

SERGIO RICARDO MASTER PENEDO

Inpainting de imagens digitais por estimação de curvas de decaimento dos coeficientes wavelet via regularidade e espaços de Besov

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutor em Engenharia Elétrica

São Paulo
2014

SERGIO RICARDO MASTER PENEDO

Inpainting de imagens digitais por estimação de curvas de decaimento dos coeficientes wavelet via regularidade e espaços de Besov

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do título de Doutor em Engenharia Elétrica

Área de Concentração: Sistemas Eletrônicos

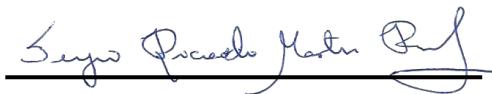
Orientador: Prof. Dr. João Francisco Justo Filho

São Paulo
2014

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 05 de junho de 2014.

Assinatura do autor



Assinatura do orientador



FICHA CATALOGRÁFICA

Penedo, Sergio Ricardo Master

Inpainting de imagens digitais por estimação de curvas de
decaimento dos coeficientes wavelet via regularidade e espaços
de Besov / S.R.M. Penedo. -- versão corr. -- São Paulo, 2014.

114 p.

Tese (Doutorado) - Escola Politécnica da Universidade de
São Paulo. Departamento de Engenharia de Sistemas Eletrô-
nicos.

1.Processamento de imagens 2.Processamento de vídeo
I.Universidade de São Paulo. Escola Politécnica. Departamento
de Engenharia de Sistemas Eletrônicos II.t.

Dedico este trabalho às duas mulheres que me ensinaram, cada uma ao seu tempo, pequenas e enormes lições de vida: à minha mãe, Vera Regina Master Penedo (*in memoriam*), que me transmitiu o valor da leitura, do estudo e da importância de se lutar até o último segundo, mesmo não estando entre nós neste momento, e à minha esposa Tatiana Martin Portela Penedo, que com seu amor e enorme tolerância, me apoiou ativamente para que eu concluisse esta jornada.

AGRADECIMENTOS

Aos meus pais, Dilson e Vera, por terem me transmitido os valores fundamentais para minha construção como ser humano.

Aos meus filhos Maurício e Fernando, por quem sempre mantive o pensamento positivo nas horas difíceis dessa caminhada, e à minha esposa Tatiana, símbolo de postura ativa diante das adversidades.

Ao Professor Doutor João Francisco Justo Filho, pelo incansável apoio e incentivo técnico, profissional e pessoal na realização deste trabalho, me fazendo voltar a acreditar em meu potencial e se revelando amigo da mais valorosa convivência, com quem enfim reaprendi muito em diversos aspectos de minha personalidade.

Ao Professor Doutor Flávio Almeida de Magalhães Cipparrone, pelas valiosas conversas sobre matemática aplicada, trazendo soluções a alguns problemas renitentes.

Aos meus distantes amigos Frederick Arthur Cornell III, em Baltimore, e Fady Zouheir Hussein, em Beirute, a quem vejo como verdadeiros irmãos e que muito me apoiaram nesta empreitada.

Ever tried. Ever failed. No matter.

Try Again. Fail again. Fail better.

(Samuel Barclay Beckett)

RESUMO

Com a crescente veiculação de sinais de informação visual nos meios de comunicação, métodos de *inpainting* de imagens passaram a ser cada vez mais importantes. Esses métodos, usualmente propostos para minimizar os efeitos de perda de blocos de dados, analisam prioritariamente a correlação espacial entre porções de uma imagem, sem explorar intensivamente a informação no domínio da frequência. Em outra vertente, técnicas que buscam avaliar perdas de blocos em imagens não descrevem localmente uma figura de mérito consistente para aferição da correlação espacial entre os *pixels* de cada bloco corrompido. No contexto de se explorar o sinal de informação visual em múltiplas resoluções, este trabalho apresenta uma técnica de *inpainting* de imagens digitais, extensível a aplicações que manipulem sinais de vídeo, a partir da análise da propriedade de regularidade das *wavelets* na decomposição multi-resolução de imagens, descrita localmente no sentido do espaço vetorial de Besov, um espaço com suavidade generalizada que comporta razoavelmente modificações nos coeficientes *wavelet* dos quadros visando à correção de informação. O trabalho descreve uma técnica de estimativa de curvas de decaimento dos coeficientes *wavelet* dos *pixels*, baseada na observação da propriedade de regularidade, objetivando a correção de blocos de *pixels* contaminados por ruído. A partir da observação de tais características, o trabalho propõe um algoritmo de *inpainting*, no domínio da transformada *wavelet*, aplicado a imagens, perfazendo tanto a detecção, baseada em um critério de limiarização de correlação cruzada dos coeficientes *wavelet*, como a correção dos blocos de *pixels* danificados, fundamentada na reconstrução de informação a partir de uma função de decaimento projetada para tais coeficientes. Resultados de simulações obtidos para a aplicação do método em imagens corrompidas por diferentes padrões de ruído mostram seu bom desempenho comparativamente a outros recentes métodos. Ao final, são apresentados tópicos que serão explorados no prosseguimento da investigação.

ABSTRACT

With the growing visual information signals transmission, methods of image inpainting acquired noticeable importance in signal processing area. These methods, usually proposed to minimize the effects of loss of data blocks, analyze primarily the spatial correlation between portions of an image, without intensively exploring the information in the frequency domain. In another aspect, techniques that seek to evaluate losses of blocks in images do not describe locally a figure of merit for consistent measurement of spatial correlation between pixels of each corrupted block. In the context of exploring the visual information signal in multiple resolutions, this work presents a technique of digital image inpainting, extensible to video applications, from the analysis of property of regularity of wavelets in multiresolution analysis of images, described locally in the sense of Besov vectorial spaces, whose generalized smoothness allows widespread changes in wavelet coefficients of image data, aiming to conceal visual information. The study describes a technique for estimation of decay curves of the wavelet coefficients from pixels, based on observation of the property of regularity, aiming to correct blocks of pixels contaminated by noise. From the observation of such characteristics, the study proposes an algorithm for inpainting, in the wavelet transform domain, applied to images, performing as the detection task, based in a criterion of threshold of cross-correlation between wavelet coefficients, as the correction of damaged pixel blocks, stated in the reconstruction of information from a designed decay function for those coefficients. Simulation results obtained for the implementation of the method in images corrupted by different patterns of noise show its good performance compared to other recent methods. At the end of this work, topics to be explored in further research are presented.

LISTA DE ILUSTRAÇÕES

FIG. 2.1 - DIAGRAMA EM BLOCOS DE UM SISTEMA USUAL DE TRANSMISSÃO DE IMAGENS	27
FIG. 2.2 - DOIS QUADROS RECONSTRUÍDOS DE UMA SEQÜÊNCIA CODIFICADA EM MPEG-2. (A) QUADRO 16, COM TRÊS PERDAS DE CÉLULA; (B) QUADRO 22, COM UMA ÚNICA PERDA DE CÉLULA. ARTEFATOS ESTÃO MARCADOS EM VERMELHO	34
FIG. 2.3 - DIAGRAMA EM BLOCOS DE UM CANAL DE COMUNICAÇÃO DE VÍDEO	35
FIG. 2.4 - RELAÇÃO ENTRE A QUALIDADE DE VÍDEO RECONSTRUÍDO NO DECODIFICADOR E A QUANTIDADE DE CORREÇÃO DE REDUNDÂNCIA EMPREGADA SOB VALORES DE ERRO DE CANAL DIFERENTES. A LARGURA DE BANDA TOTAL UTILIZADA PARA CODIFICAÇÃO DE FONTE E DE CANAL É FIXA.	37
FIG. 3.1 - ÁRVORE DE BANCO DE FILTROS: ITERAÇÃO DO ESQUEMA DE CODIFICAÇÃO EM SUB-BANDAS	51
FIG. 4.1 - BLOCO A SER ANALISADO (F_i) E SEUS 4 BLOCOS ADJACENTES (F_{i-P} , F_{i+P} , F_{i-1} E F_{i+1})	59
FIG. 4.2 - TRÊS SINAIS GERADOS A PARTIR DE AMOSTRAS, AO LONGO DO TEMPO, DOS VALORES MÉDIOS DE 3 COEFICIENTES WAVELET, CORRESPONDENTES A 3 BLOCOS ESPACIALMENTE ADJACENTES EM CADA QUADRO DE UMA SEQÜÊNCIA DE VÍDEO.....	60
FIG. 4.3 - FUNÇÃO DE CORRELAÇÃO CRUZADA ENTRE 2 DIFERENTES CURVAS DE AMPLITUDE DOS COEFICIENTES WAVELET, PARA 78 AMOSTRAS SUCESSIVAS NO TEMPO DE 2 PIXELS ADJACENTES	61
FIG. 4.4 - DIAGRAMA DO ESQUEMA PROPOSTO DE DETECÇÃO DE BLOCOS CORROMPIDOS 62	
FIG. 4.5 - DIVISÃO DO QUADRO EM BLOCOS DE 2×2 PIXELS	63
FIG. 4.6 - DECOMPOSIÇÃO WAVELET DE CADA LINHA DE BLOCOS DE PIXELS	64
FIG. 4.7 - VERIFICAÇÃO DA CORRELAÇÃO CRUZADA ENTRE COEFICIENTES WAVELET DE BLOCOS DE PIXELS SUCESSIVOS NO TEMPO. VALORES DE CORRELAÇÃO CRUZADA INFERIORES AO LIMIAR ESTABELECIDO SINALIZAM BLOCOS DE PIXELS DEGRADADOS	65

FIG. 4.8 - QUADROS SUCESSIVOS DA SEQUÊNCIA ‘FOREMAN’ (A) – (C), E INTENSIDADE DE PIXEL EM SUCESSIVAS ESCALAS DE DECOMPOSIÇÃO DA LINHA CENTRAL DE PIXELS DE CADA QUADRO. PELO FATO DE OS QUADROS POSSUÍREM REPRESENTAÇÃO EM CANAIS RGB (RED-GREEN-BLUE), SÃO ILUSTRADOS OS COMPONENTES WAVELET REFERENTES A APENAS UM CANAL – NA FIGURA, O CANAL R. PARA CADA QUADRO, SÃO CALCULADOS 4 NÍVEIS DE DECOMPOSIÇÃO WAVELET NÃO DIZIMADA	67
FIG. 4.9 - FUNÇÃO DE ANÁLISE $g_2(j)$. O REFINAMENTO DAS CARACTERÍSTICAS DE DECAIMENTO PERMITE UMA MELHOR DESCRIÇÃO DE COEFICIENTES DE ALTAS FREQUÊNCIAS, O QUE PROMOVE MELHORA NA QUALIDADE DOS BLOCOS CORRIGIDOS ...	71
FIG. 4.10 - DIAGRAMA EM BLOCOS DA ESTRUTURA DE CORREÇÃO DE BLOCOS IDENTIFICADOS COMO DEGRADADOS	72
FIG. 4.11 - FLUXOGRAMA DOS ALGORITMOS PROPOSTOS PARA DETECÇÃO E CORREÇÃO DE ERROS	74
FIG. 4.12 - CÁLCULO DA TRANSFORMADA WAVELET DOS BLOCOS DA IMAGEM ANALISADA	
75	
FIG. 4.13 - DETECÇÃO DE BLOCOS CORROMPIDOS POR LIMIARIZAÇÃO DE CORRELAÇÃO CRUZADA	76
FIG. 4.14 - CORREÇÃO DE BLOCOS CORROMPIDOS POR MEIO DA FUNÇÃO DE RECONSTRUÇÃO DEFINIDA EM (4.9)	77
FIG. 5.1 - (A) CURVAS DE INTENSIDADE DE PIXEL VS. NÚMERO DE QUADRO, EM UMA SEQUÊNCIA DE 80 QUADROS SUCESSIVOS NO TEMPO, COM UM DOS BLOCOS DEGRADADOS. (B) FUNÇÕES DE CORRELAÇÃO CRUZADA ENTRE AS CURVAS DE AMPLITUDE DE (A) E DE AMPLITUDE MÉDIA DO BLOCO: ANTES DA DEGRADAÇÃO (CURVA EM VERDE), APÓS A DEGRADAÇÃO (CURVA EM VERMELHO) E RETIRANDO-SE O BLOCO DEGRADADO (CURVA EM AZUL).	81
FIG. 5.2 - (A) QUADRO DA SEQUÊNCIA ‘FOREMAN’ CONTAMINADO POR RUÍDO. (B) QUADRO CORRIGIDO. (C) QUADRO ORIGINAL – DETALHES. (D) QUADRO CONTAMINADO POR RUÍDO – DETALHE. (E) QUADRO CORRIGIDO - DETALHE	82
FIG. 5.3 - CONJUNTO DE IMAGENS ‘LENA’. (A) ORIGINAL. (B) CORROMPIDA POR RUÍDO ALEATÓRIO. (C) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [90]. (D) CORRIGIDA PELO	

MÉTODO DA REFERÊNCIA [92]. (E) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [98]. (F) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [101]. (G) CORRIGIDA PELO MÉTODO AQUI PROPOSTO.....	84
FIG. 5.4 - CONJUNTO DE DETALHES DAS IMAGENS ‘LENA’. (A) ORIGINAL. (B) CORROMPIDA POR RUÍDO ALEATÓRIO. (C) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [90]. (D) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [92]. (E) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [98]. (F) CORRIGIDA PELO MÉTODO DA REFERÊNCIA [101]. (G) CORRIGIDA PELO MÉTODO PROPOSTO.....	86
FIG. 5.5 - QUADRO DA SEQUÊNCIA DE VÍDEO ‘SUZIE’. (A) QUADRO ORIGINAL. (B) QUADRO CORROMPIDO. (C) QUADRO RECONSTRUÍDO POR MEIO DO MÉTODO APRESENTADO EM [90]. (D) QUADRO RECONSTRUÍDO POR MEIO DO MÉTODO APRESENTADO EM [92]. (E) QUADRO RECONSTRUÍDO POR MEIO DO MÉTODO APRESENTADO EM [98]. (F) QUADRO RECONSTRUÍDO POR MEIO DO MÉTODO APRESENTADO EM [101]. (G) QUADRO RECONSTRUÍDO COM A FUNÇÃO DE ESCALONAMENTO DADA EM (4.2). (H) QUADRO RECONSTRUÍDO COM A FUNÇÃO DE ESCALONAMENTO CORRIGIDA, DADA EM (4.9).....	88
FIG. 5.6 - CURVAS DE PSNRxBER PARA O QUADRO CORROMPIDO DA SEQÜÊNCIA DE VÍDEO ‘SUZIE’, ILUSTRADO NA FIG. 5.5.....	89
FIG. 5.7 - CURVAS DE PSNRxBER PARA DIFERENTES QUADROS DA SEQÜÊNCIA ‘SUZIE’ COM $BER = 5 \times 10^{-5}$ BIT/BIT	90
FIG. 5.8 - QUADRO DA SEQUÊNCIA ‘FOOTBALL’. (A) CORROMPIDO. (B) CORRIGIDO PELO MÉTODO APRESENTADO EM [90]. (C) CORRIGIDO PELO MÉTODO APRESENTADO EM [92]. (D) CORRIGIDO PELO MÉTODO APRESENTADO EM [98]. (E) CORRIGIDO PELO MÉTODO APRESENTADO EM [101]. (F) CORRIGIDO PELO MÉTODO AQUI PROPOSTO.	91
FIG. 5.9 - CURVAS DE PARA O MÉTODO PROPOSTO E PARA OS APRESENTADOS EM OUTRAS INVESTIGAÇÕES [90, 92, 98, 101], APLICADOS SOBRE A SEQUÊNCIA ‘FOOTBALL’.	
92	
FIG. 5.10 - CURVAS DE PSNR PARA DIFERENTES QUADROS DA SEQUÊNCIA DE VÍDEO ‘FOOTBALL’, COM $BER = 5 \times 10^{-5}$ bits / bit , COMPARANDO O MÉTODO PROPOSTO AOS APRESENTADOS EM OUTRAS INVESTIGAÇÕES [90, 92, 98, 101].....	93
FIG. 5.11 - QUADRO DO FILME ‘THANK YOU FOR SMOKING’. (A) QUADRO ORIGINAL. (B) QUADRO COM LEGENDAS. (C) QUADRO RECONSTITUÍDO POR MEIO DO MÉTODO APRESENTADO EM [90]. (D) QUADRO RECONSTITUÍDO POR MEIO DO MÉTODO	

APRESENTADO EM [92]. (E) QUADRO RECONSTITUÍDO POR MEIO DO MÉTODO APRESENTADO EM [98]. (F) QUADRO RECONSTITUÍDO POR MEIO DO MÉTODO APRESENTADO EM [101]. (G) QUADRO RECONSTITUÍDO COM A FUNÇÃO DE ESCALONAMENTO DEFINIDA PELA EQUAÇÃO (4.2). (H) QUADRO RECONSTITUÍDO COM A FUNÇÃO DE ESCALONAMENTO CORRIGIDA, DEFINIDA PELA EQUAÇÃO (4.9). 94

LISTA DE TABELAS

TABELA 5.1 - VALORES DE PSNR (dB) PARA OS MÉTODOS DE CORREÇÃO COMPARADOS	
	85
TABELA 5.2 - COMPLEXIDADE COMPUTACIONAL PARA OS MÉTODOS DE CORREÇÃO COMPARADOS	95

LISTA DE ABREVIATURAS E SIGLAS

ARQ	Automatic Retransmission on reQuest
ATM	Asynchronous Transfer Mode
BER	Bit Error Ratio
CBR	Constant Bit Rate
CIF	Common Intermediate Format
DAT	Digital Audio Tape
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Code Modulation
ECC	Error Control Coding
EDA	Error Detection Algorithm
FEC	Forward Error Concealment
FT	Fourier transform
GOP	Group of Pictures
HDTV	High Definition Television
IDCT	Inverse Discrete Cosine Transform
ISDN	Integrated Services Digital Networks
ISO	International Organization for Standardization
ITU-T	International Telecommunications Union – Telecommunication Standardization Sector

JPEG	Joint Picture Experts Group
MCP	Motion Compensated Prediction
MPEG	Motion Picture Experts Group
MSE	Mean Square Error
PCM	Pulse Code Modulation
POCS	Projection Onto Convex Sets
PSNR	Peak Signal-to-Noise Ratio
QCIF	Quarter Common Intermediate Format
SPC	Spatial Predictive Concealment
TPC	Temporal Predictive Concealment
VLC	Variable Length Coding
WT	Wavelet transform
WWW	World Wide Web

LISTA DE SÍMBOLOS

$B_q^\gamma(L_p)$	Espaço de Besov associado ao espaço L_p
C_b	Sinal de crominância - diferença de cor azul
C_g	Constante de admissibilidade
C_r	Sinal de crominância - diferença de cor vermelho
$F_i(m,n)$	<i>Pixel</i> (m,n) no bloco de pixels i
$g(j)$	Função de análise para reconstrução de pixels degradados
$g_2(j)$	Função de análise refinada para reconstrução de pixels degradados
K_j	Constante de extração
$L^2(\mathbb{R})$	Espaço bidimensional dos números reais
r	Expoente de Lipschitz
s_m	Sub-banda <i>wavelet</i>
V_j	Subespaço vetorial associado ao nível de resolução j
W_j	Complemento ortogonal do subespaço V_j
Y	Sinal de luminância
α	Expoente de Hölder
α'	Expoente de Hölder estimado

$\ f\ _{B_q^\gamma(L_p)}$	Norma do espaço de Besov $B_q^\gamma(L_p)$
$\psi_{j,k}(x)$	Função-base <i>wavelet</i> com fator de escala j e fator de deslocamento k
$\Psi_{j,k}$	Coeficientes da transformada <i>wavelet</i> no nível de resolução j , com fator de deslocamento k

SUMÁRIO

1 INTRODUÇÃO	15
1.1 <i>INPAINTING</i> DE IMAGENS DIGITAIS	15
1.2 MOTIVAÇÃO DO TRABALHO	18
1.3 ESTADO DA ARTE	19
1.4 OBJETIVOS DESTE TRABALHO	24
1.5 ESTRUTURA DA TESE	25
2 CARACTERIZAÇÃO DO PROBLEMA DE DETECÇÃO E CORREÇÃO DE ERROS EM IMAGENS	26
2.1 INTRODUÇÃO	26
2.2 PADRÕES DE CODIFICAÇÃO DE IMAGENS	29
2.3 ESTRATÉGIAS DE CORREÇÃO PARA IMAGENS CODIFICADAS	32
2.4 CONCLUSÕES	43
3 WAVELETS E A PROPRIEDADE DA REGULARIDADE DE HÖLDER	44
3.1 INTRODUÇÃO	44
3.2 MOTIVAÇÃO	45
3.3 WAVELETS	46
3.4 TRANSFORMADA <i>WAVELET</i>	47
3.5 ANÁLISE MULTI-RESOLUÇÃO	50
3.6 PROPRIEDADE DA REGULARIDADE DE HÖLDER	51
3.7 CONCLUSÕES	54
4 ALGORITMO PROPOSTO DE DETECÇÃO E CORREÇÃO DE ERROS	55
4.1 INTRODUÇÃO	55
4.2 ALGORITMO DE DETECÇÃO DE ERROS	58
4.2.1 Divisão do quadro em blocos	62
4.2.2 Cálculo da transformada <i>wavelet</i> não-dizimada	63
4.2.3 Teste de correlação	64
4.3 ALGORITMO DE CORREÇÃO DE ERROS	65
4.3.1 Decomposição <i>wavelet</i> do bloco corrompido	72

4.3.2 Determinação do expoente estimado de Hölder α' e da função de extrapolação $g_2(j)$	73
4.3.3 Cálculo dos novos coeficientes <i>wavelet</i> e reconstrução do bloco.....	73
4.4 CONCLUSÕES	78
5 RESULTADOS DE SIMULAÇÕES.....	79
5.1 SIMULAÇÕES COMPARATIVAS	79
5.2 CONCLUSÕES	96
6 CONCLUSÕES GERAIS E PERSPECTIVAS DO TRABALHO	97
REFERÊNCIAS.....	100

1 INTRODUÇÃO

1.1 INPAINTING DE IMAGENS DIGITAIS

As redes de comunicação remontam dos desenvolvimentos tecnológicos ocorridos no século XIX, decorrentes da descoberta e controle do eletromagnetismo. Naquela época, a preocupação em estabelecer meios de comunicação à distância mais eficientes se justificava pela necessidade de superar barreiras geográficas e, sobretudo, superar barreiras econômicas, reduzindo o custo dos meios mais convencionais de comunicação. Assim, surgiram invenções como o telégrafo (1844), o telefone (1876) e o teletipo [1]. O advento da televisão, já em meados do século XX, permitiu a transmissão de imagens a grandes distâncias. Entretanto, restava a barreira da transmissão bidirecional de imagens.

Os desenvolvimentos posteriores permitiram então a implementação de novas aplicações, como o videofone e a teleconferência, que objetivam a comunicação visual bidirecional, e em um estágio posterior a televisão de alta definição (*high definition television - HDTV*), que tem como propósito fornecer transmissão de imagens com alta qualidade. Os avanços na área levaram à criação de organizações internacionais, como a *International Telecommunications Union* (ITU), que tem como meta padronizar, dentre outras atribuições, essas novas formas de comunicação em uma dimensão global, de modo que haja compatibilidade entre equipamentos e que seja possível a comunicação entre quaisquer pontos do planeta. A revolução da informação digital no final do século XX provocou profundas transformações na sociedade, gerando também novos desafios e oportunidades. Surgiram novos dispositivos de criação e manipulação de vídeo, exigindo canais confiáveis para a entrega e troca de informação. Como resultado, um tópico de crescente interesse está relacionado à melhoria na qualidade do vídeo transmitido. Soluções para o problema de erro na transmissão de sequências de vídeo contribuem não apenas para a compreensão da tecnologia envolvida, como também oferecem novas oportunidades econômicas a serem exploradas.

Um grande desafio se apresentou para a transmissão de informação visual em larga escala: a capacidade limitada do canal utilizado, que define a quantidade de informação que pode trafegar em uma unidade de tempo. Como o sinal de vídeo é massivo, em termos de amostras de dados, surge a necessidade de processamentos adicionais para reduzir, sem perda significativa de qualidade, a carga útil de informação. Surgiram, assim, técnicas de codificação (compressão) de vídeo a baixas taxas de transmissão, despertando grande interesse de muitos pesquisadores nas últimas décadas [2, 3].

