

8

Conclusões e Sugestões para Trabalhos Futuros

8.1

Conclusões

Este trabalho foi focado na codificação de textura intra-quadro, sem predição, para a codificação de imagens baseada em objeto. Dentre os ‘profiles’ previstos no MPEG-4 parte 2, esse tópico enquadra-se no ‘profile’ *Simple de Estúdio*, que é um ‘profile’ de alta qualidade para uso em estúdio, empregando apenas *I-frames* para facilitar a edição [18],[19],[23].

Para esses codificadores, esta tese enfatizou a exploração de indicadores morfológicos do objeto a ser codificado, englobando características de forma e textura, visando à análise de eficiência de algoritmos conhecidos na literatura e também à proposição de novos algoritmos de codificação.

Duas abordagens para a codificação de textura intra-quadros baseada em objeto foram analisadas: a DCT baseada e blocos, associada a diferentes técnicas de extrapolação dos blocos de contorno, e a DCT adaptativa à forma - SA-DCT. Dentre os algoritmos de extrapolação conhecidos na literatura, foi selecionado o algoritmo EI - *Extension/Interpolation*, por apresentar a melhor eficiência de codificação. Apesar da complexidade computacional a princípio ser grande, esse esforço pode ser reduzido através de uma implementação rápida alternativa, baseada em operações matriciais simples [80] . O algoritmo de extrapolação EI associado à DCT-2D foi denominado, neste trabalho, EI-DCT.

Inicialmente foi apresentada uma análise comparativa ente alguns algoritmos descritos na literatura para o particionamento da imagem em blocos, em termos do compromisso ‘eficiência × carga computacional’. Após a proposição de dois critérios de otimização distintos para a redução do número de blocos, os algoritmos Ortogonal Simplificado e Flexível Simplificado visando à redução do número de blocos de contorno foram apontados como os melhores no atendimento desse compromisso.

Para explorar as características de forma e textura intrínsecas ao

objeto a ser codificado, foram propostos os indicadores morfológicos AE (Área Efetiva), DNPO (Distribuição de blocos de acordo com o Número de Pixels do Objeto) e TNPO (Textura de acordo com o Número de Pixels do Objeto). O indicador AE provê a informação do percentual da área coberta pelos blocos a serem codificados que efetivamente é ocupada por pixels do objeto. Ele pode fornecer, a princípio, uma previsão de qual das abordagens de codificação é mais adequada a uma determinada faixa de bits, dependendo do objeto a ser codificado. Observou-se que o algoritmo EI-DCT é o mais eficiente a taxas baixas, mas que à medida em que a taxa sobe, o desempenho da SA-DCT tende a superar bastante o da EI-DCT no caso de objetos com baixa AE. Os indicadores DNPO e TNPO, por sua vez, classificam os blocos a serem codificados em 11 grupos, dependendo do número de pixels do objeto (de 1 a 64). O indicador DNPO informa a porcentagem de blocos existentes em cada um dos grupos e o indicador TNPO, a textura média dos blocos nesses grupos.

Este trabalho examinou duas estratégias reportadas na literatura para a escolha da direção preferencial de processamento do algoritmo EI, na codificação de imagens baseada em objeto: MALV e MILV. Essas estratégias baseiam-se nas variâncias dos comprimentos dos objetos em ambas as direções. MALV seleciona como prioritária a direção de máxima variância, ao passo que MILV escolhe a direção de mínima variância para realizar o primeiro processamento. Embora a literatura ainda não tenha apresentado um consenso sobre a adequação de uma ou outra estratégia [77],[80], este trabalho mostrou que para alguns objetos, normalmente os de alta textura, MILV chega a ser bem superior a MALV e para outros, normalmente os de baixa textura, o contrário se verifica.

