

Comparação de Diferentes Métodos de Transcodificação Espacial de Vídeo Digital com o Método Utilizando Wavelets

Josiel do Nascimento Cruz

IFPB - Campina Grande josielnc@yahoo.com.br

Leonardo Araújo Cartaxo da Costa

IFPB - Campina Grande leo.cartaxo@gmail.com

Mauricio Wanderley de Freitas Ferreira

IFPB - Campina Grande athavus@gmail.com

Vinícius Fernandes de Sousa

IFPB - Campina Grande vinifersousa@hotmail.com

Carlos Danilo Miranda Regis

IFPB - Campina Grande regis.danilo@gmail.com

Resumo: Com o amplo desenvolvimento tecnológico dos receptores móveis, a visualização e transmissão de vídeo digital necessitam de recursos de qualidade e velocidade cada vez maiores. A utilização de transcodificadores de vídeo visa criar soluções adequadas para os diversos tipos de textura, movimento e cores encontrados. Este trabalho apresenta um comparativo entre os resultados de um sistema de transcodificação espacial de vídeo CIF para obter vídeos QCIF utilizando wavelet, com outros diversos métodos de transcodificação.

Palavras-Chave: Transcodificação de vídeo, wavelet, compressão de vídeo.

Abstract: With the extensive technological development of mobile receivers, the transmission and display of digital video increase requires resources of quality and speed. The use of video transcoders aims to create solutions for the various types of texture, movement and color matches. This paper presents a comparison between the results of a system of spatial transcoding video CIF for QCIF video using wavelet, with several other methods of transcoding.

Keywords: Video transcoding, wavelet, video compression.

1. Introdução

A transmissão de vídeo digital prevê a recepção de sinais de vídeo em vários formatos, tanto por receptores fixos como móveis tais sinais de vídeo apresentam vários formatos de resoluções (espacial e temporal), codificação e taxa de bits, possibilitando assim que o conteúdo dos vídeos seja reproduzido nos mais diferentes tipos de receptores, que vão desde as convencionais televisões a PDAs (*Personal Digital Assistants*), Ipods e celulares (Alencar, 2009).

Cada forma de transmissão afeta diretamente a taxa com que os bits são transmitidos, podendo reduzir a satisfação dos usuários, já que pode apresentar uma reprodução dos vídeos com pausas.

Com intuito de atenuar tais efeitos pode-se obter vídeos de diferentes configurações para atender as imposições de todos os tipos de transmissão e recepção. Contudo, isso necessitaria de um grande espaço para armazenamento, bem como elevado custo de produção, tornando assim, esse tipo de abordagem inviável.

Uma possível solução é a utilização do transcódificadores de vídeo, que são sistemas capazes de converter uma sequência de vídeo em outra usando parâmetros diferentes (codificação, resolução temporal, resolução espacial e/ou taxa de bits) (Regis, 2009), (Christopoulos, 2003), implicando assim em uma economia de espaço e tempo de produção.

O transcódificador pode reduzir a resolução ou taxa de bits, ou ainda modificar o padrão de compressão do vídeo. Neste trabalho, Os transcódificadores utilizados nas comparações mudam a resolução espacial, reduzindo um sinal de vídeo digital no formato CIF (*Common Intermediate Format* – 352 x 288 pixels) para o formato QCIF (*Quarter Common Intermediate Format* – 176 x 144 pixels).

Esta redução pode ser muito interessante em receptores móveis apresentam diversas limitações como: duração da bateria, menor capacidade de processamento, menor capacidade de memória e tamanho reduzido do display, quando comparado com os aparelhos de TV tradicionais (Regis, 2010).

A finalidade deste trabalho é comparar um sistema de transcódificação de sinais de vídeo, utilizando a transformada *wavelet*, com outros di-

versos transcódificadores de maneira a determinar qual o método mais adequado de codificação para cada tipo de receptor móvel.

2. Tipos de Transcodificação

A transcódificação pode ser classificada como homogênea, heterogênea e funções adicionais. A transcódificação homogênea realiza a conversão da taxa de bits, da resolução espacial, da resolução temporal e mudanças da codificação VBR (*variable bit-rate*) para CBR (*constant bit-rate*), ou seja, modifica um conjunto de parâmetros do vídeo sem modificar o seu padrão de compressão. Entretanto a transcódificação heterogênea realiza a conversão de padrões de compressão ou a conversão entre os formatos entrelaçado e progressivo (Ahmad, 2005).

Além desses dois tipos, existe um outro tipo de transcódificação que consiste em incluir funções adicionais à sequência de vídeo codificada, como proporcionar resistência a erros (*error resilience*) e adicionar logomarcas ou marcas d'água invisíveis. Um diagrama com os vários tipos de transcódificação é apresentado na Fig. 1.



Figura1. Diagrama dos possíveis tipos de transcódificação

2.1 Transcodificação Espacial

O objetivo na transcódificação da resolução espacial é reduzir ou aumentar a resolução espacial do vídeo utilizando técnicas como subamostragem/superamostragem, filtragem, etc. Uma possível aplicação da redução de resolução espacial é a conversão de um vídeo em HDTV (*high definition television*) com resolução 1920x1024i, pixels, (i sendo o padrão entrelaçado) para um em SDTV (*standard definition television*) que tem resolução 720x384, pixels, como apresentado em (Xin, 2002). Nesse caso podem-se usar as

técnicas de filtragem, subamostragem e superamostragem. Como o fator de conversão não é inteiro ($1920:720 = 1024:384 = 8:3$), o vídeo passa por um subamostrador e um superamostrador em cascata (Mitra, 2006). Um filtro é utilizado entre o subamostrador e o superamostrador para garantir que os sinais resultantes não tenham distorção (*aliasing*). A seguir é feita uma breve descrição das técnicas de redução de resolução espacial no domínio dos *pixels*.

