

# Redução de Complexidade Através da Ordenação de Custo dos Processos de Codificação em Vídeo Digital

Marcos Moecke e Rui Seara

**Resumo**—O processo de estimação de movimento e a seleção do modo de codificação são responsáveis por grande parte do esforço computacional nos codificadores de vídeo, seguindo o padrão H.264 otimizado em taxa-distorção (R-D). Este trabalho propõe uma estratégia de classificação dos processos de forma a evitar a avaliação da distorção para cada vetor de movimento e modo de codificação, mantendo a condição de otimização em R-D. Essa estratégia resulta em uma redução significativa de complexidade de até 90%. Resultados experimentais verificam a abordagem proposta.

**Palavras-chave**—Codificação de vídeo, Complexidade computacional, Estimação de movimento, Otimização em taxa-distorção, H.264.

**Abstract**—Motion estimation process and coding mode selection are responsible for a large portion of the required computational effort in H.264-based video coding systems optimized in rate-distortion (R-D). This work proposes a classification strategy which avoids assessing the distortion obtained for each motion vector and coding mode, upholding the R-D optimization criterion. This strategy yields a reduction of up to 90% in the computational complexity. Some experimental results confirm the proposed approach.

**Keywords**—Video coding, Computational effort, Motion estimation, Rate-distortion optimization, H.264.

## I. INTRODUÇÃO

O desempenho dos codificadores de vídeo envolvendo estimação de movimento (EM) com escolha do modo de codificação através de um processo otimizado em taxa-distorção (R-D) tem se mostrado muito superior àqueles que otimizam apenas a distorção [1]-[4]. Quando o número de modos de codificação é pequeno, a otimização em R-D é relativamente simples de implementar; no entanto, em codificadores que possuem um grande número de modos de codificação, como por exemplo, aqueles seguindo a recomendação H.264 [5], o referido processo de otimização torna-se um tanto complexo. Visando então reduzir a complexidade envolvida, diversos trabalhos de pesquisa vêm sendo propostos na literatura. No modelo de teste TMN8 [3], tal complexidade é evitada usando um procedimento

conhecido por técnica do vetor de movimento (VM) nulo, que consiste no favorecimento do VM nulo sempre que esse produza uma distorção próxima ao VM ótimo. Em [6], é usada uma janela de procura de tamanho reduzido, limitando-se a EM em torno das posições do VM predito, VMs dos blocos vizinhos e VM nulo. Em [7], a EM com resolução de 1 *pixel* é otimizada apenas em distorção; em uma segunda fase, é efetuada a EM otimizada em R-D com resolução de *subpixel* em torno do melhor VM. Todas essas abordagens visam reduzir a complexidade do processo de EM otimizado em R-D, não afetando a complexidade do processo de seleção do modo de codificação.

Neste trabalho, tratamos o problema de complexidade tanto na EM quanto na seleção do modo de codificação. A estratégia proposta, denominada *increasing coding cost* (ICC), fundamenta-se em classificar os VMs e modos de codificação em uma ordem crescente de custo de codificação. Esse pré-ordenamento, associado às regras de parada por custo e qualidade de imagem, permite pré-eliminar VMs e modos de codificação sem que as distorções sejam avaliadas. Apesar da redução importante de complexidade, o processo permanece ainda otimizado em R-D, pois apenas os VMs, os tipos de partição de macroblocos e os modos de codificação que não podem atingir níveis ótimos são descartados. A estratégia ICC também permite pré-eliminar formas de predição espacial no modo *Intra*; no entanto, esse tópico não será tratado neste artigo [8].

## II. CUSTO DE CODIFICAÇÃO DOS VMs

Na determinação do custo de codificação do processo de EM, o custo dos VMs é um fator primordial. Além do custo dos VMs, outro fator que deve ser considerado é o tamanho do bloco. Os VMs  $p = (x, y)$  são constituídos dos componentes  $x$  e  $y$  que indicam o movimento de translação de todos os *pixels* do bloco entre o quadro atual e o quadro de referência (QR). O processo de EM é realizado iniciando pelos blocos do canto superior esquerdo da imagem, sendo assim possível obter um vetor de predição  $p_p = (x_p, y_p)$  para o VM atual  $p_a = (x_a, y_a)$ , levando em conta os VMs dos blocos vizinhos (usualmente, o VM predito correspondente ao valor mediano dos componentes  $x$  e  $y$  desses blocos). Após a predição, o vetor de diferença (VD) entre o VM atual e seu vetor predito é codificado através de tabelas de codificação de comprimento variáveis (VLC - *variable-length coding*) ou usando codificação aritmética. No padrão H.264,

Marcos Moecke, Centro Federal de Educação Tecnológica de Santa Catarina - CEFET/SC, Curso de Telecomunicações, São José, SC. E-mail: moecke@sj.cefetsc.edu.br.