Além disso, constatou-se que para a EI-DCT podem ocorrer variações nos desempenhos de MILV e MALV para grupos de blocos (definidos no indicador TNPO) distintos. Essas variações dependem dos valores do indicador TNPO, que é uma função da forma e da textura. Para grupos com texturas muito baixas ou muito altas, a utilização do indicador TNPO, com base em limiares das texturas dos grupos de blocos, constitui uma estratégia razoável para a escolha do método mais adequado (MILV ou MALV). Contudo, à medida em que os blocos de contorno tornam-se mais preenchidos por pixels do objeto, a escolha desses limiares para a determinação da estratégia mais eficiente fica dificultada, uma vez que as variâncias dos comprimentos em ambas as direções passam a assumir valores muito próximos. Além disso, para grupos de blocos apresentando um valor de textura mediano, pode haver uma grande quantidade de blocos com

baixa e alta textura, para os quais MALV e MILV, respectivamente, são mais indicados.

Portanto, foi proposta uma estratégia diferente para determinar a primeira direção de processamento do algoritmo EI nesses casos: trata-se da estratégia MACES (*MAximum Cumulative Energy Sum*), que mede a capacidade de compactação de energia dos coeficientes da DCT em ambas as direções. Essa estratégia mostrou-se melhor que MILV e MALV em termos de correlação entre o indicador TNPO e a determinação da primeira direção de processamento do algoritmo de extrapolação EI, à medida em que os blocos tornam-se mais cheios e apresentam, em média, valores de textura em uma faixa intermediária.

Esses resultados experimentais motivaram o desenvolvimento de um método híbrido que possibilitasse a escolha adaptativa da estratégia mais indicada para cada grupo de blocos de contorno (entre MALV, MILV e MACES), utilizando o indicador TNPO como a base para o chaveamento entre as estratégias. Como a seleção da melhor estratégia pode variar com a taxa de bits, os limiares de chaveamento foram calibrados de forma a aumentar a eficiência de codificação a taxas baixas. Isso foi feito porque é nessa faixa que a EI-DCT é a estratégia mais indicada, por apresentar desempenho superior ao da SA-DCT.

Foi mostrado que o esquema híbrido ou é comparável, ou supera a eficiência da melhor entre as estratégias MILV e MALV para cada objeto. Cabe ressaltar ainda que o bloco extrapolado gerado pelo método proposto é independente da taxa de bits, o que representa uma vantagem sobre o esquema que usa o critério que seleciona a direção que produz o menor produto $emq \times nbits$ [77].

Uma análise semelhante à apresentada para a EI-DCT foi desenvolvida para a SA-DCT, com o objetivo de determinar a direção prioritária para o primeiro processamento 1D. Resultados dessa análise permitiram a seleção da estratégia mais adequada a cada grupo, apontando também para algumas condições em que a relação entre o indicador TNPO e as eficiências de MILV e MALV é comprometida. Em seguida, o indicador TNPO foi aplicado como a base de um esquema híbrido para a SA-DCT que seleciona adaptativamente a melhor estratégia para um grupo específico de blocos de contorno. Preferencialmente, a estratégia selecionada é MALV ou MILV, mas há casos em que a opção é pela estratégia MACES.

A eficiência do esquema híbrido proposto para a SA-DCT mostrou-se comparável ou superior ao das estratégias MALV, MILV e ‘COLUMNS-FIRST’ (algoritmo original de Sikora [67]). Ressalta-se que também aqui a

escolha da direção preferencial para um bloco específico independe da taxa de bits requerida.

O uso conjunto de cada um dos esquemas híbridos propostos (para a EI-DCT no Capítulo 5 e para a SA-DCT no Capítulo 6) e de algoritmos de partição da imagem em blocos foi analisado. Concluiu-se que é possível obter uma melhoria de desempenho sensível do esquema híbrido para a EI-DCT ao se usar o particionamento Flexível Simplificado, devido principalmente ao aumento da Área Efetiva dos objetos. Entretanto, no caso da SA-DCT, foi verificado que não existe garantia de que esses algoritmos de partição possam melhorar a eficiência de codificação ao concentrar os pixels do objeto em um menor número de blocos, como havia sido afirmado em [76]. Na verdade, o ganho de codificação pode inclusive diminuir, qualquer que seja a estratégia usada para a determinação da direção preferencial de processamento. Dado ainda que o emprego desses algoritmos implica aumento da carga computacional tanto do codificador quanto do decodificador, mostramos que o seu uso não é, de forma alguma, indicado no caso da codificação adaptativa à forma (SA-DCT).