Com a crescente demanda de transmissão de imagens e vídeos em tempo real, um grande desafio é aperfeiçoar a qualidade da informação transmitida. A demanda para transmissão de vídeo comprimido sobre redes de dados cresce, induzindo também o crescimento da largura de banda e do volume de armazenamento requeridos. Buscando superar tais limitações, diversos algoritmos de codificação vêm sendo propostos [4, 5], concebidos para melhorar a taxa de compressão a partir da hipótese de que os canais de transmissão não adicionam quaisquer distorções ou ruídos ao sinal transmitido. Como tal hipótese não é realista, pesquisas em codificação de vídeo têm sugerido algumas modificações que buscam, a partir de considerações acerca da natureza do sinal e do meio, minimizar os efeitos do ruído sobre esse sinal [6, 7]. Tais modificações, geralmente, elevam o esforço computacional de codificação. Não obstante, durante períodos de congestionamento da rede, pacotes de vídeo podem ser descartados ou corrompidos, degradando ainda mais a qualidade do vídeo como resultado da perda de tais pacotes. Uma vez que a retransmissão não é uma opção economicamente viável para aplicações multimídia em tempo real, algoritmos de correção de erro têm sido propostos, buscando estimar a informação perdida. Tais algoritmos devem ser integrados ao *hardware* decodificador (na recepção), e ser suficientemente simples visando à implementação em tempo real.

Torna-se evidente, assim, determinar para tal problema soluções de pós-processamento que estabeleçam um compromisso satisfatório entre complexidade computacional e eficiência de correção. Tais abordagens devem, se possível, explorar alguma característica do sinal de imagem que possa tornar a tarefa de correção mais eficiente. Desta forma, o grande obstáculo que se apresenta é o fato de que, ao se lidar com sinais de vídeo codificados, deve-se considerar que

uma importante quantidade de redundância espacial e temporal tenha sido eliminada no processo de codificação, fato esse que é de suma relevância para a definição de um critério de similaridade útil no processo de correção, tornando o sinal transmitido ainda mais suscetível a erros.

Técnicas que objetivam atingir uma alta taxa de compressão dos dados, como técnicas de estimativa e compensação de movimento na codificação inter-quadros, têm buscado explorar, de alguma forma, a redundância temporal existente em vídeo, com destaque para a técnica de codificação da diferença entre quadros com compensação de movimento (*motion compensated prediction* - MCP) [2]. Essa técnica tira proveito, em termos de velocidade de processamento, do fato de que dois quadros consecutivos altamente correlacionados exibem diferenças relevantes apenas onde existam movimentos significativos, regiões essas geralmente localizadas nas bordas e contornos dos objetos. Outra técnica que tem se mostrado vantajosa na codificação de imagens é a representação de imagens por estruturas denominadas *quad-trees*. Tal abordagem foi utilizada por Yu *et al.* [8] para otimizar a quantidade de dados a ser quantizada em quadros MCP, obtendo-se resultados bastante satisfatórios. Mais recentemente, várias outras pesquisas em vídeo digital têm utilizado tal abordagem para a compressão de dados em quadros MCP [9-12]. Entretanto, tais técnicas, na tentativa de explorar a redundância de informação visual, obtêm somente uma redução da quantidade de dados a serem codificados, sem qualquer correção de eventuais erros. Uma das variações mais recentes dessas técnicas [13] trabalhou na direção de reduzir o esforço computacional dos vetores de deslocamento, graças à determinação e codificação dos vetores de deslocamento apenas em regiões selecionadas pela decomposição *quad-tree*, em uma etapa de pré-processamento anterior à estimativa de movimento. Entretanto, erros que afetem o sinal após a codificação não são corrigidos.

Esse capítulo introdutório faz uma retrospectiva das mais recentes técnicas de *inpainting* – denominação dada a qualquer processo de preenchimento de porções perdidas ou danificadas de um sinal de informação visual, apontando suas eventuais deficiências ainda existentes e apresentando uma proposta para o problema em questão.

1.2 MOTIVAÇÃO DO TRABALHO

Técnicas variadas de *inpainting* [14, 15] têm explorado, de alguma forma, a redundância da imagem a ser corrigida. Entretanto, o desempenho desses procedimentos tem sido insuficiente na reconstrução de bordas e contornos para uma grande parte das imagens analisadas. Isso se deve à forma inaceitável com que a redundância espacial é considerada nesses métodos, na caracterização do erro presente em cada bloco de *pixels*. Nesse contexto, recentes técnicas propõem utilizar de forma mais efetiva as características de redundância espacial [16], sem, no entanto, obter resultados satisfatórios com respeito ao maior esforço computacional despendido.

Nesse ínterim, embora as técnicas de compressão baseadas na transformação de sinais sejam efetivas para a tarefa de codificação, não são sempre adequadas para separar, do sinal corrompido, componentes de ruído [17-19], dada a caracterização precária de componentes de altas frequências em imagens, o que motivou a concepção de técnicas de reconstrução de sinais que operem após o estágio de codificação. Tais procedimentos são conhecidos como *algoritmos de correção de erro*. Por outro lado, diversos modelos analíticos foram recentemente desenvolvidos para estimar o impacto do efeito da propagação de erro causada por canais de transmissão com perdas, concentrando-se nas características estatísticas da imagem transmitida, possuindo assim alta complexidade computacional [20]. Assim, a exploração da redundância no domínio da transformada surge como importante ferramenta para a implementação de algoritmos eficientes de *inpainting*.

O maior problema enfrentado por técnicas de exploração de redundância, majoritariamente operando nos domínios do espaço e da transformada de Fourier (*Fourier transform* – FT), reside no fato de que, ao se explorar a redundância em frequência para se obter uma maior compressão da informação, há uma redução de redundância espacial a cada quadro, o que compromete o desempenho desse segmento. Como a transformada *wavelet* permite a representação de sinais no domínio espacial através do uso da redundância inerente de seus coeficientes de altas frequências, pode ser matematicamente explorada para *inpainting*. Tal similaridade entre os coeficientes de sucessivas escalas da decomposição *wavelet*,

caracterizada pela propriedade da regularidade de sua transformada, permite a definição de figuras de mérito robustas para a identificação e correção de blocos corrompidos em imagens, o que é escopo deste trabalho.

1.3 ESTADO DA ARTE

Diversos trabalhos de pesquisa analisaram o problema da correção de erros em imagens. Esta seção discute aqueles mais relevantes, onde são apontadas suas vantagens e desvantagens que suscitaram a proposição de um novo procedimento de correção. As diversas técnicas apresentadas são categorizadas conforme o meio com que exploram a similaridade de informação: técnicas temporais (mais adequadas para sinais de vídeo), espaciais e no domínio da transformada, além daquelas que atuam na correção de células e pacotes de transmissão.

No que se refere à exploração da redundância temporal, algumas técnicas têm se destacado na literatura. Magal *et al.* [5] apresentam uma variedade de técnicas de correção de erro aplicadas a codificadores de vídeo visando, sobretudo, a operação em canais ruidosos. São confrontadas técnicas baseadas em exploração da redundância temporal, códigos corretores de erros e códigos detectores de erros. Essa comparação tem como propósito avaliar em que situações a correção de erro no pós-processamento (após a codificação) é mais vantajosa. Bhadviya *et al.* [21] propõem uma técnica de codificação e correção de erro em vídeo baseada na exploração da redundância temporal. Essa técnica se revela satisfatória, embora apenas sob duas condições: existência de quadros em que o movimento não pode ser satisfatoriamente previsto e presença de quadros intracodificados sem qualquer informação de movimento. Ding, Liu e Fan [22] apresentam um estudo dos avanços de correção de erro com a aplicação do Algoritmo de Viterbi [23] à codificação *motion picture experts group* (MPEG) visando sequências de vídeo transmitidas via satélite, para predição temporal considerando ou não a compensação de movimento. Por meio da correção parcial de erros aleatórios no domínio da transformada de cosseno, os resultados obtidos atestam o bom desempenho da técnica de compensação de movimento para restaurar sequências de vídeo com

pouca variação interquadros, não prevendo, no entanto, o caso de variações bruscas de cena, o que evidencia uma lacuna deixada por técnicas baseadas unicamente na exploração da correlação temporal. Em [24], Kannan aprofunda o estudo da correção de erro em sequências de vídeo codificadas em MPEG transmitidas sobre redes *asynchronous transfer mode* (ATM), combinando a exploração de correlação temporal à técnica de codificação com controle de erro. Embora seja possível detectar com uma maior precisão a localização dos *pixels* afetados por erro, tarefa essa de difícil execução em técnicas que operam no domínio temporal, a melhora alcançada, em termos de parâmetros objetivos, não é significativa.

Diversas técnicas fazem uso da redundância espacial para correção. Bernardini *et al.* [25] propõem um algoritmo de restauração de informação perdida em sequências de vídeo, fazendo uso da interpolação espacial da vizinhança de *pixels* de bordas. O algoritmo faz a reconstrução de blocos de *pixels* de regiões de bordas, sem, no entanto, estender sua funcionalidade a outras regiões.

Algumas técnicas adotam esquemas híbridos, que utilizam tanto a correlação temporal como espacial. Sun *et al.* [26] desenvolveram uma técnica para a correção de erro em sistemas HDTV. Tal procedimento detecta áreas de quadros degradados por erro, buscando reparar as porções danificadas através da exploração das redundâncias temporal e espacial dos sinais de vídeo recebidos e reconstruídos. Essa abordagem possibilita ao sistema HDTV fornecer uma qualidade de imagem aceitável sob condições de recepção adversas, estando fortemente condicionada a uma detecção adequada de falhas. Em outra proposta, Wu e Gao [27] introduziram a noção de interpolação espacial ao procedimento de correção puramente temporal, buscando contornar o problema de que uma sequência de *bits* transmitidos erroneamente produz uma sequência horizontal de macroblocos danificados, limitando a interpolação a macroblocos nas vizinhanças inferior e superior. Aign [28] e Ju *et al.* [29] também fazem estudos, em transmissão de sinais de vídeo para HDTV, acerca das diferentes técnicas de manipulação de erro no lado do receptor, formulando etapas iterativas adicionais à decodificação que exploram a ressincronização da informação antes da decodificação.

Tsekeridou *et al.* [30] desenvolveram um algoritmo que alia a reconstrução temporal de quadros tipo intracodificado (I), predito (P) e bidirecionalmente predito

(B), à interpolação espacial de macroblocos danificados a partir de macroblocos vizinhos. Apesar de se obter resultados visualmente satisfatórios, a propagação de erros não é completamente evitada, sobretudo nas bordas de quadros reconstruídos. Rabiee *et al.* [31] desenvolveram uma nova aproximação para o problema da correção de erros em imagens estáticas, como também para sequências de vídeo comprimidas e transmitidas em redes ATM, especificamente com o uso de codificadores baseados em blocos. Essa aproximação busca recuperar os *pixels* perdidos através de interpolação espacial de sua vizinhança, combinada a um esquema de filtragem multidirecional não-linear recursiva. Sob a condição de que os dados comprimidos estejam definidos como de alta prioridade no cabeçalho ATM, o algoritmo consegue efetivamente reconstruir a informação temporal e espacial ausente, apresentando, no entanto, alta complexidade computacional.

Podem ser também enumeradas as técnicas baseadas na exploração de redundância no domínio da transformada. Zhai *et al.* [32] propuseram um algoritmo que corrige erros de *bits* tanto em imagens estáticas quanto em sequências de imagens codificadas usando a transformada discreta do cosseno (*discrete cosine transform* – DCT) e codificação de palavra de comprimento variável (*variable length coding* – VLC), sob um cenário de estimativa bayesiana. O algoritmo não introduz qualquer modificação na etapa de codificação, como também não provoca aumento na necessária taxa de *bits*. O algoritmo somente utiliza a informação no domínio transformado para identificar erros de *bits* e corrigi-los.

Langari, Vaseghi e Pedram [33] propuseram uma técnica de correção de erro aplicada a imagens codificadas por DCT. Blocos danificados são restaurados utilizando a propriedade de suavidade de uma imagem no limite desses blocos, aliada a uma técnica de interpolação guiada por bordas. Baseado em tal propriedade, é definida uma função custo que representa as variações temporais entre blocos adjacentes, sendo então determinados novos coeficientes DCT. A técnica fornece bons resultados para imagens estáticas clássicas apresentadas na literatura, não conduzindo a um aumento de complexidade computacional. Entretanto, tal técnica não foi estendida para sinais de vídeo.

Em outra técnica aplicada a sequências de vídeo MPEG-2, Ding, Liu e Fan [34] exploram a redundância interquadros para projetar um algoritmo de correção em

diferentes níveis de resolução na decomposição piramidal de quadros I. O movimento global da cena é analisado em níveis mais baixos, ao passo que variações locais nos quadros são processadas em níveis mais altos. Embora os resultados obtidos sejam satisfatórios em termos subjetivos, a técnica não privilegia condições mais realistas do vídeo a ser reconstruído, focando-se apenas a movimentos de um único objeto na cena. Também no domínio transformado, Sohel, Bennamoun e Hahn [35] propõem um esquema de correção de sequências codificadas em MPEG-4. Os resultados de correção apresentados são quantitativamente satisfatórios, embora um melhor resultado de correção seja obtido à custa de algoritmos que combinem técnicas das diferentes categorias. Mais especificamente em aplicações de vídeo em ATM, Cuenca *et al.* [36] fazem uma revisão do estado-da-arte em técnicas de correção de erro para sinais de vídeo sobre redes ATM, abrangendo algoritmos que operam nos domínios temporal, espacial e de frequência. Yun, Park e Hong [37] propuseram um algoritmo de correção de erro para vídeos codificados em pacotes, buscando eliminar a propagação dos erros durante a transmissão. Atuando no domínio da DCT, o algoritmo estima coeficientes perdidos segundo a técnica de projeção sobre conjuntos convexos (*projection onto convex sets* - POCS), maximizando a consistência das estimativas de intensidade de *pixel* degradadas, e assim reduzindo consideravelmente os efeitos de propagação do erro ao longo da sequência. Os resultados mostrados evidenciam o bom desempenho do algoritmo para o caso de perda de todos os coeficientes da DCT, sem ilustrar, no entanto, seu comportamento para casos intermediários.

Elad *et al.* [38] desenvolveram modelos de processamento de imagens para *inpainting*, discutindo diversas técnicas do estado-da-arte. Em outra vertente, Afonso *et al.* [39] propuseram um algoritmo que realiza o *inpainting* a partir da suposição de que é uma aproximação para um problema linear inverso mal-colocado (problema em que uma ou mais de três condições - existência, unicidade e continuidade -, com relação a erros experimentais, não são satisfeitas), em que uma função de regularização não-suave é minimizada por meio de operadores lagrangeanos aplicados a imagens.

No grupo das técnicas que buscam corrigir erros oriundos da transmissão da informação de vídeo (em aplicações formalmente derivadas da transmissão de

imagens), Jeng e Lee [40] investigaram técnicas de correção de erro que se aplicam à degradação de vídeo decorrente de erros de *bits* e perdas de células, apresentando resultados de simulações avaliados somente de forma subjetiva.

Dispensando-se maior atenção aos algoritmos de compressão definidos pelas recomendações ISO-MPEG [41] e ITU-T H.26x [2, 42], dos organismos de padronização *International Organization for Standardization* (ISO) e *International Telecommunication Union – Telecommunication Standardization Sector* (ITU-T), é nitidamente observado que o sucesso da técnica de correção está condicionado à detecção correta de macroblocos danificados, e não a erros isolados de *bits* [42], já que a maior parte das técnicas de compressão do estado-da-arte, incluindo H.263+ e MPEG-1, 2, 4, é baseada em MCP (para a qual nenhum método de correção de erro é formalmente definido). Na *et al.* [43] desenvolveram um algoritmo de pós-processamento de sequências codificadas em MPEG-2 para compensar o efeito da perda de informação devida ao ruído de canal de transmissão. Artefatos provenientes de erros não detectados são reduzidos através de um esquema de decodificação implementado em paralelo à estrutura de pós-processamento. Os resultados obtidos ilustram a versatilidade da utilização de pós-processamento na tentativa de se corrigir erros em sequências de vídeo, aliando baixa complexidade computacional e simplicidade de implementação.

Observa-se, assim, no contexto particular de algoritmos de correção de erros em vídeo que fazem uso da redundância temporal inter-quadros, que não se extrai benefício da redundância espacial entre os *pixels*, característica essa que é descartada ou ao menos comprometida. Em algumas situações, no intento de minimizar a perda de redundância espacial, alguns dos algoritmos de correção considerados para vídeo alteram o processo de codificação de fonte, elevando a complexidade computacional do procedimento de correção.

Com base em tais considerações, o presente trabalho propõe desenvolver um algoritmo que explora, por pós-processamento do sinal visual degradado, a redundância dos coeficientes *wavelet* em imagens, caracterizando através do decaimento de tais coeficientes um critério de detecção e correção de erros. Deste modo, pode ser utilizado no *inpainting* de imagens, escopo deste trabalho, ou mesmo de forma conjugada com algoritmos que operam sobre sinais de vídeo,

explorando sua redundância temporal – o que pode ser desenvolvido, a partir da contribuição aqui apresentada, em trabalhos posteriores.

1.4 OBJETIVOS DESTE TRABALHO

1.4.1 OBJETIVO GERAL

Este trabalho apresenta como objetivo principal descrever uma técnica para a detecção e correção de erros presentes em sinais de imagem, codificados ou não, baseada no emprego da propriedade de regularidade das *wavelets*.

1.4.2 OBJETIVOS ESPECÍFICOS

O presente trabalho apresenta como objetivos específicos:

- Definir uma figura de mérito para uso nas tarefas de detecção e correção de erros em imagens, útil no modelamento das características de decaimento dos coeficientes *wavelet* para sucessivos níveis de decomposição da imagem analisada, por meio da estimação da regularidade de Hölder ao longo de cada nível de decomposição. Tal aproximação é construída a partir da consideração de que funções *wavelet* se situam em um espaço vetorial aninhado a um espaço de Besov;
- Determinar, a partir da estimação da figura de mérito representativa da regularidade, uma função característica de decaimento a partir da avaliação dos níveis de decomposição para um bloco de *pixels* específico em cada imagem analisada. Esse procedimento permite obter a reconstrução de blocos de *pixels* que sejam classificados como degradados, observadas as características de decaimento através de escalas de decomposição *wavelet*. O método se mostra eficiente na identificação e correção de imagens, comparativamente a outros algoritmos de *inpainting* testados nos capítulos seguintes, conduzindo a resultados satisfatórios tanto em termos subjetivos quanto da medida de PSNR (*Peak Signal-to-Noise Ratio*).

1.5 ESTRUTURA DA TESE

O presente trabalho é organizado como segue. O Capítulo 2 categoriza algumas aproximações para o problema de detecção e correção de erro em imagens, apresentando também uma breve descrição de tais tarefas à luz das principais recomendações para codificação de imagens. São descritos também os algoritmos de correção considerados no padrão de codificação de imagens JPEG 2000, de forma que a técnica de correção de erro proposta possa ser, como contribuição posterior, implementada como um modelo de teste para um padrão vigente de codificação de imagens. O Capítulo 3 apresenta a propriedade de regularidade das *wavelets*, relacionando esta à correlação existente em diferentes escalas da decomposição de um sinal, o que motivou seu uso na implementação de uma abordagem de detecção e correção de erro em imagens. Tal abordagem explora a redundância espacial, distinguindo, no contexto do espaço de suavidade de Besov, entre porções suaves ou não-suaves em cada quadro. O Capítulo 4 apresenta uma proposta para *inpainting* de imagens, baseada no processamento espacial de quadros perdidos utilizando a estimativa de coeficientes *wavelet* de alta frequência. Para se avaliar de forma quantitativa o desempenho da técnica proposta, comparativamente a métodos recentemente apresentados na literatura, são apresentados resultados de simulações sob certas condições de erro de transmissão, obtidos para diferentes aferições da suavidade local dos coeficientes *wavelet*. O Capítulo 5 é dedicado às discussões e conclusões obtidas a partir de simulações da técnica proposta. No Capítulo 6, são apresentadas conclusões gerais e perspectivas para o trabalho.

2 CARACTERIZAÇÃO DO PROBLEMA DE DETECÇÃO E CORREÇÃO DE ERROS EM IMAGENS

2.1 INTRODUÇÃO

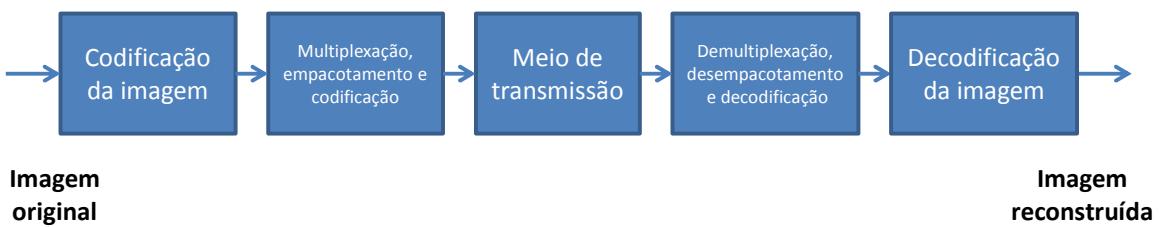
Estratégias de detecção e correção de imagens degradadas buscam realizar modificações nos sinais analisados para restaurar porções de informação visual que foram corrompidas, seja por degradação causada por fontes externas de ruído (como por exemplo, a perda de quadros transmitidos por radiodifusão devido a intempéries), seja pela inserção voluntária de informação aditiva que precisa ser removida (como por exemplo, a inserção voluntária de logotipos de emissoras e legendas). Nesse sentido, a proposição de técnicas de correção de erro em imagens toma como relevante o modo como ruídos ou distorções se fazem presentes, o que requer uma caracterização adequada de como tais adversidades se manifestam e como devem ser tratadas.

Nesse sentido, o presente capítulo traz um compêndio de técnicas de codificação de imagens que operam em condições adversas de operação (como por exemplo, sujeitas à introdução de artefatos provocados por distorções inseridas por um canal de comunicação), descrevendo seus procedimentos de detecção de erros (quando houver) e os meios para correção das degradações detectadas.

Considere-se, como exemplo formal, um sistema de comunicação através do qual uma imagem seja transmitida. Tal sistema é descrito basicamente por cinco etapas, conforme ilustrado na Fig. 2.1. A imagem a ser transmitida é inicialmente comprimida por um codificador, e os dados comprimidos são segmentados em pacotes de comprimento fixo ou variável para, a seguir, serem multiplexados com sinais de outra natureza (por exemplo, áudio). Os pacotes devem ser transmitidos diretamente através da rede de comunicação, se houver garantias de transmissão isenta a erros de *bits*. Caso contrário, tais pacotes são enviados a um estágio de codificação de canal, tipicamente contendo alguma técnica de correção direta (FEC – *forward error concealment*) que previne sua contaminação por erros de

transmissão. No lado receptor, os pacotes recebidos são decodificados em canal e desmembrados, no que a sequência de *bits* resultante é decodificada e o sinal de imagem é reconstruído.

Figura 2.1 – Diagrama em blocos de um sistema usual de transmissão de imagens.



Fonte: O Autor.

O problema de detecção correção de erros em transmissão de imagens vem se tornando um tema de pesquisa cada vez mais importante na área de processamento digital de sinais [39, 43-49]. Particularmente, a limitação de banda em canais de comunicação faz com que o fluxo de *bits* da imagem codificada se torne vulnerável a erros de transmissão, o que pode ocasionar perda de blocos de *pixels*, conduzindo a uma distorção visual no processo de decodificação. Em tal situação, a entrega de imagens sem erro poderia ser obtida através da retransmissão dos blocos de *pixels* corrompidos [39], o que, entretanto, seria inadequado em aplicações em tempo real.

Mecanismos designados para combater erros de transmissão de imagens podem ser classificados em três grupos: (i) esquemas introduzidos no codificador de fonte e de canal, para tornar os dados mais robustos a erros; (ii) esquemas aplicados após a decodificação, que detectam e corrigem erros, e (iii) esquemas que requerem interações entre codificador e decodificador, de modo que a codificação se adapte às perdas detectadas após a decodificação [50, 51]. O trabalho aqui apresentado se situa no grupo (ii), e a escolha por tal abordagem se justifica por duas razões: a modelagem estatística de fonte e canal, em transmissão de vídeo, torna difícil, se não impossível, a obtenção de uma solução ótima para correção de erro segundo (i), dada a natureza variante no tempo do ambiente de transmissão

[46]. Além disso, as altas taxas de *bits* do sinal de vídeo limitam severamente a complexidade das operações realizadas pelo codificador/decodificador, especialmente em aplicações em tempo real, o que limita o emprego de técnicas segundo (iii).