Rui Seara, LINSE - Laboratório de Circuitos e Processamento de Sinais, Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, SC, E-mail: seara@linse.ufsc.br.

Este trabalho foi parcialmente financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

as VLCs usadas são do tipo Exp-Golomb [5] e, neste caso, o custo de codificação  $R(p_d)$  do VD depende diretamente dos componentes  $x_d$  e  $y_d$ . O inverso do custo de codificação dos VDs é mostrado na Fig. 1.

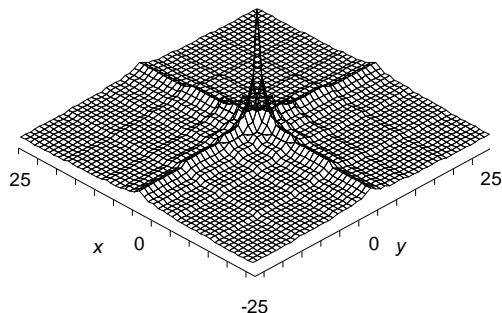


Fig. 1. Inverso do custo de codificação dos VDs.

A distribuição estatística dos componentes dos VDs em seqüências típicas de videoconferência, considerando a otimização em distorção (DO) e em taxa-distorção (RDO), é determinada em [8]. No primeiro caso, é obtida uma distribuição com característica aproximadamente Laplaciana simétrica circular [9]; enquanto ocorre, no segundo caso, uma concentração de VDs nos eixos principais  $x$  e  $y$ . Comparando a característica da distribuição obtida com RDO [Fig. 2(b)] com o inverso do custo de codificação dos VDs (Fig. 1), percebe-se que há uma forte dependência entre a frequência de ocorrência dos VDs e o inverso de seu custo.

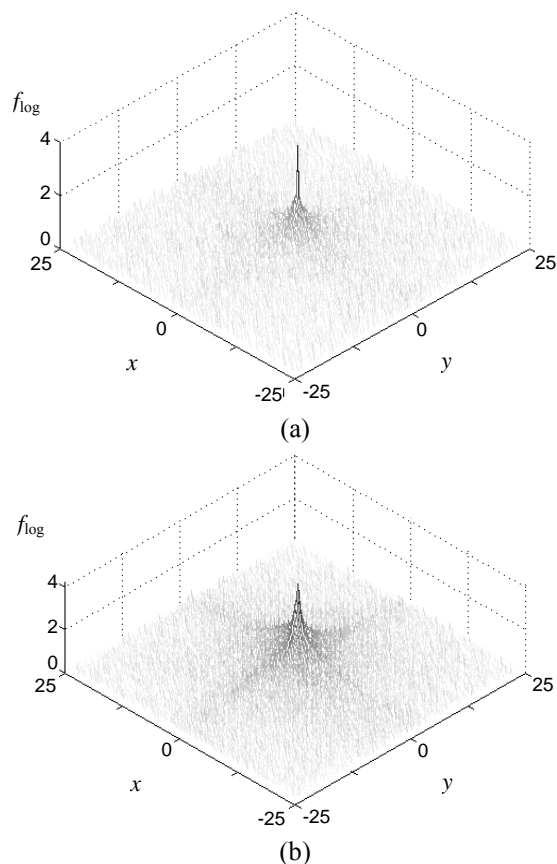


Fig. 2. Distribuição bidimensional dos VDs obtidos para o codificador JM7.3. (a) otimização em distorção. (b) otimização em taxa-distorção.

### III. ESTRATÉGIA ICC NA ESTIMAÇÃO DE MOVIMENTO

Para se obter o melhor compromisso R-D na EM, deve-se escolher o VM que leve ao menor custo para uma dada distorção. Mostraremos agora que nos codificadores RDO é vantajoso classificar os VMs em ordem crescente de custo de codificação (ICC) e avaliar em primeiro lugar os VMs com menor custo. Dessa forma, pode-se estabelecer uma regra de parada baseada no custo de codificação, permitindo reduzir a complexidade computacional do processo de EM, sem afetar o desempenho em R-D.

