

FILTRAGEM B-SPLINE APLICADA A UMA ESTRUTURA DE CODIFICAÇÃO DE VÍDEO DE DOIS ESTÁGIOS ESCALÁVEL ESPACIALMENTE PARA OPERAÇÃO EM BAIXA TAXA DE BITS

EVÂNIO RAMOS NICOLEIT E RUI SEARA

LINSE: Circuitos e Processamento de Sinais

Departamento de Engenharia Elétrica

Universidade Federal de Santa Catarina

Campus Universitário, 88040-900 – Florianópolis – SC – Brasil

Tel: (0xx48) 331-9504, Fax: (0xx48) 331-9091, evanio@linse.ufsc.br

ABSTRACT

This paper proposes the use of B-splines in a spatially scalable two-stage video coding structure for low bit-rate operation. Initially, the original frame sequence are B-spline low-pass filtered and downsampled. Motion estimation and compensation are performed on that downsampled sequence for reducing its time redundancy, with a consequent reduction on the computational complexity. At a second stage, a spatial refinement step is applied to the upsampled images from the first stage, obtained from B-spline filtering. Computer simulation results are shown, comparing the proposed methodology with the H.263-based structure for rate-distortion performance. Those results demonstrate that the new approach presents a significant potential for spatially scalable video coding.

RESUMO

Este artigo propõe o uso de *B-splines* em uma estrutura de codificação de vídeo de dois estágios escalável espacialmente para operação em baixas taxas de bits. Inicialmente, a seqüência original de quadros é filtrada por uma função passa-baixas, tipo *B-spline*, seguida de uma operação de subamostragem. Estimação e compensação de movimentos são realizadas nessa seqüência subamostrada para reduzir sua redundância temporal, com uma consequente redução na complexidade computacional. No estágio seguinte, uma técnica de refinamento espacial é aplicada às imagens interpoladas oriundas do primeiro estágio, obtidas a partir de filtragem via *B-spline*. Resultados de simulações são apresentados, comparando o desempenho taxa-qualidade da metodologia proposta com o de uma estrutura baseada na Recomendação H.263. Esses resultados demonstram que a nova abordagem apresenta um significativo potencial para codificação de vídeo escalável espacialmente.

1. INTRODUÇÃO

As técnicas de compressão de sinais de vídeo para baixas taxas de bits têm experimentado nos últimos anos um avanço significativo, impulsionado, principalmente, pela grande demanda do mercado de vídeo e aplicações na *internet*, destacando-se os sistemas de videofonia e videoconferência. Essa compressão faz-se imperativa para o armazenamento e/ou transmissão de vídeos digitais, devido ao grande volume de dados provenientes desses sinais. A codificação de vídeo é

usualmente realizada segundo um esquema híbrido implementado utilizando dois procedimentos: redução de redundância temporal e, espacial. Para a redução da redundância temporal, a técnica de estimação/compensação de movimentos (EM/CM) é freqüentemente aplicada. A redução de redundância espacial nos quadros com movimentos compensados é geralmente efetuada utilizando-se um método baseado em transformação. A grande maioria dos codificadores de vídeo padronizados [1,2] é baseada na transformada cosseno discreta (DCT - *discrete cosine transform*). Contudo os codificadores baseados em DCT introduzem alguns inconvenientes [3], tais como: surgimento de artefatos de blocos e efeitos de recobrimento em altas taxas de compressão. Por outro lado, esquemas baseados em multi-resolução espacial têm recebido uma considerável atenção no contexto de compressão de sinais de vídeo. Isso deve-se principalmente às seguintes razões:

- i) o decodificador pode reconstruir quadros de menor resolução, usando um subconjunto do *bitstream* completo, sendo útil em aplicações nas quais seja possível operação com taxa de bits variável no canal;
- ii) boas aproximações via interpolação *B-spline* podem ser alcançadas a partir de subimagens;
- iii) a propriedade de escalabilidade espacial pode ser facilmente obtida;
- iv) redução significativa do efeito de recobrimento, obtida pela escolha adequada de filtros de decomposição e reconstrução.