- Superamostragem e Subamostragem: essas são técnicas para reduzir a resolução espacial, tanto na direção horizontal quanto na vertical. O superamostrador é utilizado para aumentar a taxa de amostragem de uma sequência por um fator inteiro $L > 1$, enquanto que o subamostrador é utilizado para diminuir a taxa de amostragem por um fator inteiro $S > 1$ (Vaidyanathan, 1993), (Mitra, 2006).

- Filtragem: podem ser usados os filtros passa-baixa de média, mediana, moda, sigma e etc (Regis, 2010):

Média dos *pixels*: o cálculo da média é uma técnica para redução da resolução espacial. Nessa técnica substituem-se os valores do bloco $M \times M$ do quadro do vídeo por um único *pixel* cujo valor corresponde à média aritmética dos *pixels* dentro do bloco $M \times M$. Calcular a média de *pixels* é um dos métodos mais simples, entretanto essa técnica pode adicionar um alto nível de borrado ao quadro.

Mediana: a técnica da mediana consiste na substituição de um conjunto de *pixels* pela mediana de um bloco $M \times M$. Para o cálculo da mediana é feita uma reorganização dos valores dos *pixels* do bloco $M \times M$ de forma crescente.

3. Transformada de Wavelet

O surgimento da transformada *wavelet* deu-se na década de 80, tendo como precedentes as ideias desenvolvidas principalmente no campo de análise multi-escala. A nova ferramenta foi introduzida por Morlet e Grossmann (Pinto, 2005), cuja principal meta era contornar as limitações da análise de Fourier.

Diferente da transformada de Fourier, em que as funções de base são senoidais, a transformada *wavelet* tem como funções de base pequenas ondas chamadas *wavelets*, que apresentam variações de frequência e duração limitada no tempo.

A Fig. 2 ilustra a diferença entre ondas senoidais e *wavelets*. A Fig. 2 (a) apresenta funções senos que diferem na frequência, mas não no espaço. Por outro lado, a Fig. 2 (b) apresenta *wavelets* (neste caso a *wavelet* conhecida como Chapéu Mexicano) que diferem tanto na frequência quanto ao longo do eixo de posição.

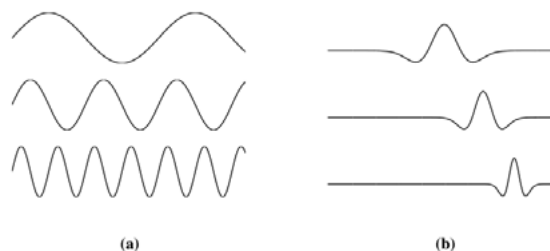


Figura 2. (a) Ondas senoidais e (b) a wavelet conhecida como Chapéu Mexicano.

As *wavelets* são funções matemáticas contendo determinadas propriedades que as tornam adequadas para decomposição de uma determinada função em outra (Silva, 2008). Uma propriedade matemática que faz com que a transformada *wavelet* seja uma ferramenta eficiente no processamento de imagens é a análise de multirresolução que, ao ser aplicado a um sinal, evidencia seus detalhes.

A análise de multirresolução, também conhecida como Algoritmo Piramidal de Mallat (Silva, 2008), é um método proposto para implementar a transformada *wavelet* discreta, e refere-se ao procedimento de se obter aproximações e detalhes de uma imagem. Uma aproximação consiste em uma representação de baixa frequência da imagem original, enquanto que um detalhe é a diferença entre duas aproximações sucessivas da imagem original (Silva, 2008).

O método proposto por Mallat aborda técnicas de filtragem, em que os filtros mais utilizados no processamento das imagens são: o filtro passa-baixa $H(t)$ e o filtro passa-alta $G(t)$ que dividem a imagem em duas partes iguais.

O filtro passa-baixa $H(t)$ tem o efeito de suavizar a imagem e reduzir o número de níveis de cinza. A suavização tende a minimizar os ruídos e apresenta o efeito de “borramento” da imagem. Já o filtro passa-alta $G(t)$ tende a realçar os detalhes, isto é, as transições entre regiões diferentes se tornam mais nítidas (Silva, 2008).

No final do processo de filtragem são geradas as saídas 1, que representa os componentes de baixa frequência, e 1, que representa os componentes de alta frequência. O processo pode ser repetido quantas vezes forem necessárias.

A função *wavelet*, denotada por $\psi(t)$, é definida como um conjunto de funções produzidas por operações matemáticas de translação e escalonamento da função escala, com propriedades particulares que a torna adequada para servir de base para a decomposição de outras funções (Silva,2008).

A função básica escala denotada por ϕ , tendo como funções básicas associadas.

$$\phi_i^j(t) = \phi(2^j t - i) \quad i = 0, \dots, 2^{j-1} \quad (1)$$

em que ϕ é a função escala, i o deslocamento, j a escala e t o tempo.

Para que uma função seja considerada uma *wavelet* é necessária satisfazer as seguintes condições básicas:

$\psi(t) \in L^2(\mathcal{R})$, ou seja, a função pertence ao espaço das funções de quadrado integrável ou, ainda, ao espaço das funções de energia, no sentido que:

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (2)$$

A Transformada de Fourier $\Psi(\omega)$ satisfaz a condição de admissibilidade.