Métodos de correção de erro buscam, geralmente, minimizar os efeitos de perda de blocos de dados, procurando restaurar a informação faltante. Como a restauração de *pixels* perdidos em uma imagem ou vídeo pode ser realizada pelo lado transmissor, sem assim alterar o processo de codificação da informação [39], alguns métodos de correção de erro buscam resolver tal problema introduzindo considerações acerca da natureza da informação codificada. Diversos pesquisadores da área têm tecido, assim, considerações sobre a natureza dos sinais de imagem e de vídeo e, dependendo de como essas considerações sejam interpretadas, diferentes procedimentos de correção têm sido propostos. Todavia, a necessidade de compressão dos sinais a serem transmitidos, considerando-se as exigências de largura de banda, cria um obstáculo: as redundâncias temporal e espacial, características potencialmente úteis no procedimento de correção, são em grande parte suprimidas, limitando severamente a capacidade de correção dos procedimentos usualmente aplicados.

Para melhor situar este trabalho, este capítulo enuncia as aproximações mais recentes para o problema da correção de erro em imagens degradadas, como também as soluções atualmente propostas para tal. A Seção 2.2 apresenta as características mais importantes do sinal de imagem, representando brevemente os padrões vigentes para a codificação desse tipo de sinal. A Seção 2.3 discute os requisitos básicos e as limitações das atuais estratégias desenvolvidas para detecção e correção de erro em imagens. A Seção 2.4 apresenta o efeito da propagação de erros de bits para imagens transmitidas. A Seção 2.5 apresenta o algoritmo de codificação de imagens JPEG2000, que se baseia em uma concepção de representação de imagem em *wavelets* similar à proposta desse trabalho. Por fim, são apresentadas conclusões que justificam o uso da técnica de correção proposta, segundo a análise das características de decaimento dos coeficientes da decomposição *wavelet* de uma imagem degradada a ser corrigida.

2.2 PADRÕES DE CODIFICAÇÃO DE IMAGENS

Embora o presente trabalho tenha como foco resolver o problema de degradação de imagens estáticas, sua aplicação pode ser estendida para solucionar o paradigma de que a redundância temporal exigida para o tratamento de outros sinais, como os de vídeo, não faz uso da redundância espacial, no que é útil a compreensão de como canais de comunicação podem afetar sinais de informação visual, seja imagem ou vídeo.

Nesse contexto, pode-se definir um sinal de vídeo como constituído de várias imagens estáticas dispostas em sequência no tempo, denominadas quadros, e que, apresentadas a uma taxa suficientemente rápida, causam ao olho humano a impressão de movimento contínuo. Esse recurso é produzido mediante o aproveitamento da limitação de velocidade do olho humano para perceber alterações nas imagens. Dado o grande volume de informação que uma sequência de vídeo apresenta, é possível que alguma parte da informação seja deteriorada ou totalmente perdida durante, por exemplo, o processo de transmissão. Para que se definam técnicas que controlem sua transmissão, em termos de suprimir ou atenuar eventuais distorções, apresenta-se aqui a padronização das diferentes categorias de sinais de vídeo, já que há diferentes requisitos para cada aplicação na manipulação desses sinais. Quatro categorias podem ser enumeradas:

- Sinais de videofone – são sinais que devem possuir a menor largura de banda possível, para permitir o tráfego em linhas telefônicas convencionais. Suas características são: tamanho de quadro reduzido e baixa taxa de quadros, tipicamente 176×144 pixels, com 5 a 10 quadros/s. Os algoritmos dessa categoria se preocupam basicamente com a quantidade de informação a ser transmitida que pode trafegar em canais de 64 kbits/s, por exemplo.
- Teleconferência – necessita de um maior tamanho de quadro, bem como maior resolução temporal. São usuais quadros de 352×288 pixels, transmitidos a taxas de 10 a 15 quadros/s. O canal ocupado será maior do que o ocupado por sinal de videofone, sendo necessário o uso de vários canais telefônicos simples ou o uso de um canal de maior capacidade, como

os proporcionados pela rede digital de serviços integrados (*integrated service digital network* - ISDN), em que canais da ordem de 384 kbits/s são utilizados.

- Aplicações que primam pela qualidade, como multimídia e vídeo CD. As taxas podem chegar a até 2.0 Mbits/s. O tamanho dos quadros é da ordem de uma imagem de televisão (TV) convencional, 520×480 pixels, e a taxa de quadros é de 15 a 30 quadros/s.
- Televisão digital de alta definição (HDTV), com taxas elevadas e alta qualidade. A maior parte dos sistemas digitais de televisão de alta definição trabalha com um aumento aproximado de 100% do número de pixels horizontais e verticais em relação aos sistemas de televisão de definição padrão (SD) - é proposto 1MB por quadro, com aproximadamente 1000 linhas de 1000 pontos horizontais cada, e taxas de quadros de 30, 60 120 e até mesmo 240 quadros/s.

De forma a possibilitar a comunicação entre diferentes países, é necessária a definição de padrões internacionais. Para tal, organismos internacionais, tais como ITU-T e ISO têm elaborado recomendações de forma a permitir a comunicação entre dispositivos de diferentes características e procedências, definindo um formato comum de codificação e armazenamento digital, e estabelecendo padrões para representações codificadas de imagem, vídeo e áudio associado, para armazenamento em mídia digital, TV a cabo, dentre outras aplicações. Desses esforços, tais organismos têm recomendado a definição de padrões de codificação, tais como ISO-JPEG e ISO-MPEG (que abrangem respectivamente, para a codificação de imagens, os padrões JPEG, e, para vídeo, os padrões MPEG-1, MPEG-2 e MPEG-4 - *Moving Picture Experts Group Phase 1, Phase 2, Phase 4* [41]) e, pela ITU-T, as recomendações ITU-T.8x e ITU-H.26x (representando, respectivamente, codificação de imagens estáticas e codificação de vídeo [2, 42, 53, 54]).

No que diz respeito às recomendações contidas no padrão de codificação de imagens ISO-JPEG, estas especificam (tal como o fazem similarmente as recomendações ITU-T T.8x) uma estratégia de codificação de imagens estáticas. O JPEG utiliza um formato de compressão com perdas, baseado na transformada discreta de cosseno (*discrete cosine transform* – DCT). Um modelo perceptual

baseado no sistema psicovisual humano descarta informações de altas frequências, como por exemplo, transições bruscas de intensidade. Nesse sentido, o JPEG busca codificar a informação visual de *pixels* segundo um critério de compressão mais efetiva para menores valores de coeficientes no domínio da transformada, o que reduz consideravelmente o volume de dados, sem prejuízo notável na qualidade perceptual da imagem comprimida. Os coeficientes remanescentes são então ordenados e empacotados em uma sequência de bits, admitindo-se um controle externo da taxa de compressão (e consequentemente da qualidade visual da imagem comprimida em função inversa ao volume de informações remanescentes). O método de compressão JPEG traz usualmente perdas, significando que alguma porção da informação original é perdida e não pode ser recuperada, o que possivelmente afeta a qualidade da imagem original.

Em uma extensão do padrão JPEG para vídeo, a ISO definiu o conjunto de recomendações ISO-MPEG. O padrão MPEG possui quatro especificações distintas, sendo três delas ainda existentes [2, 41]. A primeira fase, denominada MPEG-1, é voltada a aplicações de multimídia. A segunda fase, chamada MPEG-2, destina-se à normatização de aplicações com HDTV. Uma terceira fase, que se chamaría MPEG-3, foi cancelada e englobada na especificação MPEG-2, levando ao surgimento da fase MPEG-4, direcionado à padronização de comunicação visual pessoal (sistemas de videoconferência) em ambientes com baixíssima disponibilidade de banda passante.

O MPEG foi originalmente concebido para atender à faixa de 1,5 Mbits/s, correspondente aos dispositivos de armazenamento digital, tais como CD-ROM, DAT (*Digital Audio Tape*), unidades de fita e de disco rígido, dentre outras. O MPEG trata ainda da codificação do sinal de áudio e sua sincronização com o sinal de vídeo. O padrão tem como principais características: o acesso aleatório, permitindo a decodificação de uma sequência a partir de qualquer ponto; a busca rápida, que permite a procura de dados a partir de poucos quadros; a reprodução em sentido reverso; sincronismo de áudio e vídeo; robustez a erros e editabilidade da sequência de vídeo. O algoritmo de codificação, definido no padrão MPEG, se baseia fundamentalmente em duas técnicas: estimativa e compensação de movimentos, para a extração da redundância temporal, e codificação por DCT, para a extração da redundância espacial.

A próxima seção apresenta a categorização do erro em sinais de imagem, apresentando-se sua classificação e seus efeitos sobre a estrutura de um sistema de transmissão em tempo real. Adicionalmente, destaca-se a importância da redundância na informação como variável que pode ser explorada para a concepção de procedimentos de correção.

2.3 ESTRATÉGIAS DE CORREÇÃO PARA IMAGENS CODIFICADAS

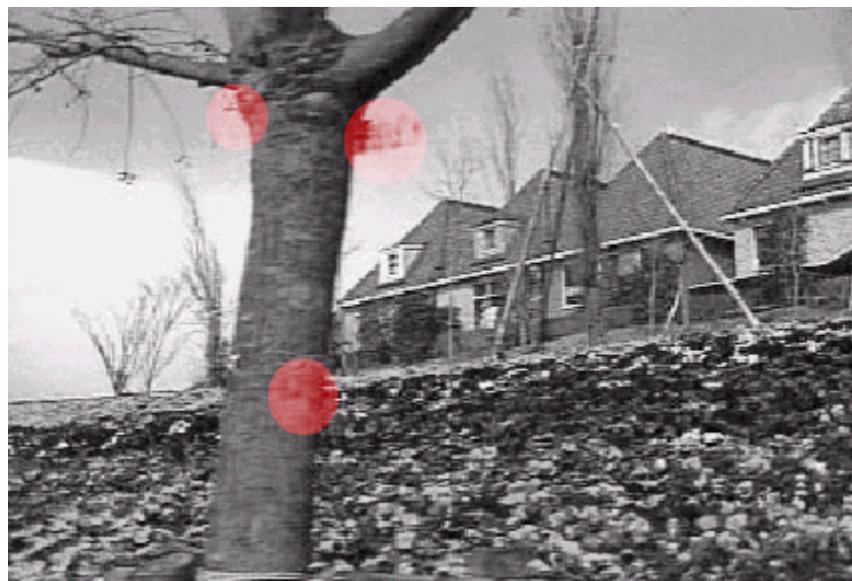
Erros em transmissão de imagens podem ser classificados basicamente em duas categorias: erros aleatórios de *bits* e erros sistemáticos (ou falhas). Erros aleatórios de *bits* são causados por imperfeições dos canais físicos, resultando em inversão, inserção e deleção de *bits* contidos em uma imagem codificada. Dependendo dos métodos de codificação e do conteúdo de informação afetado, o impacto de erros de *bits* aleatórios pode ser desprezível ou até mesmo imperceptível. Quando é utilizada alguma técnica de codificação de comprimento fixo de palavra, um erro aleatório de *bit* afetará somente uma palavra código, e o dano provocado será geralmente tolerável. Todavia, se alguma técnica de codificação VLC – por exemplo, a codificação de Huffman – for utilizada, erros aleatórios de *bits* podem comprometer o sincronismo da informação codificada de tal forma que muitos *bits* seguintes serão não-decodificáveis até que a próxima palavra código de sincronização apareça. Em alguns casos, mesmo após a sincronização ser obtida, a informação decodificada pode ainda ser inútil, já que não há modo de determinar quais localizações espaciais ou temporais correspondem à informação decodificada.

Falhas, por outro lado, podem ser causadas por perdas de pacotes em redes de comunicação, interrupções momentâneas no processo de transmissão, dentre outras causas. Erros aleatórios de *bits* em VLC podem também provocar falhas, uma vez que um erro em um único *bit* pode conduzir a muitos *bits* não-decodificáveis, por efeito da perda de sincronismo entre os *bits*, até que uma próxima palavra de sincronização seja encontrada. O efeito de falhas é mais destrutivo do que erros aleatórios de *bits* devido à perda ou dano de um segmento contíguo de *bits*.¹

¹ Uma vez que quase todas as técnicas de compressão de vídeo no estado-da-arte utilizam VLC de um ou outro modo, não há necessidade de tratar erros aleatórios de *bits* e falhas separadamente. O termo genérico “erros de transmissão” pode ser utilizado para se referir tanto a erros de *bits* quanto a falhas.

Para ilustrar tais artefatos provocados por erros de transmissão, a Fig. 2.2 ilustra dois quadros de vídeo reconstruídos de uma sequência de vídeo codificada em MPEG-2 e entregue a uma rede de transmissão sem fio. O vídeo é codificado a 6 Mbps, e a sequência de vídeo é dividida em GOPs, com cada GOP consistindo de 15 quadros. O primeiro quadro em cada GOP é intracodificado (quadro I), enquanto que os quadros remanescentes são codificados como preditos (quadros P). Cada quadro é particionado em fatias (*slices*), cada qual contendo 16×16 macroblocos na mesma linha. Um código é inserido ao início de cada *slice* de modo que o erro em um *slice* não afeta a decodificação do próximo. Qualquer perda no interior em um macrobloco de um *slice* tornará os demais macroblocos não-decodificáveis. Assim, os blocos danificados em um quadro I provocarão erros de reconstrução nos quadros P seguintes.

Fig. 2.2 - Dois quadros reconstruídos de uma sequência codificada em MPEG-2. (a) Quadro 16, com três perdas de célula; (b) Quadro 22, com uma única perda de célula. Artefatos estão marcados em vermelho.



(a)



(b)

Fonte: Tsekereidou et al., 1997.

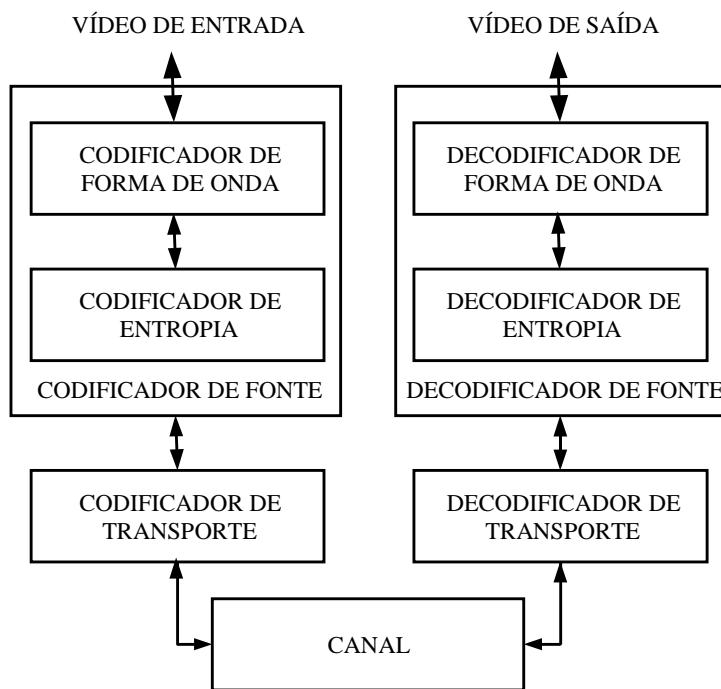
Neste exemplo, um macrobloco danificado é simplesmente substituído pelo macrobloco correspondente no quadro reconstruído anterior, o que provoca uma descontinuidade visível quando o bloco danificado pertence a uma região de rápido movimento. A primeira imagem mostrada na Fig. 2.2 é um quadro I, em que três perdas de célula conduzem a três *slices* danificados. As primeiras duas perdas ocorrem em dois *slices* consecutivos na parte superior do tronco da árvore; a terceira perda ocorre na parte inferior do tronco da árvore. A segunda imagem é um quadro P, que possui uma única perda de célula, sobre a localização da terceira perda no quadro 15. Pode-se observar o efeito da perda deste quadro bem como naqueles propagados do quadro 15. Quando o vídeo reconstruído é reproduzido em tempo real, essas distorções são visualmente desagradáveis e certamente não aceitáveis para aplicações de entretenimento.

No contexto da inviabilidade de retransmissão de sinais devido a requisitos de tempo real, técnicas para corrigir erros de transmissão em comunicação de imagens têm sido desenvolvidas ao longo de duas vertentes. Por um lado, esquemas tradicionais de recuperação e controle de erro para comunicações de dados têm sido estendidos para transmissão de imagens, objetivando a recuperação sem perdas. Exemplos de tais esquemas incluem correção de erro direto (*forward error correction* - FEC) ou, mais geralmente, codificação com controle de erro (*error control coding* - ECC), e requisição de retransmissão automática (*automatic retransmission on*

request - ARQ) [54]. Em uma abordagem distinta, técnicas de reconstrução de sinais e de correção de erros têm sido propostas para obter uma aproximação razoável do sinal original ou tentar fazer os erros no sinal de saída no decodificador menos perceptível para a visão humana. Ao contrário da transmissão de dados, para o qual o requisito ‘sem perdas’ é incondicional, a visão humana pode tolerar certo grau de distorção em sinais de imagem e de vídeo. Essa segunda vertente requer a detecção das posições de erro dentro de cada quadro, ou o conhecimento dos vetores movimento e coeficientes DCT, o que pode resultar em quadros borrados. Por outro lado, a primeira linha sugere a inserção de atrasos e de informação adicional (palavras-código), conduzindo a um maior uso de largura de banda.

Para melhor compreensão do efeito do erro inserido em um processo de transmissão de imagens, é necessário conhecimento pormenorizado da estrutura de transmissão sujeita a tais distorções. A Fig. 2.3 mostra um diagrama de bloco funcional de um sistema de comunicação de vídeo em tempo real. O vídeo de entrada é comprimido pelo codificador de fonte ao valor de *bit* desejado.

Fig. 2.3 - Diagrama em blocos de um canal de transmissão de vídeo.



Fonte: O Autor.

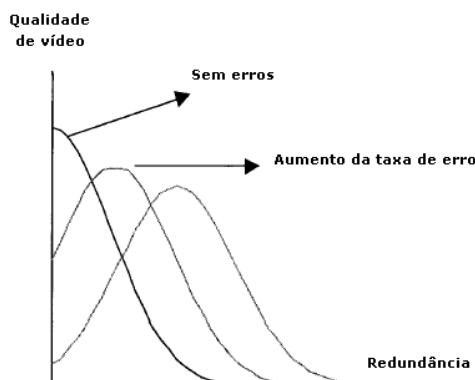
O *codificador de transporte* na figura anterior refere-se a um arranjo de dispositivos desempenhando codificação de canal, empacotamento e/ou modulação, e controle de nível de transporte usando um protocolo particular. Esse codificador de transporte é utilizado para converter o *bit* de saída corrente do codificador de fonte em unidades de dados convenientes para transmissão. No lado de receptor, as operações inversas são desempenhadas para obter o vídeo reconstruído de forma adequada para a reprodução. O mesmo processo pode ocorrer em sentido inverso, atuando o emissor como receptor e vice-versa, como ilustrado na Fig. 2.3 (flechas de duplo sentido) para enfatizar o fato de que, em algumas aplicações, ocorre fluxo bidirecional de informações, necessário em determinadas estratégias de controle e correção de erros.

A codificação de fonte pode ser realizada por forma de onda ou por entropia. O codificador de forma de onda é um dispositivo com perdas que representa o vídeo original no qual a intensidade de *pixels* ou uma simples diferença entre *pixels* é codificada. A codificação por entropia, por outro lado, é representada por um dispositivo sem perdas que não consideram a natureza da informação a ser comprimida, isto é, tratando todos os dados como sequências de *bits*, sem aperfeiçoar a compressão por meio do tipo de informação que se comprime, ignorando, assim, a semântica de tal informação. Exemplos de codificadores de entropia incluem codificadores de Huffman e codificadores aritméticos. Embora o codificador de forma de onda possa utilizar qualquer método de codificação conhecido, é geralmente implementado a partir de técnicas híbridas que utilizam DCT e predição por compensação de movimento. Tal esquema de codificação provou ter sido o mais efetivo para uma ampla faixa de aplicações, e é a base para todos os padrões atuais de codificação de vídeo [2, 42, 55, 57-59]. O codificador de transporte pode sofrer alterações para diferentes aplicações. Exemplos de protocolos de transmissão em tempo real incluem o H.221 na especificação H.320, o protocolo H.223 na especificação H.324, e o protocolo H.225 na especificação H.323 [56-62].

Em geral, para auxiliar na tarefa de detecção de erro e correção no decodificador, uma porção de redundância deve ser adicionada a cada nível de codificação, denominada *redundância de correção*. A Fig. 2.4 ilustra qualitativamente a dependência da qualidade do vídeo reconstruído com a redundância de correção e

com a taxa de erro do canal. Aqui, assume-se que a taxa total de *bits* é mantida fixa para a codificação de fonte e de canal. A figura em questão mostra que à medida que a taxa de erro do canal aumenta, uma maior porcentagem da largura de banda total deve ser alocada para a redundância de correção, de forma a se obter a melhor qualidade de vídeo. Elevando-se a redundância, diminui-se a informação relevante transmitida. O problema de correção de erro pode ser formulado de modo simplista como o projeto de um par codificador/decodificador de fonte e um par codificador/decodificador de transporte, de forma que a distorção de sinal no decodificador seja minimizada para um dado modelo de fonte, largura de banda total de canal, e características de erro do canal.

Fig. 2.4 - Relação entre a qualidade de vídeo reconstruído no decodificador e a quantidade de correção de redundância empregada sob valores de erro de canal diferentes. A largura de banda total utilizada para codificação de fonte e de canal é fixa.



Fonte: O Autor.

O problema apresentado é de difícil, se não impossível, resolução, devido às muitas variáveis envolvidas e o fato de que é frequentemente difícil modelar ou descrever de forma adequada tais variáveis. Primeiramente, o projeto de um codificador de fonte requer um modelo adequado da fonte para melhorar seu desempenho tanto em termos de eficiência de codificação quanto de robustez a erros de transmissão. Entretanto, fontes naturais de imagem e vídeo são fortemente não-estacionárias, e nenhum modelo efetivo ainda tem sido obtido para tais fontes. Além disso, as características de erro em alguns canais de transmissão são também não-estacionários, podendo variar significantemente durante uma sessão de serviço.

Por exemplo, uma rede ATM pode se tornar congestionada com o uso de multiplexação estatística para um grande número de fontes, dentre outras razões. Um videofone móvel pode operar a taxas de erros bastante diferentes, dependendo das condições meteorológicas, velocidades de movimento dos veículos, dentre outras. Além disso, outros fatores, tais como atraso de processamento, complexidade computacional e configuração de aplicativos tornam a solução do problema ainda mais difícil.

Muitas técnicas vêm sendo propostas na literatura para combater o problema do erro de transmissão sob diferentes pontos de vista. Na maior parte, se não em todos os casos, algumas variáveis são inicialmente fixadas e, então, uma solução localmente ótima é obtida. Há, em particular, três categorias de técnicas consideradas [63, 64]:

- Correção direta de erro – Esse termo agrupa técnicas em que o codificador desempenha o papel principal. O algoritmo de codificação de fonte e/ou mecanismos de controle de transporte são projetados para minimizar os efeitos dos erros de transmissão sem requerer qualquer procedimento de correção de erro no decodificador, ou para tornar a tarefa de correção de erro no decodificador mais efetiva. Exemplos de técnicas de correção direta de erro incluem FEC, codificação conjunta de fonte e canal e codificação em camadas.
- Correção de erro por pós-processamento – Abrange técnicas em que o decodificador desempenha a tarefa de correção de erro. Em geral, tais métodos tentam recuperar a informação perdida por estimativa e interpolação sem requerer qualquer informação adicional do codificador. Alisamento espacial e temporal, interpolação e filtragem compõem tal categoria.
- Correção interativa de erro – Codificador e decodificador trabalham cooperativamente para minimizar o impacto dos erros de transmissão. São exemplos de técnicas nessa categoria ARQ e codificação preditiva seletiva, baseada na realimentação advinda do decodificador.