A representação em diferentes escalas é um caminho interessante quando aplicada à codificação de vídeo. Em uma primeira etapa, o princípio básico por trás dessa filosofia consiste na construção de uma representação grosseira da imagem que está sendo codificada. Um dos possíveis caminhos para melhorar essa imagem pode ser através de uma filtragem baseada em funções do tipo *B-splines*. O aprimoramento obtido dessa forma deve-se às restrições de continuidade impostas para a *spline* de ordem n e suas derivadas até ordem $n-1$, nos pontos de amostragem, gerando, assim, um alto grau de concordância entre amostras no processo de reconstituição da imagem.

Este artigo propõe o uso de *B-splines* em uma estrutura de codificação de vídeo de dois estágios escalável espacialmente para operação em baixas taxas de bits. Nessa estrutura, o primeiro estágio codifica a seqüência de imagens em baixa resolução espacial, com consequentes taxa de bits parcial e complexidade computacional reduzidas. O segundo estágio

detecta as perdas de informação provenientes da codificação do primeiro estágio, aplicando um procedimento de refinamento espacial sobre uma seqüência de imagens com resolução idêntica à original. O desempenho do esquema para compressão de vídeo usando filtragem baseada em funções *B-splines* é competitivo em taxa-qualidade com o esquema de codificação baseado na Recomendação H.263 [2]. Resultados de simulações mostram que a metodologia de codificação usando *B-splines* leva a um sinal de vídeo codificado com alta qualidade para operação em baixas taxas de bits.

2. ESTRUTURA DE CODIFICAÇÃO

A estrutura de codificação utilizada é mostrada na Fig. 1. Essa estrutura consiste de dois estágios. O primeiro opera em uma baixa resolução espacial e utiliza o mesmo princípio de codificação do esquema baseado na Recomendação H.263. O segundo realiza um refinamento espacial que detecta a perda de informação (filtragem passa-baixas, subamostragem e interpolação, e quantização e codificação dos coeficientes) do primeiro estágio [4,5].

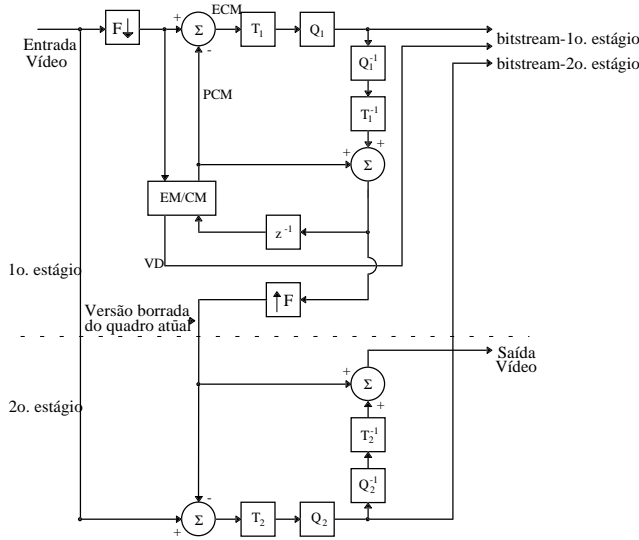


Fig. 1 - Codificador de vídeo de dois estágios escalável espacialmente.

O princípio de codificação desta estrutura é descrito a seguir. A primeira etapa é composta de uma filtragem linear passa-baixas do quadro atual seguida por uma subamostragem da imagem (em cada dimensão por um fator de dois) implementadas pelo operador $[F\downarrow]$. A subamostragem no primeiro estágio permite que essa estrutura seja espacialmente escalável [6]. A predição do quadro atual subamostrado é obtida usando estimação/compensação de movimentos (EM/CM) em uma resolução espacial reduzida, conduzindo a uma maior eficiência computacional na codificação. No processo de codificação, após a obtenção da versão subamostrada codificada da imagem, estima-se uma predição do quadro atual na escala original da imagem. Isso é realizado através de um processo de expansão da imagem (interpolação e filtragem subsequente) por um fator 2, efetuado pelo operador $[\uparrow F]$ em cada uma das dimensões da imagem [6]. Essa predição é uma versão (grosseira) passa-baixas

da imagem na resolução original. Ela é geralmente afetada pelo efeito de recobrimento, uma vez que é obtida pela versão expandida da imagem passa-baixas subamostrada [7]. A imagem erro do segundo estágio é processada por uma transformação T_2 , reduzindo assim a redundância espacial, e quantizada pelo bloco Q_2 . O bloco Q_2^{-1} reconstrói os coeficientes quantizados e T_2^{-1} realiza a transformação inversa desses coeficientes. A imagem erro reconstruída é então somada ao quadro atual gerado pelo primeiro estágio, resultando no quadro codificado atual, que é então reconstruído da mesma forma no decodificador.