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (3)$$

Segue da condição de admissibilidade que:

$$\lim_{\omega \rightarrow 0} \Psi(\omega) = 0 \quad (4)$$

assim, se $\Psi(\omega)$ é contínua então, $\Psi(0)=0$,

ou seja, $\int_{-\infty}^{\infty} \psi(\omega) d\omega = 0$.

O escalonamento utilizado na Transformada Wavelet tem o mesmo significado da escala utilizada nos mapas geográficos, em que as escalas maiores fornecem visões globais enquanto que os

menores permitem observar os detalhes. Escalas grandes permitem analisar os trechos longos do sinal, enquanto escalas menores permitem analisar trechos mais curtos de forma mais detalhada.

As principais funções *wavelets* são as de Haar, a Sombreiro, a de Densidade Gaussiana, a complexa de Molet, a de Shannon, a Generalizada de Shannon, de Meyer e a Daubechies (Oliveira,2007). Neste trabalho foram usadas as *wavelets* de Haar e Daubechies.

3.1 Sombreiro

Esta wavelet é a derivada segunda da função Gaussiana $f(t) = \frac{-1}{e^{2t^2}}$ pertence à família das wavelets não ortogonais, na Fig. 3 temos a sua representação, esta wavelet é definida como (Protázio,2002):

$$\psi(t) = \frac{2}{\sqrt{3}} \pi^{-0,3} (1-t^2)e^{-0,5t^2} \quad (5)$$

Esta wavelet é muito utilizada principalmente na detecção de bordas.

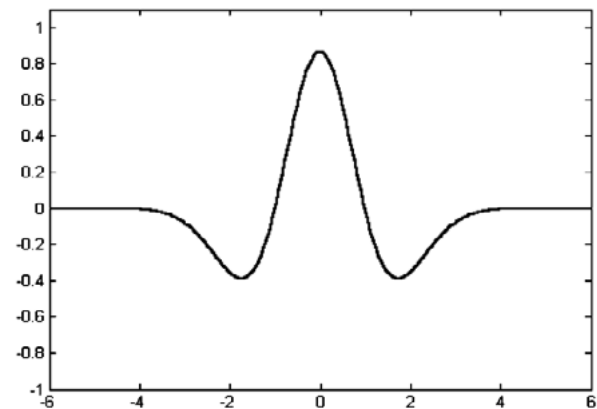


Figura 3. Visualização da wavelet Sombreiro.

3.2 Wavelet de Morlet

A wavelet de Morlet pertence à família das wavelets não ortogonais e é definida como (Protázio,2002):

$$\psi(t) = \pi^{-\frac{-1}{4}} (e^{-i\omega_0 t} - e^{-\frac{-1}{2a_0^2}}) e^{\frac{-1}{2t^2}} \quad (6)$$

que é geralmente aproximada por

$$\psi(t) = \pi^{-\frac{1}{4}} e^{-i\omega_0 t} e^{\frac{-1}{2t^2}}, \text{ para } \omega_0 \geq 5 \quad (7)$$

e é muito utilizada no estudo de sinais geofísicos, a Fig. 4 representa graficamente a função *wavelet* de Morlet.

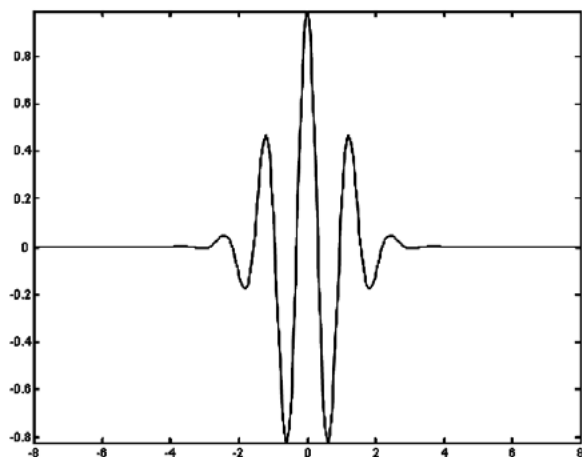


Figura 4. Função wavelet de Morlet.

3.3 Wavelet de Daubechies

Ingrid Daubechies baseou-se no trabalho de Mallat para construir o conjunto de *wavelets* de base ortonormal mais utilizado atualmente. A *wavelet* de Daubechies na verdade é uma família de funções baseadas em funções polinomiais.

A *wavelet* de Daubechies representa as funções ϕ e ψ pelos coeficientes dos filtros k e k , que são respectivamente os filtros passa-baixa e passa-alta. Quanto maior o número de coeficientes, mais suave é a *wavelet*. A função *wavelet* de Daubechies, denominada por DbN possui uma família composta por nove membros, de Db2 a Db10, em que N representa o número de momentos nulos que a função *wavelet* de Daubechies possui.