Finalmente, o indicador morfológico TNPO foi utilizado em uma análise comparativa de eficiência de codificação dos esquemas híbridos propostos para a EI-DCT e para a SA-DCT. Os resultados dessa análise levaram à proposição de um esquema de codificação orientada por objeto adaptativo em função da taxa de bits, da textura e da forma do objeto a ser codificado. O codificador adaptativo proposto foi capaz de escolher, para cada grupo de blocos, a estratégia que, em média, proporciona os maiores ganhos de codificação.

Em geral, a SA-DCT é mais eficiente que a EI-DCT nos blocos vazios. Os blocos medianamente preenchidos são melhor codificados a taxas baixas utilizando-se a EI-DCT. Porém, à medida em que a taxa sobe, a SA-DCT tende a tornar-se mais eficiente. O valor da taxa fixada para o chaveamento entre as estratégias foi definido, para esses blocos, com base no indicador TNPO do objeto a ser codificado. Nos blocos mais cheios, a alocação de bits utilizada influenciou bastante no desempenho final de codificação, e portanto foi analisada em conjunto com a capacidade de suavização de textura proporcionada pela estratégia EI-DCT. Foi mostrado que essa capacidade depende, essencialmente, da textura média nesses grupos e de quão cheios apresentam-se os blocos.

Para a faixa de 0 a 8 bppo, o esquema adaptativo mostrou-se comparável ou superior à melhor dentre as estratégias reconhecidas na literatura como as mais eficientes para a codificação a taxas baixas (EI-DCT) e a taxas

altas (SA-DCT), tanto em valores de RPR, quanto na análise da qualidade subjetiva das imagens.

Uma outra vantagem do esquema adaptativo é que o conjunto final de coeficientes não necessita ser computado a cada mudança da taxa de bits. Ao contrário, ele pode ser previamente calculado para quatro faixas de taxas de bits específicas: menores que 2 bppo, 2 a 4 bppo, 4 a 6 bppo e maiores que 6 bppo. Essa característica também aumenta o conjunto de aplicações possíveis para o esquema adaptativo proposto.

8.2

Sugestões para Trabalhos Futuros

Atualmente, o padrão MPEG-4, como relatado na Seção 2.2, prevê dois algoritmos para a codificação de textura dos *P-frames* e dos *B-frames*: a DCT baseada em blocos com a técnica de extrapolação ‘zero-padding’ e a SA-DCT [17],[18],[82]. Um tópico que daria prosseguimento direto ao desenvolvimento desse trabalho seria a extensão dos algoritmos propostos nos Capítulos 5, 6 e 7 para a codificação de textura inter-quadros (com compensação de movimento). O sinal de entrada, nesse caso, é a imagem de erro de predição, que tem características bastante distintas das imagens usadas nas simulações desta tese. Além das texturas baixas, ela apresenta um grande número de coeficientes nulos. Portanto, para a realização deste tópico, inicialmente deverão ser escolhidos indicadores morfológicos que sejam representativos do sinal de entrada e que possam constituir-se como a base de chaveamento dos algoritmos híbridos. A sugestão é implementar um algoritmo híbrido que escolha a primeira direção de processamento da SA-DCT, ao mesmo tempo em que seja capaz de selecionar a melhor estratégia de codificação para cada grupo de blocos de contorno, dentre a DCT baseada em blocos e o algoritmo híbrido da SA-DCT, dependendo da taxa de bits e dos indicadores morfológicos a serem definidos.