#### A. Regra de Parada por Custo

Inicialmente, é necessário classificar o conjunto dos possíveis VMs  $\mathbf{P} = \{p_1, p_2, \dots, p_{last}\}$  em ordem ICC. Para tal, divide-se o referido conjunto em subconjuntos  $\mathbf{S}_k$  que contêm apenas os vetores com custo de codificação  $k$

$$\mathbf{S}_k = \{p_1, p_2, \dots\} \mid p_i \in \mathbf{S}_k \Leftrightarrow R(p_i) = k. \quad (1)$$

Estes subconjuntos  $\mathbf{S}_{k_i}$  são organizados de forma que  $k_1 < k_2 < \dots < k_n$ , obtendo-se o conjunto ordenado

$$\mathbf{P} = \{\mathbf{S}_{k_1} \prec \mathbf{S}_{k_2} \prec \mathbf{S}_{k_3} \prec \dots \prec \mathbf{S}_{k_n}\} \mid \mathbf{S}_{k_i} \prec \mathbf{S}_{k_j} \Leftrightarrow k_i < k_j \quad (2)$$

onde  $\prec$  indica uma relação de precedência entre os subconjuntos [10].

Quando tabelas VLCs são consideradas para codificar os VMs, pode-se ordenar os VMs *a priori*. Nos casos em que se utiliza uma predição do VM, deve-se considerar o custo dos VDs.

No processo de EM/RDO, o VM ótimo de cada bloco é obtido minimizando

$$J(p) = D(p) + \lambda R(p) \quad (3)$$

$$p_{opt} = \min_{p \in \mathbf{P}} \{J(p)\}.$$

Uma vez definido o critério de otimização (3) e organizados os VMs em ordem ICC (2), o processo de EM consiste na análise dos VMs de cada subconjunto  $\mathbf{S}_k$ , sendo selecionado o VM  $p_{kopt}$  que resulte na menor distorção.

Então, seu custo Lagrangiano é dado por

$$J_k(p_{kopt}) = D(p_{kopt}) + \lambda k. \quad (4)$$

O VM ótimo provisório  $\tilde{p}_{opt}$  é aquele que possui o menor custo Lagrangiano  $J(\tilde{p}_{opt})$  dentre os subconjuntos já avaliados. Antes de iniciar a avaliação dos VMs de cada novo subconjunto  $\mathbf{S}_{k+1}$ , testa-se a regra de parada (5), verificando se o custo  $J(\tilde{p}_{opt})$  é menor ou igual à penalidade de distorção para os VMs do novo subconjunto  $\mathbf{S}_{k+1}$ . Assim, se

$$J(\tilde{p}_{opt}) \leq \lambda(k+1) \quad (5)$$

o processo de EM é terminado e o custo Lagrangiano ótimo é então dado por  $J(p_{opt}) = J(\tilde{p}_{opt})$ . O uso dessa regra de

parada reduz o espaço de procura dos VMs, preservando o critério de RDO.

Quando múltiplos QRs [11] são usados, o processo de EM passa a ser a escolha do melhor conjunto  $\{p_{\text{opt}}, r_{\text{opt}}\}$ , onde  $p_{\text{opt}}$  é o VM ótimo e  $r_{\text{opt}}$  é a referência temporal desse vetor. Nesse caso, para implementar a estratégia ICC, é necessário incluir o custo de codificação da referência  $R(r)$  antes de realizar o processo de ordenação dos subconjuntos  $S_k$ . Assim,

$$R(p, r) = R(r) + R(p) \quad (6)$$

e a otimização em R-D é então expressa por

$$J(p, r) = D(p, r) + \lambda R(p, r) \\ \{p, r\}_{\text{opt}} = \min_{p \in \mathbf{P}, r \in \mathbf{R}} \{J(p, r)\}. \quad (7)$$

Como conseqüência do uso de múltiplos QRs em um codificador ICC, dependendo do custo de codificação  $R(p, r)$ , o processo de EM alterna a procura dos VMs nos diversos quadros de referência  $r$ . Nesse caso, a introdução da regra de parada reduz o espaço de procura tanto nas dimensões  $x$  e  $y$  quanto na dimensão  $r$ . O grau de redução do espaço de busca depende do valor de  $\lambda$  e do tipo de imagem a ser codificada.

### B. Regra de Parada por Qualidade

Conforme o bloco de imagem a ser codificado e qualidade mínima requerida em cada aplicação, é possível estabelecer um limiar de distorção  $D_L$ , que pode ser usado para parar o processo de EM quando uma qualidade mínima é atingida, reduzindo, assim, ainda mais a complexidade do processo. Usando a estratégia ICC, pode-se parar o processo de EM quando

$$J(\tilde{p}_{\text{opt}}) \leq D_L \quad (8)$$

pois  $\tilde{p}_{\text{opt}}$  é o VM de menor custo para o limiar de distorção  $D_L$  considerado.