A seguir é apresentado um exemplo da *wavelet* mais simples de Daubechies, a Db4, gerada a partir de apenas quatro coeficientes.

$$(h_0, h_1, h_2, h_3) = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}} \right) \quad (8)$$

A partir desses coeficientes pode-se construir a função escala:

$$\phi(t) = \sqrt{2} \sum_{k=0}^{2N-1} h_k \phi(2t-k) \quad (9)$$

e calcular n :

$$(g_0, g_1, g_2, g_3) = \left(\frac{1-\sqrt{3}}{4\sqrt{2}}, \frac{-3+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{-1-\sqrt{3}}{4\sqrt{2}} \right) \quad (10)$$

Assim, a *wavelet* de Daubechies é dada por:

$$\psi(t) = \sqrt{2} \sum_{k=0}^{2N-1} g_k \phi(2t-k) \quad (11)$$

As *wavelets* Daubechies usam janelas sobrepostas, de modo que os coeficientes de alta frequência do espectro refletem todas as mudanças nas frequências altas. Portanto as *wavelets* de Daubechies são úteis na compressão e remoção de ruído no processamento de sinais de áudio (Mahmoud,2007).

3.4 Transformada Wavelet de Haar

A Transformada de Haar (TH) é uma das mais simples e básicas transformações. Ela decompõe cada sinal em duas componentes, uma chamada de média (aproximação) e outra conhecida como diferença (detalhes) (Bhardwaj,2009). Uma fórmula precisa para os primeiros valores do

subsinal de aproximação, $a = (a_1, a_2, \dots, a_{N/2})$ em um nível com sinal de tamanho N , isto é,

$$f = (f_1, f_2, \dots, f_N) \text{ é}$$

$$a_n = \frac{f_{2n-1} + f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2 \quad (12)$$

e os primeiros valores do subsinal de detalhe, $d = (d_1, d_2, \dots, d_{n/2})$, no mesmo nível são

$$d_n = \frac{f_{2n-1} - f_{2n}}{\sqrt{2}}, \quad n = 1, 2, 3, \dots, N/2 \quad (13)$$

3.5 Transformada Rápida de Haar

A transformada rápida de Haar também consiste na decomposição da imagem em duas componentes: uma de média (aproximação) e outra referente à diferença (detalhes) (Stollnitz, 1996). Diferenciando da Haar por em suas

duas amostras f_{2n-1} e f_{2n} (vizinhas) serem aplicadas uma transformação linear para gerar valores

a_n (aproximação) e d_n (detalhes), sendo que a aproximação é representada pela equação:

$$a_n = (f_{2n-1} + f_{2n})/2, \quad n = 1, 2, 3, \dots, N/2 \quad (14)$$

e o detalhe é baseado na amostra f_{2n-1} , com a seguinte equação

$$d_n = f_{2n-1} - f_{2n}, \quad n = 1, 2, 3, \dots, N/2 \quad (15)$$

A generalização para uma imagem digital se dá com o cálculo das duas componentes citadas, para cada par de *pixels* da imagem. Quanto às componentes obtidas, pode-se dizer que os valores de média representam as componentes de baixa frequência da imagem e os valores obtidos como diferença, isolam as componentes de alta frequência (detalhes da imagem).

4. Métodos de Avaliação

A qualidade do vídeo pode ser medida utilizando métricas de qualidade objetivas ou de forma subjetiva por meio de experimentos com observadores humanos.

As métricas objetivas como SNR (*signal to noise ratio*), PSNR (*peak signal to noise ratio*) ou MSE (*Mean Squared Error*) faz medições em tempo real, comparando a imagem sob teste com a

imagem original. Essas medidas mostram resultados não coerentes com a percepção de observadores humanos, por consistirem de uma comparação matemática entre o sinal original e o sinal processado. Por exemplo, um quadro deslocado de um *pixel* tem uma PSNR baixa, mas um observador não nota a diferença entre a imagem deslocada e a original.

Para avaliação dos sistemas, é comum o uso da PSNR, ou da WSNR (*weighted signal to noise ratio*). A PSNR é definida pela Equação 16. Para vídeos codificados com 8 *bits*.

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right] \quad (16)$$

O MSE é dado pelo valor médio dos erros quadráticos entre os *pixels* do quadro original e do quadro decodificado. O MSE é dado pela seguinte equação:

$$MSE = \frac{1}{P} \sum_k \sum_{x,y} (f(x,y,k) - h(x,y,k))^2 \quad (17)$$

sendo P o número total de *pixels*, x e y os índices das linhas e colunas, respectivamente, k o número de quadros, f e h representando o quadro original e o modificado, respectivamente.

Testes mostram que há pouca correlação de medidas baseadas no MSE e PSNR com as observações subjetivas (Wang, 2004). Assim, nas três últimas décadas, têm-se estudado maneiras de medir a qualidade de vídeo e de imagens objetivamente de forma que essas métricas levem em consideração as características do sistema visual humano (HVS – *Human Visual System*). Em (Muntean, 2005) são apresentadas diversas outras métricas, dentre elas pode-se citar a métrica de similaridade estrutural (SSIM – *Structural Similarity*).

O SSIM é um algoritmo que utiliza a estatística da imagem para a avaliação de qualidade de imagens usando janelas. Os índices SSIM são calculados pelas janelas deslizantes, que se movem *pixel-a-pixel* do topo esquerdo até o canto inferior direito do quadro.

Em (Wang,2002) é proposta uma indexação da imagem, na qual f e h são dois sinais não negativos alinhados um com o outro, e existe média e desvio padrão dos sinais f e h , $\mu_f, \mu_h, \sigma^{2f}, \sigma^{2h}$ e σ_{fh} . A média e o desvio padrão são considerados estimativas aproximadas da luminância e do

contraste do sinal. A covariância (σ_{fh}) pode ser pensada como a medida de quanto um sinal é diferente do outro (Bovik,2002), (Vranjes,2007).