É conveniente, classificações à parte, definir que critérios são empregados para atribuir maior relevância a alguma técnica de correção em particular. Obviamente, o mais importante critério de efetividade de uma técnica repousa sobre a qualidade de imagem que proporciona, após a correção. O atraso requerido é também uma variável crítica em transmissão bidirecional multiponto. O terceiro fator a ser considerado para avaliação é o *overhead* adicionado à informação pela redundância de correção no nível de codificação de fonte e/ou de transporte. Finalmente, a complexidade computacional é também uma importante característica de qualquer sistema. Deve-se observar que a prioridade de tais critérios pode depender da aplicação envolvida. Por exemplo, o atraso é muito menos importante para transmissão unidirecional de vídeo, como o *streaming* de vídeo na Internet e vídeo sob demanda, do que para videoconferência bidirecional. Além disso, algumas técnicas podem operar satisfatoriamente apenas para uma dada aplicação, enquanto que outras podem ser usadas ou adaptadas a uma ampla gama de aplicações. Por exemplo, a retransmissão pode ser útil para transmissão ponto-a-ponto, mas é de difícil emprego em aplicações multi-ponto. Por outro lado, a correção de erro por pós-processamento no decodificador pode ser considerada em quase todas as aplicações.

Aproximações para correção de erro podem ser classificadas como ativas ou passivas. Em correção ativa de erro, tanto retransmissão como códigos de controle de erro são utilizados, o que assegura reconstrução perfeita, no decodificador, do sinal corrompido, se a perda de dados não for severa. Uma vez que a perda de pacotes pode resultar na perda de linhas inteiras de uma imagem, técnicas de empacotamento baseadas no entrelaçamento de dados são também classificadas como ativas. Na categoria de técnicas passivas, encontram-se todas as técnicas de correção de erro por pós-processamento no decodificador (também denominada “correção passiva”), que utilizam dados espaciais, temporais ou ambos [65, 66]². Para essas técnicas, macroblocos corrompidos são substituídos por outros macroblocos estimados a partir de coeficientes, no domínio transformado, dos macroblocos adjacentes.

Antes que qualquer técnica de correção de erro seja aplicada ao decodificador, algumas técnicas exigem primeiramente a detecção da localização de

² O presente trabalho situa-se nesse grupo de técnicas.

eventuais erros de transmissão. Uma forma de se realizar a detecção de erro se faz pela adição de uma informação de cabeçalho. Por exemplo, em uma transmissão de vídeo realizada em pacotes, cada pacote deve conter um cabeçalho e um campo de informação útil [64]. O cabeçalho contém um campo com o número sequencial do pacote, que pode ser utilizado no decodificador para detecção de perdas de pacotes. O padrão de multiplexação H.223 utiliza tal estratégia para detecção de pacotes [66].

Outro método de detecção de erro se baseia no uso de FEC. Neste método, a correção de erro é aplicada a segmentos da saída do codificador. No decodificador, alguns erros de *bits* são detectados e possivelmente corrigidos. Por exemplo, o padrão H.223 utiliza FEC tanto para multiplexar cabeçalhos de pacotes e informação útil, quanto para detectar erros em tais campos dos pacotes [66, 67]. No padrão H.261, um código FEC de 18 *bits* é aplicado a cada quadro de transporte de vídeo de 493 *bits*, para detecção e correção de erro [42].

Para executar a detecção de erro no decodificador, características inerentes aos sinais de vídeo naturais também têm sido exploradas. Nos métodos propostos nas referências [68] e [69], valores de diferenças de *pixel* entre duas linhas vizinhas são utilizados para detectar erros de transmissão em modulação por codificação de pulso (*Pulse Code Modulation* – PCM) e codificação de pulso diferencial (*Differential Pulse Code Modulation* – DPCM). Quando a diferença é superior a um determinado limiar, o segmento de quadro atual é marcado como danificado. Em [70], Mitchell e Tabatabai propõem, em um algoritmo clássico, detectar a falha em um único coeficiente DCT examinando a diferença entre os *pixels* vizinhos em um bloco e seus quatro blocos vizinhos. No decodificador, quatro vetores-diferença separados são formados pelos valores de diferença entre o bloco atual e seus blocos adjacentes, sobre uma superfície de espessura de cerca de 1 *pixel*, para as 4 direções, respectivamente. A seguir, a DCT unidimensional é aplicada a tais vetores-diferença. Assumindo-se que a transição entre blocos é suave, os valores dos coeficientes DCT devem ser relativamente pequenos na ausência de erros de transmissão. Além disso, a posição do coeficiente danificado é também estimada. Uma importante contribuição ao tema foi dada por Zai *et al.*, ao estudarem o problema da detecção de erro no domínio da frequência [32]. Com tal abordagem, uma palavra-código de sincronização é inserida ao fim de cada linha de varredura dos blocos. Quando uma palavra-código de sincronização é capturada ao fim de

uma linha de varredura, o número de blocos decodificados é verificado. Quando é encontrada uma discrepância de valores, é assumido um erro e a posição do bloco com erro é determinada a partir do cálculo do erro quadrático médio entre os coeficientes de cada bloco na linha atual e os da linha anterior, para um bloco de 8×8 pixels. Um peso maior é dado para os coeficientes de baixas frequências, e um peso menor é atribuído para os coeficientes de altas frequências, de modo que a medida de distorção se aproxima ao máximo da percebida pelo sistema visual humano. O bloco com maior erro é identificado como bloco com erro, e dividido em dois blocos ou combinado a um bloco adjacente, dependendo de o número de blocos decodificados ser menor ou maior que um número preestabelecido. O procedimento se repete para todos os blocos danificados.

Conforme mencionado anteriormente, quando a VLC é utilizada no codificador de fonte, qualquer dano a um único bit pode provocar a perda de sincronismo de toda a sequência de vídeo, o que resulta em bits subsequentes não-decodificáveis. Entretanto, essa condição pode ser usada como um meio para detectar erros de transmissão. Na maior parte dos casos, a VLC não representa um código completo, ou seja, nem todas as possíveis palavras-código são legítimas. Assim, uma vez que o decodificador detecte uma palavra-código que não esteja presente em sua tabela de decodificação, é assumido um erro de transmissão. Além disso, a sintaxe incorporada à cadeia de bits pode ser também utilizada para detecção de erro. Por exemplo, se o tamanho do passo de quantização for igual a zero, ou o número de coeficientes DCT decodificados for maior que o número máximo de coeficientes (por exemplo, 64 coeficientes para um codificador DCT com blocos de 8×8 pixels), é detectado um erro de transmissão.

Geralmente, a detecção de erro por adição de informação de cabeçalho e/ou códigos FEC no nível de transporte é mais usual, mesmo que exija maior largura de banda de transmissão. O benefício de técnicas de detecção de erro no decodificador de vídeo que se baseiam na propriedade de suavidade dos sinais advém do fato de que tais técnicas não adicionam quaisquer bits além dos alocados pelo codificador de fonte. O uso de palavras-código de sincronização e/ou códigos VLC incompletos oferecem um compromisso: ao manter um pequeno grau de redundância ao processo de codificação, facilitam a detecção de erro no decodificador. Em situações realísticas, as técnicas FEC e VLC podem ser implementadas de forma conjunta.

Finalmente, é necessário dar especial atenção aos esquemas de correção de quadros de vídeo que exploram sua redundância temporal. Usualmente, tais técnicas, destacando-se nesse grupo aquelas baseadas em compensação de movimento, exercem um importante papel na tarefa de codificação de sequências de vídeo, em particular na codificação de sequências a baixas taxas de *bits*, tal como em videofone e HDTV [71]. Nos últimos anos, foram desenvolvidas várias técnicas de compensação de movimento, que, em conjunto com esquemas de codificação, têm sido utilizadas na codificação eficiente de sequências de vídeo.

A diminuição da propagação de erros em sinais de imagem é talvez o maior desafio à correção, o que se relaciona intrinsecamente à elevação da robustez de codificação. Trata-se de uma característica intrínseca à natureza dos sinais visuais: variações abruptas no plano de fundo provocadas por ruído podem ser interpretadas como variações abruptas de objetos ou mesmo de iluminação ambiente, e vice-versa, o que são desconsideradas na codificação. Isto sugere uma diferente abordagem do problema de distinção ruído/variação abrupta de cada quadro: a transformação da imagem, em blocos, do domínio espacial para o domínio transformado, que explora a alta correlação entre blocos espacialmente adjacentes. Embora mais eficientes para a tarefa de codificação, concentrando em poucas amostras a maior parcela de energia de cada quadro, boa parte das técnicas recentes de *inpainting* baseadas em tal artifício, como as representadas no Capítulo 5, não consegue realizar a separação precisa entre componentes de ruído e informação útil. Tal lacuna decorre da ausência da definição de uma figura de mérito no domínio transformado para aferição e determinação de como se deve proceder à correção de *pixels*.

Nesse contexto, o presente trabalho traz sua contribuição, apresentando uma técnica de correção de erro por pós-processamento para a reconstrução de blocos corrompidos e/ou perdidos, que propõe uma função de reconstrução formulada a partir da propriedade de regularidade, aplicada tanto no domínio espaço-frequência como no domínio do tempo. O desempenho do algoritmo proposto é comparado, por meio da PSNR, a diferentes métodos de correção de erro que operam no domínio transformado e no domínio híbrido espacial-temporal.

2.4 CONCLUSÕES

Este capítulo apresenta os conceitos fundamentais para a compreensão do problema de correção de erro em imagens. As limitações decorrentes da má caracterização de componentes de altas frequências, apresentadas pelos métodos de correção do estado-da-arte, motivam a concepção de um procedimento de reconstrução de imagens que possa ser aplicado após a codificação, de forma a privilegiar elementos de detalhes que sejam degradados por algum tipo de ruído.

Uma importante propriedade de sinais de imagem, no domínio da transformada *wavelet*, será explorada no próximo capítulo: a propriedade de regularidade. Explorando-se a redundância espacial (ou, de modo mais consistente, no domínio espaço-frequência) por meio do uso de tal propriedade, apresenta-se nos capítulos seguintes deste documento um procedimento de correção para imagens que, caracterizando de forma mais apropriada os componentes de altas frequências da informação visual, obtém resultados de detecção e correção de erro subjetivamente mais eficientes que os obtidos com as técnicas que se baseiam tão unicamente na exploração da informação espacial ou frequencial.

3 WAVELETS E A PROPRIEDADE DA REGULARIDADE DE HÖLDER

3.1 INTRODUÇÃO

O processamento de sinais é hoje essencial em praticamente qualquer área do conhecimento humano. Nas áreas exatas, o processamento de sinais já era uma realidade desde os primórdios dos sistemas analógicos. Nas áreas biológicas, a necessidade de se extrair informações valiosas sobre processos fisiológicos, através de seus sinais vitais, leva ao uso de técnicas cada vez mais sofisticadas para processá-los. Nas áreas humanas, a abordagem de imagens, sons, comportamentos e percepções, vistos como processos representáveis sob a ótica de sinais, traz a necessidade de se lançar mão de processamento de sinais para auxiliar no processo de reorganizar e compreender melhor o significado da informação. Dessa forma, o processamento de sinais é uma ferramenta fundamental de demanda crescente no mundo moderno. Particularmente em Processamento de Imagens, as *wavelets* encontram aplicações relevantes em compressão de imagens [72], reconhecimento facial [73] e reconhecimento de retina [74], entre outras aplicações.

O processamento digital de sinais traz várias vantagens em relação ao analógico. Isto porque a manipulação de um sinal discreto é realizada através de algoritmos computacionais, ou seja, por *software*, dispensando-se a necessidade de uma grande quantidade de módulos físicos (*hardware*), bancos de filtros, dentre outros. Além disso, a precisão alcançada no processamento é superior à analógica, e o processamento algorítmico é relativamente mais simples de se modelar e compreender. Nesse sentido, a transformada *wavelet* (*wavelet transform* – WT) faz parte de uma teoria relativamente nova, revelando-se uma poderosa e vantajosa ferramenta no processamento e análise de sinais para diversas aplicações [75, 76], com um vasto campo para exploração e experimentação, até que se estabeleça como uma ferramenta clássica.

Este capítulo apresenta uma breve introdução da teoria *wavelet*, evidenciando uma importante propriedade: a regularidade de Hölder [77]. Comenta-se inicialmente

a motivação para se explorar as *wavelets*, conceituando-a a seguir através de definições e exemplos. Define-se então a transformada *wavelet*, introduzindo-se o tema por meio das técnicas de análise e síntese, como também a análise multi-resolução de Mallat [78], intimamente ligada à teoria de bancos de filtros. Ao fim, apresenta-se formalmente o conceito de regularidade de Hölder e sua relação com a transformada *wavelet*, para a qual é dada ênfase à característica de correlação entre as sub-bandas. Tal propriedade servirá de motivação ao algoritmo de correção de erro proposto neste trabalho, que busca preservar o comportamento de decaimento de coeficientes através das escalas de decomposição *wavelet* de uma imagem.

3.2 MOTIVAÇÃO

Por meio das características de um sinal, é possível extrair informações relevantes sobre o processo ao qual está associado, tanto no domínio do tempo (espaço) como no domínio da frequência. Atualmente, uma ferramenta clássica na análise de sinais é a transformada de Fourier, que permite realizar o tratamento de sinais transformados no domínio da frequência. Assim, a transformada de Fourier mapeia um sinal, por exemplo $s(t)$, descrito no domínio do tempo (ou $s(x)$, no domínio da sequência contínua x) em um sinal $S(f)$, descrito no domínio da frequência. Entretanto, uma adequada resolução temporal (espacial) implica uma inadequada resolução em frequência, e vice-versa, de modo que é desejável o uso de transformadas que estabeleçam um compromisso entre localização temporal (espacial) e em frequência. Nessa classe de transformadas, situa-se a transformada *wavelet*, definida matematicamente na próxima seção, que possibilita representar um sinal em diversas escalas, introduzindo assim uma análise espaço-escala de sinais. Essa propriedade intrínseca à transformada *wavelet* a torna particularmente atraente para aplicações em análise multi-resolução.

À medida que forem sendo apresentadas certas propriedades da TW, ficarão mais evidentes as vantagens de se utilizá-la, em especial para os casos em que a TF não representa de forma adequada os resultados desejados.

3.3 WAVELETS

Wavelets são um conjunto de funções obtidas de uma função protótipo (dita *wavelet* "mãe") por meio de dilatações (operações em escala) e translações. A forma geral de uma "família" de wavelets é dada por [79]:

$$\psi_{c,b}(t) = \frac{1}{\sqrt{c}} \psi\left(\frac{t-b}{c}\right), \quad (3.1)$$

onde $\psi(t)$ é a *wavelet* "mãe" e $c, b \in \mathbb{R}^*$, com \mathbb{R}^* denotando o conjunto dos números reais não-nulos, c denotando o fator de escala e b denotando o fator de translação.

Existem vários tipos de wavelets na literatura [75, 76, 79], sendo que a escolha de um tipo particular de *wavelet* está quase sempre associada à aplicação em questão. Regras de construção de wavelets estão sendo propostas segundo as restrições e necessidades que cada aplicação específica impõe. Isso permite concluir que é possível gerar uma infinidade de wavelets diferentes e, particularmente, construir um conjunto de wavelets adequado a um processamento específico, levando à obtenção de resultados cada vez melhores.

Para uma função ser considerada uma *wavelet*, ela deve ser admissível. Assim, uma função $\psi(t) \in L^2(\mathbb{R})$ é dita admissível se

$$C_g = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty, \quad (3.2)$$

onde $\Psi(\omega)$ é a transformada de Fourier de $\psi(t)$ e $L^2(\mathbb{R})$ é o conjunto de todos os sinais integráveis quadraticamente (ou seja, de energia finita) e C_g é a constante de admissibilidade. A imposição de C_g finita é condição para que a integral da transformada inversa converja, ou seja, a operação anti-transformada só será possível se $C_g < \infty$.

Para uma dada função $g(t)$, contínua na vizinhança da origem, a condição $C_g < \infty$ implica que $\Psi(0) = 0$, ou, similarmente, que $\int_{-\infty}^{\infty} \psi(t)dt = 0$, isto é, $\psi(t)$ possui valor médio (componente DC) nulo.

A forma geral de uma família de *wavelets* é

$$\psi_{j,k}(x) = \psi(2^j x - k), \quad j, k \in \mathbb{Z}, \quad (3.3)$$

onde o parâmetro j é denominado fator de escala, e o parâmetro k , fator de translação. Tais funções são seccionalmente contínuas, constituindo uma base para $L^2(\mathbb{R})$. Adicionalmente, para o produto interno vale a regra [77, 78]

$$\int_{-\infty}^{\infty} \psi_{j,k}(x)\psi(2^j x - k)dx = 0, \quad j, k \in \mathbb{Z}, \quad (3.4)$$

ou seja, as dilatações e translações de $\psi_{j,k}(x)$ são mutuamente ortogonais para todos j, k inteiros.

É importante considerar que as funções de base da transformada de Fourier $e^{j\omega t}$ também formam uma base ortogonal em $L^2(\mathbb{R})$. No decorrer deste capítulo, as vantagens dessa propriedade ficarão mais evidentes.

3.4 TRANSFORMADA WAVELET

Uma forma de se conceituar a transformada *wavelet* se faz pela decomposição de sinais em funções de base ortogonais. Assim, um sinal $x(t)$ pode ser decomposto através de um conjunto ortogonal de funções de base, que correspondem a *wavelets* dilatadas e transladadas $\psi_{j,k}(t)$, $j, k \in \mathbb{R}$, $j \neq 0$ [78]. O operador transformada pode ser então definido:

$$\Psi_x(j, k) = \int_{-\infty}^{\infty} x(t) \psi_{j,k}^*(t) dt . \quad (3.5)$$

Esta é a forma da transformada *wavelet* contínua (TWC). A operação em (3.5) pode ser interpretada como o produto interno do sinal $x(t)$, que se quer analisar, com a família de *wavelets* $\psi_{j,k}(t)$:

$$\Psi_x(j, k) = \langle x(t), \psi_{j,k}^*(t) \rangle \quad (3.6)$$

Como todas as *wavelets* derivam de um mesmo protótipo $\psi(t)$, nenhuma escala específica é privilegiada na análise. A transformada $\Psi_x(j, k)$ pode ser representada graficamente em um plano bidimensional, onde j denota a escala/frequência e k , a dimensão temporal.

Outra visão alternativa da TW é baseada na interpretação de bancos de filtros [77-79]. A transformada equacionada em (3.5) pode ser reescrita como

$$\Psi_x(j, k) = \sqrt{j} \int_{-\infty}^{\infty} x(jt) \psi^* \left(\frac{t-k}{j} \right) dt . \quad (3.7)$$

Na equação (3.5), quando o fator de escala cresce, a resposta ao impulso do filtro associado se alarga no tempo, analisando o sinal sobre um intervalo de tempo maior. Equivalentemente, na equação (3.7), quando o fator de escala cresce, o sinal $x(jt)$ é analisado através de uma janela de comprimento constante.

A transformada *wavelet* é uma operação linear. Entretanto, quando confrontada com a transformada de Fourier, a análise em frequência possui uma cobertura logarítmica, devido ao operador de escala 2^j . O banco de filtros de análise é composto de filtros passa-banda com largura de banda relativa constante, isto é,

$$\Delta f / f = \text{constante} . \quad (3.8)$$

Em outras palavras, ao invés da resposta em frequência do filtro ser regularmente espaçada no eixo das frequências, ela é espalhada em uma escala

logarítmica. O processo de análise resulta, portanto, em um conjunto de coeficientes *wavelet* que exprimem a “proximidade” ou similaridade do sinal original em relação àquelas funções de base. Na condição de que é possível representar qualquer sinal genérico como uma decomposição em *wavelets*, é natural que se queira recuperar a forma de onda original somando blocos construtivos elementares.

A síntese de $x(t)$ a partir dos coeficientes *wavelet*, isto é, sua reconstrução, consiste na soma de todas as projeções (ortogonais) do sinal decomposto. Isso pode ser equacionado como:

$$x(t) = \frac{1}{C_g} \iint \Psi_x(j, k) \psi_{j,k}(t) \frac{djd\kappa}{j^2}. \quad (3.9)$$

Em resumo, um amplo potencial de aplicações da transformada *wavelet* decorre do fato de que ela supera a TF em termos de resolução: enquanto a TF possui resolução fixa para qualquer posição, ou seja, sinais de frequências diferentes são analisados com uma mesma resolução temporal e vice-versa, a TW apresenta o conceito de multi-resolução (que será detalhado mais tarde). Quando tal conceito é analisado de forma análoga à operação de bancos de filtros, a resolução no tempo cresce com a frequência (central) do filtro, de forma que Δt e Δf variam com a frequência central do filtro (passa-banda). A resolução temporal Δt torna-se melhor nas frequências altas, enquanto a resolução Δf torna-se melhor nas baixas frequências.

Fazendo a conexão destes conceitos de resolução variável com o conceito de escala na transformada *wavelet*, observa-se que em escalas maiores (j maior) o sinal $x(t)$ é analisado globalmente, isto é, sem detalhes, logo com menor resolução temporal. Se o que se deseja é magnificar $x(t)$, reduz-se a escala (j menor) e obtém-se melhor resolução temporal. Na visão global (j maior) a resolução em frequência é melhor, enquanto que em escalas menores perde-se o poder de resolução em frequência. Quanto maior o fator de escala j , para uma frequência alta a resolução em frequência é também alta (e baixa no tempo).

3.5 ANÁLISE MULTI-RESOLUÇÃO

A análise multi-resolução, concebida por Mallat [78], estabelece uma sequência de subespaços agrupados $V_j \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$, onde $L^2(\mathbb{R})$ é o espaço de todas as funções integráveis quadraticamente, com as seguinte propriedades:

1. $\dots V_{-1} \subset V_0 \subset V_1 \subset V_2 \dots$
2. $f(x) \in V_j \Rightarrow f(2x) \in V_{j+1}$
3. $f(x) \in V_j \Rightarrow f(x - 2^{-j}k) \in V_j$

Define-se ainda o subespaço W_j como o complemento ortogonal de V_j em V_{j+1} :

$$V_{j+1} = V_j \oplus W_j, \quad (3.10)$$

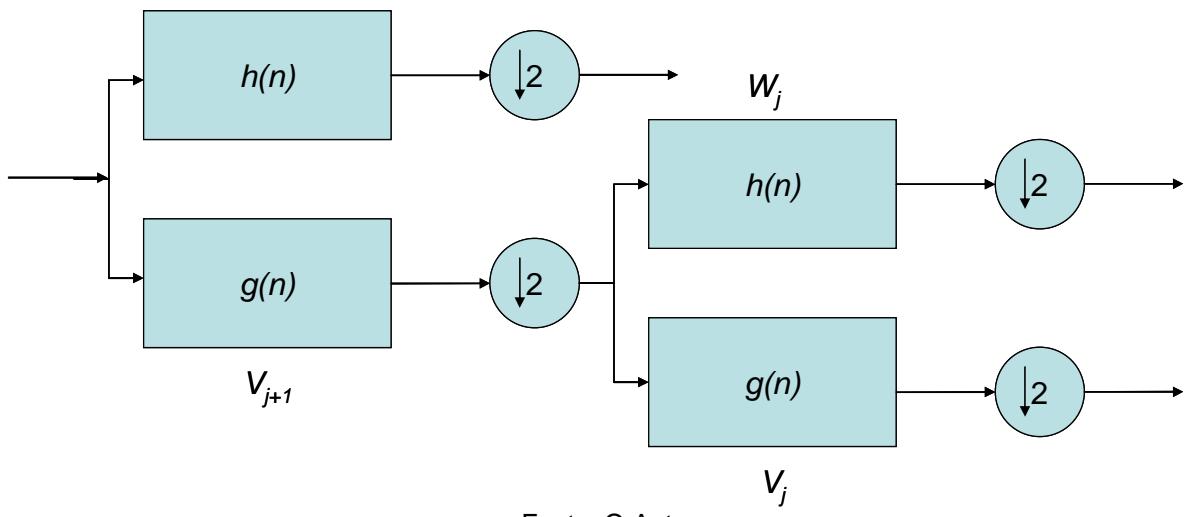
onde ‘ \oplus ’ representa a união de subespaços ortogonais. Logo, os subespaços W_j contém os detalhes necessários para se ir de V_j a V_{j+1} , o que por recursão conduz a

$$V_{j+1} = W_j \oplus W_{j-1} \oplus W_{j-2} \oplus W_{j-3} \oplus W_{j-4} \dots \quad (3.11)$$

A análise multi-resolução de um sinal pode ser visualizada como a separação desse sinal em sub-bandas, como ilustrado na Fig. 3.1. A cada estágio de decomposição, a porção passa-altas (equivalente a W_j) corresponde à diferença entre a porção passa-baixas corrente (V_j) e a porção passa-baixas anterior (V_{j+1}). O esquema ilustrado representa uma “estrutura de banco de filtros em árvore”, equivalente a uma transformação *wavelet* discreta implementada utilizando uma estrutura de filtros digitais com um fator de sub-amostragem de 2 [70,71]. As respostas ao impulso $h(n)$ e $g(n)$ representam, respectivamente, estágios de

filtragem passa-altas e passa-baixas do esquema de banco de filtros, enquanto que o operador $\downarrow 2$ simboliza a operação de dizimação (remoção de amostras ímpares de um sinal). Assim, a cada etapa de filtragem, reduz-se à metade a largura de banda do sinal.