Essa estrutura é espacialmente escalável. O primeiro estágio é independente com relação ao *bitstream* do segundo estágio. Para a estrutura de codificação da Fig. 1, a perda parcial ou total de informação no segundo estágio não degrada o desempenho de qualidade para os próximos quadros codificados.

3. B-SPLINES

Splines são funções constituídas por seções de polinômios que são conectadas entre si através de transições suaves. Para uma *spline* de grau n , cada segmento é um polinômio de grau não maior do que n . A transição suave é alcançada impondo-se restrições de continuidade da função e suas derivadas até a ordem $(n-1)$ nos pontos de concatenação. No contexto deste trabalho, somente estão sendo consideradas *splines* para amostras uniformemente espaçadas. As *B-splines* (*basis-splines*) podem redefinir uma *spline* como combinação linear de funções de base [8], cada uma apresentando um pequeno suporte (intervalo em que a função é diferente de zero). Assim, uma *spline* pode ser caracterizada através de uma combinação linear de *B-splines* da forma

$$s(x) = \sum_{k \in \mathbf{Z}} c(k) \beta^n(x-k) \quad , x \in \mathbf{R} . \quad (1)$$

A representação (1) envolve deslocamentos inteiros da *B-spline* centrada, de grau n , denotada por $\beta^n(x)$. Os parâmetros $c(k)$ da representação são os coeficientes *B-spline*. Uma forma eficiente de atender a tais requisitos é construir as *B-splines* a partir de n operações de convolução entre $n+1$ pulsos retangulares, $\beta^0(x)$ [9]:

$$\beta^0(x) = \text{rect}(x) = \begin{cases} 1, & \text{for } |x| \leq 1/2 \\ 0, & \text{for } |x| > 1/2 \end{cases} \quad (2)$$

e

$$\beta^n(x) = \beta^0(x) * \beta^{n-1}(x) = \underbrace{\beta^0(x) * \beta^0(x) * \dots * \beta^0(x)}_{(n+1) \text{ termos}} , \quad (3)$$

resultando em funções simétricas e bem comportadas.

4. FILTRAGEM BASEADA EM B-SPLINES

Dadas as amostras de um sinal discreto $s(k)$, busca-se determinar os coeficientes $c(k)$ de sua representação, Eq. (1). Esse problema pode ser descrito através de uma abordagem via filtragem digital. Para tal, necessita-se introduzir o operador

B-spline discreto $b_m^n(k)$, obtido por amostragem da função *B-spline* de grau n expandida por um fator m :

$$b_m^n(k) = \beta^n \left(\frac{x}{m} \right) \Big|_{x=k}, \quad k \in \mathbf{Z}. \quad (4)$$

O sinal $s(k)$ pode ser agora representado, para $m = 1$, por

$$s(k) = \sum_{l \in \mathbf{Z}} c(l) b_1^n(k-l) = b_1^n(k) * c(k). \quad (5)$$

Definindo-se o operador de convolução inverso por $(b_1^n)^{-1}(k)$, os coeficientes de $s(k)$ podem ser determinados por

$$c(k) = (b_1^n)^{-1}(k) * s(k). \quad (6)$$

Na representação de sinais em diferentes níveis de resolução, normalmente, são utilizados fatores de escala potências de dois (escalas diádicas). Isso implica que os subespaços *splines*, \mathbf{S}_m^n , com $m = 2^i$, são aninhados:

$$\mathbf{S}_1^n \supset \mathbf{S}_2^n \supset \dots \supset \mathbf{S}_{2^i}^n \dots$$

