

Propagação de Qualidade para Sequências de Vídeo de Baixo Movimento na Codificação H.264

Bruno Macchiavello, Alexandre Zaghetto e Ricardo L. de Queiroz

Resumo—Neste trabalho é proposto uma técnica de codificação de vídeo capaz de explorar a propagação de previsões de alta qualidade em sequências de baixo ou médio movimento. São propostos dois algoritmos de baixa complexidade para seleção do modo de codificação do macrobloco e para o passo de quantização. Tais algoritmos foram implementados utilizando como codec de vídeo o padrão H.264/AVC. As técnicas aqui apresentadas são capazes de aumentar o desempenho em termos de taxa-distorção e possuem o potencial de reduzir a complexidade computacional do H.264. Os resultados mostram que o método apresentado pode superar a codificação convencional H.264/AVC por até 1,0 dB para sequências de baixo movimento.

Palavras-Chave—codificação de vídeo, H.264, redução de complexidade.

Abstract—In this paper, we explore the propagation of high-quality predictions in low-motion or no-motion macroblocks and we propose a simplified mode selection and quantizer parameter adaptation algorithms that increase rate-distortion performance and may potentially reduce encoding complexity of H.264/AVC. Results show that the proposed scheme can outperform regular H.264/AVC by up to 1.0 dB in low-motion sequences.

Keywords—video coding, H.264, complexity reduction.

I. INTRODUÇÃO

O codificador H.264/AVC tem sido bem explicado na literatura [1], [2]. Inúmeros trabalhos têm ilustrado o seu desempenho mostrando resultados comparativos contra codificadores tais como MPEG-2 [3] e H.263 [4]. Mesmo que os resultados apontam para pelo menos uma melhora em fator de dois em relação aos padrões anteriores, há cenários específicos, tais como a codificação de sequências de baixo movimento, onde o H.264/AVC pode ser melhorado. Além disso, as muitas melhorias na codificação trazidas pelo H.264/AVC tem também adicionado uma quantidade significativa de complexidade computacional. Entre essas melhorias, estima-se que os algoritmos de estimação movimento inteira e sub-pixel são responsáveis por cerca de 90% da carga computacional do codificador [5].

No padrão H.264/AVC, cada macrobloco (MB) em um *slice* é codificado usando uma previsão a partir de dados previamente codificados (exceto no modo I_PCM)[2]. No modo intra, um MB pode ser predito a partir de amostras que já foram codificadas e reconstruídas no *slice* atual. No modo inter, a previsão usa um ou mais quadros ou campos de vídeo codificados anteriormente para realizar estimação e compensação de movimento. Em ambos os casos, um MB de 16×16 pixels pode ser dividido em blocos e sub-blocos de partições de um modo *quad-tree* [1]. Uma opção disponível para a previsão do inter é o modo *skip*. O padrão H.264/AVC também permite variar o parâmetro de quantização (QP) de

cada macrobloco, a fim de alcançar um melhor desempenho em termos de taxa-distorção.

Neste trabalho, estudamos o comportamento de sequências de vídeo de baixo movimento, como sequências *head-and-shoulders* e aplicações remotas de *desktop*. Apresentamos um método de pré-classificação e seleção de QP que se baseia na propagação de qualidade de uma boa referência. Um algoritmo inicial pré-classifica cada MB de acordo com uma pontuação de movimento e, dependendo da classe de MB atribuída, restringe os modos selecionáveis para um subconjunto dentre todos os modos possíveis definidos pelo padrão H.264/AVC. O objetivo principal deste algoritmo é identificar uma série de MBs que vão ser codificados sem partições. desta forma é possível alcançar uma redução na complexidade de codificação. É importante mencionar que, como um esforço para reduzir a complexidade na codificação de vídeo, trabalhos anteriores utilizam algoritmos para a decisão rápida entre modo inter e intra [6], [7], enquanto outras abordagens utilizam uma arquitetura de finalização antecipada [8], [9]. Aqui redução da complexidade, é conseguida através de pré-seleção rápida de modos disponíveis para cada MB. Depois, cada de MB ter sido classificado, selecionamos o QP utilizados para cada MB. A referência inicial da série de MB será codificada com uma alta qualidade, a qual será propagada para os quadros futuros.