Fig. 3.1 - Árvore de banco de filtros: iteração do esquema de codificação em sub-bandas.



Fonte: O Autor.

Uma vez adquirido o conhecimento mínimo acerca das *wavelets*, a seção seguinte trata de uma importante propriedade dessas funções, a regularidade de Hölder, que permitirá o desenvolvimento matemático da técnica de correção de erro proposta nesse trabalho.

3.6 PROPRIEDADE DA REGULARIDADE DE HÖLDER

A transformada *wavelet* fornece um meio pelo qual a suavidade local de um sinal pode ser quantificada. Do ponto de vista matemático, a suavidade é representada pela taxa de decaimento dos coeficientes *wavelet* em sucessivas escalas de decomposição do sinal, que pode ser analisada pela observação da

propriedade de regularidade de Hölder (caso geral da propriedade de continuidade de Lipschitz [79]).

A regularidade local de uma função $f(x)$ em um ponto a pode ser analisada segundo a noção de continuidade de Lipschitz: uma função $f(x)$ é dita r -Lipschitz em a , $0 \leq r < 1$, se

$$|f(x) - f(a)| \leq K|x - a|^r, \quad (3.12)$$

onde K é uma constante arbitrária. Uma descontinuidade em degrau, por exemplo, possui expoente de Lipschitz $r = 0$. A consideração $r > 1$ é possível se (3.12) for satisfeita para derivadas de mais alta ordem.

As seguintes condições valem para o comportamento local de coeficientes wavelet próximos a um ponto r -Lipschitz [79]: se $f(x)$ é r -Lipschitz em $x = a$, $r < N$ e $\psi(x)$ possui ao menos N momentos³ desvanecentes, então

$$\max |\psi_{j,k}|_{(j,k) \in A} \leq K 2^{-j(r+1/2)}, \quad (3.13)$$

onde A contém os pares (j,k) para os quais a pertence ao suporte compacto de $\psi_{j,k}(x)$.

Uma extensão do conceito de continuidade de Lipschitz ($r > 1$) é feita requerendo que (3.12) seja satisfeita para $f(x)$ e suas derivadas de n -ésima ordem. Uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ possui regularidade de Hölder (ou expoente de Hölder)

$$\alpha = n + r, \text{ com } n \in \mathbb{N} \text{ e } 0 \leq r < 1, \quad (3.14)$$

se existe uma constante $K < \infty$ tal que

³ O termo “momento” deriva do fato de que seu desvanecimento equivale a dizer que as primeiras p derivadas da transformada de Fourier do filtro wavelet são todas nulas quando avaliadas na origem. Isto é perfeitamente análogo à ideia probabilística de uma “função geradora de momentos” de uma variável aleatória, que é basicamente a transformada de Fourier de sua função densidade de probabilidade, e a n -ésima derivada avaliada na origem fornece o n -ésimo momento dessa variável (por exemplo, seu valor esperado, o valor esperado de seu quadrado, etc.).

$$|f^{(n)}(y) - f^{(n)}(x)| \leq K |y - x|^r \quad \forall x, y \in \mathbb{R}, \quad (3.15)$$

onde $f^{(n)}$ indica a derivada de n -ésima ordem de f , e r corresponde ao expoente de Lipschitz [79]. O expoente de Hölder estabelece o número de derivadas contínuas de uma função. Funções com valores grandes de expoente de Hölder são matematicamente e visualmente suaves, enquanto que funções com valores pequenos de expoente de Hölder são associadas a funções mais irregulares [80].

Isto permite, agora, estimar matematicamente o expoente de Hölder. Considere-se que a transformada wavelet não-dizimada (em que a etapa de filtragem não é sucedida pela dizimação) é um eficiente meio para medir a suavidade de uma função, permitindo sintetizar um sinal $f(x)$ através de operações de translações e dilatações de uma função wavelet “mãe” $\psi(x)$ [78]. Logo, considerando (3.3), um sinal apresenta expoente de Hölder α se existe, para cada fator de escala inteiro j , uma constante K_j tal que os coeficientes da transformada wavelet $\Psi_{j,k} = \langle f, \psi_{j,k} \rangle$ obedeçam à inequação

$$|\Psi_{j,k}| \leq K_j 2^{-j(\alpha+1/2)} \quad (3.16)$$

para qualquer $j \in \mathbb{Z}^+$, $k \in \mathbb{R}$ [79, 81-83]. Esse teorema caracteriza a regularidade de uma função pelo decaimento da magnitude dos coeficientes de sua transformada wavelet através das sucessivas escalas de decomposição, definindo a similaridade entre escalas decorrentes de um processamento multi-resolução.

No contexto da caracterização da suavidade de uma função por meio da regularidade, pesquisas em análise multiresolução de imagens têm indicado a presença de características similares em forma, porém distintas em suporte espacial para escalas diversas [77, 80, 83, 84]. Funções mais suaves exibem maior similaridade entre escalas, matematicamente definida pelo teorema do decaimento dado por (3.16). Pode-se mostrar que a correlação entre coeficientes de sub-bandas wavelet s_m e s_n , para as respectivas escalas 2^m e 2^n , é limitada numericamente por [82]:

$$|\langle s_m, s_n \rangle| \leq K_{m+n} 2^{-(m+n)(\alpha+1/2)}, \quad (3.16)$$

onde K_{m+n} é uma constante correspondente ao par (m,n) e α é o expoente de Hölder. Tal inequação indica que a similaridade entre as escalas decresce exponencialmente à medida que a regularidade da função analisada cresce, observação corroborada por dados experimentais e pela literatura da área [78-83].

3.7 CONCLUSÕES

A regularidade de Hölder mostra-se uma útil propriedade matemática para a análise de sinais correlacionados. Com base nessa perspectiva, este trabalho sugere sua utilização como um critério matematicamente consistente de decaimento de coeficientes correspondentes para separação da informação corrompida em imagens, já que os algoritmos recentemente desenvolvidos que se dedicam a tal tarefa, quando operam simultaneamente nos domínios do espaço e da freqüência, não definem de maneira formal uma grandeza relacionada à função de correlação dos coeficientes de cada quadro no domínio transformado. Nesse contexto, a regularidade se apresenta como elemento indicativo de similaridade de informação através das escalas da transformada *wavelet*.

O próximo capítulo aborda detalhadamente a utilização da propriedade da regularidade para a proposição de uma técnica de *inpainting* de erro em imagens.

4 ALGORITMO PROPOSTO DE DETECÇÃO E CORREÇÃO DE ERROS

4.1 INTRODUÇÃO

Muitos autores [24, 26, 30, 34, 84, 85] descrevem basicamente duas abordagens para correção de erro em imagens: a correção por predição temporal (*temporal predictive concealment* - TPC) e correção por predição espacial (*spatial predictive concealment* – SPC). A TPC explora a redundância entre quadros sucessivos no tempo, permitindo a recuperação de blocos danificados em áreas sem movimento nos quadros. A SPC realiza a interpolação de áreas danificadas utilizando a vizinhança espacial dos dados perdidos, apresentando bom desempenho para regiões com movimento. Assim, um algoritmo de correção que incorpore mutuamente a funcionalidade de algoritmos baseados em TPC e SPC é um tópico de interesse em pesquisa na área de processamento de vídeo.

Diversos algoritmos SPC foram propostos de forma clássica na literatura. Sun e Kwok propõem um algoritmo de interpolação espacial baseado na técnica de POCS [86]. Lee *et al.* propuseram um esquema de correção baseado em filtragem interpoladora espacial e recuperação de coeficientes DCT através de Lógica Fuzzy [87]. Wang *et al.* apresentaram uma técnica de otimização em que os coeficientes DCT são estimados pela imposição dos limites de suavidade entre os valores de intensidade de *pixels* adjacentes [88-90, 92]. Entretanto, a complexidade computacional de tais algoritmos não é aceitável para aplicações em tempo real. Hemami *et al.* propuseram um algoritmo que reconstrói blocos corrompidos a partir de uma combinação linear de blocos adjacentes [91, 93, 94]. Utilizando um critério similar ao usado em [87], são obtidos valores ótimos de pesos que minimizam tal critério. Entretanto, resultados de simulação mostram que o desempenho do algoritmo não é satisfatório em termos subjetivos.

A tarefa de correção de erro necessita, para um satisfatório desempenho, de que a detecção de blocos corrompidos seja eficaz (geralmente implementada por um algoritmo de detecção de erro – *Error Detection Algorithm*, EDA). Desta forma, a tarefa de correção requer testes em ambientes realísticos de transmissão para avaliação de sua eficácia, como por exemplo, em redes ATM [95-97]. Entretanto, há muitas aplicações em que erros de *bits* são mais críticos do que perdas de células como, por exemplo, em comunicação sem fio. Assim, para se desenvolver um algoritmo de correção de erro genérico⁴, é essencial haver um procedimento prévio de detecção de erro que localize os blocos corrompidos. Chen [98] propôs um sistema de correção de erro genérico, denominado decodificador de vídeo MPEG não-supervisionado, que inclui um EDA.

Há, na literatura, trabalhos que não exploram a correlação entre blocos para a correção de erro. Por exemplo, Hadar *et al.* [99] desenvolveram uma técnica de correção de erro por pós-processamento que combina métodos de correção espacial e temporal, perfazendo, a partir de uma árvore de decisão, a correção do bloco corrompido segundo uma ou outra técnica. Embora o algoritmo obtenha bons resultados em comparação a outros que operam apenas nos domínios espacial ou temporal isoladamente, seu desempenho é restrito a condições de movimento contínuo ou ausência de movimento, não respondendo bem em situações de pouco deslocamento ou mesmo movimento de fundo. Essa lacuna poderia ser satisfatoriamente coberta, se na tarefa de correção fosse considerada, de alguma forma, a correlação da informação visual em algum domínio transformado [100].

Há diversas técnicas de correção de quadros baseadas nos domínio espacial ou da frequência que possuem um embasamento teórico bem desenvolvido [100-102]. Entretanto, essas técnicas falham pelo fato de que os filtros projetados para correção capturam informação do quadro inteiro, em vez de capturar informações de componentes locais de frequência. Adicionalmente, erros de aproximação, decorrentes da utilização da representação de sinais no domínio da frequência, podem ser mensurados apenas no espaço bidimensional, não incorporando qualquer informação acerca da suavidade de cada quadro

⁴ Por genérico, denota-se o sistema de correção de erro que não depende de qualquer informação, fornecida por algum dispositivo externo ao sistema, acerca da localização do erro.

(diferenciabilidade). Isso significa que no espaço L_2 não é possível capturar, por exemplo, a diferença entre uma borda aguda e uma borda suave.

Como a informação de vídeo não é perdida no processo de transformação linear, dados altamente correlacionados no domínio espacial estão aptos a serem processados no domínio transformado, representando fielmente o sinal de vídeo no domínio temporal. O uso de transformação linear do sinal de vídeo implica, entretanto, a redução da redundância temporal e espacial da informação, o que não é desejável quando se busca exatamente realizar a separação da informação desejável de sua contraparte indesejada. Com tal objetivo, este capítulo apresenta a proposição de um algoritmo de *inpainting* de fácil implementação, a partir da análise multi-resolução da informação visual inerente às imagens analisadas. O algoritmo explora a redundância de informação contida nas características de decaimento dos coeficientes *wavelet* da imagem, assumindo que não se tem qualquer conhecimento da localização de blocos corrompidos, implementando-se isoladamente um EDA: a detecção é realizada através da análise do decaimento dos coeficientes de blocos em escalas sucessivas da decomposição *wavelet*, determinando-se a localização espacial de cada bloco corrompido. Tal informação é extraída pela observação da função de correlação cruzada entre as características de decaimento médio dos coeficientes daquele bloco. Na etapa específica de correção, essa curva de decaimento médio é reconstruída no bloco corrompido, e os novos coeficientes são determinados.

A grande diferença da técnica aqui proposta em relação a outras no estado-da-arte [90, 92, 98, 101] está na utilização do expoente de Hölder (definido no capítulo anterior), medido em altas frequências do bloco corrompido, como figura de mérito para a correção que admite um ajuste preciso das amplitudes dos *pixels* reconstruídos. Isto é corroborado pelo fato de que *wavelets* podem construir bases em uma grande variedade de espaços vetoriais bidimensionais, levando em conta, para os cálculos, a suavidade de uma função. Tais espaços, chamados Espaços de Besov, podem ser interpretados como espaços generalizados L_p que também consideram a diferenciabilidade de uma função [84, 85].

Como o expoente de Hölder caracteriza de forma consistente as sub-bandas *wavelet* de uma imagem, os blocos corrompidos que apresentam, em sucessivas

escalas, uma função de correlação cruzada variável entre seus coeficientes de sucessivas sub-bandas *wavelet*, são identificados como blocos degradados por ruído, uma vez que tal função é representativa da característica de decaimento dos próprios coeficientes, como foi observado pela relação direta entre (3.16) e (3.17).

O capítulo é organizado como segue. A Seção 4.2 apresenta o princípio de operação do algoritmo proposto de detecção de erros (EDA), descrevendo-se, com exemplos qualitativos, o papel fundamental da análise da função de correlação cruzada entre as características de decaimento dos coeficientes de blocos ao longo de sucessivas escalas *wavelet*. A Seção 4.3 apresenta o algoritmo de correção, detalhando-se o processo de reconstrução dos coeficientes a partir da detecção prévia dos blocos de *pixels* danificados. Simulações executadas utilizando o software Matlab, para quadros de duas sequências de vídeo clássicas na literatura e para dois tipos de ruído aplicados aos sinais de vídeo, são apresentadas na Seção 4.4, sendo avaliadas tanto em termos da PSNR quanto subjetivamente. Conclusões são apresentadas ao fim do capítulo, com comentários acerca dos resultados obtidos.

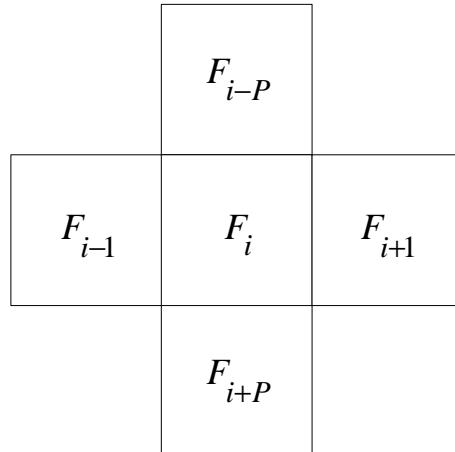
4.2 ALGORITMO DE DETECÇÃO DE ERROS

Para o caso de codificação de imagens, propõe-se definir uma figura de mérito que viabilize a estimativa da informação deteriorada por erro. Dessa forma, o estudo da evolução das características dos *pixels* é útil para a estimativa do comportamento de tais coeficientes em várias escalas.

Um estudo preliminar da correlação entre as características de decaimento pode ser realizado de forma a se estabelecer um modelo de quantificação matemática dos coeficientes *wavelet* ao longo da decomposição *wavelet*. Considere-se uma imagem dividida em $P \times Q$ blocos de $N \times N$ *pixels* cada. Denotando-se um bloco dentro desta imagem por F_i , $1 \leq i \leq P \times Q$, os blocos esquerdo e direito são representados, respectivamente, por F_{i-1} e F_{i+1} . Dentro de cada bloco, o símbolo $F_i(m, n)$ representa o n -ésimo *pixel* horizontal da m -ésima

linha.

Fig. 4.1 - Bloco a ser analisado F_i e seus 4 blocos adjacentes F_{i-P} , F_{i+P} , F_{i-1} e F_{i+1} .



Fonte: O Autor.

Representando os elementos de cada bloco em termos de seus coeficientes *wavelet* no campo discreto, tem-se:

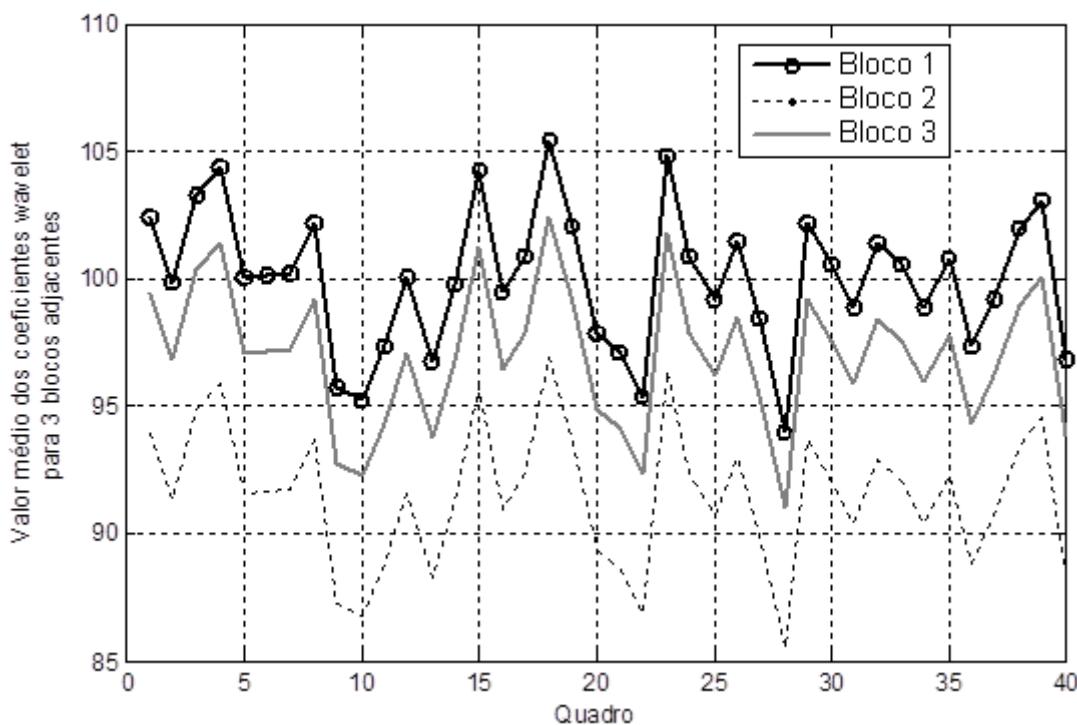
$$F_i(m, n) = \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} \Psi_{i,j,k,r,s} \psi_r(m) \psi_s(n), \quad (4.1)$$

onde $\Psi_{i,j,k,r,s}(m, n)$ são os coeficientes do bloco F_i na frequência espacial (r, s) , dados os parâmetros de escala e de translação (j, k) e $\psi_r(m), \psi_s(n)$ são as funções-base *wavelet* nas direções vertical e horizontal, respectivamente. Por simplicidade de notação, a frequência espacial (r, s) é omitida nas expressões seguintes.

A representação de blocos no domínio transformado permite a observação do comportamento dos coeficientes em diferentes escalas. Tomando-se no bloco F_i o valor médio dos coeficientes $\Psi_{i,j,k}(m, n)$ no primeiro nível de decomposição ($j=1$), bem como para os blocos vizinhos anterior, posterior, superior e inferior F_{i-1} , F_{i+1} , F_{i-P} e F_{i+P} , respectivamente, os valores médios de $\Psi_{i-1,j,k}(m, n)$, $\Psi_{i+1,j,k}(m, n)$,

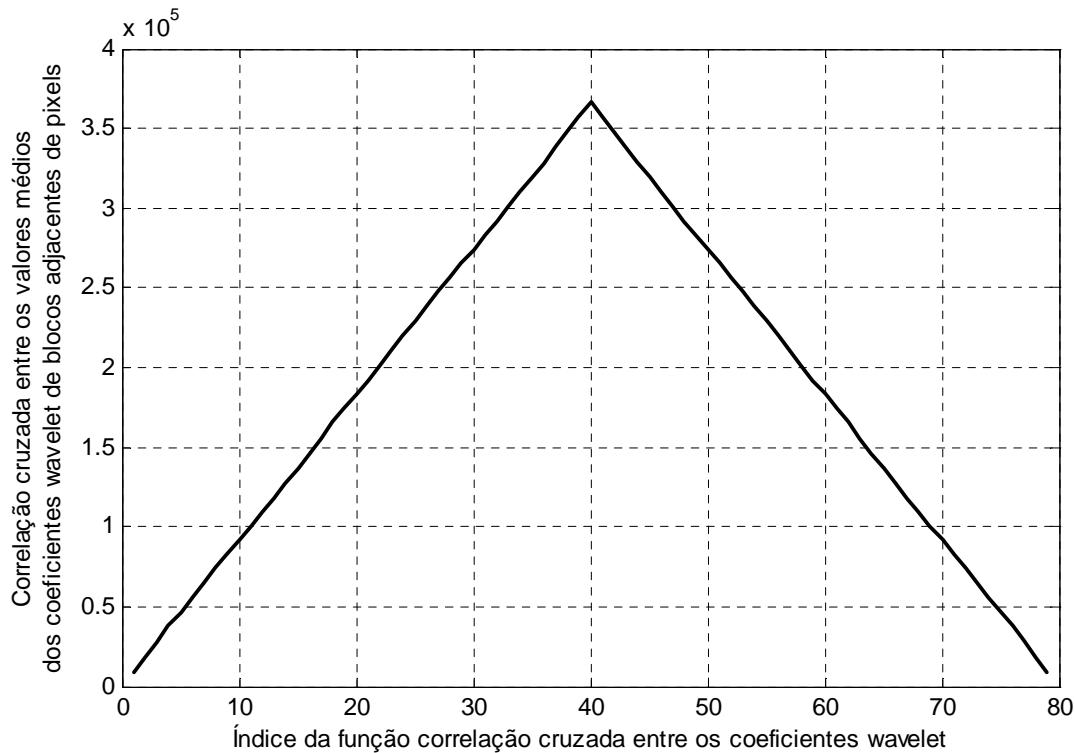
$\Psi_{i-P,j,k}(m,n)$ e $\Psi_{i+P,j,k}(m,n)$, observa-se o elevado grau de correlação cruzada entre esses coeficientes. Considerem-se, por exemplo, 3 sinais correlacionados, ao longo de 40 quadros sucessivos no tempo da sequência de vídeo ‘Foreman’, constituídos cada um dos valores médios dos coeficientes *wavelet* no primeiro nível de decomposição de 3 blocos espacialmente adjacentes em cada quadro F_{i-1} , F_i e F_{i+1} , para a direção horizontal de decomposição. A Figura 4.2 apresenta tais curvas de valores médios de coeficientes, em que se observa a forte correlação matemática entre as mesmas.

Fig. 4.2 - Três sinais gerados a partir de amostras, ao longo do tempo, dos valores médios de 3 coeficientes *wavelet*, correspondentes a 3 blocos espacialmente adjacentes em cada quadro de uma sequência de vídeo.



Fonte: O Autor.

Fig. 4.3 - Função de correlação cruzada entre duas diferentes curvas de amplitude dos coeficientes wavelet, para 78 amostras sucessivas no tempo de 2 *pixels* adjacentes.

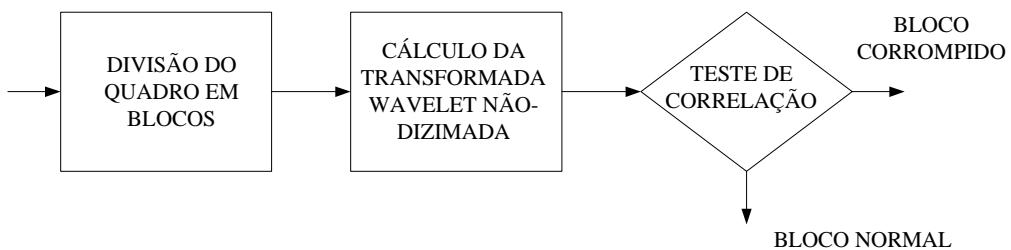


Fonte: O Autor.