Seja definida por $P_{2^i} s = s_i$ uma aproximação de mínimo erro (MSE – *mean square error*) do sinal $s(k) \in \mathbf{L}^2$ [9], com \mathbf{L}^2 denotando o espaço de funções com energia finita na escala $m = 2^i$. Assim, $P_{2^i} s$ pode ser escrita por

$$P_{2^i} s = \sum_{k \in \mathbf{Z}} c_{2^i}(k) \varphi \left(\frac{x}{2^i} - k \right), \quad (7)$$

onde os termos $\varphi(x/2^i - k)$ são funções de base (*B-spline* ou outras) na escala $m = 2^i$. Assim, os coeficientes $c_{2^i}(k)$ são obtidos pelo seguinte produto interno

$$c_{2^i}(k) = \frac{1}{2^i} \left\langle s(k), \overset{\circ}{\varphi} \left(\frac{x}{2^i} - k \right) \right\rangle, \quad (8)$$

onde $\overset{\circ}{\varphi}(x) \in \mathbf{S}_1^n$ é a função dual de $\varphi(x)$, no sentido de que

$$\left\langle \overset{\circ}{\varphi}(x-k), \varphi(x-l) \right\rangle = \delta(k-l).$$

Os coeficientes $c_{2^i}(k)$ podem ser determinados iterativamente, usando-se uma combinação de operações de filtragem discreta e subamostragem. Assim, pode-se obter $P_{2^i} s = s_i$ via filtragem de s_{i-1} , onde s_i é uma versão mais grosseira de s_{i-1} . Portanto os coeficientes $c_{2^i}(k)$, para a conjectura $P_{2^i} s = s_i \equiv P_{2^i} s_{i-1} = s_{i-1}$, podem agora ser obtidos por

$$c_{2^i}(k) = \frac{1}{2^i} \left\langle \sum_{l \in \mathbf{Z}} c_{2^{i-1}}(l) \varphi \left(\frac{x}{2^{i-1}} - l \right), \overset{\circ}{\varphi} \left(\frac{x}{2^i} - k \right) \right\rangle. \quad (9)$$

Considerando-se escalas diádicas, pode-se então definir o seguinte operador de filtragem com resposta ao impulso $\overset{\circ}{h}(k)$ por:

$$\overset{\circ}{h}(k) = \frac{1}{2^i} \left\langle \varphi \left(\frac{x}{2^{i-1}} + k \right), \overset{\circ}{\varphi} \left(\frac{x}{2^i} \right) \right\rangle = \left\langle \varphi(x+k), \overset{\circ}{\varphi} \left(\frac{x}{2} \right) \right\rangle \quad (10)$$

Agora, os coeficientes $c_{2^i}(k)$ podem ser obtidos filtrando

$c_{2^{i-1}}(k)$ por $\overset{\circ}{h}(k)$, assim

$$c_{2^i}(k) = \sum_{l \in \mathbf{Z}} \overset{\circ}{h}(l) c_{2^{i-1}}(2k-l), \quad (11)$$

para $c_1(k) = s(k)$.

Filtros baseados em *B-spline* resultam em um bom compromisso entre localização espacial e recobrimento, visto que eles contemplam tanto o filtro de Haar quanto o filtro de ordem zero, como também um interpolador do tipo $\text{sinc}(\cdot)$ de ordem infinita, tendendo a um filtro passa-baixas ideal [6,9]. Usualmente, o núcleo gerador $h(k)$ é simétrico e centrado na origem.

A escolha dos operadores $[F \downarrow]$ e $[\uparrow F]$ é fundamental para a obtenção de uma predição de qualidade razoável do quadro atual, como saída do primeiro estágio. Uma restrição que pode ser imposta para esses operadores é que a cascata $[F \downarrow]$ e $[\uparrow F]$ resulte no operador identidade [6]. Essa característica assegura que efeitos de filtragem não distorçam a predição grosseira. A interpolação da versão codificada subamostrada da imagem deve aproximar-se ao máximo (segundo um definido critério) do quadro original atual. Um critério prático de implementação, adotado neste trabalho, é que os operadores $[F \downarrow]$ e $[\uparrow F]$ sejam ótimos no sentido do erro quadrático mínimo [10,11].