### C. Resultados das Regras de Parada

A estratégia ICC usada em conjunto com a regra de parada por custo reduz a complexidade computacional em até 90% em relação ao algoritmo de busca exaustiva (FS). Usando a estratégia ICC com diferentes limiares de qualidade é possível controlar a complexidade computacional da EM. Considerando que a distorção seja medida através da soma absoluta das diferenças (SAD - *sum of absolute difference*) [3], usando valores de limiar de até 200, obtém-se uma redução adicional de 10% na complexidade com um comprometimento de qualidade inferior a 0,1 dB.

A Fig. 3 mostra também que a influência do limiar na complexidade é maior para uma qualidade de imagem melhor. Se qualidades inferiores de imagem são consideradas (baixas taxas), o fator de maior influência na complexidade passa a ser a estratégia ICC.

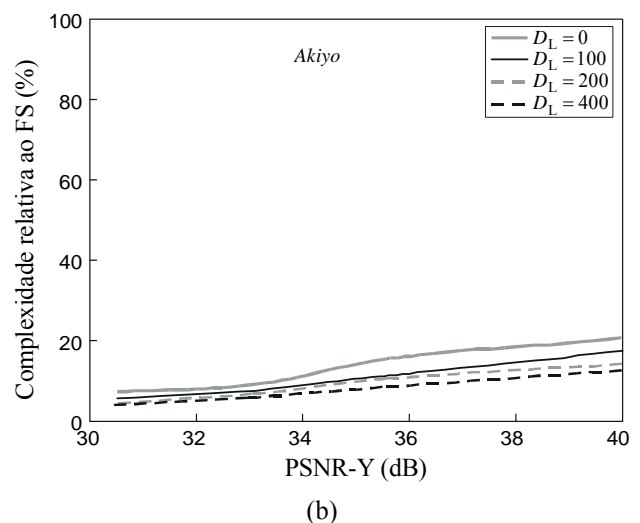
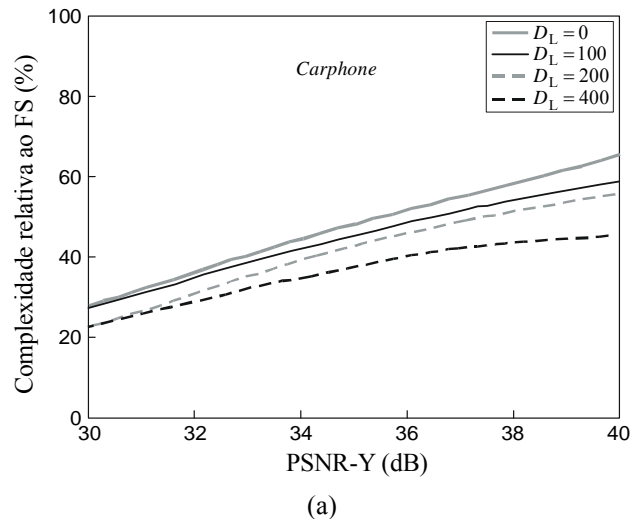


Fig. 3. Influência do limiar de qualidade na complexidade da EM do codificador ICC. Sem usar limiar (linha cinza contínua). Com o uso de limiar: SAD = 100 (linha escura contínua); SAD = 200 (linha cinza tracejada); SAD = 400 (linha escura tracejada). Sequências de vídeo: (a) *Carphone*; (b) *Akiyo*.

## IV. ESTRATÉGIA ICC NA SELEÇÃO DO MODO DE CODIFICAÇÃO

Os codificadores de vídeo híbridos admitem múltiplos modos de codificação (*Skip*, *Inter*, *Intra*) para cada bloco de imagem. Vários tamanhos de bloco são normalmente suportados para os modos *Inter* e *Intra*. Nos padrões H.263, MPEG-2 e MPEG-4, os tamanhos de blocos considerados podem ser de  $16 \times 16$  ou  $8 \times 8$  *pixels*. Esses padrões de vídeo usam uma unidade de codificação chamada macrobloco (MB) que reúne blocos de diferentes tamanhos. Já no padrão H.264, existem diversas opções de partição de MB produzindo blocos de tamanhos  $16 \times 16$  até  $4 \times 4$  *pixels*. Quando submacrobloco (SMB) de  $8 \times 8$  *pixels* são usados, cada um deles pode ser particionado ainda em blocos menores de até  $4 \times 4$ , podendo-se explorar as vantagens de uso de blocos de tamanhos variáveis [12], melhorando, dessa forma, o desempenho em R-D.

