

Dithering aplicado à Transcodificação de Vídeo

Luiz Rômulo Mendes e Max H. M. Costa

Resumo—Este artigo apresenta resultados preliminares de um estudo da utilização de *dithering* no processo de transcodificação de vídeo digital para redução da taxa de bits. É apresentada uma técnica simples que possibilita, para algumas faixas de valores da razão entre passos de quantização, ganhos que alcançam a ordem de 2 dB em relação à transcodificação convencional, sem a utilização de *dithering*.

Palavras-Chave—Processamento Digital de Imagens, Vídeo Digital, *Dithering*, Transcodificação.

Abstract—This paper presents preliminary results of a study on the application of dithering in digital video transcoding. Basic methods of dithering are compared and a simple method is proposed that provides, in some cases, 2dB gain with respect to transcoding without dithering.

Index Terms—Digital Image Processing, Digital Video, Dithering, Video Transcoding.

I. INTRODUÇÃO

À medida que aumenta a diversidade de serviços multimídia, tipos de dispositivos e formatos de representação de conteúdo, a interoperabilidade entre diferentes redes e sistemas ganha crescente importância. Neste cenário de convergência digital, transcodificação de vídeo contribui para promover esta interoperabilidade. Por exemplo, possibilita que um sinal de vídeo digital codificado originalmente em formato HDTV (*High Definition Television*) possa ser convertido para taxas de bit e resoluções temporais e/ou espaciais reduzidas, de forma a atender especificações do canal de transmissão e do dispositivo de visualização utilizados. Desta forma, um sinal de vídeo de alta definição pode ser transcodificado para ser transmitido e reproduzido em uma TV com resolução padrão (SDTV – *Standard Digital Television*), um telefone celular ou terminal PDA (*Personal Digital Assistant*) com resoluções CIF (*Common Intermediate Format*) ou QCIF (*Quarter CIF*).

II. TRANSCODIFICAÇÃO DE VÍDEO

A transcodificação de vídeo pode ser definida como um processo de conversão de um sinal de vídeo previamente codificado para outro sinal de vídeo, tipicamente, com um dos seguintes objetivos:

- Conversão entre taxa de bits constante (CBR - *Constant Bit Rate*) e taxa de bits variável (VBR - *Variable Bit Rate*);

- Redução de resolução temporal e/ou espacial;
- Redução de sensibilidade a erros de canal;
- Redução da taxa de bits;
- Conversão de formatos.

A. Conversão entre CBR e VBR

A conversão entre taxa de bits constante (CBR) e taxa variável (VBR) tem como objetivo tornar mais eficiente o transporte de vídeo, especialmente em redes baseadas em multiplexação estatística, por exemplo, redes ATM (*Asynchronous Transfer Mode*).

B. Redução de resolução temporal e/ou espacial

Um dos propósitos da redução da resolução temporal e/ou espacial do sinal de vídeo é atender requisitos de redes e dispositivos móveis que possuem limitações de banda de recepção, tamanho de tela e capacidade de processamento. Outra importante aplicação é na conversão de vídeos de alta resolução (HDTV) para formato de resolução padrão (SDTV).

C. Redução de sensibilidade a erros de canal

Neste caso, a transcodificação tem como objetivo proteger o sinal contra erros de transmissão tornando-o mais robusto. Para isso, são exploradas opções disponíveis nas técnicas e padrões de codificação de fonte do sinal de vídeo. Um dos objetivos é reduzir a propagação de erros, que pode ocorrer quando se utilizam técnicas típicas de compressão de vídeo, por exemplo, codificação de Huffman e codificação diferencial (preditiva) de quadros. Exemplos de técnicas de redução de sensibilidade a erros são apresentadas em [2] e [3]. Entre outras, citam-se: o envio de vetores de deslocamento de macroblocos intracodificados; a redução de número de blocos por fatia (*slice*); e o aumento do número de macroblocos intracodificados em cada quadro.

Outra técnica utilizada é a conversão de um sinal de vídeo codificado originalmente em uma única camada (*single layer*) para um sinal de vídeo escalonado em duas ou mais camadas. Estas camadas podem ser protegidas de forma diferenciada de modo a privilegiar a camada de alta prioridade, essencial para a decodificação do vídeo [4]. Transcodificação objetivando redução de sensibilidade a erros tem aplicações principalmente em sistemas de comunicações móveis, cujos canais tipicamente possuem bandas menores e são mais suscetíveis a erros, quando comparados com sistemas que utilizam meios de transmissão guiados.