A similaridade indexada da medida entre f e h é dada em (Wang,2002) e corresponde a

$$S(f, h) = l(f, h)c(f, h)s(f, h) = \frac{4\mu_f\mu_h\sigma_{fh}}{(\mu_f^2 + \mu_h^2)(\sigma_f^2 + \sigma_h^2)} \quad (18)$$

O problema com a Equação 18. é que quando $(\mu^{2f} + \mu^{2h})(\sigma^{2f} + \sigma^{2h})$ for igual a zero a medida resultante não está definida. Por esse motivo uma modificação da Equação 18. é feita (Bovik,2002), resultando na seguinte equação:

$$SSIM(f, h) = \frac{2(\mu_f\mu_h + C_1)(2\sigma_{fh} + C_2)}{(\mu_f^2 + \mu_h^2 + C_1)(\sigma_f^2 + \sigma_h^2 + C_2)} \quad (19)$$

em que as constantes 1 e 2, são dadas por:

$$C_1 = (K_1B)^2, \quad C_2 = (K_2B)^2 \quad (20)$$

em que B é o valor dinâmico do *pixel* (para 8 *bits/pixel*, $B = 255$), 1 e 2 são duas constantes pequenas, dadas por $K_1 = 0,01$ $K_1 = 0,0$ e $K_2 = 0,03$ $K_2 = 0,0$, respectivamente, para este trabalho (Wang,2004).

O SSIM tem as seguintes propriedades:

1. SSIM (f, h) = SSIM (h, f);
2. SSIM (f, h) ≤ 1;
3. SSIM (f, h) = 1, se e somente se $f = h$.

5. Análise do Sistema Proposto

5.1 Sistema Proposto

A técnica para redução da resolução espacial tanto na posição vertical quanto na posição horizontal no domínio dos *pixels* é mostrada na Fig. 5.

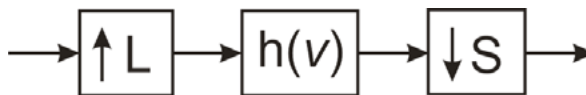


Figura 5. Interpolador fracional com taxa de L/S.

O sistema de transcodificação proposto realiza a conversão da resolução espacial, no domínio dos *pixels*, de vídeos com resolução CIF (*Common Intermediate Format – 352x288 pixels*) para vídeos QCIF (*Quarter Common Intermediate Format – 176x144 pixels*). Para isto, foi aplicado o subamostrador com valor 2 e um superamostrador com fator 1, como ilustrado na Fig. 5. ($L=1, S=2$), em que $h(v)$ representa um filtro passa-baixa. Esses fatores foram escolhidos de acordo com a taxa de redução espacial dos vídeos ($352/176 = 288/144 = 2$) e os filtros são posicionados entre o superamostrador e o subamostrador para garantir que o sinal de saída não apresente *aliasing* (Xin,2002). Foram escolhidos os formatos CIF e QCIF por existirem vídeos de testes em ambos os formatos. Esses vídeos são necessários para realização dos testes subjetivos. Um outro ponto na escolha desses formatos é que para mudar do formato SDTV para CIF também é usado o subamostrados $2=1$ e o subamostrador com $S=2$.

Os filtros, $h(v)$, utilizados neste trabalho foram construídos usando os algoritmos de Wavelet de Haar, Haar rápida e Daubechies 4.

5.2 Análise dos Resultados

Esta seção apresenta os resultados da avaliação da transcodificação usando filtros com algoritmos de Wavelets, sendo denominada Transformada de Haar, Transformada Haar Rápida e Transformada de Daubechies.

Os vídeos originais usados nos testes foram *Mobile*, *News* e *Container*. Cada um dos vídeos tem dez segundos e eles estão disponíveis para *download* público (Trace,2008). Estes vídeos

foram escolhidos porque contêm uma boa mistura de textura, movimento e cores.

5.2.1 PSNR

A Tabela 1 e a Fig. 6 mostram os resultados da PSNR dos vídeos transcódificados obtidos usando os algoritmos Haar, Haar rápida e Daubechies 4. Os valores 1, 2 e 3 representam os vídeos *News*, *Container* e *Mobile*

Tabela 1. Resultados da transcodificação espacial com a PSNR.

Filtro	News	Mobile	Container
Haar	33,3930	23,6794	32,0046
Haar Rápida	34,4056	23,9492	32,4282
Daubechies 4	26,95	22,8	25,8

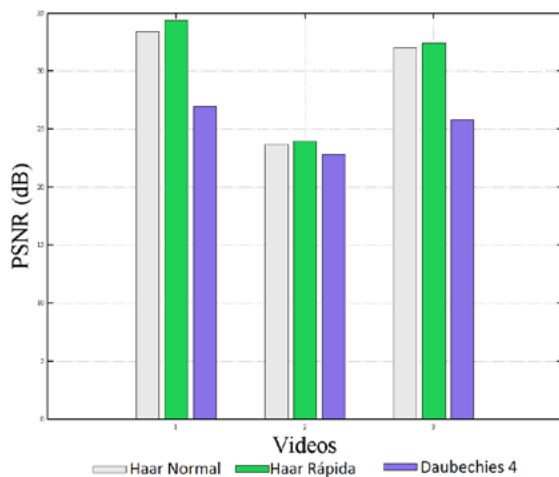


Figura 6. Resultados da transcodificação espacial com a PSNR

O vídeo *News* obteve os melhores resultados de PSNR utilizando todos os filtros, para a de Haar e Haar rápida a PSNR foi superior a 30 dB. Para o Haar e o Haar Rápida a PSNR ficou acima de 30 dB. Dentre os filtros utilizados, o filtro Haar rápida obteve os melhores resultados, porém nos vídeos *Container* e *Mobile* com valores foram bem próximos dos obtidos com o Haar.