O correspondente modelo bidimensional *B-spline* é separável. Dessa forma, a filtragem pode ser implementada através de sucessivas filtrações unidimensionais com operações de subamostragem e interpolação das linhas e colunas da imagem.

5. ASPECTOS DE IMPLEMENTAÇÃO E RESULTADOS DE SIMULAÇÕES

Os resultados descritos nesta seção são baseados na seqüência padrão QCIF *Miss America*, com 176×144 pixels, 256 níveis de cinza e amostrada a 30 quadros/s.

Algumas definições relativas à implementação da metodologia de codificação proposta se fazem necessárias e são apresentadas a seguir.

Dentro da família das *splines* polinomiais, as *splines* cúbicas são, sem dúvida, as mais populares e utilizadas [8,9]. Neste trabalho, são considerados filtros *B-spline* de terceira ordem, com resposta ao impulso centrada e simétrica. A escolha dos filtros dos operadores $[F \downarrow]$ e $[\uparrow F]$ é feita segundo o critério do erro quadrático mínimo [10,11].

Para a implementação da etapa de EM/CM, as imagens são divididas em blocos. O algoritmo BMA (*block-matching algorithm*) [12] é usado para estimar o movimento de cada bloco dentro do quadro anterior para a sua nova posição dentro do quadro atual. O tamanho de bloco utilizado para as seções de EM/CM é 4×4 pixels. Nesse caso, esta escolha tem representado um bom compromisso entre a taxa de bits requerida e a qualidade da estimação para as imagens subamostradas. O resultado dessa operação, para cada quadro, é um conjunto de 11×9 vetores de deslocamento (VD). A seção CM aplica os VD ao quadro anteriormente codificado para a obtenção de uma imagem de

predição com compensação de movimentos (PCM). A janela de busca utilizada no BMA é de ± 7 pixels em ambas as direções (horizontal e vertical). O critério de busca utilizado para encontrar a melhor representação é o MAD (*mean absolute difference*) [12]. O campo de VD é codificado adotando-se os mesmos códigos utilizados pela Recomendação H.263. Para uma representação eficiente das imagens erro com compensação de movimentos (ECM), estas imagens são transformadas através de decomposição *wavelet* (DWT – *discrete wavelet transform*). A DWT utiliza o banco de filtros Antonini-Daubechies *7/9-tap* [13,14] com três níveis de decomposição para o componente de luminância. Para alcançar um desempenho taxa-qualidade competitivo, a codificação dos coeficientes da DWT é realizada usando o algoritmo EZW (*embedded zerotree wavelet*) [15].

No segundo estágio, as imagens diferença são transformadas através de decomposição DWT em quatro níveis para o sinal luminância, empregando um banco de filtros idêntico ao usado no primeiro estágio. A codificação dos coeficientes da DWT é realizada usando novamente o algoritmo EZW.

A alteração da taxa de bits é obtida pela variação do número de passadas dominante/subordinada através do algoritmo EZW em ambos os estágios, dando-se maior ênfase ao primeiro estágio.

No conjunto de simulações, somente predição do tipo P é utilizada para o modo INTER [2]. Contudo outros modos de predição de quadro podem ser facilmente incorporados à estrutura proposta para efeito de comparação com a Recomendação H.263.

Para o codificador baseado na Recomendação H.263, foi utilizada a implementação TMN-5 [2,16].

A figura de mérito objetiva usada para medir a qualidade de reconstrução dos quadros codificados é a razão sinal ruído de pico (PSNR), definida por

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{255^2}{\text{EQM}} \right), \quad (12)$$

com o erro quadrático médio (EQM), considerando R quadros, definido por

$$\text{EQM} = \frac{1}{M \cdot N \cdot R} \sum_{r=0}^{R-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [I_r(m,n) - \hat{I}_r(m,n)]^2, \quad (13)$$

onde $I_r(m,n)$ representa o nível correspondente ao quadro r original e $\hat{I}_r(m,n)$, o nível correspondente ao quadro r reconstruído, ambos com dimensões $M \times N$ pixels.

