito do Aumento da Precisão dos Vetores de Movimento

Inicialmente, avaliamos o uso da precisão de $\frac{1}{2}$ pixel da abordagem original da recomendação H.263-baseline e comparamos com a de precisão de $\frac{1}{4}$ pixel da abordagem proposta. Nesse caso, um único vetor de movimento é utilizado para cada bloco de 16×16 pixels de luminância e o vetor de movimento usado na CM dos blocos 8×8 pixels dos componentes de crominância é derivado diretamente do vetor de movimento Y.

Os resultados da Tabela 3 mostram que, confrontada com a abordagem estabelecida pela recomendação H.263, a nossa proposta resulta em ganhos significativos de qualidade da compensação de movimento em quase todas as seqüências testadas. Esses ganhos são obtidos exclusivamente em decorrência do aumento de precisão na CM da crominância, uma vez que os VMs usados já estavam codificados no fluxo de bits e, portanto, não acarretaram acréscimo na quantidade de bits. Essa abordagem é, assim, completamente diferente de aumentar a qualidade da CM da luminância através do aumento da precisão dos VMs [6], o que implica, nesse caso, em aumento da taxa de bits. Na prática, a única modificação necessária para melhorar a qualidade da CM dos componentes de crominância é manter a precisão dos VMs dos componentes de crominância em $\frac{1}{4}$ pixel e incluir a interpolação bilinear de $\frac{1}{4}$ pixel no processo de CM.

B. Efeito da Redução do Tamanho do Bloco

Quando o modo de predição avançado da recomendação H.263 é usado, quatro vetores de movimento são estimados para cada macrobloco de luminância de 16×16 pixels e ganhos ainda superiores são obtidos com a abordagem

proposta. Nesse caso, além de tirar proveito do aumento da precisão dos VMs, pode-se também melhorar a CM reduzindo o tamanho dos blocos de crominância de 8×8 para 4×4 pixels e usando-se quatro VMs no lugar de um VM médio.

A Fig. 5 mostra dois exemplos de seqüência de vídeo, nos quais pode-se perceber que uma melhora de qualidade da CM dos componentes Cb e Cr está ocorrendo para todos os quadros. Na seqüência *Miss America* o ganho obtido no componente Cb é pequeno, no entanto, no componente Cr ele é consideravelmente relevante, chegando a 6,1 dB para o quadro 95. Outro aspecto muito importante, mostrado na Fig. 5, é o elevado aumento da qualidade mínima obtida nas seqüências testadas. Esse ganho é muito mais significativo do que o aumento médio de qualidade, uma vez que para a avaliação subjetiva, os quadros com menor qualidade influenciam muito mais no julgamento do que aqueles cuja qualidade já é superior a um nível considerado "bom" [7].

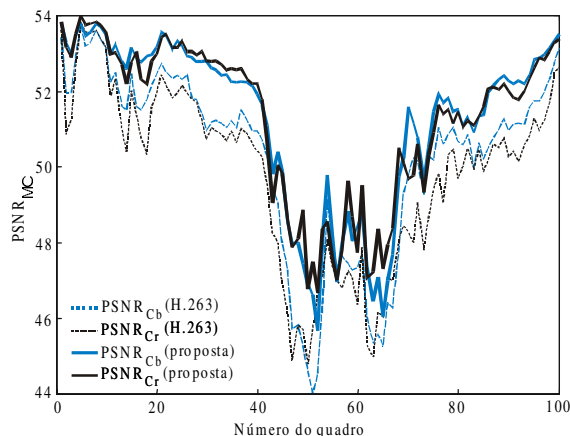
A Fig. 6 mostra um detalhe da imagem do quadro 95 da seqüência *Miss America*, no qual é possível perceber um dos motivos da diferença de qualidade entre as abordagens em questão. Ao ser feita a estimação de movimento com a informação de luminância usando blocos de 8×8 pixels, quatro vetores de movimento são determinados. No quadro ilustrado na Fig. 6(a), os vetores encontrados para os blocos marcados com 1, 2, 3 e 4 são diferentes e, por isso, o uso do valor médio desses vetores na compensação de movimento da crominância resulta em um erro de cor, destacado por uma circunferência na Fig. 6(b). No caso do uso da abordagem proposta, são considerados quatro vetores de movimento na CM da crominância, melhorando a compensação, conforme mostra a Fig. 6(c). É importante destacar que como a CM da luminância é feita da mesma forma nos dois casos, não existe qualquer diferença perceptível entre as imagens se observadas sobre um suporte monocromático.