Para selecionar a melhor partição de MB, um codificador necessita determinar tanto a taxa de bits quanto a distorção de cada opção, implicando em uma alta carga computacional se muitas opções são consideradas. Com a abordagem ICC é possível eliminar antecipadamente algumas possibilidades (modos) de codificação. Para tal, considera-se no cálculo do custo Lagrangiano os bits que são usados para indicar o modo de codificação  $R(m)$  e o custo do conjunto de VMs  $R(p)$  necessário para realizar a CM dos blocos que constituem cada MB.

De acordo com a sintaxe do fluxo de bits definida para o codificador, pode-se obter um custo inicial de codificação  $R_{\text{init}}(m)$  para cada modo de partição  $m$ , o qual corresponde ao número de bits obrigatórios para a codificação do MB. Usando esse custo, uma penalidade de distorção inicial pode ser determinada através de

$$D_{\text{init}}(m) = \lambda R_{\text{init}}(m). \quad (9)$$

Com os valores de penalidade de distorção  $D_{\text{init}}(m)$  obtidos, pode-se arranjar os modos em ordem ICC

$$\mathbf{M} = \{m_1 < m_2 < \dots < m_N\} | m_i < m_j \Leftrightarrow D_{\text{init}}(m_i) \leq D_{\text{init}}(m_j). \quad (10)$$

O modo *Skip*, que tem normalmente a menor penalidade de distorção inicial, é considerado o primeiro modo  $m_1$ , enquanto os modos do tipo *Intra* são os últimos do arranjo  $\mathbf{M}$ . Para os modos *Inter*, o ordenamento é determinado pelo número de blocos que formam o macrobloco.

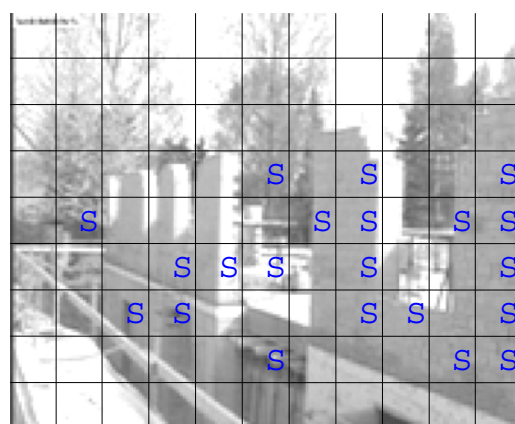
#### A. Pré-seleção do Modo Skip

Na codificação de vídeo com otimização em R-D, a frequência de uso do modo *Skip* aumenta à medida que a taxa de bits total é reduzida e uma menor qualidade é admitida. O percentual de MBs codificados no modo *Skip* varia conforme a quantidade de movimento e complexidade da imagem. Em imagens com fundo estático, nas quais o movimento existente é devido a um ou dois interlocutores (por exemplo, *Akiyo* e *Mother&Daughter*), o percentual de MBs codificados no modo *Skip* é muito alto. Por outro lado, em cenas com movimento de câmera (*CoastGuard*, *Foreman* e *Carphone*), esse percentual é bem menor. O uso de uma quantização mais grosseira (menor taxa de bits) eleva também esse percentual.

A grande quantidade de MB-*Skip* pode ser justificada a partir da análise de alguns exemplos mostrados na Fig. 4. Na Fig. 4(a), ocorrem apenas pequenos movimentos em relação ao quadro anterior e, por isso, 94% dos MBs são codificados no modo *Skip*. Na imagem da seqüência representada na Fig. 4(b), apesar de existir movimento de câmera, ainda encontramos 24% dos MBs codificados no modo *Skip*. No primeiro caso, um grande número de MB-*Skip* é naturalmente esperado; no entanto, no segundo caso, o percentual obtido surpreendeu. Esse percentual só ocorreu porque o padrão H.264 associa o VM predito aos MB-*Skip*, ou seja, o bloco de imagem do quadro anterior é copiado levando em conta o deslocamento indicado pelo VM predito.



(a)



(b)

Fig. 4. Exemplos de MB-*Skip* (assinalados com “S”) com codificador JM7.3-RDO. (a) Quadro #210 de *Mother&Daughter*-10Hz,  $Q = 30$ . (b) Quadro #237 de *Foreman*-10Hz,  $Q = 28$ .

No modelo de referência JM, a opção pelo modo *Skip* é postergada até que sejam determinados os custos de codificação de todos os outros modos (*Inter* e *Intra*) e constatado que esses custos são maiores do que aquele do modo *Skip*. Usando a abordagem ICC, uma parte do processo de EM (modos *Inter*) e da avaliação dos modos *Intra* pode ser evitada fazendo uma escolha antecipada do modo *Skip*. A redução de complexidade obtida depende do parâmetro de quantização  $Q$  usado, do custo Lagrangiano considerado para o modo *Skip* e do tipo de imagem a ser codificada.

