

Quantização Vetorial Binária de Imagens Codificadas por BTC

ABRAHAM ALCAIM
LUCIANO VEREDA OLIVEIRA

IBM Brasil-Centro Científico Rio
Av. Presidente Vargas, 844 - VB22
20071-001 - Rio de Janeiro - RJ
alcaim@riosc.bitnet

Abstract. This paper describes the application of a simple method of classified vector quantization to encode the bit plane generated by the BTC image compression technique. Simulation results show that this method achieves a better performance if compared to vector quantization without using classification of image blocks. Moreover, the classified vector quantization approach significantly reduces the complexity of the search procedure in the encoding process.

1 Introdução

Uma das grandes limitações enfrentadas pelo algoritmo BTC (Block Truncation Coding) [Delp-Mitchell (1979)] para compressão de imagens é o alto número de bits necessário para representar o plano de bits associado a cada bloco de tamanho $n \times n$ da imagem. Sem um procedimento adicional de codificação é necessário 1 bit/pixel para representar essa informação. Uma das técnicas propostas para contornar esta dificuldade é a quantização vetorial (QV) do plano de bits. A proposta do presente trabalho é reduzir as degradações introduzidas por esta técnica, através da utilização de QV classificada. A idéia básica desta proposta é tratar de forma distinta cada bloco da imagem, utilizando um dicionário de códigos especificamente projetado para a classe a que o bloco pertence. O objetivo é obter um melhor desempenho, com a vantagem adicional de reduzir a complexidade do procedimento de busca no processo de codificação.

Na Seção 2 deste trabalho é feita uma breve descrição do algoritmo BTC. A Seção 3 apresenta uma das técnicas existentes para a redução do número de bits necessário para codificar o plano de bits - a quantização vetorial - e mostra os resultados obtidos e suas limitações. Uma maneira de contornar as dificuldades apresentadas na Seção 3 - a quantização vetorial classificada do plano de bits - é proposta na Seção 4. Seu desempenho é analisado para várias imagens de teste e comparado aos resultados obtidos na Seção 3. Também são discutidos os aspectos relativos à redução de complexidade proporcionada por esta técnica. Finalmente, na Seção 5, são apresentadas as principais conclusões do trabalho.

2 Codificação em Blocos por Truncamento

O algoritmo BTC (Block Truncation Coding) tem como primeiro passo a divisão da imagem em blocos contendo $N = n \times n$ pixels X_1, X_2, \dots, X_N , onde n é usualmente igual a 4 (este é o valor escolhido também para este trabalho). Cada pixel pertencente a um bloco é, então, quantizado para um nível Y_0 se seu valor for inferior a um certo limiar X_{th} , ou para Y_1 caso contrário.

Os níveis de quantização são escolhidos de forma a atender diferentes critérios de preservação de momentos. Isso pode ser particularmente útil [Nunes *et al.* (1992)] para imagens comprimidas que passarão por um posterior processo de classificação. No algoritmo inicial proposto por Delp e Mitchell [Delp-Mitchell (1979)] a escolha visava preservar a média μ e o desvio padrão σ do bloco. Outros parâmetros frequentemente escolhidos são o par (μ, μ_i) , onde μ_i é o valor médio dos pixels abaixo do limiar de quantização X_{th} , ou o par (μ, μ_a) , onde μ_a corresponde ao primeiro momento absoluto central.

O conjunto de parâmetros estatísticos escolhido é quantizado e transmitido para o decodificador como uma informação paralela. Cada bloco $n \times n$, então reduzido a um conjunto de pixels de valores Y_0 ou Y_1 , é representado por um plano de bits, onde os 0's e 1's indicam se cada pixel foi quantizado para o nível Y_0 ou Y_1 , respectivamente. O plano de bits obtido é também transmitido para o decodificador.

No decodificador BTC a imagem é reconstruída a partir da recuperação do plano de bits e dos parâmetros que compõem a informação paralela de cada bloco. Com base nessa informação é possível determinar os níveis de quantização Y_0 e Y_1 e asso-

ciar estes níveis aos 0's e 1's do plano de bits.