TABELA 3

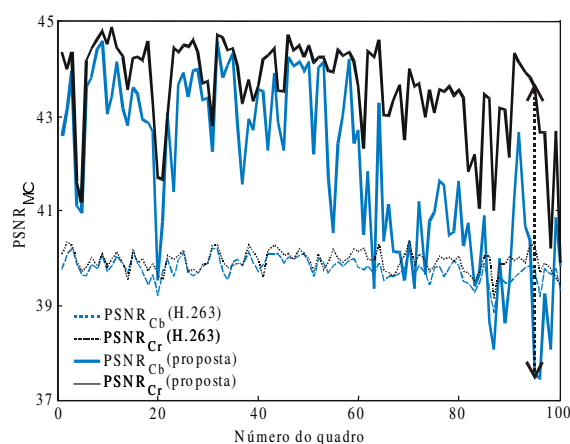
VALORES DE GANHO DE QUALIDADE EM PSNR DA ABORDAGEM PROPOSTA EM RELAÇÃO À ESTABELECIDA PELA H.263-BASELINE NOS COMPONENTES DE CROMINÂNCIA Cb E Cr

Seqüência de Vídeo	PSNR Y*	Abordagem	PSNR Cb*	Ganho Cb	PSNR Cr	Ganho Cr
<i>Suzie</i>	37,56	H.263 proposta	50,30 51,12	0,82	49,99 51,06	1,07
<i>Miss America</i>	43,11	H.263 proposta	39,95 40,04	0,09	42,72 43,85	1,13
<i>Mother and Daughter</i>	38,66	H.263 proposta	44,96 45,25	0,30	45,84 46,33	0,49
<i>Salesman</i>	40,69	H.263 proposta	46,50 46,58	0,08	48,57 48,69	0,12
<i>Foreman</i>	35,79	H.263 proposta	45,15 47,04	1,89	45,78 48,00	2,22
<i>Carphone</i>	36,03	H.263 proposta	46,64 48,32	1,69	46,49 48,35	1,85

* Valores médios de PSNR após a CM.



(a)

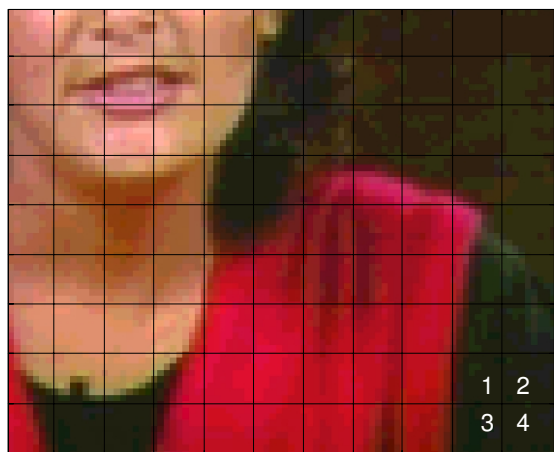


(b)

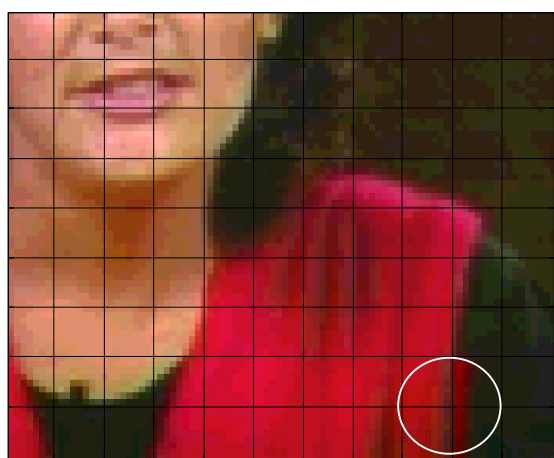
Fig. 5. Comparação de qualidade da CM de Cb e Cr entre a abordagem estabelecida no modo de predição avançado da H.263 e a abordagem proposta. (a) Seqüência *Suzie*; (b) seqüência *Miss America*.