II. PROPAGAÇÃO DE QUALIDADE

Durante a codificação H.264 de uma sequência de vídeo de baixo movimento, espera-se que a maioria dos MBs serão codificados sem ser dividido em blocos ou sub-blocos. Além disso, muitos deles provavelmente serão codificados usando o modo *skip* ou com vetor de movimento nulo (*zero motion vector* - ZMV). Logo, se o codificador de vídeo fosse capaz de identificar uma série de MBs de quadros contíguos codificados sem partições, que podem representar uma parte de fundo ou um objeto estático, a qualidade na codificação do MB inicial seria muito importante. Uma referência inicial de alta qualidade pode gerar menos resíduos nos MBs futuros. Uma primeira boa referência também poderia aumentar drasticamente o número de MBs codificados com o modo *skip*.

Por isso, se podemos determinar quais MBs serão utilizados como referência inicial e codificá-los com alta qualidade, esta qualidade deve-se propagar, de modo que o resíduo dos outros MBs possa ser codificado com um QP alto. Este cenário é ilustrado na Fig. reffig:mbs.

Note que, diferentemente do algoritmo de otimização taxa-distorção, a propagação de qualidade é baseado na análise

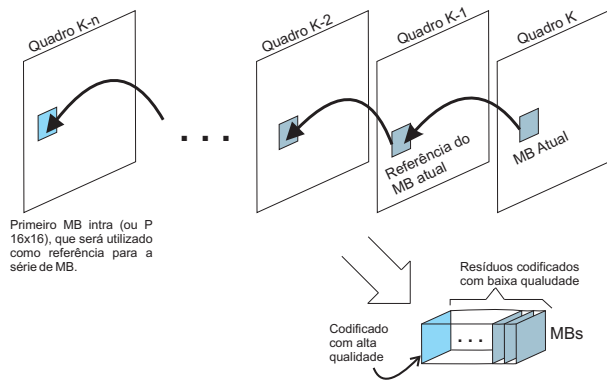


Fig. 1. Série de MBs codificados sem partições

de quadros futuros e anteriores. Algumas implementações possíveis de um algoritmo de codificação que identifica sequências de MBs codificados sem partições, pode ser feito usando um buffer de quadros ou técnicas de olhar para frente (*look-ahead*). No entanto, o buffer de quadros aumenta os requisitos de memória do sistema e apresenta atraso.

III. MÉTODO PROPOSTO

Neste trabalho, propomos uma rotina simples, a fim de identificar uma sequência de MBs de quadros adjacentes codificados sem partições. Essa rotina não requer um buffer de quadros e pode ser realizada online. O método é baseado em um algoritmo de pré-classificação para restringir os modos disponíveis de codificação e para selecionar o QP apropriado. Este algoritmo tem complexidade computacional baixa já que a classificação e seleção QP é baseado na qualidade do bloco de referência ZMV.

Seja $MB_{(x,y)}^k$ o MB atual, no k -ésimo quadro centrado nas coordenadas (x, y) . Portanto, $MB_{(x,y)}^{k-1}$ representa a MB co-localizado no quadro anterior. Além disso, vamos definir que SSD representa uma função que calcula a soma dos quadrados diferença entre dois MBs. Então, para cada quadro o método de pré-classificação é o seguinte:

ALGORITMO DE PRÉ-CLASSIFICAÇÃO

```

1  for todos os valores possíveis de x
2  for todos os valores possíveis de y
3  if  $SSD(MB_{(x,y)}^k, MB_{(x,y)}^{k-1}) < T_1$ 
4  then  $mode\_MB_{(x,y)}^k \leftarrow 0$ 
5  elseif  $SSD(MB_{(x,y)}^k, MB_{(x,y)}^{k-1}) < T_2$ 
6  then  $mode\_MB_{(x,y)}^k \leftarrow 1$ 
7  elseif  $SSD(MB_{(x,y)}^k, MB_{(x,y)}^{k-1}) > T_3$ 
8  then  $mode\_MB_{(x,y)}^k \leftarrow 3$ 
9  else
10     if  $mode\_MB_{(x,y)}^{k-1} = 3$ 
11     then  $mode\_MB_{(x,y)}^k \leftarrow 1$ 
12     else  $mode\_MB_{(x,y)}^k \leftarrow 2$ 
    
```

Como pode ser visto, o método de pré-classificação é uma seleção com base em limiares (T_1, T_2 e T_3). Utilizando apenas a SSD computada do MB co-localizada no quadro anterior, cada modo ($mode_MB$) selecionado no algoritmo representa na verdade uma grupo de modos de codificação. Se o modo

de 0 é selecionado, então o MB tem apenas duas opções disponíveis coding: codificação *skip* ou ZMV. Se o modo de 1 é escolhido, então o MB será codificado como *skip* ou 16×16 sem partições ($P_{16} \times 16$). Se foi selecionado o modo de 2, então o MB pode ser codificado apenas com todas as opções inter preditivas e bi-preditivos disponíveis no H.264/AVC (incluindo partições). Se o modo de 3 foi selecionado então a MB só pode ser codificado usando qualquer um dos modos disponíveis do tipo intra. Note que os modos 0 e 1 podem ser vistos como pontos de finalização antecipada para a previsão do inter. No entanto, a classificação também decide entre predição inter e intra. Pode-se esperar que, se muitos blocos estão mal classificados, a taxa irá aumentar gerando uma perda significativa na codificação de desempenho. No entanto, para aplicações remotas de *desktop*, ou de sequências de baixo movimento, este algoritmo simples será preciso o suficiente e reduzir o tempo de codificação.

Para a seleção de QP é proposto um método simples o qual é baseado no modo selecionado anteriormente. O algoritmo é mostrado a seguir, onde $REF(MB_{(x,y)}^k)$ indica o MB usado como referência para $MB_{(x,y)}^k$, QP_MB indica o QP usado para um determinado MB e QP_I é o QP inicial (que define a qualidade desejada).

QP SELECTION IN PRE-CLASSIFICATION

```

1  for todos os valores possíveis de x
2  for todos os valores possíveis de y
3  if  $mode\_MB_{(x,y)}^k = 0$  or  $mode\_MB_{(x,y)}^k = 1$ 
4  then  $l \leftarrow k, i \leftarrow x, j \leftarrow y$ 
5  while  $QP\_REF(MB_{(i,j)}^l) \neq QP_I - \delta_1$  and  $l \geq 0$ 
6  then
7  (i, j) ← coordinates of  $REF(MB_{(i,j)}^l)$ 
8  l ← l - 1
9  if  $l \geq 0$ 
10 then  $QP\_MB_{(x,y)}^k \leftarrow QP_I + \delta_2$ 
11 else  $QP\_MB_{(x,y)}^k \leftarrow QP_I$ 
12 elseif  $mode\_MB_{(x,y)}^k = 2$ 
13 then  $QP\_MB_{(x,y)}^k \leftarrow QP_I$ 
14 else  $QP\_MB_{(x,y)}^k \leftarrow QP_I - \delta_1$ 
    
```