Neste trabalho o limiar X_{th} é fixado como a média do bloco, definida por:

$$X_{th} = \frac{1}{N} \sum_{i=1}^N X_i \quad (1)$$

Os momentos a serem preservados foram escolhidos como sendo a média μ e o primeiro momento absoluto central μ_a , definido por:

$$\mu_a = \frac{1}{N} \sum_{i=1}^N |X_i - \mu| \quad (2)$$

Pode-se mostrar [Lema-Mitchell (1984)] que os valores de Y_0 e Y_1 que preservam os momentos μ e μ_a para cada bloco de imagem reconstruída são dados por:

$$Y_0 = \mu - \frac{N}{2\beta} \mu_a \quad (3)$$

$$Y_1 = \mu + \frac{N}{2(N - \beta)} \mu_a \quad (4)$$

onde β é o número de pixels abaixo do limiar X_{th} . Esta opção é justificada pelo fato da utilização destes parâmetros minimizar o valor médio quadrático do erro entre a imagem original e a reconstruída. Além disso, esses parâmetros são bastante resistentes aos efeitos da quantização [Oliveira-Alcain (1993)].

A informação paralela foi quantizada neste trabalho com dois quantizadores escalares de 8 bits: um para codificar μ e outro para codificar μ_a . Esta estratégia praticamente elimina os erros na quantização da informação paralela, dando ênfase aos efeitos da quantização vetorial do plano de bits, objetivo primordial do presente trabalho.

3 Quantização Vetorial do Plano de Bits

Como foi visto na seção anterior, os pixels pertencentes a cada bloco da imagem são representados por 0's ou 1's, dependendo de serem maiores ou menores que X_{th} . Isto significa um gasto de $n \times n$ bits apenas para codificar o plano de bits, o que equivale ao comprometimento de 1 bit/pixel. É, portanto, de interesse investigar formas eficientes para codificação do plano de bits, de modo a reduzir este comprometimento já bastante elevado. Uma abordagem atraente

é a utilização de quantização vetorial [Udpikar-Raina (1987)].

O plano de bits é uma matriz composta apenas de 0's e 1's. É possível mostrar [Udpikar-Raina (1987)] que os elementos desta matriz binária ainda guardam parte da correlação existente nos pixels pertencentes ao bloco $n \times n$ original, ou seja, a correlação que existia entre pixels adjacentes do bloco original é parcialmente conservada na matriz binária obtida a partir deste. A consequência deste fato é a distribuição de forma não-uniforme das matrizes binárias de tamanho $n \times n$ no espaço binário N -dimensional, o que incentiva a utilização de quantização vetorial (QV) para codificar estas matrizes. O emprego da quantização vetorial possibilita usar menos de 1 bit/pixel para codificar o plano de bits.

Para quantizar as matrizes binárias de tamanho $n \times n$ com l bits ($l < n^2$) é necessário projetar um quantizador vetorial de 2^l níveis e dimensão $N = n^2$. Para obter as 2^l matrizes binárias mais representativas (o dicionário de códigos ótimo) utilizou-se uma técnica proposta em [Udpikar-Raina (1987)]. Esta técnica usa um procedimento iterativo para chegar a um quantizador vetorial ótimo baseado em uma longa seqüência de matrizes binárias. Aqui essas matrizes foram obtidas a partir de 5 imagens de treinamento: ZELDA, BARRA, JHOTELY, JGOLDY e IBOATSY. Como os elementos das matrizes são binários, a medida de distorção e o centróide precisam ser redefinidos. A distorção é definida pelo número de posições que duas matrizes diferem. Cada posição do centróide associado a uma determinada célula é ocupada pelo dígito 1 se naquela posição a maioria das matrizes que compõem a célula contém o dígito 1; caso contrário, a posição é ocupada pelo dígito 0.

A Tabela 1 mostra o desempenho obtido em função do número de bits l utilizado para a QV do plano de bits. Foram utilizadas 4 imagens de teste, distintas das imagens usadas no treinamento. O desempenho foi medido através da razão pico-ruído, definida como

$$PRSR = 10 \log_{10} \frac{255^2}{\frac{1}{L} \sum_{i=1}^L (S_i - S'_i)^2} \quad (5)$$

onde L representa o número total de pixels da imagem, e S_i e S'_i são os valores dos pixels nas imagens original e reconstruída, respectivamente.

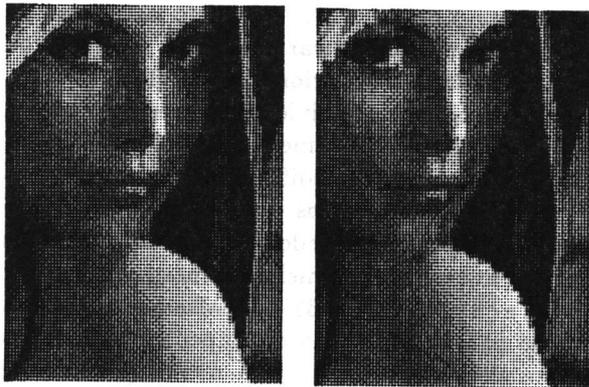
Para efeito de comparação são apresentados na terceira coluna da Tabela 1 os resultados obtidos sem QV do plano de bits, o que corresponde a $l = 16$.