Fig. 1. Diagrama de blocos simplificado de transcodificador de malha aberta.

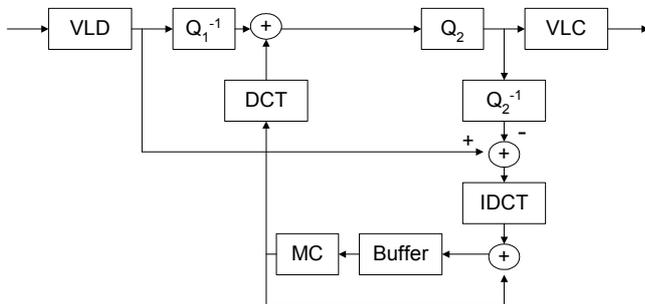


Fig.2 – Diagrama de blocos simplificado de transcodificador de malha fechada.

D. Redução da taxa de bits

A transcodificação pode ser também utilizada simplesmente para redução da taxa de bits, mantendo as resoluções temporal e espacial originais. A redução da taxa de bits tem por objetivo adequar o sinal à capacidade do canal ou do meio de armazenamento disponível. Uma forma direta de se obter redução de taxa é através da requantização dos coeficientes da Transformada Discreta do Cosseno (DCT). Uma técnica alternativa usa o descarte de coeficientes de alta frequência.

Dois arquiteturas básicas podem ser utilizadas no processo de transcodificação: o sistema de malha aberta, apresentado no diagrama da Fig. 1, e o sistema de malha fechada, apresentado na Fig. 2. No sistema de malha aberta, após a decodificação de entropia em um VLD (*Variable Length Decoder*), os coeficientes da DCT são recuperados em Q_1^{-1} , reconstrutor do quantizador original Q_1 , que utiliza passo de quantização q_1 . Em seguida são quantizados pelo quantizador Q_2 , utilizando um passo de quantização $q_2 > q_1$, e são novamente codificados por um VLC (*Variable Length Coder*), utilizando, por exemplo, codificação de corridas de zeros e de Huffman. Observa-se que os coeficientes da DCT originais são submetidos a uma concatenação em cascata de dois quantizadores: Q_1 e Q_2 . Este arranjo de duas operações não lineares em série pode reforçar o ruído de quantização observado no sinal reconstruído. O valor do passo de quantização q_2 é ajustado de forma a se obter a taxa de bits desejada.

Outra possibilidade para controle de taxa é o descarte de coeficientes de alta frequência. Apesar de sua simplicidade, tipicamente, esta técnica não é tão eficiente quanto a técnica de requantização. Observa-se que estas duas técnicas são semelhantes às opções de escalonamento SNR e partionamento de dados (*data partitioning*), previstos no padrão MPEG-2 [8].

A arquitetura de malha aberta tem como vantagem a simplicidade de implementação. Entretanto, os erros de quantização gerados nos quadros de referência I e P (intracodificados e preditos) são propagados para os

correspondentes quadros P e B (preditos e bidirecionalmente interpolados), caracterizando descasamento (*drift*) entre os preditores de quadros do codificador e do decodificador [5]. Para evitar este problema, pode-se utilizar o sistema de malha fechada apresentado no diagrama de blocos da Fig. 2. Neste sistema, os quadros são recuperados no domínio espacial e novas estimativas de vetores de deslocamento são realizadas. Para evitar o cálculo da DCT e da IDCT, e assim reduzir o número de operações, têm sido propostas arquiteturas que utilizam compensação de movimento no domínio da transformada DCT, baseadas em métodos apresentados em [7].