Dado os algoritmos descritos anteriormente se a referência tem uma alta qualidade (codificados com QP baixo), esta qualidade será propagada para amostras futuras, enquanto estas amostras sejam semelhantes à referência. Em outras palavras, MBs de referência de boa qualidade também irá produzir previsões de boa qualidade, especialmente quando o modo ZMV ou *skip* são selecionados. A flutuação de QP vai potencialmente melhorar o desempenho RD do codec. A flutuação QP é baseada em duas constantes previamente definidas: δ_1 e δ_2 . Note-se que todos os blocos intra codificados têm a mais alta qualidade, já que eles sempre são codificados com um QP baixo. Esta alta qualidade certamente vai propagar para blocos co-localizado em quadros futuros que são codificados com o modo 0. Para os blocos classificados com o modo 1, o mais provável é que o bloco do quadro anterior codificado como intra será escolhido como referência. É por isso que, no algoritmo de pré-classificação o modo 2 não é uma opção válida, se o MB co-localizado no quadro anterior foi codificado com o modo 3. Quando isto acontece, o modo de 1 é selecionado, favorecendo assim, o uso de uma referência

de alta qualidade. Note que, a fim de reduzir a complexidade computacional do algoritmo de seleção de QP, uma variável que propaga o valor do QP do MB de referência inicial pode ser utilizada durante a implementação do método proposto. Isso evita a necessidade de pesquisar o QP utilizado em todas as referências anteriores.

IV. RESULTADOS

A fim de testar o método proposto, foram utilizados seis diferentes sequências de teste de diferentes resoluções. Usamos três sequências clássicas com baixo movimento: “Akiyo” e “Mother-daughter” em resolução CIF e ‘Hall-objects” em resolução QCIF. Foram criadas algumas sequências de vídeo para aplicações *desktop*. A Figura 2 mostra o quadro inicial de cada sequências criada. Duas estão em resolução de alta definição (1080p), uma é um *desktop* com uma aplicação de webcam (“Dsk+Cam”), a outra também tem uma aplicação de webcam e um editor de texto simples (“Dsk+Cam+Txt”). A última (“Dsk+Vid”) está em resolução CIF e representa uma área de trabalho com a sequência de vídeo clássica “Football” e uma calculadora (operações são feitas na calculadora durante a sequência). Os testes foram realizados usando o modo de codificação *IPPPP...*, com dois quadros de referência, janela de busca de 16 pixels, sem otimização de taxa-distorção, estimação de movimento completa (*full search*) e codificação de entropia CABAC. Para a seleção de QP, os limiares foram $\delta_1 = \delta_2 = 6$, o que significa que o passo QP real será duplicado ou reduzido pela metade para a propagação de qualidade. Os limiares $T_1 = 150$, $T_2 = 1000$ e $T_3 = 350.000$ foram determinados empiricamente utilizando uma outra sequência de teste (“Foreman”).

Na Tabela I apresentamos o $\Delta PSNR$, calculado usando o método Bjontegard [10], com $QP_I = \{18, 22, 26, 30\}$. Quando o valor do $\Delta PSNR$ calculado é positivo significa que o método proposto supera o padrão H.264/AVC convencional, enquanto um valor negativo indica que o H.264/AVC tem um melhor desempenho global. Como pode ser visto apenas uma das sequências apresentadas mostram uma perda muito pequena no desempenho taxa-distorção geral, enquanto todos os outros mostram uma melhora significativa. Também apresentados na tabela é o percentual de quantas vezes cada modo de pré-classificação foi selecionado durante a codificação. Observe que, para todas as sequências a maioria de todos os MBs são codificados com os modos 0 ou 1, enquanto apenas menos de metade dos blocos são codificados com o modo 2 (o qual é o único que testa todas as partições disponíveis no H.264). Este comportamento potencializa uma redução significativa no tempo de codificação.