Uma vez que se está interessado nas características em regime permanente de codificação, operando sobre uma estrutura preditiva em laço fechado, o primeiro quadro da seqüência não é codificado. Portanto, para o primeiro quadro, a taxa de bits e a PSNR não são incluídas nas medidas taxa-qualidade.

O desempenho de taxa-qualidade para os dois codificadores é apresentado na Fig. 2. Os resultados de PSNR são mostrados para a estrutura da Fig. 1, utilizando a metodologia proposta e a estrutura baseada na Recomendação H.263. Através dos resultados apresentados na Fig. 2, pode ser inferido que o esquema baseado nos operadores $[F\downarrow]$ e $[\uparrow F]$, usando filtragem *B-spline*, é superior em características taxa-qualidade ao codificador baseado na Recomendação H.263 para taxas de bits

inferiores a 50 kbits/s. Nessa figura, pode ainda ser notado que, em torno de 28 kbits/s, a estrutura em questão apresenta uma PSNR de aproximadamente 0,9 dB maior que a do codificador baseado na Recomendação H.263.

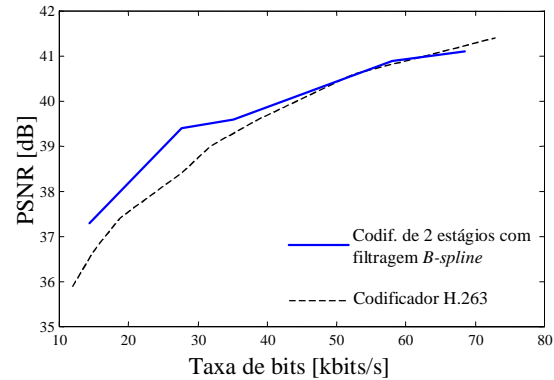


Fig. 2 - Características taxa-qualidade para os codificadores de dois estágios com filtragem *B-spline* e, baseado na Recomendação H.263

A Fig. 3 mostra as curvas de PSNR para cada quadro reconstruído usando as seguintes estruturas: uma baseada nos operadores $[F\downarrow]$ e $[\uparrow F]$ com filtragem *B-spline* e a outra baseada na Recomendação H.263. São apresentados os resultados para a codificação da seqüência *Miss America* à taxa de 28 kbits/s. A estrutura baseada na filtragem *B-spline* supera em qualidade de reconstrução o codificador da Recomendação H.263 para quase todos os quadros codificados.

A Fig. 4 apresenta os resultados obtidos para o quadro #121 da seqüência *Miss America* codificada a 28 kbits/s. A Fig. 4(a) mostra o quadro original. A Fig. 4(b) apresenta o quadro reconstruído a partir do codificador baseado na Recomendação H.263, com uma PSNR de 37,99 dB. A Fig. 4(c) mostra o quadro reconstruído para a estrutura usando a abordagem por filtragem *B-spline*, resultando em uma PSNR de 39,07 dB. Pode também ser notado que o quadro reconstruído pela estrutura baseada na Recomendação H.263 apresenta-se subjetivamente muito mais degradado do que aquele reconstruído pela abordagem proposta.

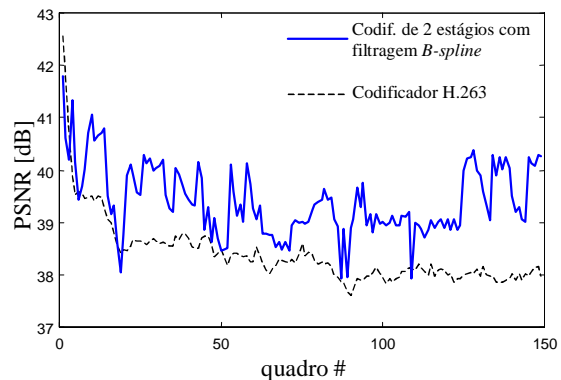


Fig. 3 - Resultados de PSNR para a seqüência *Miss America* codificada à taxa de 28 kbits/s usando: (a) estrutura com filtragem *B-spline* e (b) codificador baseado na Recomendação H.263.