III. DITHERING APLICADO À REQUANTIZAÇÃO

Codificadores de vídeo baseados na transformada DCT, entre eles o MPEG-2 [8], utilizam um conjunto de técnicas que possibilitam comprimir o sinal digital original de forma eficiente. Parte destas técnicas envolve operações inversíveis, por exemplo, a própria DCT e a codificação de Huffman. Essencial para o objetivo de codificação eficiente do sinal é a operação de quantização dos coeficientes da transformada DCT, uma operação não linear na qual o sinal de entrada, que é contínuo no domínio da amplitude, é mapeado em valores discretos na saída. No caso de transcodificação, o sinal de entrada já está quantizado com certo passo, e é submetido a um novo quantizador de passo maior. A combinação em série de duas quantizações pode ter o efeito de reforçar o ruído de quantização, dependendo da razão entre os passos q_2/q_1 . O uso de *dithering* tende a atenuar este efeito.

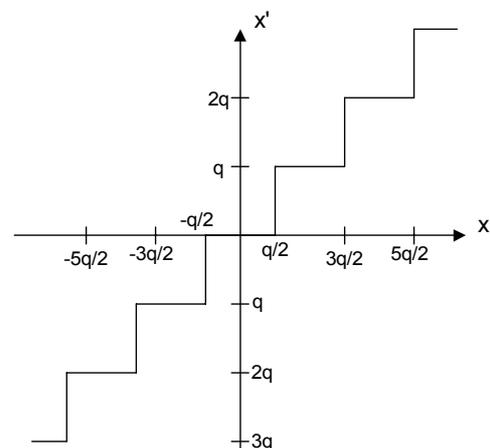


Fig. 3 – Função característica que relaciona entrada e saída de um quantizador uniforme.

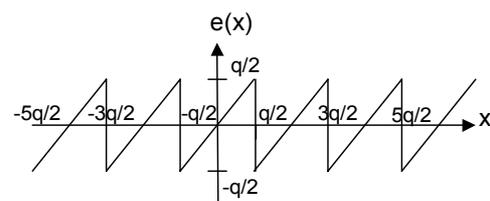


Fig. 4 – Erro de quantização em função do sinal de entrada do quantizador uniforme.

O gráfico da Fig. 3 apresenta a relação entre entrada e saída de um quantizador uniforme cujo passo de quantização é q . A diferença entre o sinal original e o sinal quantizado é o erro de quantização. A Fig. 4 apresenta o erro de quantização em função do sinal de entrada de um quantizador uniforme. Pode-se observar que há um relacionamento determinístico entre o sinal de entrada e o erro de quantização, que pode resultar em uma correlação significativa entre estes dois sinais, especialmente no caso de quantização grosseira, quando o passo de quantização é relativamente grande em comparação à variância do sinal de entrada.

Apesar da quantização ser uma operação não linear, é possível obter uma relação de linearidade entre as estatísticas do sinal de entrada e do sinal da saída [9][10], onde a função de densidade de probabilidade da saída do quantizador é obtida por um processo de amostragem linear da função de densidade de probabilidade da entrada. Esta relação entre as estatísticas também pode ser obtida através da adição de ruído branco uniforme (*dither*) ao sinal original antes da quantização. A função densidade de probabilidade deste ruído é apresentada no gráfico da Fig. 5. Esta técnica, conhecida

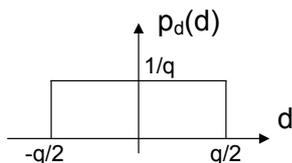


Fig. 5 – Função densidade de probabilidade do ruído uniforme d .

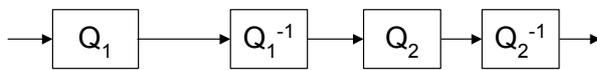


Fig. 6 – Diagrama de blocos simplificado de transcodificador que não utiliza dithering.

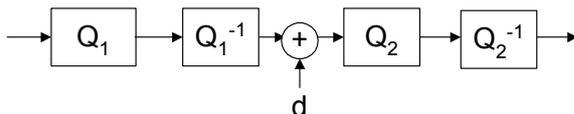


Fig. 7 – Diagrama de blocos simplificado de transcodificador com dither aditivo.

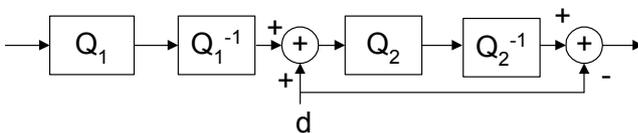


Fig. 8 – Diagrama de blocos simplificado de transcodificador com dither subtrativo.