Como pode ser observado na Figura 4.3, a correlação não-normalizada entre os coeficientes wavelet correspondentes a blocos de *pixels* espacialmente adjacentes, para quadros de uma sequência de vídeo, é elevada – da ordem de grandeza de 10^5 (valor que foi observado experimentalmente para mais de uma centena de observações, para níveis de intensidade de *pixel* de 0 a 255). A observação da conformação dessa correlação cruzada sugere, assim, uma forma de se detectar falhas ou descontinuidades severas de amplitude na informação útil de um sinal de vídeo – que se relacionam a descontinuidades igualmente severas no perfil de correlação cruzada ilustrado na Figura 4.3. Em uma análise generalizada, blocos de *pixels* contaminados por ruído, pertencentes a uma imagem, podem ser detectados por meio da observação da curva de correlação cruzada entre seus coeficientes wavelet para distintas escalas: se tal curva não satisfizer ao padrão de amplitudes observado para o caso da mesma imagem isenta de ruído, classificam-se tais blocos como corrompidos.

Deve-se, naturalmente, estabelecer um compromisso entre complexidade computacional e resolução de análise (para localizar mais precisamente os componentes de ruído). Uma escolha intuitiva, que reduz consideravelmente o esforço computacional sem comprometer a precisão na correção, consiste em considerar a divisão de quadros em blocos os menores possíveis, mas de dimensão superior a 1 *pixel* de área, caso em que a escolha natural é a de 2×2 *pixels*. O diagrama em blocos da estrutura é ilustrado na Fig. 4.4.

Fig. 4.4 - Diagrama do esquema proposto de detecção de blocos corrompidos.



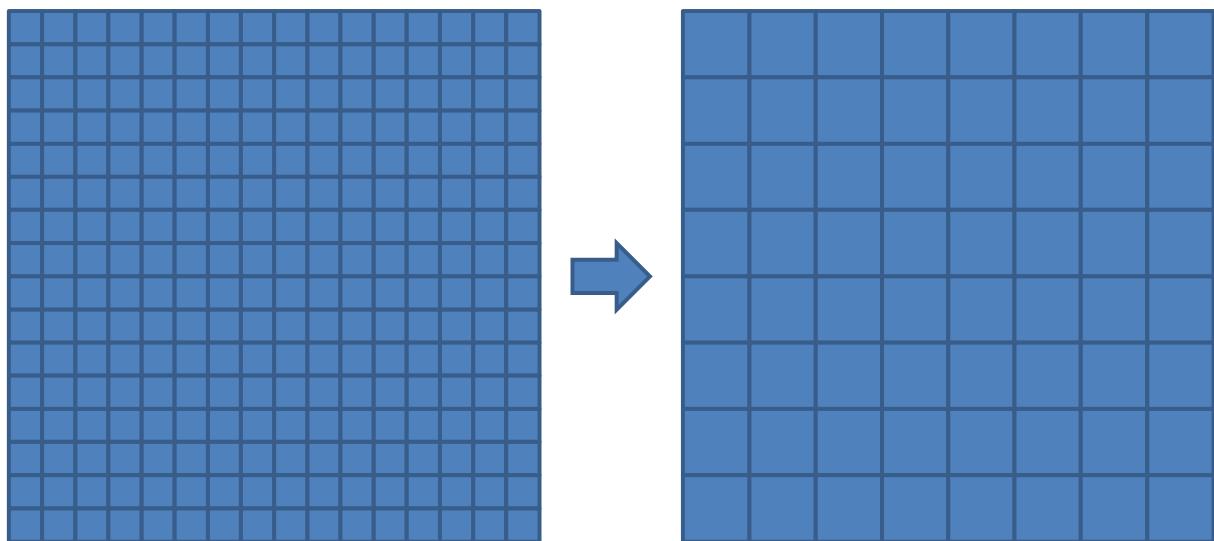
Fonte: O Autor.

O algoritmo de detecção de erros é pormenorizado nas próximas sub-seções.

4.2.1 Divisão do quadro em blocos

Cada imagem (ou quadro) analisada é dividida em blocos quadrados de 2×2 *pixels*, o que reduz notavelmente o esforço computacional das etapas seguintes (Fig. 4.5). Além de ser inviável tratar toda a imagem em uma única etapa, detalhes menos significativos podem ser eliminados na substituição de um *pixel* isolado por um bloco contendo o valor médio de quatro *pixels*, sem grande prejuízo visual.

Fig. 4.5 - Divisão do quadro em blocos de 2×2 pixels.

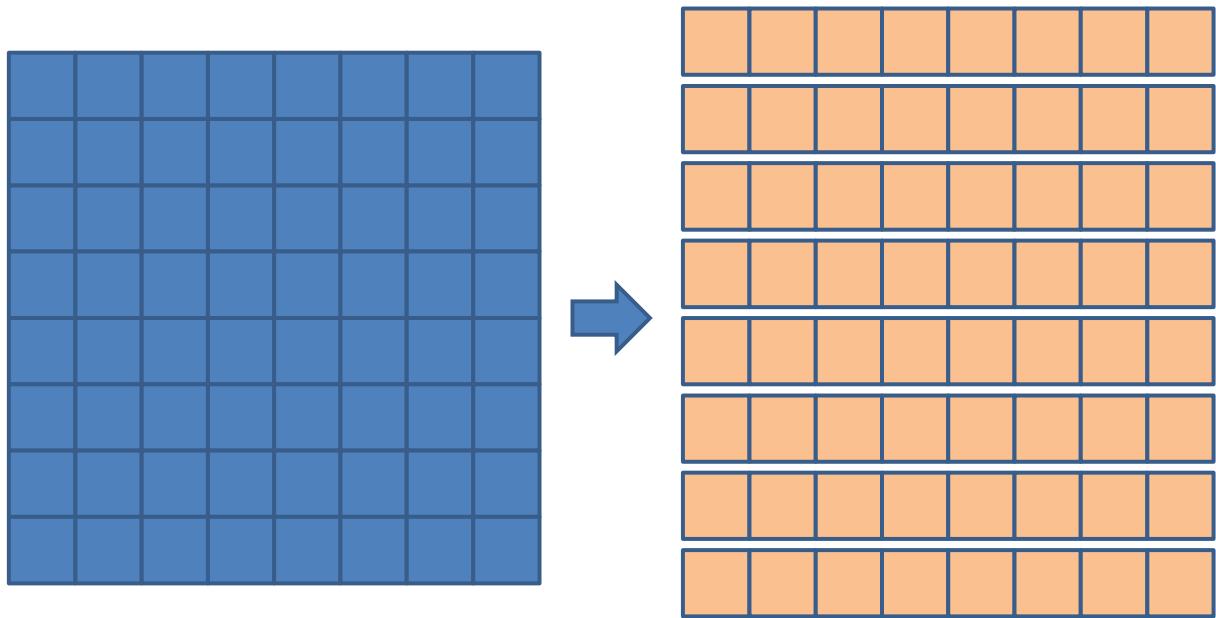


Fonte: O Autor.

4.2.2 Cálculo da transformada *wavelet* não-dizimada

Para cada linha de blocos da etapa anterior, é computada a transformada *wavelet* não-dizimada (que mantém o suporte espacial da imagem analisada), em um nível de decomposição. Assim, uma imagem de 256×256 pixels é dividida em 128×128 blocos de 2×2 pixels (Fig. 4.6). Como descrito no passo anterior, reduz-se com a blocagem o tempo de processamento para o cálculo da transformada *wavelet*, sem comprometer a eficácia do resultado, pois a divisão do quadro em blocos de pixels não produz falsos positivos ou negativos no que diz respeito à detecção de porções corrompidas.

Fig. 4.6 - Decomposição *wavelet* de cada linha de blocos de *pixels*.



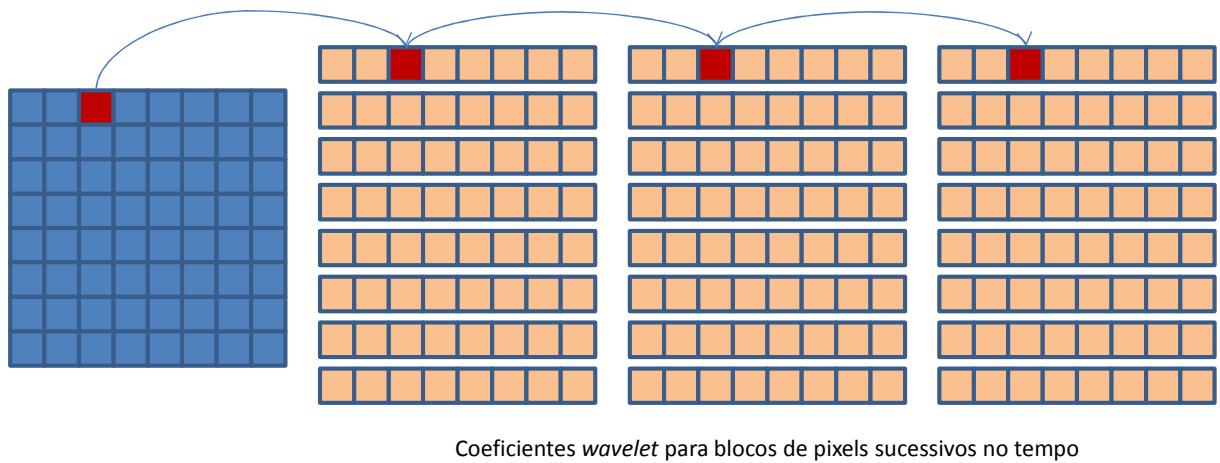
Fonte: O Autor.

4.2.3 Teste de correlação

Nesta etapa, é realizado o cálculo de correlação cruzada entre os valores médios dos coeficientes *wavelet* de cada bloco de *pixels* do quadro analisado e os blocos correspondentes à mesma posição espacial, para sequências sucessivas no tempo (Fig. 4.7). Armazenando-se o valor máximo da correlação cruzada obtido, a ocorrência de valores de correlação cruzada inferiores à metade deste valor, a título de observação, sinaliza um bloco de *pixels* contaminado por ruído. Há de se ressaltar que não há uma estimativa quantitativa precisa para limiarização da correlação cruzada considerada indicativa de degradação – a consideração de que valores de pico inferiores à metade do máximo valor observado caracterizam corrupção de blocos, embora meramente empírica, se mostrou eficiente, sem que a otimização deste parâmetro, testado em simulações na faixa de 0,25 a 0,75, tenha

se mostrado mais eficaz no processo de detecção aqui descrito.

Fig. 4.7 - Verificação da correlação cruzada entre coeficientes *wavelet* de blocos de pixels sucessivos no tempo. Valores de correlação cruzada inferiores ao limiar estabelecido sinalizam blocos de pixels degradados.



Fonte: O Autor.

Uma vez detectados os blocos corrompidos, cabe à etapa seguinte de processamento proceder à determinação dos valores reais dos coeficientes, a partir da observação do seu decaimento médio dentro do bloco sinalizado como corrompido.

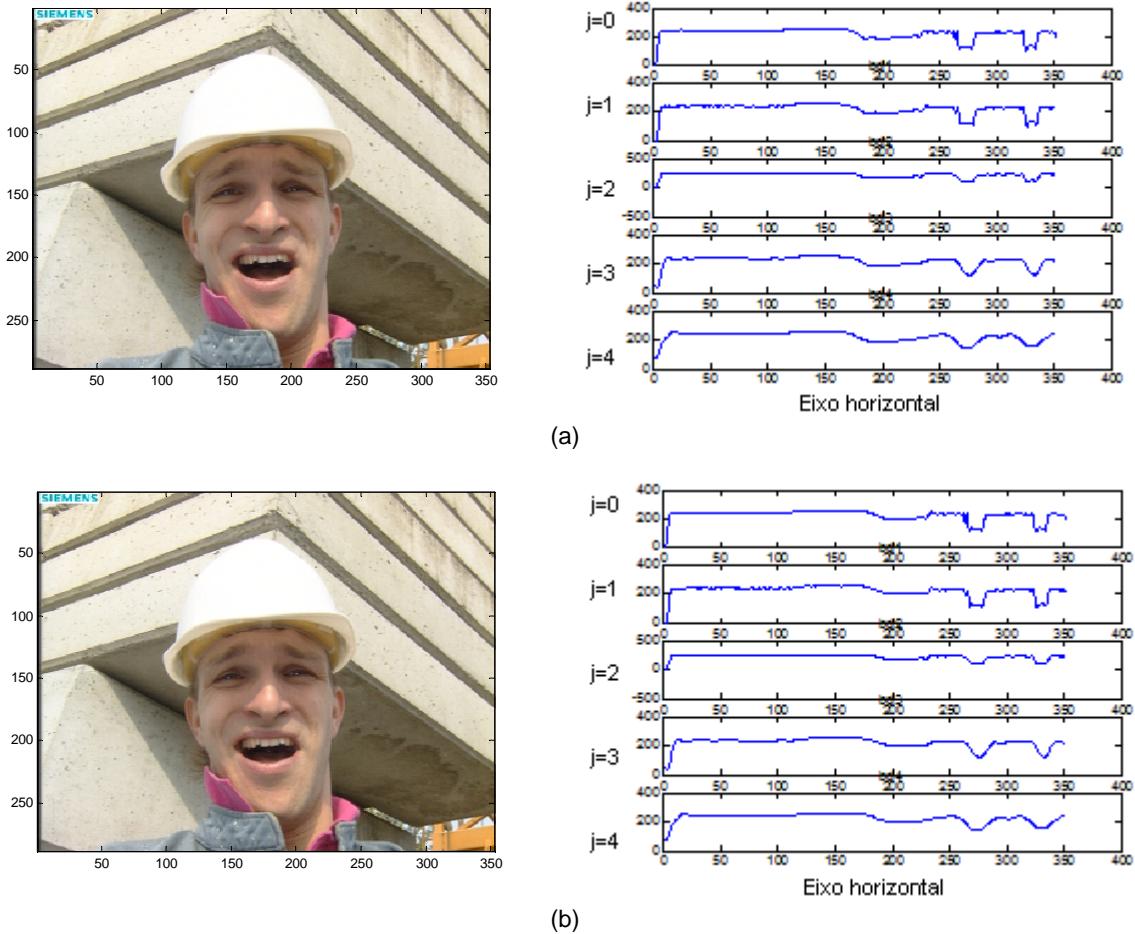
4.3 ALGORITMO DE CORREÇÃO DE ERROS

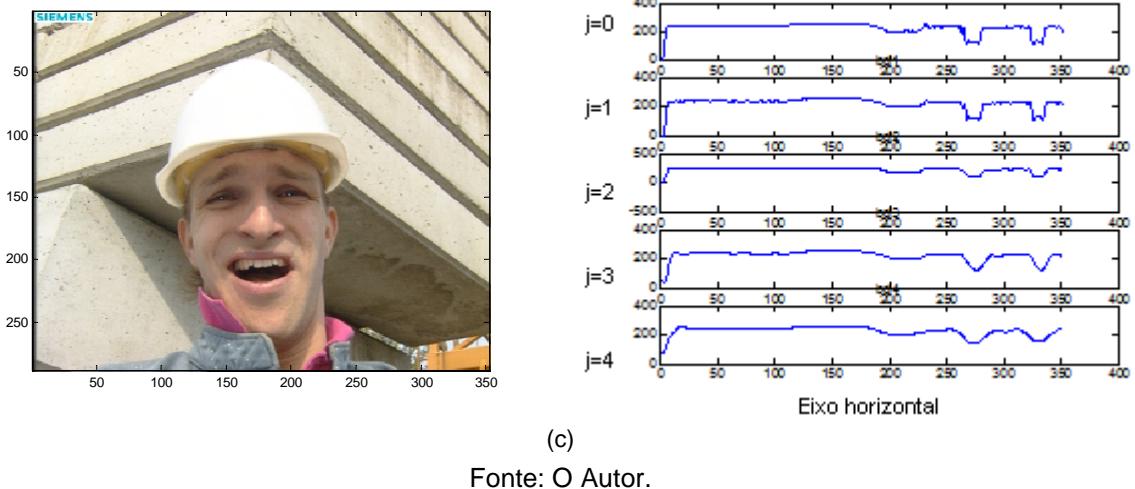
A etapa proposta de cálculo dos novos coeficientes *wavelet* parte da premissa que tais coeficientes carregam algum indício acerca da natureza da informação de ruído, através da observação de características de decaimento inconsistentes. Não existem, na literatura da área, trabalhos que se proponham a realizar o *inpainting* de imagens digitais a partir da análise multi-resolução (no domínio espaço-frequência),

em termos da análise do decaimento de coeficientes *wavelet* através de sucessivas escalas, o que constitui o escopo central deste trabalho.

A similaridade através das escalas da decomposição *wavelet* de imagens pode ser estendida para o caso de quadros de uma sequência de vídeo. Como mostrado na Fig. 4.8, tomando-se a sequência ‘Foreman’, os sinais referentes à decomposição *wavelet* não-dizimada de uma linha arbitrária de *pixels*, para diferentes níveis de decomposição j , têm preservada a localização de máximos e mínimos.

Fig. 4.8 - Quadros sucessivos da sequência ‘Foreman’ (a) – (c), e intensidade de *pixel* em sucessivas escalas de decomposição da linha horizontal central de *pixels* de cada quadro. Pelo fato de os quadros possuírem representação em canais RGB (RED-GREEN-BLUE), são ilustrados os componentes *wavelet* referentes a apenas um canal – na figura, o canal R. Para cada quadro, são calculados 4 níveis de decomposição *wavelet* não-dizimada.





Fonte: O Autor.

Sob tais condições, busca-se agora utilizar o comportamento de decaimento dos coeficientes *wavelet* na reconstrução dos blocos corrompidos que tenham sido detectados no passo anterior.

Os coeficientes da transformada *wavelet* de um bloco F_i com $m \times n$ pixels são limitados por

$$\begin{aligned} |\Psi_{i,j,k}(m,n)| &= K_j 2^{-j(\alpha+1/2)} \\ &= K_j 2^{-j\alpha} \cdot 2^{-j/2} \\ &= K_j 2^{-j\alpha} \cdot g(j), \end{aligned} \quad (4.2)$$

onde j corresponde ao parâmetro de escala e k ao fator de translação da transformada, α é o expoente de Hölder (caso particular em que tal expoente é idêntico ao expoente de Lipschitz), K_j é obtido por extrapolação dos coeficientes no domínio da transformada *wavelet*, referentes a um mesmo bloco de pixels, para cada parâmetro de escala j , e $g(j)$ é uma função de escalonamento. Em outras palavras, pode-se dizer que, para um único nível de decomposição, sua curva de decaimento é escalonada de um fator de $g(1) = 2^{-1/2}$, conforme a equação (4.2).

Entretanto, na estimativa de um expoente médio de Hölder $\bar{\alpha}$, a função $g(j)$ pode ser refinada de modo a elevar a contribuição de componentes de altas frequências no cálculo desse expoente, como descrito a seguir. Na Eq. (3.15), o

termo $2^{-j\alpha}$ provê o decaimento de amplitude através das escalas [79, 80]. Sob tal condição, é desejável que exista uma função $g'(j)$, obtida da modificação de $g(j) = 2^{-j/2}$, que atenda pontualmente às restrições:

- a) $g(j) = 1$ para $j = 0$, situação particular em que os expoentes de Hölder e de Lipschitz são iguais;
- b) $g(j) \leq 2^{-1/2}$ para $j = 1$, que obedece ao decaimento de amplitude através das escalas da decomposição *wavelet* (Eq. 3.15);
- c) $\lim_{j \rightarrow \infty} g(j) = 0$, já que o decaimento de um sinal através das escalas de decomposição *wavelet* ocorre para qualquer grau de suavidade de um sinal e suas derivadas.

Uma das mais importantes propriedades de bases ortogonais *wavelet* é a de que tais bases pertencem a uma ampla classe de espaços ortogonais de suavidade, denominada de classe de espaços de Besov [84]. Espaços de Besov agrupam funções que possuem um grau específico de suavidade em suas derivadas. O espaço de Besov $B_q^\gamma(L_p)$ consiste, em uma visão simplificada, de funções que possuem γ derivadas “limitadas” em L_p . O parâmetro q é um parâmetro de refinamento adicional.⁵

Dada a decomposição *wavelet* de uma função f em coeficientes de escalamento e *wavelet*, computada pelo uso de *wavelets* biortogonais⁶ representadas pelo par (ϕ, ψ) com N momentos desvanecentes, como estabelecido na equação (3.13), há uma caracterização baseada em *wavelets* acerca da norma de Besov em f , dada por:

⁵ A definição precisa de espaços de Besov no domínio espacial é muito complexa. Sem perda de generalidade, pode-se afirmar que a suavidade de uma função pertencente ao espaço $B_q^\gamma(L_p)$ é elevada conforme o aumento do parâmetro γ [75].

⁶ *Wavelets* biortogonais são *wavelets* em que a transformada *wavelet* associada é inversível, mas não necessariamente ortogonal, permitindo mais graus de liberdade em seu projeto do que para *wavelets* ortogonais.

$$\|f\|_{B_q^\alpha(L_p)} = \|\Phi_0\|_{L_p} + \left(\sum_{j>0} (2^{j(\gamma p + p/2+1)} \|\Psi_{i,j,k}\|_{L_p})^q \right)^{1/q} \quad (4.3)$$

Uma função é assim dita pertencente ao espaço de Besov $B_q^\gamma(L_p)$ se a norma de Besov é finita [84].