Note-se que, com $l = 16, 10$ e 8 as taxas de bits associadas ao plano de bits são 1 bit/pixel; $0,625$ bit/pixel e $0,5$ bit/pixel, respectivamente.

Imagem	$l = 8$	$l = 10$	$l = 16$
LENA	27.79	28.66	30.06
LAKE	25.73	26.68	28.05
BABOON	23.52	24.30	26.97
BARBY	26.82	27.71	29.73

Tabela 1 : PRSR em função do número de bits l da QV do plano de bits.

Da Tabela 1 pode-se observar que a QV do plano de bits provocou uma queda de $2,3$ a $3,5$ dB para $l = 8$ e de $1,4$ a $2,7$ dB para $l = 10$. Os dois detalhes da imagem LENA exibidos na Figura 1 ilustram os efeitos visuais da QV do plano de bits. Pode-se notar que as regiões de detalhes são as que mais sofrem com a adoção da QV do plano de bits. Nessas regiões é natural esperar que os valores dos níveis de quantização Y_0 e Y_1 sejam bastante afastados, pois o desvio padrão destes blocos tende a ser alto. Assim, quando a matriz binária que representa um bloco de detalhes é quantizada, a troca de um elemento igual a 0 por outro igual a 1 , ou vice-versa, implica em um erro $\varepsilon = Y_1 - Y_0$ de valor alto. Já o contrário ocorre nos blocos da categoria restante, que apresentam uma certa homogeneidade. Nestes blocos o desvio padrão é mais baixo e, conseqüentemente, Y_0 e Y_1 estão mais próximos.



SEM QV

COM QV

Figura 1 : Detalhes da imagem LENA codificada, sem e com QV do plano de bits.

4 Quantização Vetorial Classificada

Conforme foi visto, a categoria composta por blocos de detalhes é a mais afetada pela quantização vetorial do plano de bits. Isto nos leva a buscar dar uma maior proteção a estes blocos, de grande importância para a qualidade da imagem reconstruída. Uma proposta simples a ser investigada nesta seção consiste em projetar um dicionário de mapas binários especial para esta categoria, reservando outro dicionário para os mapas binários oriundos de blocos pertencentes à categoria restante. Espera-se, com isso, diminuir a freqüência de trocas de 0 's (1 's) por 1 's (0 's) na quantização vetorial dos mapas binários associados a blocos de detalhes, aumentando conseqüentemente a qualidade global da imagem reconstruída.

Para aplicar esta estratégia é necessário buscar respostas razoáveis a duas questões: (1) Como distinguir de forma simples os blocos pertencentes à categoria de detalhes daqueles pertencentes à categoria restante? (2) Dado que desejamos codificar o plano de bits utilizando um dicionário de tamanho 2^l , qual a fração deste dicionário que deve ser dedicada à categoria detalhes?

Uma resposta à primeira pergunta foi sugerida em [Nasiopoulos *et al.* (1991)]. Naquele trabalho foi proposto o seguinte teste:

- Calcula-se a diferença D entre os pixels de maior e menor valor pertencentes ao bloco.
- Se $D > L$, sendo L um limiar de comparação, o bloco pertence à categoria detalhes; caso contrário, pertencerá a categoria restante.

A Figura 2 mostra a imagem LENA original e a imagem LENA representada da seguinte forma: os pixels pertencentes a blocos da imagem cuja diferença D entre os pixels de maior e menor valor supera um limiar $L = 60$ são representados em negro e os demais estão em branco. Da Figura 2 pode-se constatar que os blocos negros em geral indicam regiões de detalhes da imagem, enquanto que os brancos estão associados às regiões que apresentam maior homogeneidade.

Para determinar um bom valor para o limiar L procedemos da seguinte maneira: primeiramente o tamanho do dicionário utilizado para codificar o plano de bits foi fixado em 256 níveis. Esse número de níveis corresponde a uma taxa de $0,5$ bit/pixel. Dos 256 níveis foram atribuídos 128 ao dicionário associado à categoria detalhes e os outros 128 à categoria restante. Em seguida, o limiar foi variado entre 40 e 70 . Para cada um desses valores, foram geradas as seqüências de treinamento correspondentes a cada categoria e projetados os dois quantizadores vetoriais a elas associadas. O quantizador vetorial final é

o resultado da união dos quantizadores individuais. No processo de codificação é determinada a que categoria pertence o bloco a ser quantizado e buscado no QV correspondente qual a matriz binária mais próxima da original, ou seja, qual a palavra-código que melhor representa o plano de bits da entrada. O critério de busca consiste em minimizar a distância de Hamming entre o plano de bits da entrada e as matrizes correspondentes a cada nível do dicionário.