As imagens dos quadros de diferenças dos componentes de crominância (DFD_C) mostradas na Fig. 7 ilustram, através de tons de cinza, as diferenças entre o componente Cr do quadro original e o do respectivo quadro após a compensação de movimento. Pode-se perceber, através da ausência de regiões com tons claros e escuros na DFD_C , que a compensação de movimento desse componente através da abordagem proposta é bastante superior à obtida pela abordagem estabelecida pela recomendação H.263. Assim, para codificar a DFD_C , é necessário um número de bits maior na H.263 do que na abordagem proposta.

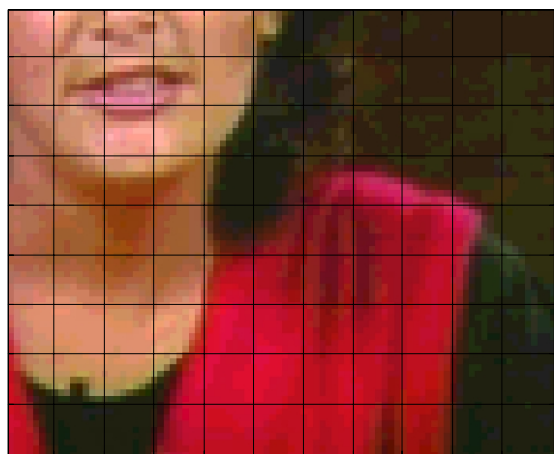
Avaliações também foram realizadas com outras seqüências de vídeo, sendo que, em todas as seqüências consideradas, o uso da abordagem proposta resultou sempre em um certo ganho de qualidade, conforme mostram os valores médios de PSNR da Tabela 4. Para a maioria das seqüências, os ganhos médios são superiores a 1 dB. No caso da seqüência *Foreman*, um ganho médio de 2,6 dB é obtido para ambos os componentes Cb e Cr.



(a)



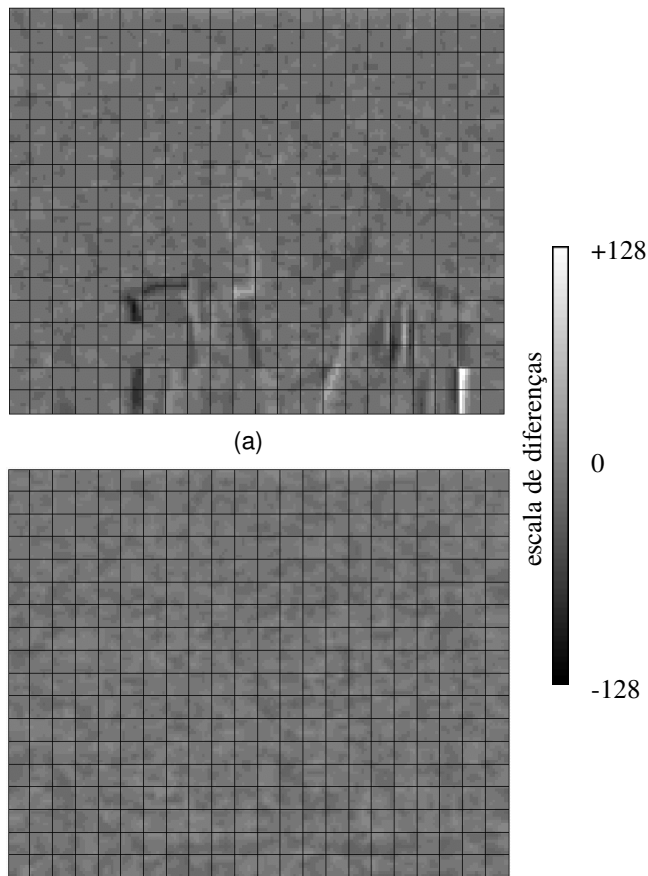
(b)



(c)

Fig. 6. Detalhe do quadro 95 da seqüência *Miss America*. (a) Imagem original; (b) imagem após a CM de Cb e Cr pela abordagem da H.263¹ (c) imagem após CM de Cb e Cr pela abordagem proposta.