Uma inspeção mais acurada da Eq. (4.3) mostra que a norma de Besov pode ser interpretada como uma norma L_p com ponderação distinta dos coeficientes de detalhes. Em contraste com a formulação complexa da norma de Besov no domínio espacial, a representação de uma função por *wavelets* fornece um meio prático e rápido de computar seu valor. A norma na Eq. (4.3) é definida apenas para $0 < \gamma < \alpha$, onde α reflete a regularidade dos sinais ortogonais envolvidos. Se α_m representa o valor máximo do expoente de Hölder, a regularidade das *wavelets* é dada por

$$\alpha = \min(\alpha_m, N) \quad (4.4)$$

como já delimitado pela Eq. (3.3). De (4.3), pode-se reescrever a Eq. (3.13) na forma:

$$\begin{aligned} |\Psi_{i,j,k}(m,n)| &= K_j \cdot 2^{-j(\alpha p + p/2+1)} \\ &= K_j \cdot 2^{-j\alpha p} \cdot 2^{-(jp/2+j)} \end{aligned} \quad (4.5)$$

com $q=1$, por simplicidade de notação. Para cada *pixel* analisado através das escalas, é intuitivo que $p=1$, no que a Eq. (4.5) pode ser expressa como:

$$\begin{aligned} |\Psi_{i,j,k}(m,n)| &= K_j \cdot 2^{-j\alpha} \cdot 2^{-3j/2} \\ &= K_j \cdot 2^{-j\alpha} \cdot g(j) \end{aligned} \quad (4.6)$$

Esta aproximação, $g(j) = 2^{-3j/2}$, obedece às restrições (a)-(c) dadas, e, satisfazendo a Eq. (4.2) para todos os valores de j , é menos severa no que diz

respeito à atenuação dos valores dos coeficientes (e assim da energia de componentes referentes a erros de *pixels*). Assim, tais coeficientes *wavelet* podem ser estimados como:

$$\begin{aligned}\Psi'_{i,j,k}(m,n) &= K \cdot 2^{-j\alpha'} \cdot 2^{-3j/2} \\ &= K \cdot 2^{-j(\alpha'+3/2)},\end{aligned}\quad (4.7)$$

em que $\Psi'_{i,j,k}(m,n)$ e α' representam, respectivamente, estimativas para $\Psi_{i,j,k}(m,n)$ e α . No caso mais extremo, o maior valor admissível para K (valor estimado para a constante de extração) é dado por $K = K_j \cdot 2^{j/2}$, que é obtido considerando-se $r = 0$ e $K_j = \max |\Psi_{i,j,k}(m,n)|$. Uma vez que tal valor de K_j equivale ao limite superior de $\Psi'_{i,j,k}(m,n)$, a Eq. (4.7) pode ser reescrita como

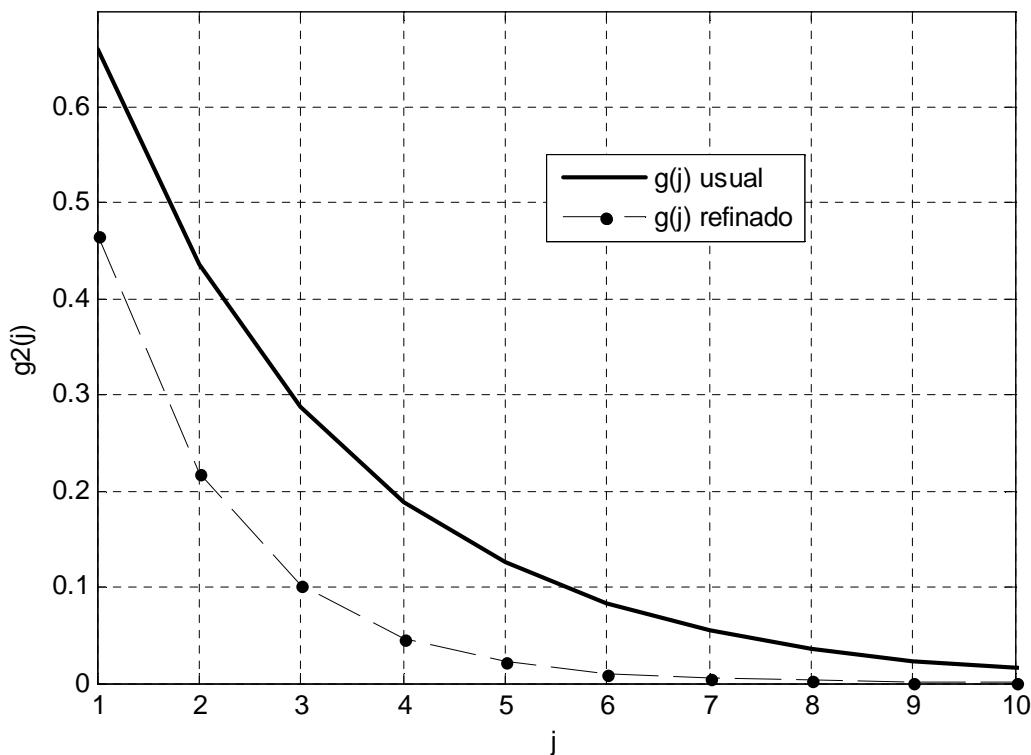
$$\begin{aligned}\Psi'_{i,j,k}(m,n) &= K_j \cdot 2^{j/2} \cdot 2^{-j\alpha'} \cdot 2^{-3j/2} \\ &= K_j \cdot 2^{-j(\alpha'+1)}.\end{aligned}\quad (4.8)$$

Pode-se definir assim, a partir da Eq. 4.8, uma função de decaimento matemático dos coeficientes *wavelet*.

$$g_2(j) = \frac{\Psi'_{i,j,k}(m,n)}{\max |\Psi_{i,j,k}(m,n)|} = 2^{-j(\alpha'+1)}, \quad (4.9)$$

que descreve o perfil de decaimento dos coeficientes estimados, em função da escala j observada. A Figura 4.9 apresenta o comportamento de $g_2(j)$ versus j : sem o refinamento de $g(j)$, a curva se restringiria a um valor de $2^{-j\alpha'}$, decrescente com o aumento das escalas.

Fig. 4.9 - Função de análise $g_2(j)$. O refinamento das características de decaimento permite uma descrição mais acurada de coeficientes de altas frequências, o que, por conseguinte, promove melhoria na qualidade dos blocos corrigidos.

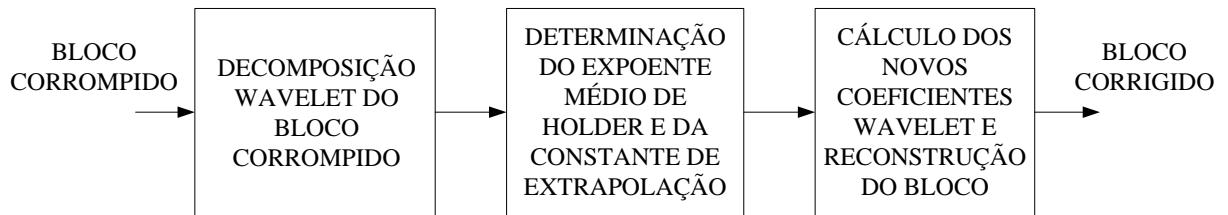


Fonte: O Autor.

Tal correção assegura uma descrição mais rica de componentes de altas frequências, no que foi feita a melhor escolha para $g(j)$, e consequentemente para $g_2(j)$. A Fig. 4.10 ilustra o diagrama em blocos do algoritmo de correção⁷.

⁷ Não se deve associar detecção e correção de erro a técnicas de remoção de ruído por supressão de coeficientes no domínio da transformada (*wavelet denoising*, por exemplo). As últimas não se propõem a determinar valores “ótimos” para os coeficientes, eliminando tão somente, por limiarização, coeficientes de altas frequências associados à informação de ruído presente nos sinais analisados.

Fig. 4.10 - Diagrama em blocos da estrutura proposta para correção de blocos corrompidos.



Fonte: O Autor.

O algoritmo de correção de erros é pormenorizado nas próximas sub-seções.

4.3.1 Decomposição *wavelet* do bloco corrompido

Tal como na etapa de detecção descrita anteriormente, para cada linha de blocos corrompidos é computada a transformada *wavelet* não-dizimada, em mais de um nível de decomposição (simulações realizadas consideraram 3 níveis de decomposição *wavelet*). Embora se possam armazenar os valores de coeficientes obtidos naquela etapa para uso na etapa de correção, um novo procedimento de cálculo de coeficientes é ainda mais vantajoso do que a necessidade de se armazenar valores de coeficientes para todos os blocos, corrompidos ou não. Além disso, o tempo de processamento requerido para o cálculo da transformada *wavelet* não é significativo dentro do algoritmo de correção, não representando parcela considerável dentro do custo computacional global do algoritmo.

4.3.2 Determinação do expoente estimado de Hölder α' e da função de extração $g_2(j)$

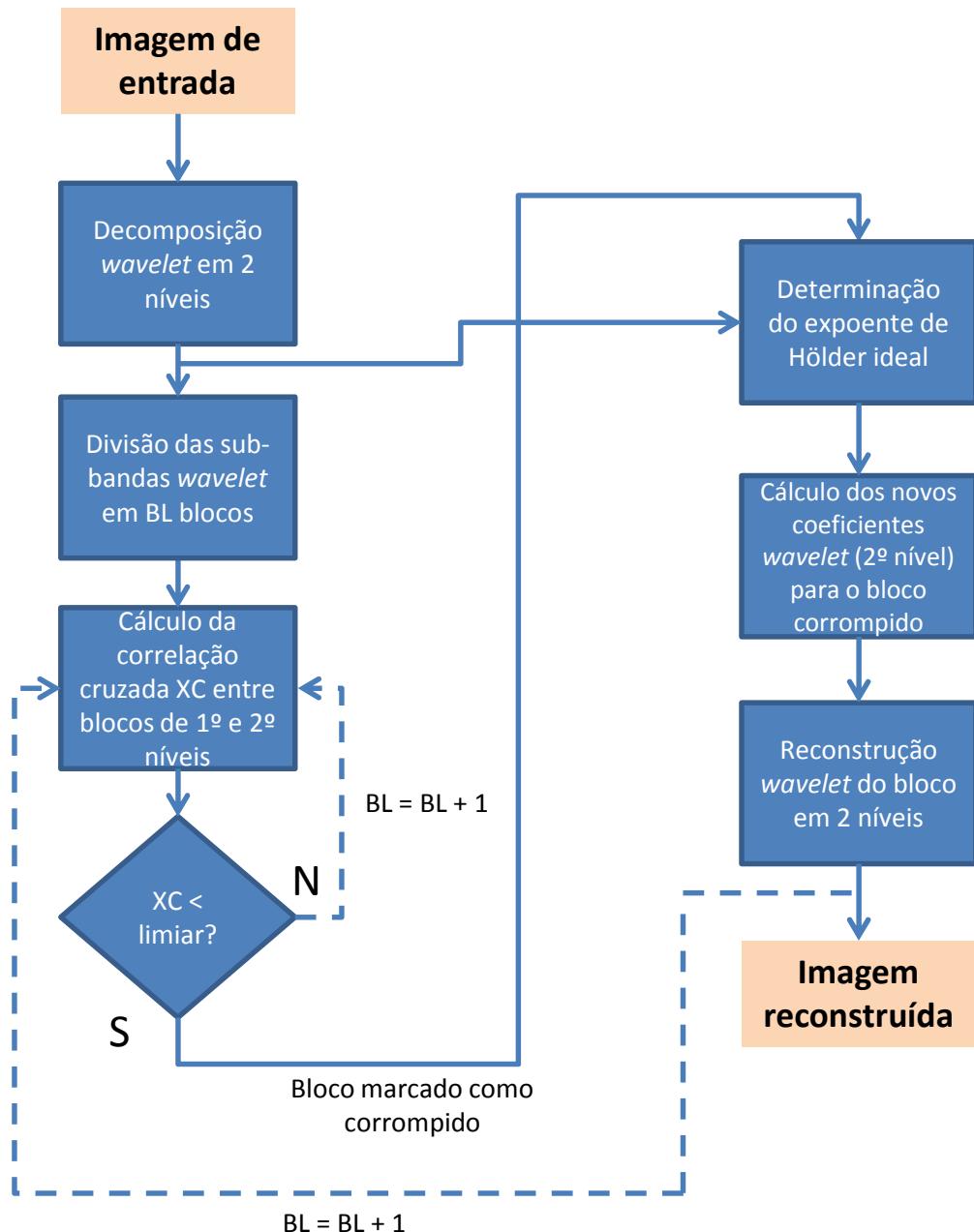
A observação dos coeficientes *wavelet* determinados no passo anterior permitem a determinação da função de análise $g'(j)$, dependente da escala de decomposição, que por sua vez conduz à determinação do expoente estimado de Hölder α' , a partir da utilização da Eq. (4.7). A partir dessa estimativa, por sua vez, constrói-se a função de análise refinada $g_2(j)$ segundo a Eq. (4.9), a ser aplicada de forma multiplicativa sobre cada coeficiente *wavelet* do respectivo nível de decomposição j .

4.3.3 Cálculo dos novos coeficientes *wavelet* e reconstrução do bloco

A função de análise refinada $g_2(j)$ é utilizada para determinar, a partir dos valores de intensidade dos blocos de *pixels* classificados como degradados, os valores de decaimento estimado dos coeficientes *wavelet*, o que permite iterativamente a reconstrução de cada bloco considerado degradado, até que toda a área do quadro analisado seja varrida.

A Figura 4.11 apresenta a sequência de passos realizada pelos algoritmos de detecção e correção de erros.

Fig. 4.11 – Fluxograma dos algoritmos propostos para detecção e correção de erros.



As Figuras 4.12 – 4.14 apresentam em mais detalhes as etapas de cálculo da transformada wavelet, detecção de blocos corrompidos por limiarização de sua correlação cruzada no domínio da transformada e reconstrução do bloco corrompido de acordo com a função de reconstrução definida pela Eq. 4.9.

Fig. 4.12 – Cálculo da transformada *wavelet* dos blocos da imagem analisada.



Fig. 4.13 – Detecção de blocos corrompidos por limiarização de correlação cruzada.

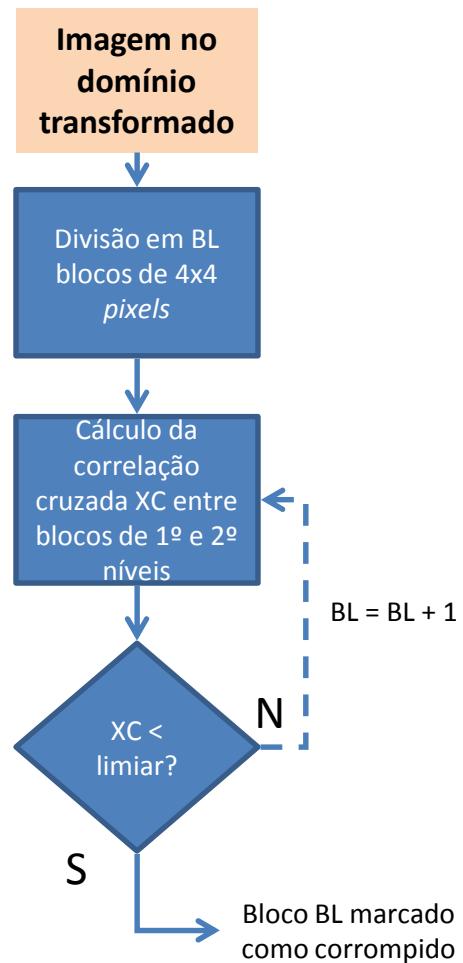
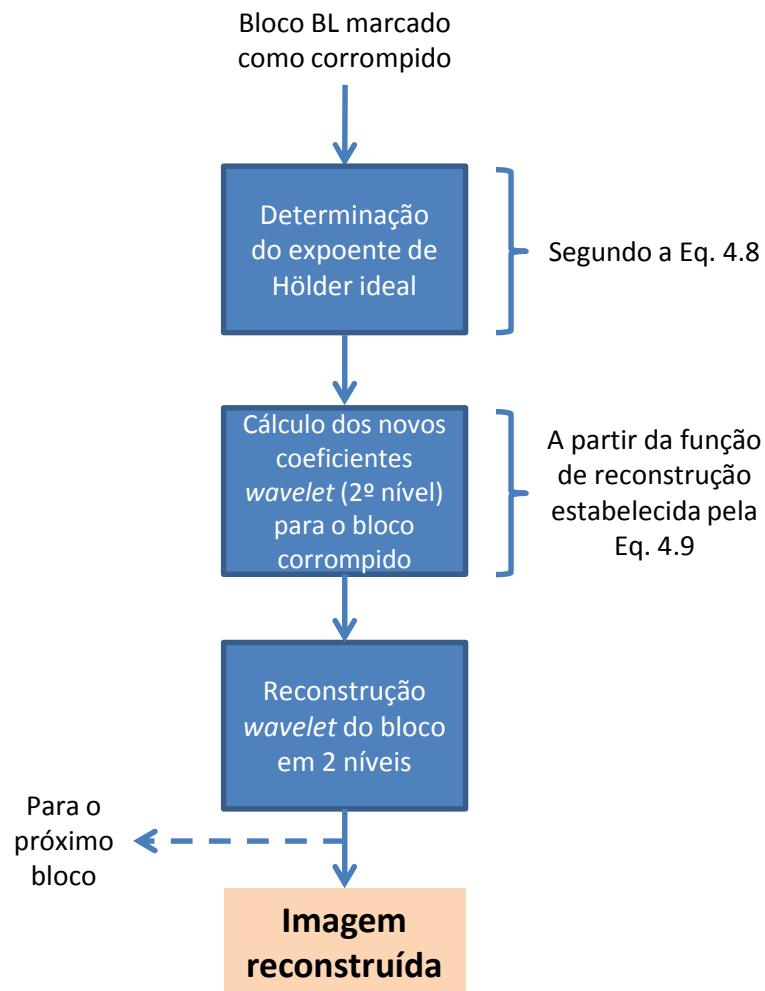


Fig. 4.14 – Correção de blocos corrompidos por meio da função de reconstrução definida em (4.9).



O capítulo seguinte apresenta simulações do método de correção aqui proposto, utilizando imagens e quadros de sequências de vídeo clássicas da literatura.

4.4 CONCLUSÕES

Este capítulo apresentou uma técnica de detecção e correção de erros em imagens que faz uso da propriedade de regularidade de seus coeficientes *wavelet*, constituído de duas etapas distintas: um algoritmo de detecção de erros (EDA) e um algoritmo de correção de erros, independentes entre si. Foi mostrado que os coeficientes contaminados por ruído podem ser estimados através da reconstrução das características de decaimento, através da observação da similaridade existente entre as escalas da decomposição *wavelet*, explorando, assim, a redundância de informação remanescente do processo de transformação linear.

5 RESULTADOS DE SIMULAÇÕES

5.1 SIMULAÇÕES COMPARATIVAS

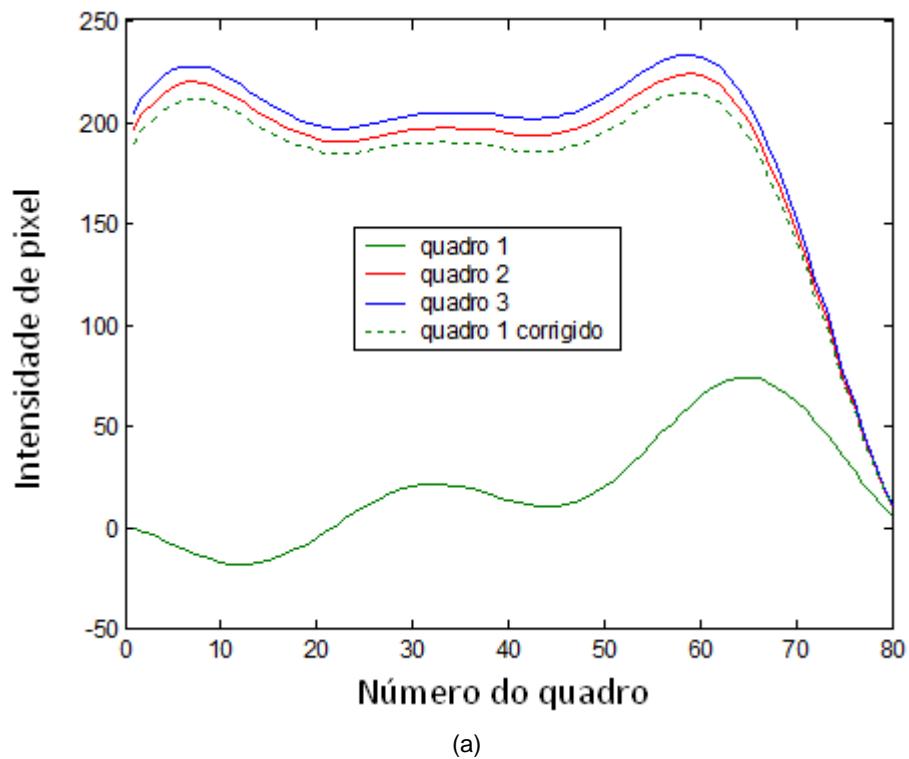
A validação de estratégias de correção de informação visual pressupõe a aferição de figuras de mérito quantitativas (p.ex. erro absoluto, razão sinal-ruído de pico, entre outras) e qualitativas (p.ex. métricas de sensibilidade ao contraste), na ideia de caracterizar completamente seu grau de eficiência. Dentro desta premissa, e para verificar a validade do algoritmo de *inpainting* proposto neste trabalho, descrito no domínio da transformada *wavelet*, foram consideradas situações de contaminação por ruído tanto de imagens estáticas como de quadros pertencentes a sequências de vídeo. Estas duas situações permitem explorar mais detalhadamente o mecanismo de correção visto que o algoritmo em si trabalha consecutivamente em dois domínios: o do espaço-frequência, quando executa os procedimentos de detecção de erros (em termos de sua localização), e o temporal, quando executa a correção propriamente dita dos *pixels* assinalados como corrompidos. No caso específico da aplicação do método sobre imagens estáticas, constrói-se uma sequência artificial de vídeo a partir da repetição, como quadros, da imagem estática considerada. Tal artifício é utilizado para verificar a eficiência da etapa de detecção de erros, dependente da informação intraquadros, sem anular a validade da etapa de correção de erros, dependente da informação interquadros e intimamente relacionada à função de correlação entre as cenas.

Considere-se, inicialmente, a aplicação do método de correção proposto a um quadro da sequência de vídeo ‘Foreman’, sem compressão, composta por 400 quadros e de resolução 352×288 *pixels* (formato QCIF), em que foi superposto ao quadro 27 dessa sequência um bloco de 20×20 *pixels* totalmente escuro⁸. O algoritmo de detecção, dividindo o quadro em blocos menores de 2×2 *pixels*, realiza a decomposição *wavelet* em 4 níveis, utilizando a família de *wavelets* Daubechies biortogonal 9-7 [77]. A análise da correlação cruzada entre blocos adjacentes de 2×2 *pixels* permite delimitar regiões espaciais corrompidas,

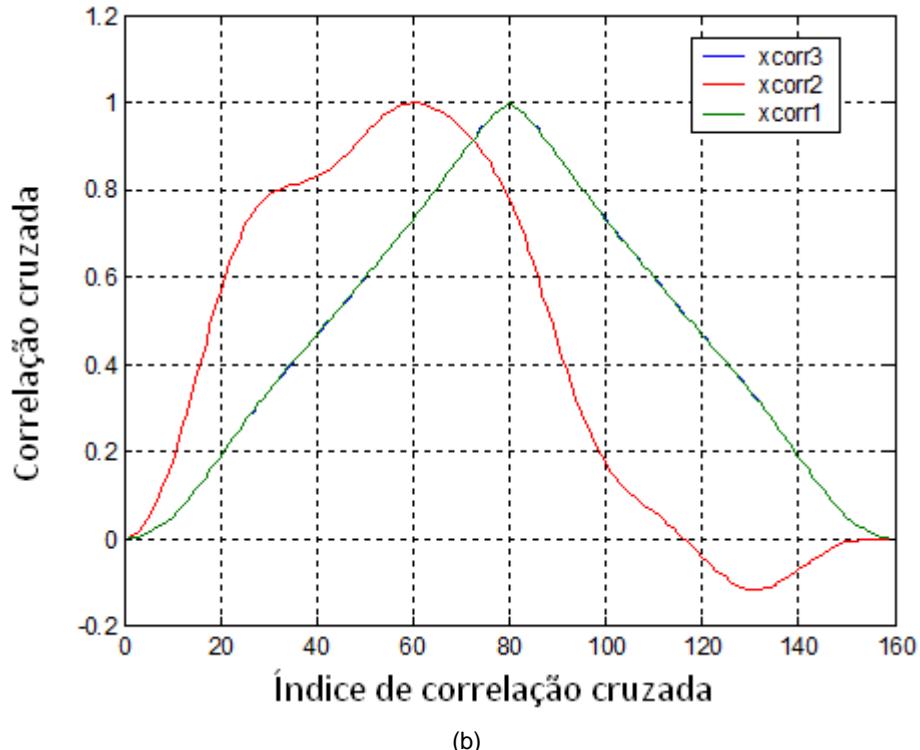
⁸ Foram testados somente blocos superpostos com ausência de informação, para simular o pior caso em termos de perda de informação.

permitindo a fácil identificação de coeficientes espúrios ao quadro analisado. A Fig. 5.1(a) ilustra as curvas de amplitude média dos coeficientes *wavelet*, no 1º nível de decomposição, para 3 blocos adjacentes de *pixels* em um grupo de 80 quadros sucessivos no tempo, bem como o comportamento da função de correlação cruzada entre tais curvas. Pode-se observar, na Fig. 5.1(b), a discrepância na curva de correlação cruzada entre os blocos do quadro corrompido ('quadro 1') e os blocos adjacentes não afetados por ruído ('quadro 2'), independente de qual quadro seja tomado como referência.

Fig. 5.1 - (a) Curvas de intensidade de *pixel* vs. número de quadro, em uma sequência de 80 quadros sucessivos no tempo, com um dos blocos degradados. (b) Funções de correlação cruzada entre as curvas de amplitude de (a) e de amplitude média do bloco: antes da degradação (curva em verde), após a degradação (curva em vermelho) e retirando-se o bloco degradado (curva em azul).