Na Fig. 6 também pode ser observado que não há diferença significativa no vídeo *Container* entre os filtros devido à imagem ser bastante uniforme (baixa quantidade de textura).

No artigo (Avelar,2011) é apresentada uma qualificação para a PSNR, a partir dessa classificação pode-se observar que os filtros Haar e Haar Rápida obtiveram classificação boa para os vídeos *News* e *Container* e pobre para o vídeo *Mobile*.

5.2.2 SSIM

A Tabela 2 e a Fig. 7 apresentam o resultado do SSIM obtido pelos vídeos transcódificados pela Transformada Wavelet e os valores 1, 2 e 3 representam os vídeos *News*, *Container* e *Mobile*, respectivamente:

Tabela 2. Resultados da transcodificação espacial com a SSIM

Filtro	News	Mobile	Container
Haar Normal	0,9751	0,8773	0,9634
Haar Rápida	0,9826	0,8831	0,9640
Daubechies 4	0,932	0,849	0,888

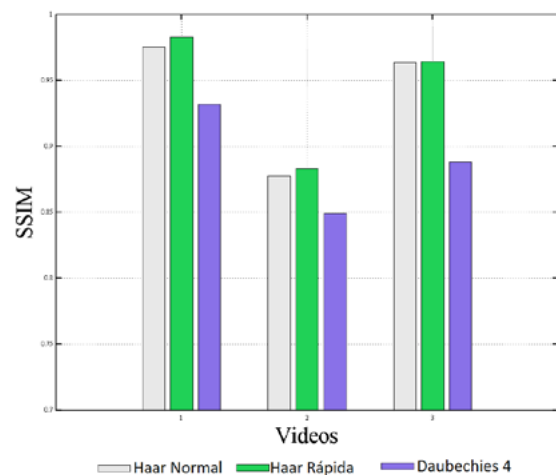


Figura 7. Resultados da transcodificação espacial com a SSIM.

O vídeo *News* como na PSNR obteve os melhores resultados de SSIM, para todos os filtros. Pela métrica SSIM é possível verificar que os filtros de Haar se sobressaem sobre o Daubechies para o vídeo *Container*, ao contrário do que ocorreu para a PSNR.

Observando os resultados obtidos verifica-se que os filtros da Haar e Haar Rápida obtiveram os resultados mais satisfatórios. Para uma

escolha correta entre os filtros deve ser verificado o tempo de processamento de cada um.

5.2.3 Tempo de Processamento

Outro importante fator que deve ser considerado quando se compara diferentes algoritmos de transcodificação é o tempo processado ou complexidade computacional do algoritmo. Esse fator é apresentado na Tabela 3. e na Fig. 8. Os valores 1, 2 e 3 representam os vídeos *News*, *Container* e *Mobile*, respectivamente:

Tabela 3. Resultados do Tempo de Processamento da Transcodificação Espacial em Segundos.

Filtro	News	Mobile	Container	Tempo Médio
Haar Normal	20,14	22,42	22,44	21,66
Haar Rápida	7,13	7,66	5,63	6,81
Daubechies 4	15,00	14,31	14,55	14,64

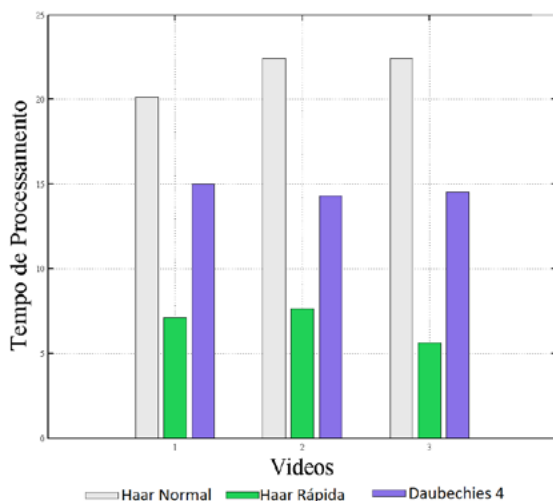


Figura 8. Resultados do Tempo Médio de Processamento da Transcodificação Espacial.

Avaliados os resultados obtidos para o tempo de processamento verificou-se que o filtro de Haar obteve um tempo elevado, acima de 20 segundos, se comparado com a Haar Rápida, que obteve um tempo próximo a 7 segundos.

Observando a PSNR, a SSIM e o tempo de processamento é possível identificar que o filtro

de Haar rápida tem os melhores resultados e é o filtro mais apropriado dentre os avaliados para a transcodificação espacial.

5.3 Comparação com outras técnicas

Para verificar os algoritmos implementados foi realizada uma comparação com os resultados obtidos no artigo “*Objective and Subjective Evaluation of Spatially Transcoded Videos for Mobile Receivers*” (Regis, 2009). Neste artigo a transcodi-ficação é realizada de treze formas diferentes.

A Média foi usada em quatro procedimentos, com a janela deslizante de 1x1, 2x2, 3x3 e 4x4. Três foram encontrados usando a Mediana, com a janela deslizante de 2x2, 3x3 e 4x4. Mais três foram encontrados usando o Sigma, com a janela deslizante de 2x2, 3x3 e 4x4. E por ultimo foram usados três formas de média ponderada.