Para a decisão da pré-seleção do modo *Skip*, inicialmente, é calculado o custo Lagrangiano para o MB-*Skip*. Assim,

$$J(\text{Skip}) = D(\text{Skip}) + \lambda R(\text{Skip}) \quad (11)$$

onde  $D(\text{Skip})$  denota a distorção do MB do quadro atual em relação aos *pixels* do bloco  $16 \times 16$  apontados pelo VM predito no quadro de referência, onde  $R(\text{Skip})$  é o custo de codificação do MB no modo *Skip*.

Se a condição

$$J(\text{Skip}) \leq D_{\text{init}}(m_2) + D_L \quad (12)$$

é verdadeira, então o modo *Skip* é pré-selecionado e evita-se todo o processo de EM, bem como de avaliação dos modos *Intra*.

No caso do H.264, a indicação do modo *Skip* é feita através de um código *run-length* que corresponde ao número de MBs seguidos, codificados no modo *Skip* [5]. Por isso, o número de bits por MB-*Skip* é no máximo 1. No modo  $m_2 = \text{Inter}_{16 \times 16}$ , são necessários no mínimo 2 bits para a indicação do modo e 2 bits para codificação do VM predito, logo  $R_{\text{init}}(m_2) = 4$ . Assim, o modo *Skip* é pré-selecionado se

$$D(\text{Skip}) \leq 0,85 \times [4 - R(\text{Skip})] 4^{(Q-12)/3} + D_L. \quad (13)$$

A pré-seleção do modo *Skip* é controlada diretamente pelo parâmetro de quantização  $Q$ , privilegiando-o em baixas taxas de bits. Essa pré-seleção traz como principal consequência a redução de complexidade da EM, sem que haja perda de desempenho em R-D, quando  $D_L = 0$ .

Usando a estratégia ICC em seqüências típicas de videoconferência, uma significativa parcela dos MBs (até 80%) pode ser pré-selecionada no modo *Skip* [ver Fig. 5 (a)]. No entanto, em seqüências com movimento de câmera (tais como, *Carphone* e *Foreman*), essa percentagem é menor; porém, ainda significativa [ver Fig. 5(b)].

É importante ressaltar que o uso da estratégia de pré-seleção do modo *Skip* não impede que o modo *Skip* seja selecionado após a avaliação do custo Lagrangiano dos demais modos de codificação.

#### B. Definição da Partição do MB no Modo Inter

Nos codificadores de vídeo para agrupar diferentes tamanhos de bloco, é usada uma estrutura de macrobloco (MB), o qual pode ser definido como um conjunto de blocos

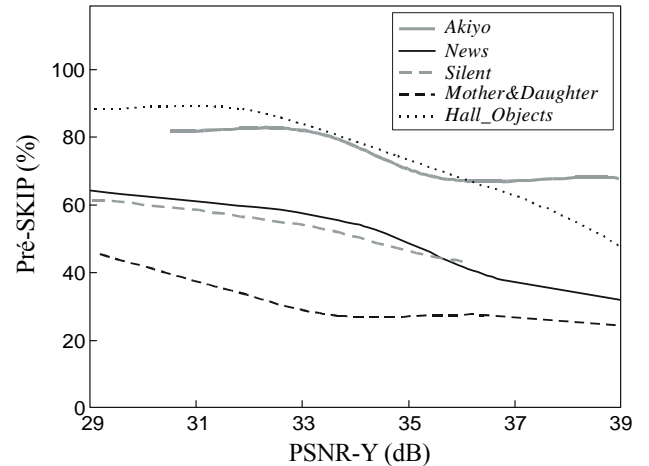
$$\text{MB} = \{B_1, B_2, \dots, B_M\} \quad (14)$$

onde  $B_i$  são os blocos agrupados pelo MB. No caso de codificadores que seguem o padrão H.264, são possíveis 52 diferentes partições.

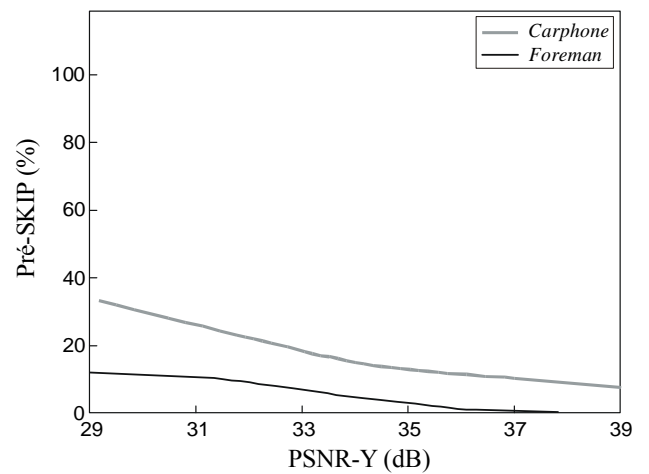
Para a otimização do desempenho em R-D, é necessário que o codificador realize o processo de EM para cada tamanho de bloco, obtendo VMs ótimos  $p_{B_i, \text{opt}}$  para cada bloco  $B_i$ . Finalizado o processo de EM para cada MB, o custo Lagrangiano  $J_{\text{MB}}(m)$  é calculado para cada possibilidade de partição do MB. Assim,