Fig. 4 - Resultados de codificação: (a) quadro #121 original; (b) baseada na Recomendação H.263; (c) abordagem usando filtragem *B-spline*.

6. CONCLUSÕES

Neste artigo, foi apresentado o uso de filtragem *B-spline* em uma estrutura de codificação de vídeo de dois estágios escalável espacialmente para operação em baixas taxas de bits. A abordagem proposta apresenta resultados taxa-qualidade superiores aos da Recomendação H.263 para taxas de bits inferiores a 50 kbits/s. Para taxas superiores, a metodologia apresentada é competitiva em taxa-qualidade à da Recomendação H.263. Baseado nos resultados apresentados e discutidos, pode ser inferido que a utilização de filtragem *B-spline*, para a mudança de resolução espacial em imagens, apresenta um potencial significativo para aplicações em codificação de vídeo escalável. Os resultados obtidos sugerem um maior aprofundamento no estudo de filtragem *B-spline* para as operações $[F\downarrow]$ e $[\uparrow F]$ em aplicações de codificação de vídeo para baixas taxas de operação.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ITU-T. "Recommendation H.261 - Video Codec for Audiovisual Services at $p \times 64$ kbit/s", Mar. 1993.
- [2] ITU-T. "Recommendation H.263 - Video Coding for Low Bitrate Communication", Jul. 1995.
- [3] Mandal M. K., Chan E., Wang X. and Panchanathan S., "Multiresolution Motion Estimation Techniques for Video Compression". *Optical Engineering*, vol. 35, no. 1, pp. 128-136, Jan. 1996.
- [4] Kwon H., Venkatramam M. and Nasrabadi N. M., "Very Low Bit-Rate Video Coding Using Variable Block-Size Entropy-Constrained Residual Vector Quantizers". *IEEE J. Selected Areas in Communications*, vol. 15, pp. 1714-1725, Dec. 1997.
- [5] Nicoleit E. R., Seara R. and Bermudez J. C. M., "Compressão de Vídeo com Alta Qualidade e Baixa Taxa de Bits Usando Codificação Multi-Estágio". *XV Simpósio Brasileiro de Telecomunicações*, Recife, PE, pp. 590-593. Sep. 1997.
- [6] Illgner K. and Müller F., "Spatially Scalable Video Compression Employing Resolution Pyramids". *IEEE J. Selected Areas in Communications*, vol. 15, pp. 1688-1703, Dec. 1997.
- [7] Burt P. J. and Adelson E. H., "The Laplacian Pyramid as a Compact Image Code". *IEEE Trans. on Communications*, vol. Com-31, no. 4, pp. 532-540, Apr. 1983.
- [8] Buchanan J. L. and Turner P. R., "Numerical Methods and Analysis", McGraw Hill, New York. 1992.
- [9] Unser M., "Splines - A Perfect Fit for Signal and Image Processing". *IEEE Signal Processing Magazine*, pp. 22-38, Nov. 1999.
- [10] Unser M., Aldroubi A. and Eden M., "B-Spline Signal Processing - Part I - Theory". *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821-833. Fev. 1993.
- [11] Unser M., Aldroubi A. and Eden M., "B-Spline Signal Processing - Part II - Efficient Design and Applications". *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 834-848. Fev. 1993.
- [12] Musmann H. G., Pirsch P. and Grallert H. J., "Advances in Picture Coding". *Proceedings of the IEEE*, vol. 73, no. 4, pp. 523-548. Apr. 1985.
- [13] Antonini M., Barlaud M., Mathieu P. and Daubechies I., "Image Coding Using Wavelet Transform". *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 205-220. Apr. 1992.
- [14] Villasenor J. D., Belzer B. and Liao J., "Wavelet Filter Evaluation for Image Compression". *IEEE Trans. on Image Processing*, vol. 4, no. 8, pp. 1053-1060. Aug. 1995.
- [15] Shapiro J. M., "Embedded Image Coding Using Zerotrees of Wavelet Coefficients". *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445-3462. Dec. 1993.
- [16] Telenor R&D, "Tmn (h.263) encoder/decoder, version 1.7", <ftp://bonde.nta.no/pub/tmn>, TMN codec. 1996.