V. CONCLUSÕES

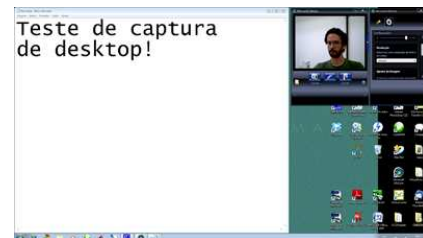
Neste trabalho foram apresentados algoritmos de pré-classificação para definir o modo de codificação e o parâmetro de quantização para cada MB. A pré-classificação decide quais os modos de codificação estão disponíveis para um determinado MB. A variação do QP, é feita de tal maneira que um bloco de referência inicial é codificado com alta qualidade, enquanto blocos futuros que utilizam esse bloco



(a) “Dsk+Vid”, CIF



(b) “Dsk+Cam”, Full HD



(c) “Dsk+Cam+Txt”, Full HD

Fig. 2. Quadro exemplo das sequências de aplicações de *desktop* remoto.

TABELA I
PERCENTUAL DE MODOS SELECIONADAS E $\Delta PSNR$ MÉDIO
COMPARANDO COM O H.264/AVC CONVENCIONAL.

Sequencia	% dos Blocos codificados com um determinado modo				$\Delta PSNR$
	0	1	2	3	
Akiyo	59.02	27.00	12.97	0.01	0.27
Mother-Daughter	4.30	63.09	32.56	0.04	-0.05
Hall	0.03	59.70	40.22	0.03	0.4
Dsk+Vid	89.99	0.69	8.22	1.10	0.92
Dsk+Wcam	93.56	3.87	2.34	0.23	1.02
Dsk+Wcam+Txt	97.95	1.20	0.80	0.04	0.90

específico como referência são codificados com um QP alto. Trabalhos futuros podem incluir um estudo a fim de determinar as constantes (δ_1 , δ_2) ótimas para a seleção de QP. No entanto, os resultados obtidos já mostram um desempenho superior em termos de taxa-distorção para todas as aplicações de remotas de *desktop* testadas, com uma potencial de redução significativa no tempo de codificação. Para sequências de baixo movimento (não *desktop*), apenas uma delas apresentou uma perda muito pequena no desempenho, para os demais testes o método proposto mostrou uma melhora de qualidade objetiva e potencial redução na complexidade de codificação.

REFERÊNCIAS

- [1] I. E. G. Richardson, *H.264 and MPEG-4 video compression*, Wiley, EUA, 2003.
- [2] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A.; Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. on CSVT*, vol. 13, pp. 560–576, July 2003.
- [3] B. G. Haskell, A. Puri, and A. N. Netravalli, *Digital Video: An Introduction to MPEG-2*, Chapman and Hall, EUA, 1997.
- [4] ITU-T, "Video Coding for Low Bit Rate Communication. ITU-T Recommendation H.263," Version 1: Nov. 1995, Version 2: Jan. 1998, Version 3: Nov. 2000.
- [5] Y.W. Huang, B.Y. Hsieh, S.Y. Chien, S.Y. Ma, and L.G. Chen, "Analysis and complexity reduction of multiple reference frames motion estimation in H.264/AVC," *IEEE Trans. on CSVT*, vol. 16, no. 4, pp. 507–522, 2006.
- [6] C. Kim and C.-C.J. Kuo, "A feature-based approach to fast h.264 intra/inter mode decision," in *Proc. of the IEEE ISCAS 2005*, May 2005, vol. 1, pp. 308–311.
- [7] M. Jafari and S. Kasaei, "Fast intra-prediction mode decision in h.264 advanced video coding," in *Proc. of the IEEE ICCS 2006*, 2006, pp. 1–6.
- [8] C.Y. Su and S.L. Chang, "Adaptive early termination for fast h.264 video coding," in *Proc. of the IEEE ISM 2007*, 2007, pp. 72–77.
- [9] L. Yang, K. Yu, J. Li, and S. Li, "An effective variable block-size early termination algorithm for h.264 video coding," *IEEE Trans. on CSVT*, vol. 15, no. 6, pp. 784–788, 2005.
- [10] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," *presented at the 13th VCEG-M33 Meeting, Austin, TX*, Apr. 2001.