$$J_{\text{MB}}(m) = \sum_{\forall B_i \in \text{MB}} D(p_{B_i, \text{opt}}) + \lambda \left[ \sum_{\forall B_i \in \text{MB}} R(p_{B_i, \text{opt}}) + R(m) \right] \quad (15)$$

onde  $R(p_{B_i, \text{opt}})$  é o custo dos VMs ótimos de cada bloco  $B_i$ ,  $R(m)$  são os bits usados para indicar o tipo de partição e  $D(p_{B_i, \text{opt}})$  é a distorção de cada bloco  $B_i$  após a compensação do movimento. A partição do MB que resultar no menor custo Lagrangiano é aquela então selecionada.



(a)



(b)

Fig. 5. Percentagem de MBs pré-selecionados no modo *Skip* usando o codificador JM7.3 e a estratégia ICC. (a) Seqüências com câmera fixa. (b) Seqüências com câmera móvel.

#### C. Pré-eliminação de Tamanhos de Bloco

Para a pré-eliminação de alguns tamanhos de bloco, adotamos a estratégia ICC no codificador e seguimos no processo de EM a ordem de teste dos modos e tamanhos de bloco definida em (10). A determinação da penalidade de distorção inicial  $D_{\text{init}}(m)$  é obtida por

$$D_{\text{init}}(m) = \lambda \left[ \sum_{\forall B_i \in \text{MB}} R(p_{B_i, \text{pred}}) + R_{\text{init}}(m) \right] \quad (16)$$

onde  $R_{\text{init}}(m)$  denota o número de bits obrigatório para a codificação do modo *Inter* com a partição  $m$  e  $R(p_{B_i, \text{pred}})$  é o custo de codificação do VM predito do bloco  $B_i$ .

A penalidade de distorção obtida torna-se então proporcional ao número de blocos que compõem o MB, pois depende da soma dos bits usados para a codificação dos VMs preditos de cada bloco. Por isso, o processo de EM deve ser iniciado sempre com partições em blocos maiores, evitando a EM de blocos menores se a seguinte condição é satisfeita:

$$\min_{m_i \in \mathbf{M} | i < j} \{J_{\text{MB},\text{min}}(m_i)\} \leq D_{\text{init}}(m_j) + D_L \quad (17)$$

onde  $m_j$  corresponde à partição sob avaliação e

$\min_{m_i \in \mathbf{M} | i < j} \{J_{\text{MB},\text{min}}(m_i)\}$  é o custo Lagrangiano mínimo para o MB dentre todos aqueles que o antecedem de acordo com (10).

Quando a condição (17) for verdadeira, nenhuma nova partição *Inter* necessita ser avaliada e a EM é finalizada, visto que o custo Lagrangiano obtido já é inferior à penalidade de distorção mínima para qualquer outra forma de partição *Inter* ainda não avaliada no conjunto ordenado  $\mathbf{M}$ .

Para aplicar a estratégia ICC de seleção dos modos, em codificadores que seguem o padrão de vídeo H.264, deve-se determinar o número de bits obrigatórios para cada tipo de partição *Inter* do MB, como apresentado em [8].

## V. RESULTADOS EXPERIMENTAIS

Os resultados obtidos em termos de redução de complexidade computacional, relativa ao algoritmo FS, com o uso da estratégia ICC são mostrados na Fig. 6. Aqui são avaliadas simultaneamente a regra de parada por custo, a pré-seleção do modo *Skip* e a pré-eliminação de partições *Inter*. Para tal, é usada uma configuração de codificação IPPP (10Hz), com janela de procura de  $\pm 16$  pixels e com 5 quadros de referência. A proposta mostrou-se eficiente na redução de complexidade tanto com seqüências típicas de videoconferência (*Mother&Daughter*, *News*, *Akiyo*), quanto em seqüências com movimento de câmera (*Foreman*, *Carphone*, *CoastGuard*).