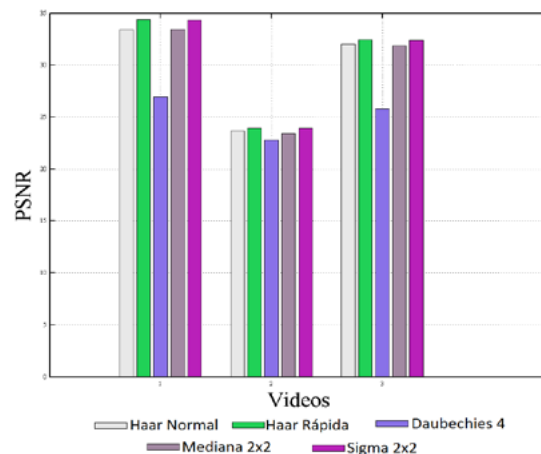


Figura 9. Resultados da transcodificação espacial com a PSNR.

Os melhores resultados encontrados em (Regis, 2009) foram a Mediana 2x2 e o Sigma 2x2. A comparação entre esses filtros e os filtros *wavelet* são apresentados na Fig. 9. e Fig. 10.

Pela avaliação da PSNR (Figura 9.) pode ser observado que para todos os vídeos o filtro Haar Rápida obteve valores próximos ao melhor vídeo encontrado no artigo (Regis, 2009). Já o filtro Haar tem o mesmo desempenho que o filtro da Mediana 2x2, enquanto que o filtro Daubechies 4 obteve o menor desempenho

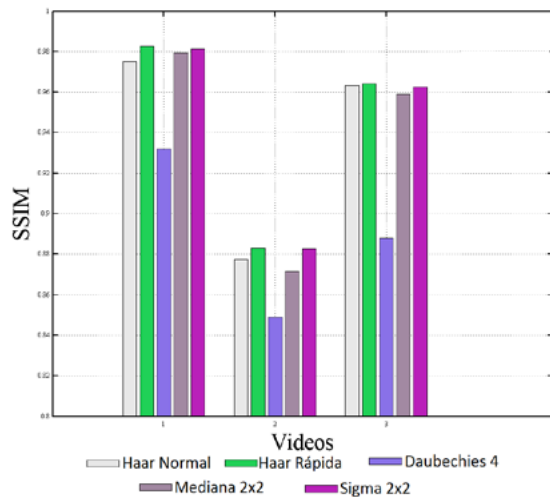


Figura 10. Resultados da transcodificação espacial com a SSIM.

Na avaliação do SSIM (Figura 10.) pode ser observado que o filtro Haar Rápida obteve valores um pouco superiores aos demais filtros. Já o filtro Haar tem o mesmo desempenho que o filtro Sigma 2x2 para o vídeo *Mobile* e inferior ao filtro da Mediana 2x2 para o vídeo *Container*, visto que o filtro sempre Sigma 2x2 é melhor que o filtro da Mediana 2x2. Novamente o filtro Daubechies 4 obteve o menor desempenho.

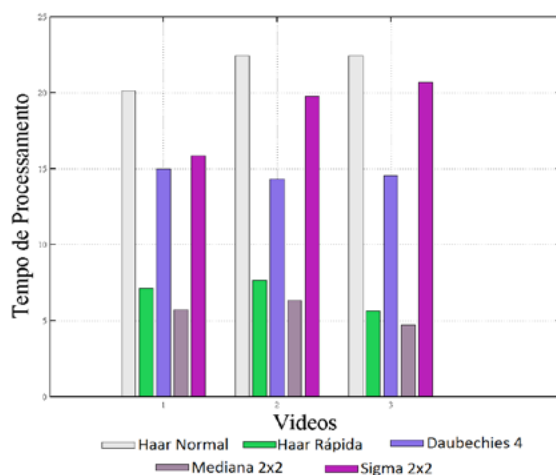


Figura 11. Resultados do Tempo Médio de Processamento da comparação da transcodificação espacial.

Observando o tempo de processamento pode ser observado que o filtro Haar tem o pior desempenho em relação aos filtros avaliados, com valor médio do tempo de processamento de 21,66 segundos, como apresentado na Tabela 3. e na

Fig. 11. O melhor resultado obtido foi com o filtro da Mediana 2x2, com o tempo de médio de 5,59 segundos, com o filtro da Haar Rápida obtendo resultados próximos.

Após a avaliação da PSNR, da SSIM e do tempo de processamento verificou-se que o filtro da Mediana 2x2 tem o menor tempo de processamento e o filtro da Haar Rápida obteve o melhor resultado para as avaliações objetivas. A partir desses resultados pode se recomendar o uso do filtro de Haar rápida na transcodificação espacial dos vídeos já que ele tem o melhor resultado na avaliação da qualidade e um tempo de processamento próximo ao do filtro da Mediana 2x2.

6. Conclusões

Neste Projeto é apresentada uma avaliação da qualidade objetiva de vídeos transcodificados utilizando *wavelet*. A operação consiste em mudar a resolução espacial de vídeos digitais utilizando a superamostragem, a filtragem e a subamostragem. Para a avaliação foram utilizados os filtros de *wavelet*, Haar, Haar Rápida e Daubechies 4.

A avaliação da qualidade objetiva dos vídeos foi realizada com métodos PSNR e SSIM, que são métricas muito utilizadas na literatura e que foram usados no artigo (Regis, 2009).