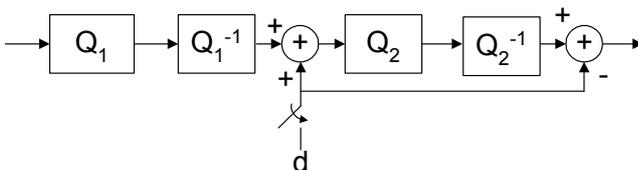


Fig. 9 – Diagrama de blocos simplificado do transcodificador com dither subtrativo aplicado a coeficientes não nulos.

como *dithering*, pode ser utilizada para decorrelacionar o ruído de quantização do sinal. A propriedade de decorrelação entre ruído e sinal torna a distorção provocada pela quantização subjetivamente mais tolerável. Ademais, pode-se aplicar técnicas de redução de ruído, por exemplo, através de filtragem de Wiener e de Kalman [11], para atenuar os efeitos da quantização.

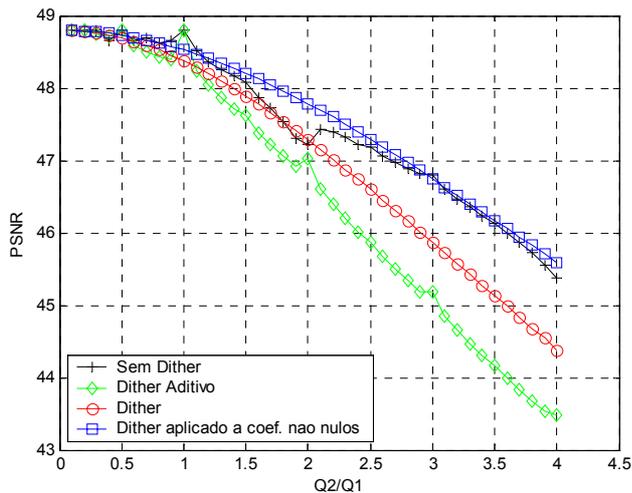
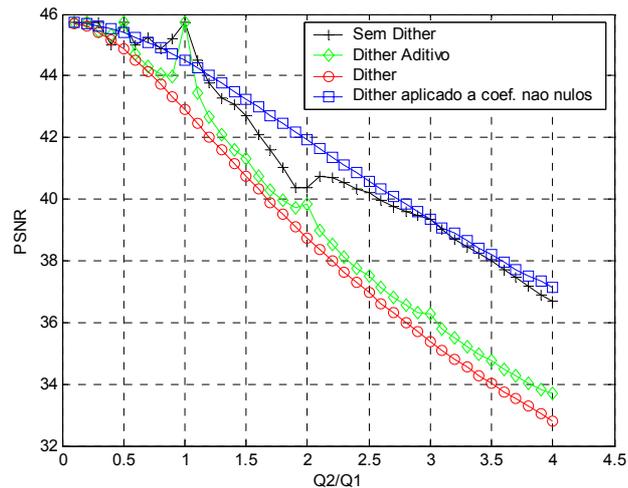
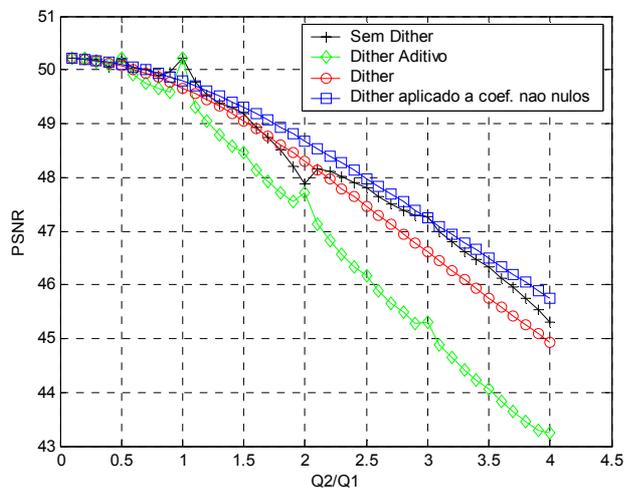
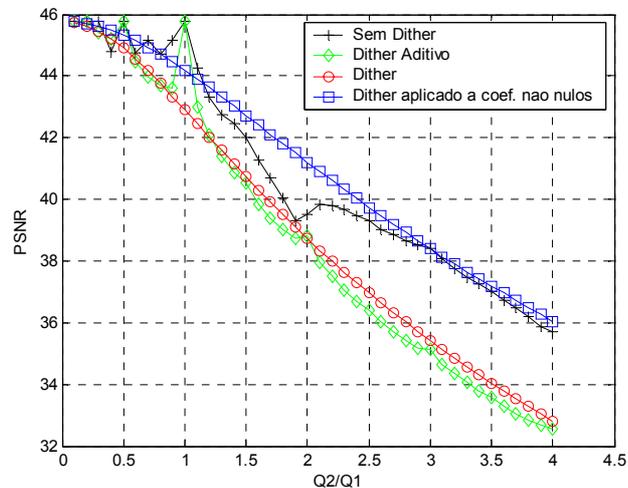
O *dither* é também utilizado para conversão de imagens multinível (*continuous tone*) em imagens binárias (técnicas de *digital halftoning*), e tem tido um papel fundamental na redução de preços e na expansão do mercado de impressoras a jato de tinta [12]. Outra aplicação importante é na área de instrumentação e controle, onde o *dither* é utilizado para melhorar a precisão de instrumentos de medida, e de conversores A/D [13][14].