Figura 2 : LENA original e com os detalhes ressaltados.

O desempenho do codificador BTC usando QV classificada do plano de bits é apresentado na Tabela 2 para diversos valores de L e diferentes imagens de teste (não incluídas na seqüência de treinamento).

Comparando os resultados da Tabela 2 com aqueles mostrados na Tabela 1 vemos que a utilização da QV classificada, com divisão do dicionário em duas partes, é vantajosa para as imagens de teste LENA e LAKE qualquer que seja o valor de L. Para a imagem BABOON o desempenho é praticamente inalterado e para a imagem BARBY o desempenho é um pouco degradado para valores de L diferentes de 60. Para L = 60 o desempenho da imagem BARBY é praticamente inalterado com a adoção da nova estratégia. Tendo em vista esses resultados decidimos fixar o limiar L em 60.

Imagem	L = 40	L = 50	L = 60	L = 70
LENA	28.04	28.03	28.11	28.15
LAKE	26.21	26.25	26.17	26.22
BABOON	23.70	23.60	23.47	23.66
BARBY	26.37	26.57	26.71	26.53

Tabela 2 : PRSR do codificador BTC usando QV classificada do plano de bits (2 dicionários de 128 níveis).

Uma vez especificado o limiar L buscou-se determinar a melhor partição do dicionário de matrizes binárias associado ao quantizador vetorial do plano de bits. Dos 2^l níveis deste dicionário, X níveis serão dedicados à codificação dos planos de bits correspondentes aos blocos da categoria detalhes e Y níveis à codificação dos blocos oriundos da categoria restante. Obviamente $X + Y = 2^l$. Variando os valores de X e Y buscamos encontrar uma boa partição, ou seja, uma partição cujo dicionário final (formado pela união dos dicionários individuais correspondentes a cada categoria) ofereça o melhor desempenho na recuperação das imagens codificadas.

A Tabela 3 mostra os resultados obtidos para a imagem LENA assumindo $l = 8$ bits. Dentre as partições examinadas as melhores foram (X = 192 , Y = 64) e (X = 224 , Y = 32). Escolhemos a primeira partição para a análise do desempenho com outras imagens de teste. A Tabela 4 mostra a melhoria obtida com a adoção da QV classificada do plano de bits para várias imagens de teste. Nesta tabela é apresentado o desempenho obtido sem a utilização da QV do plano de bits, com a adoção desta técnica e, finalmente, com a aplicação da QV classificada.

X	Y	PRSR
32	224	27.01
64	196	27.72
128	128	28.11
192	64	28.33
224	32	28.33
240	16	28.20

Tabela 3 : Desempenho a 0,5 bit/pixel em função dos números de níveis X e Y dos dicionários associados à categoria detalhes e não-detalhes, respectivamente.

Imagem	Técnica de Quantização		
	sem quantização	QV	QV classificada
LENA	30.06	27.79	28.33
LAKE	28.05	25.73	26.41
BABOON	26.97	23.52	23.85
BARBY	29.73	26.82	26.91

Tabela 4 : PRSR em função da técnica de quantização do plano de bits a 0,5 bit/pixel.

As Tabelas 5 e 6 fornecem informações análogas às Tabelas 3 e 4 porém para $l = 10$ bits, ou seja, 0,625 bit/pixel. Neste caso escolhemos a partição ($X = 768$, $Y = 256$) para simulações com outras imagens de teste.

X	Y	PRSR
576	448	28.94
640	384	28.97
768	256	28.97
896	128	28.87
960	64	28.83

Tabela 5 : Desempenho a 0,625 bit/pixel em função dos números de níveis X e Y dos dicionários associados à categoria detalhes e não-detalhes, respectivamente.

Imagem	Técnica de Quantização		
	sem quantização	QV	QV classificada
LENA	30.06	28.66	28.97
LAKE	28.05	26.68	26.96
BABOON	26.97	24.30	24.71
BARBY	29.73	27.71	27.74

Tabela 6 : PRSR em função da técnica de quantização do plano de bits a 0,625 bit/pixel.