O filtro da Haar Rápida apresentou os melhores resultados entre os filtros de *wavelet* principalmente na comparação com a PSNR. Considerando os filtros obtido em (Regis, 2009), ele também obteve os melhores resultados.

Considerando o tempo processamento o filtro da Haar Rápida apresentou o melhor resultado em relação aos filtros de *wavelet*, enquanto o filtro Haar obteve o pior resultado. Em relação aos filtros utilizados em (Regis, 2009) o melhor resultado foi encontrado com filtro da Mediana 2x2, e o pior desempenho é o da Haar.

O filtro Haar Rápida possui um bom desempenho no tempo de processamento e o melhor resultado no ponto de vista de qualidade objetiva dos vídeos, obtendo um tempo inferior em relação aos outros algoritmos que possuem resultado de PSNR e SSIM similares.

Como trabalhos futuros podem ser avaliados os algoritmos *Quick Sort* e *Quick Select* que faz com que o algoritmo da mediana se torne mais rápida (Gurwitz,1992).

7. Referências

- AHMAD, I.; WEI, X.; SUN, Y.; ZHANG, Y. Q. "Video transcoding: an overview of various techniques and research issues", **IEEE Transactions on Multimedia**, vol. 7, no. 5, pp. 793–804, Out. 2005.
- ALENCAR, M. S. **Digital Television Systems**. Cambridge Univ Press, 2009.
- AVELAR, E.; MARQUES, L.; BEMERGUY, T.; DIAS, K. "On teaching median-finding algorithms", **IEEE Latin America Transactions**, vol. 9, no. 2, pp. 230–232, Abr. 2011.
- BHARDWAJ, A.; ALI, R. "Image compression using modified fast Haar wavelet transform", **World Applied Sciences Journal**, 2009.
- GURWITZ, C. "On teaching median-finding algorithms", **IEEE Transactions on Education**, vol. 35, no. 3, pp. 230–232, Ago. 1992.
- MAHMOUD, H.; DESSOUKY, M. "Comparison between haar and daubechies wavelet transformions on fpga technology", **World Academy of Science, Engineering and Technology**, 26, 2007.
- MITRA, S. K. **Digital Signal Processing: A Computer Based Approach**. MacGraw-Hill, 2006.
- MUNTEAN, G.; PERRY, P.; MURPHY, L. "Objective and Subjective Evaluation of QOAS Video Streaming over Broadband Networks", **IEEE Electronic Transactions on Network and Service Management**, vol. 2, no. 1, pp. 19–28, Nov. 2005.
- OLIVEIRA, H. M. de. **Análise de Sinais para Engenheiros: uma abordagem via Wavelet**. Brasport, 2007.
- PINTO, S. C. D. **Análise de formas 3d usando a wavelets 1d, 2d e 3d**, Tese de Doutorado, Universidade de São Paulo, São Carlos, Out. 2005.
- PROTÁZIO, J. M. B. **Análise wavelet aplicada a sinais geofísicos**, Master Thesis, Universidade Estadual de Campinas, Campinas, Brasil, Fev. 2002.
- REGIS, C. D.; ROCHA, R. R.; FARIAS, M.; ALENCAR, M. S. "Objective and subjective evaluation of space-transcoded videos for mobile receivers," in **Journal of Communications Software and Systems**, vol. 6, no. 2, 2010.
- REGIS, C.; MORAIS, D.; ALENCAR, M.; FARIAS, M. "Objective and subjective assessment of space transcoded videos for mobile receivers," in **IEEE International Symposium on Broadband Multimedia Systems and Broadcasting**, 2009. BMSB '09., pp. 1–6, Maio, 2009.
- SILVA, J. F. da. **Sistema de armazenamento de imagens comprimidas através da transformada wavelet**, Dissertação de Mestrado, Universidade Estadual Paulista, Jun. 2008.
- STOLLNITZ, E. J.; DEROSE, T. D.; SALESIN, D. H. **Wavelets for Computer Graphics: theory and applications**. Morgan Kaufmann publishers, 1996.
- TRACE, W. **YUV Video Sequences**, Abr. 2008.
- VAIDYANATHAN, P. P. **Multirate systems and filter banks**. New Jersey: Prentice Hall/ Englewood Cliffs, 1993.
- VETRO, A.; CHRISTOPOULOS, C.; SUN, H. "Videotranscoding architectures and techniques: an overview", **Signal Processing Magazine, IEEE**, vol. 20, no. 2, pp. 18–29, Mar. 2003.
- VRANJES, M.; RIMAC-DRLJE, S.; ZAGAR, D. "Objective video quality metrics," **49th International Symposium ELMAR**, 2007, pp. 45–49, Set. 2007.
- WANG, Z.; LU, L.; BOVIK, A. C. "Video quality assessment using structural distortion measurement", in **Proc. IEEE Int. Conf. Image Proc**, 2002, pp. 65–68.
- WANG, Z.; BOVIK, A. "A universal image quality index", **Signal Processing Letters, IEEE**, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- WANG, Z.; BOVIK, A. C.; SHEIKH, H. R.; SIMONCELLI, E. P. "Image quality assessment: from errorvisibility to structural similarity", **IEEE Transactions on Image Processing**, vol. 13, no. 4, pp. 600–612, 2004.
- XIN, J.; SUN, M.-T.; CHOI, B.-S.; CHUN, K.-W. "An HDTV-to-SDTV spatial transcoder", **IEEE Transactions on Circuits and Systems for Video Technology**, vol. 12, no. 11, pp. 998–1008, Nov. 2002.