IV. SIMULAÇÕES

Foram realizadas simulações para comparar técnicas de transcodificação utilizando dithering. Os resultados apresentados neste artigo foram obtidos através da transcodificação da componente de luminância de quadros intracodificados (quadros I), de dimensões 720x480, das seqüências *Mobile* e *Flower Garden*. A utilização exclusiva de quadros intracodificados possibilita analisar o efeito da concatenação em cascata de dois quantizadores de forma independente dos efeitos relacionados à codificação preditiva de quadros. As seguintes técnicas foram comparadas:

- Transcodificação sem dithering, apresentada no diagrama de blocos da Fig. 6.
- Transcodificação com dither não-subtrativo (identificada nos gráficos como dither aditivo). Nesta técnica, simplesmente adiciona-se ruído branco uniforme ao sinal antes da quantização Q_2 , como apresentado na Fig. 7.
- Transcodificação com dither subtrativo (identificada nas legendas dos gráficos simplesmente como *dither*). Nesta técnica, o mesmo ruído de *dither* (d) que é adicionado ao sinal antes da etapa de quantização Q_2 , é também subtraído após a recuperação dos coeficientes (Q_2^{-1}), como ilustrado na Fig. 8.
- Transcodificação com dither subtrativo aplicado somente aos coeficientes não nulos. Esta é uma variação da técnica anterior que tem como objetivo reduzir o ruído de quantização em torno da origem. O diagrama de blocos da Fig. 9 ilustra como opera esta técnica. A chave que conecta o ruído de *dither* (d) ao sistema só se fecha quando o coeficiente na entrada do somador é não nulo.

Os gráficos das Figs. 10 a 15 apresentam a variação da relação sinal ruído de pico (PSNR), em função das razões entre os passos de quantização q_2 e q_1 (q_2/q_1), utilizados nos quantizadores Q_2 e Q_1 , respectivamente. O passo de quantização q_1 , utilizado no processo original de codificação do sinal, apresenta um valor fixo para cada gráfico. Diferentes valores do passo de quantização q_2 são utilizados para construir as várias curvas. A redução da taxa de bits através da transcodificação é maior quanto maior for o valor de q_2 . Observa-se que os resultados apresentados nos gráficos não consideram a codificação de entropia. Vale também ressaltar que valores de q_2/q_1 inferiores a unidade não são de

Fig. 10 – Sequência Flower Garden, com $q_1=2$.Fig. 12 – Sequência Flower Garden, com $q_1=8$.Fig. 11 – Sequência Mobile, com $q_1=2$.Fig. 13 – Sequência Mobile, com $q_1=8$.