A partir dos resultados apresentados nas Tabelas 4 e 6, concluímos que a adoção da técnica QV classificada para a codificação do plano de bits é vantajosa. Em termos de desempenho objetivo, quando codificamos o plano de bits utilizando um dicionário de 256 níveis conseguimos uma melhoria de até 0,7 dB. Se o dicionário utilizado possui 1024 níveis, a melhoria é de até 0,4 dB.

A Figura 3 mostra a mesma área da imagem LENA exibida na Figura 1, porém desta vez aplicando a QV classificada com 0,5 bit/pixel. Esse exemplo ilustra como esta técnica consegue lidar melhor com os detalhes pertencentes à imagem, proporcionando, com isso, uma qualidade subjetiva global superior.



Figura 3 : Detalhe da imagem LENA codificada através da técnica QV classificada.

Além da vantagem no desempenho é necessário ressaltar a vantagem da redução da complexidade. É fato bastante conhecido que uma das maiores dificuldades da quantização vetorial é a busca da palavra-código mais adequada, especialmente quando os di-

cionários utilizados são grandes. Com base em medidas realizadas sobre a seqüência de treinamento verificou-se que a QV classificada consegue reduzir o esforço computacional em 68%.

Para justificar este número é só atentar para algumas características da seqüência de treinamento utilizada para a geração dos dicionários. Foram obtidos 110528 blocos de tamanho 4×4 a partir de duas imagens de 512 linhas e 512 colunas mais três imagens de 576 linhas e 720 colunas. Destes blocos, 16401 foram associados à categoria de detalhes a partir do teste empregado com $L = 60$, o restante sendo atribuído à categoria de não-detalhes. A fim de codificar o plano de bits destes blocos sem adotar a técnica QV classificada teríamos de fazer 110528×256 comparações utilizando um dicionário de 8 bits e 110528×1024 comparações utilizando um dicionário de 10 bits. Com a QV classificada estes números passam a ser $16401 \times 192 + 94127 \times 64$ (se o dicionário for composto por 64 palavras relativas à categoria de não-detalhes e 192 palavras para a categoria de detalhes) e $16401 \times 768 + 94127 \times 256$ (o que significa 768 palavras para a categoria de detalhes e 256 palavras para a categoria de não-detalhes). A razão entre o número de comparações exigidos com a utilização da QV classificada e o mesmo número ao se usar a QV direta expressa a redução da complexidade de busca. Tanto para o dicionário de 8 bits quanto para o dicionário de 10 bits essa razão é de 0,324. Isso representa uma redução de complexidade de aproximadamente 68%.

5 Conclusões

Este trabalho comparou o desempenho e a complexidade de dois esquemas empregados para a redução da taxa de bits em codificadores BTC. No primeiro, o plano de bits é codificado por um quantizador vetorial binário (elementos do vetor são 0's e 1's) a fim de explorar as correlações existentes entre elementos vizinhos. No segundo, a codificação também considera a importância dos planos de bits oriundos de blocos da imagem onde se observa a existência de detalhes. Para isso, cada bloco é classificado em detalhe ou não-detalhe, através de um algoritmo simples, e quantizado por um quantizador vetorial projetado especificamente para a categoria do bloco. Comparando os dois métodos concluiu-se que o segundo apresenta uma melhoria tanto no desempenho objetivo (de até 0,7 dB, dependendo da imagem de teste e do tamanho do dicionário utilizado) como na qualidade visual da imagem. Além disso, a QV classificada apresenta a vantagem adicional de reduzir sensivelmente a complexidade do algoritmo (em torno de 68%).

6 Referências

- E. J. Delp and O. R. Mitchell, Image compression using block truncation coding, *IEEE Transactions on Communications* **27** (1979) 1335–1342.
- M. D. Lema and O. R. Mitchell, Absolute moment block truncation coding and its application to colour images, *IEEE Transactions on Communications* **32** (1984) 1148–1157.
- P. Nasiopoulos, R. K. Ward and D. J. Morse, Adaptive compression coding, *IEEE Transactions on Communications* **39** (1991) 1245–1254.
- P. R. R. L. Nunes, A. Alcaim and M. R. L. F. da Silva, Compression of satellite images for remote sensing applications, *Proceedings of the XVII Congress of the International Society for Photogrammetry and Remote Sensing* Washington, D.C. (1992) 479–483.
- L. V. Oliveira e A. Alcaim, Sensibilidade dos parâmetros estatísticos em codificadores de imagem BTC, *Anais do 11o. Simpósio Brasileiro de Telecomunicações* Natal, RN (1993) (aceito para publicação).
- V. R. Udpikar and J. P. Raina, BTC image coding using vector quantization, *IEEE Transactions on Communications* **35** (1987) 352–356.