Sugere-se também a incorporação do método proposto em 2000 por Acocella em [44] no esquema híbrido para a SA-DCT apresentado no Capítulo 6 desta tese, visando à codificação de textura intra-quadros. Essa medida irá minimizar os efeitos da distorção do valor médio e da parcela AC da imagem (relatados no Capítulo 2), garantindo que a imagem reconstruída permaneça com média nula. Uma segunda proposta de continuidade do trabalho [44] é a utilização do alinhamento dos coeficientes pela fase [74] (em contraposição ao alinhamento pelo índice [67]) após o primeiro processamento 1D da SA-DCT.

A incorporação de abordagens para a remoção do efeito de blocagem ao esquema híbrido da EI-DCT também poderá melhorar a qualidade subjetiva das imagens codificadas a taxas baixas. Em 2003, o trabalho [83] descreveu algumas técnicas para a redução dessa distorção, classificando-as como métodos de filtragem espacial, métodos de filtragem baseados em DCT e métodos híbridos. Uma das preocupações das técnicas conhecidas é de que não haja suavização das texturas reais da imagem, devido à característica de filtragem passa-baixas. Foi proposto em [83] um método para a redução do efeito de blocagem no domínio da transformada que não apresenta esse inconveniente. Em regiões de textura suave, o método explora a correlação entre blocos vizinhos para reduzir a descontinuidade nos pixels de fronteira entre eles. Nesse caso, aplica-se uma filtragem DCT à região 8×8 em torno da fronteira desses blocos. Nas regiões de textura ou que contenham bordas, é aplicado um filtro suavizante, mas que conserva as bordas.

Sugere-se, em seguida, a implementação dos algoritmos de codificação desenvolvidos em linguagem compatível com o padrão MPEG-4, de forma a utilizar os mesmos quantizadores, codificadores por entropia e códigos de comprimento de corrida desse padrão.

Em 1993, Shapiro propôs um codificador de imagem embutido para coeficientes da Transformada Wavelet [91]. O algoritmo *EZW - Embedded Zerotree Wavelet* - pode terminar a codificação a qualquer ponto, permitindo a medida exata da distorção. Em 1996, Said e Pearlman apresentaram uma implementação diferente do algoritmo *EZW*, baseada no *SPIHT - Set Partitioning In Hierarchical Trees* [92], proporcionando desempenho ainda melhor que o algoritmo anterior, pois a parte crucial do processo de codificação - a forma pela qual os sub-conjuntos de coeficientes são particionados - é fundamentalmente diferente. Ainda em 1996, Xiong et al apresentaram uma extensão do algoritmo *SPIHT* para a DCT baseada em blocos, tratando cada bloco 8×8 no domínio da transformada como uma árvore de profundidade 3, a qual pode ser vista como uma decomposição em 64 subbandas, conforme mostra a Figura 8.1. Foram identificadas relações de descendência entre os coeficientes da DCT como se segue: o conjunto dos quatro descendentes do coeficiente j é $\{4j, 4j + 1, 4j + 2, 4j + 3\}$, para $0 \leq j \leq 15$. O coeficiente DC é a raiz da árvore e possui apenas 3 descendentes: $\{1, 2, 3\}$. A seguir, foi aplicado o mesmo codificador de [92]. A implementação de um quantizador SPIHT para a SA-DCT, baseado nos trabalhos relatados, seria de grande interesse para aplicações em codificadores de textura a baixas e médias taxas de bits.

0	1	4	5	16	17	20	21
2	3	6	7	18	19	22	23
8	9	12	13	24	25	28	29
10	11	14	15	26	27	30	31
32	33	36	37	48	49	52	53
34	35	38	39	50	51	54	55
40	41	44	45	56	57	60	61
42	43	46	47	58	59	62	63

Figura 8.1: Um bloco 8×8 de DCT, tratado como uma árvore de coeficientes de profundidade 3, com 64 subbandas.