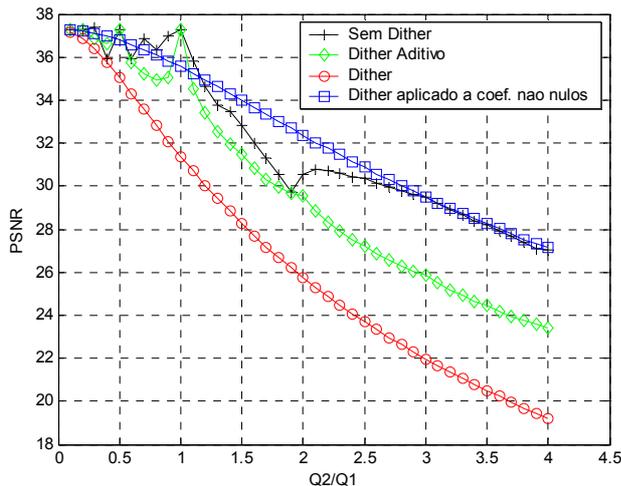
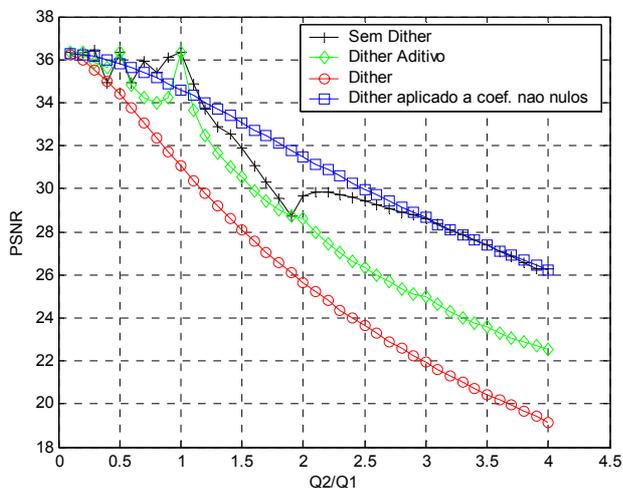
interesse em implementações práticas, já que o objetivo é a redução da taxa. Entretanto, esta faixa de valores foi considerada para ilustrar o comportamento não monotônico das curvas nos casos “sem dither” e “dither aditivo”, e também a descontinuidade das curvas de PSNR nas proximidades das razões entre passos dadas por $q_2/q_1=1$ e $q_2/q_1=0,5$.

As Figs. 12 e 13 apresentam gráficos comparativos para o caso em que $q_1=8$, enquanto que as Figs. 14 e 15 apresentam resultados para $q_1=32$. Observa-se que, em todos os gráficos, a utilização de dither subtrativo aplicado a coeficientes não nulos é superior às outras técnicas para a faixa de valores de q_2/q_1 entre 1 e 2,5. Para valores de q_2/q_1 entre 2,5 e 4, o dither aplicado a coeficientes não nulos apresentou desempenho não inferior às outras técnicas. A técnica proposta alcança ganhos da ordem de 2 dB, para $q_2/q_1=2$ em relação a requantização sem *dither*.

Os gráficos das Figs. 14 e 15 são de resultados obtidos a partir de um sinal que foi originalmente quantizado com um passo de quantização maior ($q_1=32$) quando comparado com os outros gráficos. Neste caso, é possível observar que a utilização de *dither* subtrativo apresenta desempenho

significativamente inferior até mesmo quando comparada ao esquema com *dither* aditivo. Isto ocorre porque a variância do *dither* torna-se proporcionalmente maior quando comparada à variância do sinal. Além disso, quando *dither* aditivo é utilizado na transcodificação, os coeficientes que são zerados em Q_1 não alteram seu valor quando são recuperados após terem sido requantizados em Q_2 , ou seja, os coeficientes recuperados são nulos. Já no caso do *dither* subtrativo, os coeficientes nulos são recuperados no receptor com valores da magnitude iguais aos do ruído associado ao *dither*, pois o mesmo é subtraído no processo de reconstrução dos coeficientes. Isto sugere que pode-se ter um ganho de $PSNR_q$ restringindo-se a aplicação do dither somente aos coeficientes não nulos. Como apresentado anteriormente, os resultados gráficos apresentados neste artigo confirmam esta análise.

Os gráficos apresentados nas Figs. 10 e 11 foram obtidos utilizando-se um passo de quantização muito refinado, $q_1=2$, que dificilmente é utilizado na prática. Novamente o *dither* aplicado aos coeficientes não nulos apresentou resultado superior às outras técnicas na faixa de $1 < q_2/q_1 < 2,5$. Observa-se que, neste caso, o *dither* subtrativo apresentou resultado superior ao não subtrativo (aditivo).