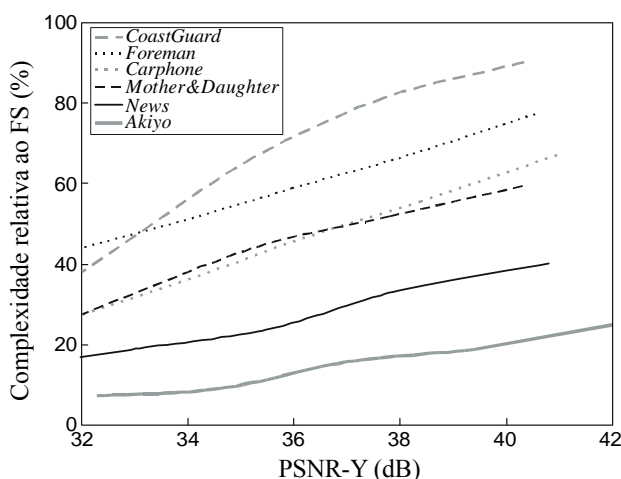


Fig. 6. Redução de complexidade usando a estratégia ICC na seleção do modo de codificação.

## VI. CONCLUSÕES

Neste artigo, foi mostrado que a estratégia ICC proposta permite reduzir significativamente a complexidade do processo de seleção dos modos *Inter* em um codificador de vídeo otimizado em R-D. Na seleção dos modos de

codificação, tal estratégia possibilita tanto a pré-seleção do modo *Skip* quanto a pré-eliminação de diversos tamanhos de bloco durante o processo de EM.

Em codificadores que usam memória de longo prazo, além de reduzir o espaço de busca nas duas dimensões dos VMs, o emprego da estratégia ICC também reduz o espaço de busca na dimensão temporal.

As simulações realizadas, integrando a estratégia ICC ao codificador de referência JM7.3, mostraram uma redução de complexidade significativa no processo de EM (de até 90%), sem perda significativa no desempenho em R-D para seqüências de vídeo com pouco movimento. Com a pré-seleção do modo *Skip*, foi obtido sucesso em até 86% dos MBs (sem perda de desempenho), sem que qualquer avaliação dos modos *Inter* e *Intra* fosse realizada nesses casos.

## REFERÊNCIAS

- [1] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Mag.*, vol. 15, no. 11, pp. 74-90, Nov. 1998.
- [2] ITU-T, Test model 12 / Appendix II of H.263, version 3 (TMN12), Int. Telecomm. Union, Portland, OR, USA, Aug. 2000.
- [3] ITU-T, *Video Codec Test Model, Near-Term, version 8 (TMN8) revision 1*, Int. Telecomm. Union, Sunriver, OR, USA, Sept. 1997.
- [4] ISO/IEC JTC1/SC29/WG11 Doc. AVC-491, "Coded Representation of Picture and Audio Information—MPEG-2 Test Model 5," International Organization for Standardization, Sydney, Australia, Apr. 1993.
- [5] ITU-T Recommendation H.264, *Advanced Video Coding for Generic Audiovisual Services*, Int. Telecomm. Union, Geneva, Switzerland, 2003.
- [6] M. C. Chen and A. N. Willson, Jr., "Rate-distortion optimal motion estimation algorithms for motion-compensated transform video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 2, pp. 147-158, Apr. 1998.
- [7] G. M. Schuster and A. K. Katsaggelos, "A theory for the optimal bit allocation between displacement vector field and displaced frame difference," *IEEE J. Select. Areas Comm.*, vol. 15, no. 9, pp. 1739-1751, Dec. 1997.
- [8] M. Moecke, "Custo de Codificação na Otimização Taxa-Distorção em Compressão de Vídeo com Baixa Taxa de Bits," Exame de Qualificação (Doutorado em Engenharia Elétrica), Universidade Federal de Santa Catarina, Florianópolis, 2002.
- [9] B. Zeng, R. Li, and M. L. Liou, "Optimization of fast block motion estimation algorithms," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 6, pp. 833-844, Dec. 1997.
- [10] M. Moecke and R. Seara, "Rate-distortion optimized video coding with stopping rules: Quality and complexity," in *Proc. IEEE Int. Conf. Image Processing (ICIP'04)*, Singapore, Oct. 2004, pp. 753-756.
- [11] T. Wiegand, X. Zhang, and B. Girod, "Long-term memory motion-compensated prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 70-84, Feb. 1999.
- [12] J. Zhang, M. O. Ahmad, and M. N. S. Swamy, "A new variable size block motion compensation," in *Proc. IEEE Int. Conf. Image Processing*, St. Barbara, CA, USA, vol. 2, Oct. 1997, pp. 164-167.