¹ Para a observação do erro de cor é necessário utilizar uma impressão colorida da imagem ou observá-la em um terminal colorido, uma vez que não existe nenhuma diferença entre as imagens quando observadas em tons de cinza.



(a)

(b)

Fig. 7. Imagem da DFD_C da crominância Cr do quadro 95 da seqüência *Miss America*. (a) Abordagem original da H.263 - Anexo F; (b) abordagem proposta.

TABELA 4
VALORES DE GANHO DE QUALIDADE EM PSNR DA ABORDAGEM PROPOSTA EM RELAÇÃO A ESTABELECIDA PELA H.263-(ANEXO F) NOS COMPONENTES DE CROMINÂNCIA Cb e Cr

Seqüência de Vídeo	PSNR Y*	Abordagem	PSNR Cb*	Ganho Cb	PSNR Cr	Ganho Cr
<i>Suzie</i>	38,85	H.263 proposta	50,28 51,29	1,01	49,87 51,29	1,43
<i>Miss America</i>	43,72	H.263 proposta	39,78 39,95	0,17	41,95 43,62	1,67
<i>Mother and Daughter</i>	39,61	H.263 proposta	44,38 45,36	0,98	44,86 46,47	1,61
<i>Salesman</i>	42,07	H.263 proposta	46,48 46,63	0,15	48,56 48,76	0,20
<i>Foreman</i>	36,71	H.263 proposta	44,43 47,10	2,67	45,54 48,16	2,62
<i>Carphone</i>	37,27	H.263 proposta	46,02 48,40	2,38	46,08 48,46	2,38

* Valores médios de PSNR após a CM.

V. CONCLUSÕES

O ganho de qualidade obtido na CM dos componentes de crominância usando a abordagem proposta é bastante significativo (superior a 1 dB). Como não há aumento na

quantidade de *bits* utilizada para a codificação dos vetores de movimento e o aumento de complexidade decorrente do uso de uma interpolação bilinear de $1/4$ *pixel* no lugar de $1/2$ *pixel* é mínimo, sugere-se a substituição da abordagem prevista na recomendação H.263 pela abordagem proposta neste artigo. É importante ressaltar que os ganhos de qualidade dos componentes de crominância podem ser convertidos tanto em redução de taxa total de *bits* quanto na melhoria da qualidade geral da imagem. As alterações essenciais sugeridas neste trabalho são: incorporação da CM com precisão de $1/4$ *pixel*, CM com blocos de 4×4 *pixels* e modificação do cálculo dos VM para os componentes de crominância. Nenhuma alteração no fluxo de *bits* dos VMs é necessária, apenas uma indicação do uso dessa abordagem de CM da crominância deve ser sinalizada através de algum meio externo (por exemplo, considerando-se a recomendação H.245 [8]).

REFERÊNCIAS

- [1] F. Dufaux and F. Moscheni, "Motion estimation techniques for digital TV: a review and a new contribution," *Proceedings of the IEEE*, vol. 83, Jun. 1995, pp. 858-876.
- [2] D. H. Kelly, "Spatiotemporal Variation of Chromatic and Achromatic Contrast Thresholds," *J. Opt. Soc. Am.*, vol. 73, no. 6, 1983, pp. 742-750.
- [3] ITU-R Recommendation BT.601-5, *Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios*, Int'l Telecommunications Union, Geneva, 1995.
- [4] ITU-T Recommendation H.263, *Video Coding for Low Bit Rate Communication*, Int'l Telecommunications Union, Geneva, 1998.
- [5] ITU-T TMN11, *Test Model 11*, Int'l Telecommunications Union, Monterey, 1999.
- [6] J. Ribas-Corbera and D. L. Neuhoff, "Optimizing motion-vector accuracy in block-based video coding," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 11, no. 4, Apr. 2001, pp. 497-511.
- [7] K. T. Tan, M. Ghanbari, and D. E. Pearson, "An Objective Measurement Tool for MPEG Video Quality," *Signal Processing*, vol. 70, no. 3, 1998, pp. 279-294.
- [8] ITU-T Recommendation H.245, *Control Protocol for Multimedia Communication*, Int'l Telecommunications Union, Geneva, 1997.