Fig. 14 – Seqüência Flower Garden, com $q_1=32$.Fig. 15 – Seqüência Mobile, com $q_1=32$.

O desempenho dos métodos de transcodificação com *dithering* pode ser melhorado com a utilização de técnicas de redução de ruído, já que com a aplicação de *dither* uniforme, o ruído de quantização torna-se descorrelacionado do sinal. Outra possibilidade que está sendo investigada é a utilização de ruídos coloridos, mais especificamente *blue noise* e *green noise* que possuem espectros com energias concentradas em frequências altas e em frequências intermediárias, respectivamente. *Blue noise* e *green noise* têm sido utilizados com sucesso em técnicas de *halfoning*, por exemplo, com aplicações em impressoras a jato de tinta [12].

V. CONCLUSÕES

Dithering é uma técnica que aplicada à quantização torna o erro de quantização uniformemente distribuído e independente do sinal. Um erro de quantização independente do sinal torna a distorção da quantização mais tolerável e facilita a aplicação de técnicas de redução de ruído, por exemplo, filtros de Wiener e Kalman [11] que estão sendo exploradas na continuidade deste trabalho.

A técnica de *dithering* subtrativo aplicado somente a coeficientes não nulos apresentada neste artigo pode melhorar o desempenho da transcodificação de quadros intracodificados em até 2dB, em relação aos outros métodos.

REFERÊNCIAS

- [1] A. Vetro, C. Christopoulos, and H. Sun, "Video Transcoding Architectures and Techniques: An Overview," *IEEE Sig. Proc. Mag.*, vol. 20, no. 2, pp. 18-29, March 2003.
- [2] G. de los Reyes, A. R. Reibman, S. F. Chang, and J. C. I. Chuang, "Error-resilience transcoding for video over wireless channels," *IEEE J. Select Areas Commun.*, vol. 18, pp. 1063-1074, June 2000.
- [3] S. Dogan, A. Cellatoglu, M. Uyguroglu, A. H. Sadka, and A. M. Kondoz, "Error-resilient video transcoding for robust inter-network communications using GPRS," *IEEE Trans. Circuits Syst. Video Technol.* vol. 12, pp. 453-464, June 2002.
- [4] L. R. Mendes, and Max H. M. Costa, "Two-scale Transmission of DCT-Coded Video over Lossy Packet Networks," ITS '98 Proceedings. SBT/IEEE International, Vol.: 2, pp. 510-515, Aug. 1998
- [5] P. A. A. Assunção, and M. Ghanbari, "Transcoding of single-layer MPEG video into lower rates," *IEE Proc. - Vis. Image Sig. Process.*, vol. 144, no. 6, Dec. 1997.
- [6] O. Werner, "Requantization for Transcoding of MPEG-2 Intraframes," *IEEE Trans. Imag. Proc.*, vol 8, no2, Feb. 1999.
- [7] S. F. Chang, and D. G. Messerschmidt, "Manipulation and Compositing of MC-DCT Compressed Video," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1-11, Jan. 1995.
- [8] ITU-T Recommendation H.262, ISO/IEC 13818-2, "Information Technology – Generic coding for moving pictures and associated audio information: Video," Aug. 1995.
- [9] B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory," *IRE Trans. Circuit Theory*, pp. 266-276, Dec. 1956.
- [10] B. Widrow, I. Lökkár, and M. C. Liu, "Statistical Theory of Quantization," *IEEE Trans. Inst. Meas.*, vol.45, no.2, April 1996.
- [11] S. M. Bozic, *Digital and Kalman Filtering*, Edward Arnold, 1979.
- [12] D. L. Lau, and G. R. Arce, *Modern Digital Half-toning*. New York: Marcel Dekker, 2001.
- [13] P. Carbone, and D. Petri, "Effect of Additive Dither on the Resolution of Ideal Quantizers," *IEEE Trans. Instrum. Meas.*, vol. 43, no. 3, pp. 389-396, June 1994.
- [14] P. Carbone, C. Narduzzi, and D. Petri, "Dither Signal Effects on the Resolution of Nonlinear Quantizers," *IEEE Trans. Instrum. Meas.*, vol. 43, no. 2, pp. 139-145, April 1994.