(a)



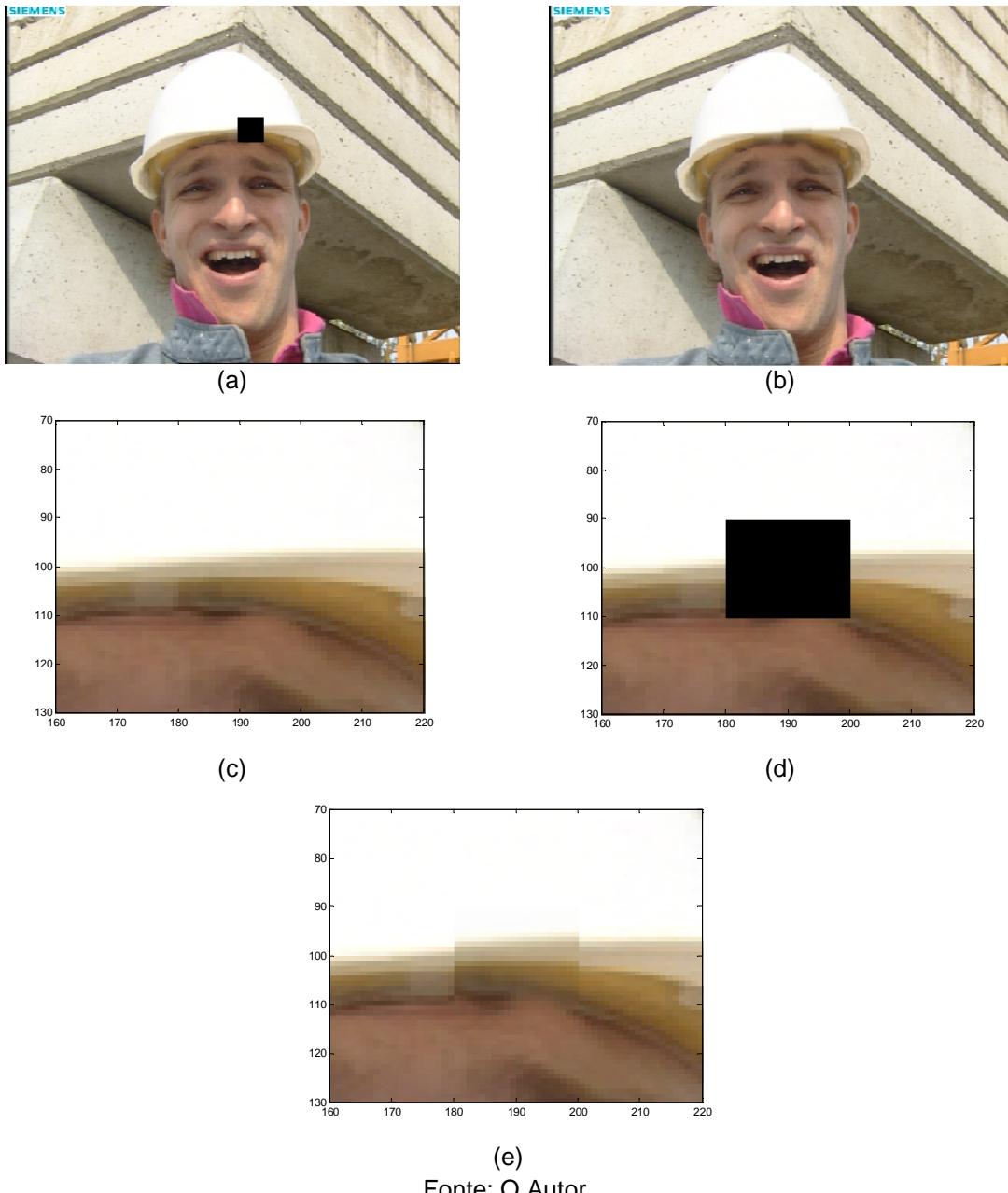
(b)

Fonte: O Autor.

Em uma etapa seguinte, a amplitude dos coeficientes do bloco contaminado é calculada a partir da estimativa do expoente médio de Hölder [conforme a Eq. (3.17)] de cada sub-bloco de 2×2 pixels, a partir de uma janela de observação de 20×20 pixels em sua vizinhança⁹, varrendo-se o bloco total da esquerda para a direita e de cima para baixo. A Fig. 5.2 mostra os resultados gráficos obtidos para o algoritmo de correção proposto neste trabalho.

⁹ A dimensão da janela de observação arbitrada é uma escolha extrema, pois o ruído inserido elimina totalmente a informação original. Componentes de ruído que produzam blocos cujas características de decaimento apresentem uma correlação acima de determinado limite podem exigir janelas de observação menores.

Fig. 5.2 - (a) Quadro da sequência ‘Foreman’ contaminado por ruído. (b) Quadro corrigido. (c) Quadro original – detalhes. (d) Quadro contaminado por ruído – detalhe. (e) Quadro corrigido - detalhe.



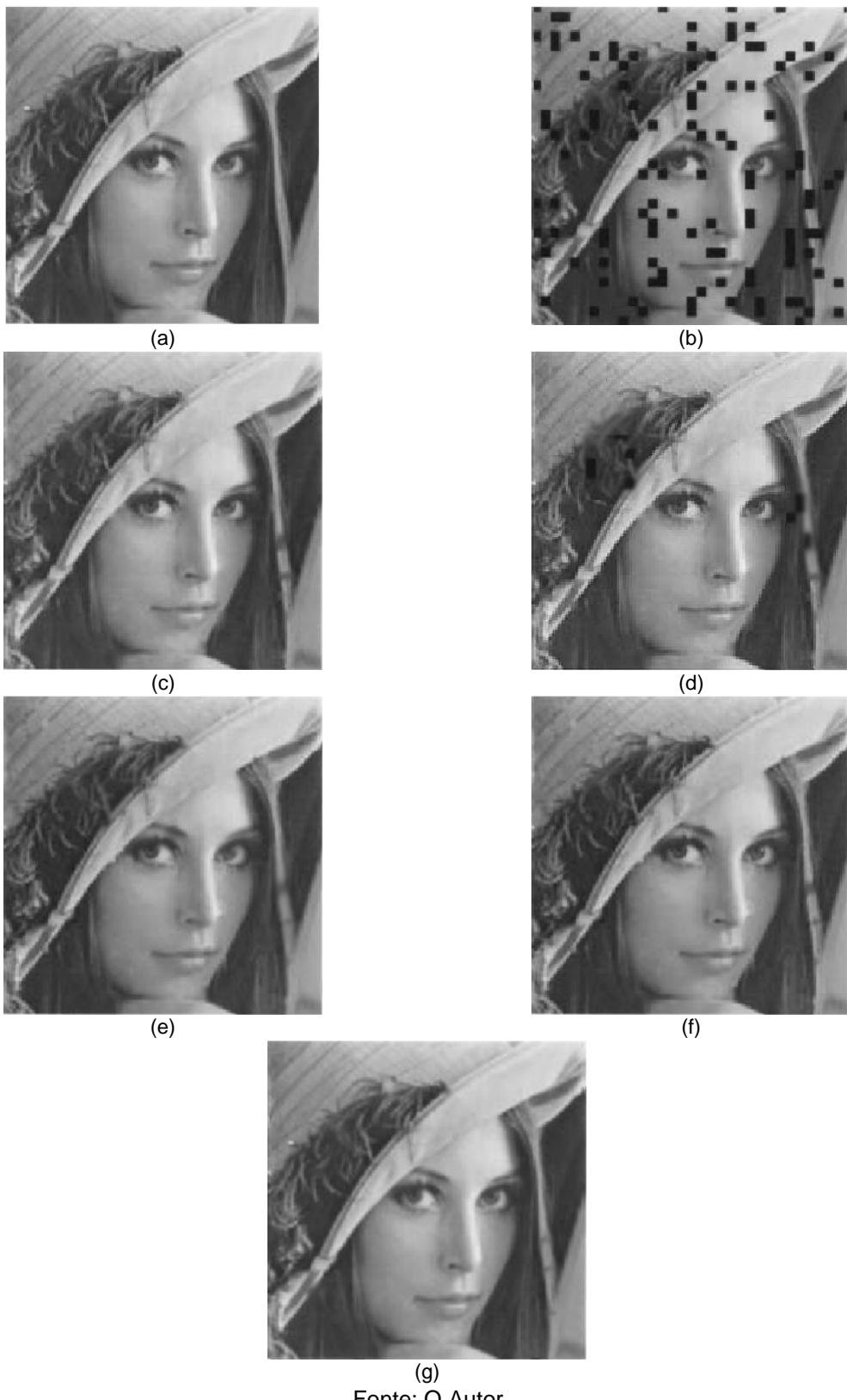
Fonte: O Autor.

Para avaliar o desempenho do algoritmo de *inpainting* proposto, a etapa de detecção é de fundamental importância, uma vez que o algoritmo de correção pode ser aplicado somente aos blocos detectados como danificados. Muitos pesquisadores, até então, tentaram resolver o problema por três diferentes procedimentos. A primeira solução considera que blocos corrompidos obedecem a um padrão pré-definido ou a informação sobre a localização de blocos corrompidos é

previamente conhecida [88, 90, 92, 95]. Embora essas sejam condições suficientes para avaliar o desempenho do algoritmo de correção, não se pode afirmar que mantenham sua validade para situações reais. A segunda solução consiste em se transmitir informação adicional para detectar ou limitar a ocorrência de erros de *bits*. Por exemplo, Lie *et al.* [100] utilizam *bits* de paridade para isolar erros de *bits*, Nasiopoulos *et al.* [102] inserem a cada macrobloco transmitido o número de *bits* transmitido. Tais soluções e outras na literatura [103, 104], no entanto, requerem a modificação da estrutura de codificação. A terceira solução consiste em utilizar a funcionalidade da rede de transmissão como, por exemplo, no esquema de transmissão ATM [96, 97], em que a perda de células pode ser detectada pelo protocolo da camada de adaptação ATM. Assim, os blocos pertencentes a células danificadas e todos os blocos subsequentes antes da resincronização são descartados, à custa de elevado esforço computacional. Em todas as soluções destacadas neste parágrafo para *inpainting*, a adição de etapas traria para a correção do quadro danificado na Fig. 5.2(a) elevado esforço computacional ou presumiria condições específicas de degradação do quadro que não são sempre realísticas, o que agrega ao método aqui proposto vantagens relativas à simplicidade de implementação e utilização sob uma ampla variedade de imagens degradadas.

Para ilustrar mais claramente tais diferenças, a Fig. 5.3 apresenta uma porção da imagem ‘Lena’ corrompida a uma taxa de erro de *bits* de 10%, e imagens reconstruídas a partir do método proposto e de outros 4 métodos de correção do estado-da-arte, 3 deles operando no domínio transformado (mais especificamente, através da DCT – *discrete cosine transform*) [90, 92, 98], e o último realizando interpolação espacial de limites de objetos por meio de *splines* cúbicas [101]. Na imagem corrompida, valores de intensidade de *pixel* de blocos são fixados em zero de forma a serem facilmente identificados.

Fig. 5.3 - Conjunto de imagens ‘Lena’. (a) Original. (b) Corrompida por ruído aleatório. (c) Corrigida pelo método da referência [90]. (d) Corrigida pelo método da referência [92]. (e) Corrigida pelo método da referência [98]. (f) Corrigida pelo método da referência [101]. (g) Corrigida pelo método aqui proposto.



Fonte: O Autor.

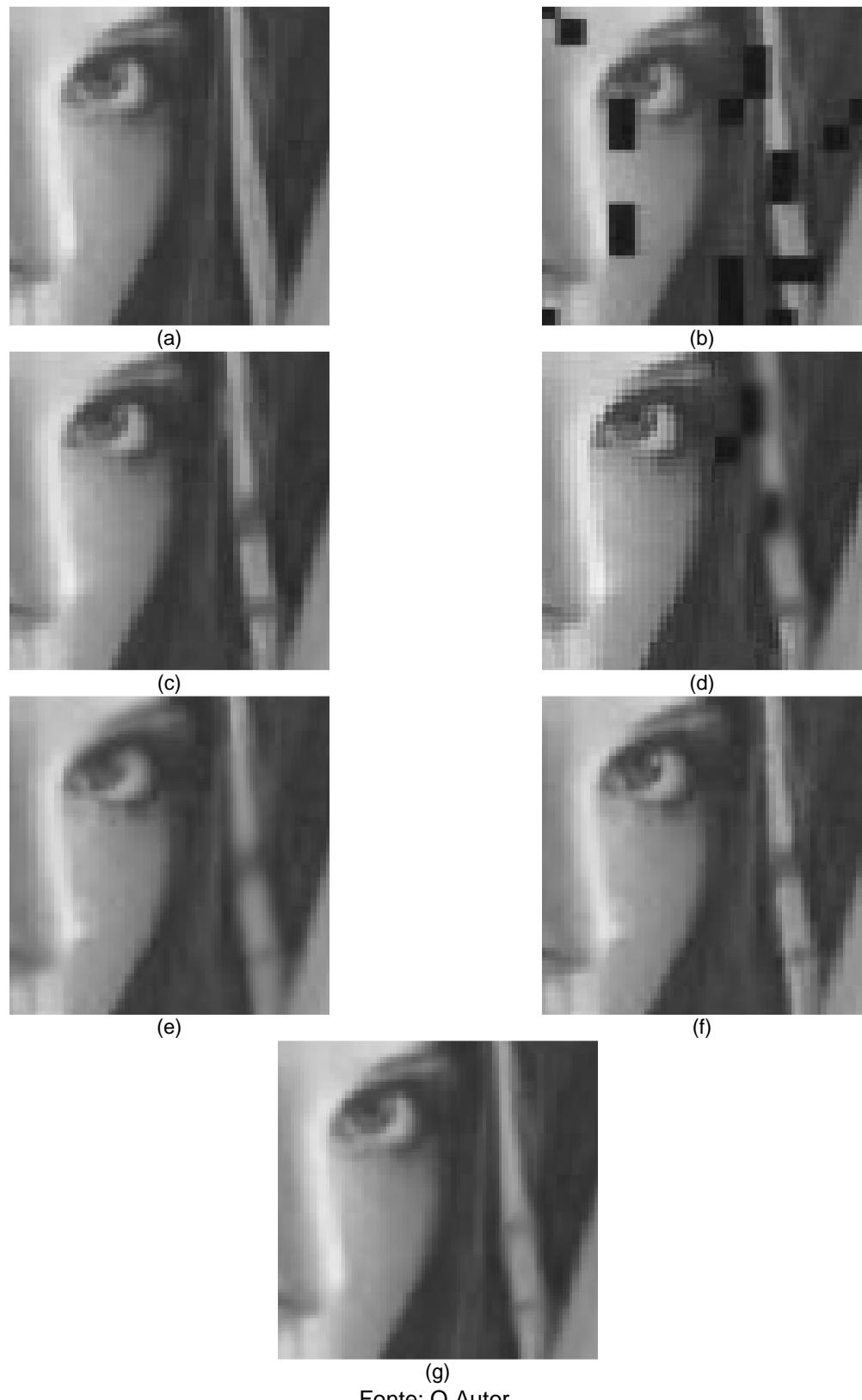
Como pode ser observado nas áreas mais suaves da imagem (figura 5.3), como a região dos ombros, da face e do plano de fundo, o desempenho dos três métodos é similar em termos subjetivos. Nas áreas que contêm bordas agudas, tais como na região do lado direito dos cabelos, o desempenho do algoritmo proposto é subjetivamente superior, suprimindo-se o efeito de blocagem adicionado pelos outros dois métodos comparados. Finalmente, nas áreas mais detalhadas, como na região dos enfeites do chapéu, os detalhes são mais bem reconstruídos pelo método proposto. Apesar da pequena melhora em termos da PSNR, em comparação aos outros métodos, a melhora subjetiva é considerável. A Tabela 5.1 apresenta os valores de PSNR obtidos para os métodos testados.

Tabela 5.1 – Valores de PSNR (dB) para os métodos de correção comparados.

Imagen	PSNR (dB)
Imagen ‘Lena’ corrompida	18.71
Imagen reconstruída pelo método de [90]	31.61
Imagen reconstruída pelo método de [92]	30.83
Imagen reconstruída pelo método de [98]	32.68
Imagen reconstruída pelo método de [101]	32.98
Imagen reconstruída pelo método proposto	34.42

Para evidenciar a qualidade das imagens obtidas pelo algoritmo proposto, a Fig. 5.4 ilustra regiões ampliadas retiradas das imagens anteriormente mostradas.

Fig. 5.4 - Conjunto de detalhes das imagens ‘Lena’. (a) Original. (b) Corrompida por ruído aleatório. (c) Corrigida pelo método da referência [90]. (d) Corrigida pelo método da referência [92]. (e) Corrigida pelo método da referência [98]. (f) Corrigida pelo método da referência [101]. (g) Corrigida pelo método proposto.



Fonte: O Autor.

Em nova simulação, a Fig. 5.5 mostra quadros reconstruídos para a aplicação dos cinco diferentes algoritmos testados para correção de erro, utilizando como sinal de entrada um quadro da versão corrompida da sequência de vídeo ‘Suzie’, por um traço diagonal preto inserido no quadro #100 (que modifica assim a correlação em linhas e colunas). Adicionalmente, a Fig. 5.6 ilustra as curvas de PSNR para diferentes valores de taxa de erro de *bit* (BER – *bit error ratio*), também fazendo a comparação dos métodos anteriormente mencionados, no que se observa que o método de *inpainting* proposto obtém melhores resultados quantitativos de correção à medida que se elevam as taxas de erro nas imperfeições inseridas nos quadros, a um menor custo computacional. A Fig. 5.7 ilustra os valores de PSNR para uma taxa de erro de *bit* igual a $BER = 5 \times 10^{-5}$ *bit/bit*, em que são levantados valores de PSNR para uma porção fixa de distorção inserida no quadro inicial da sequência tomada.

Fig. 5.5 - Quadro da sequência de vídeo ‘Suzie’. (a) Quadro original. (b) Quadro corrompido. (c) Quadro reconstruído por meio do método apresentado em [90]. (d) Quadro reconstruído por meio do método apresentado em [92]. (e) Quadro reconstruído por meio do método apresentado em [98]. (f) Quadro reconstruído por meio do método apresentado em [101]. (g) Quadro reconstruído com a função de escalonamento dada em (4.2). (h) Quadro reconstruído com a função de escalonamento corrigida, dada em (4.9).



(a)



(b)



(c)



(d)



(e)



(f)



(g)



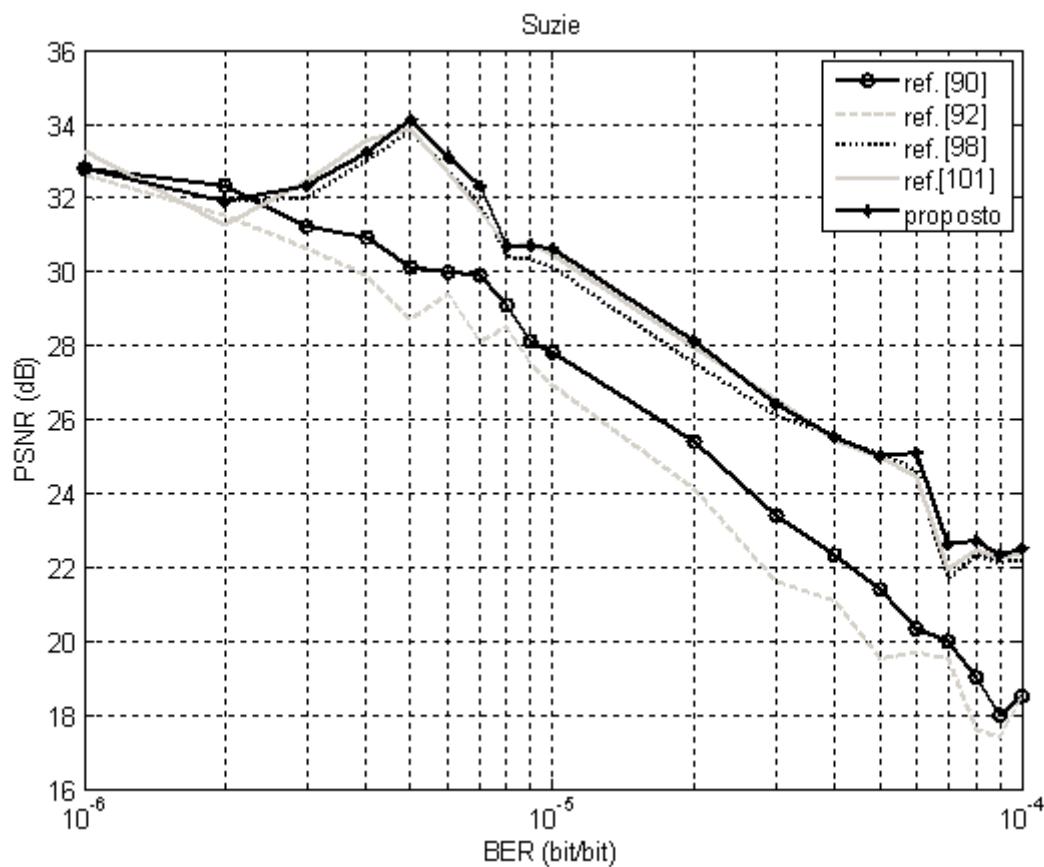
(h)

Fonte: O Autor.

Uma análise cuidadosa dos resultados permite enfatizar a qualidade de correção obtida pelo algoritmo proposto, em comparação aos demais, especificamente no que diz respeito a caracterizações precisas das contribuições de

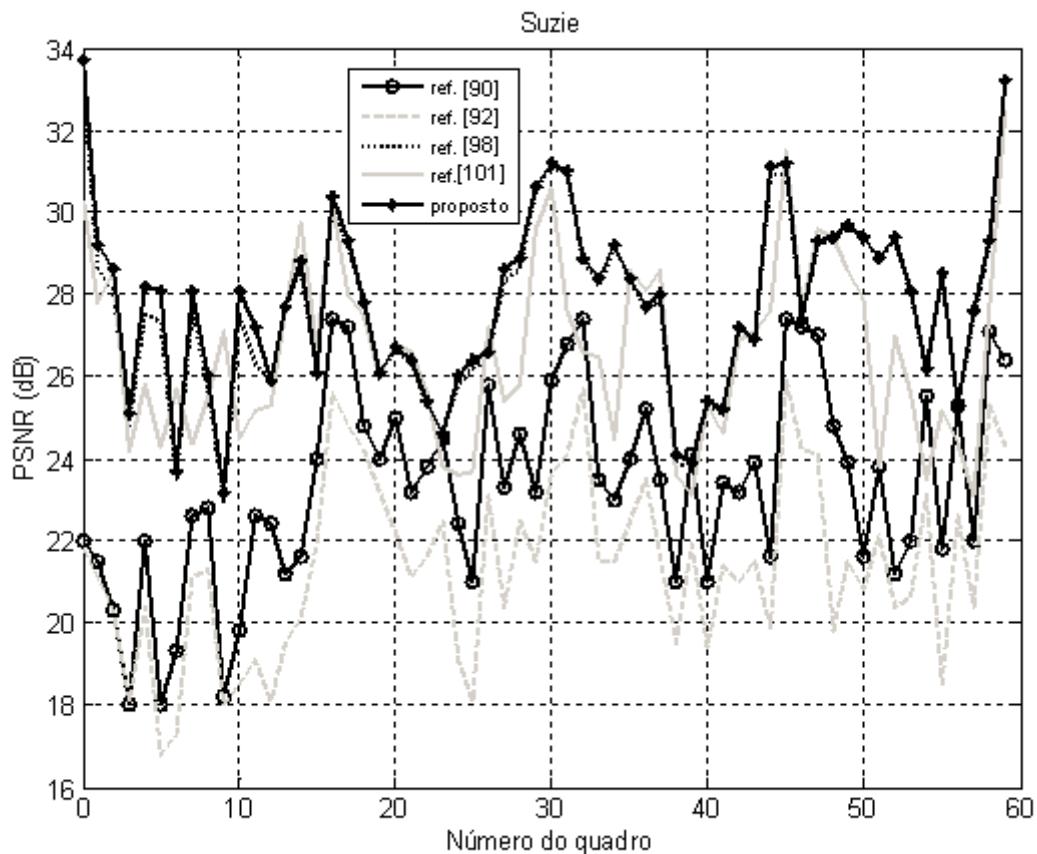
componentes de altas frequências (subjugadas quando do uso de técnicas de correção de erro puramente temporais): o conceito de estimação acurada do expoente de Hölder, como exposto nas observações teóricas introduzidas na literatura em [78] concernentes à correlação entre coeficientes *wavelet* de sub-bandas de altas frequências, permite uma definição de $g(j)$ que assegura uma estimativa mais acurada do expoente de Hölder. Assim, uma ampla variedade de aplicações que objetivam melhorar a qualidade da informação visual analisada pode ser desenvolvida a partir do conceito de exploração da correlação entre sub-bandas *wavelet* de altas frequências.

Fig. 5.6 - Curvas de PSNRxBER para um quadro corrompido da sequência de vídeo 'Suzie', ilustrado na Fig. 5.5.



Fonte: O Autor.

Fig. 5.7 - Curvas de PSNRxBER para diferentes quadros da seqüência ‘Suzie’ com $BER = 5 \times 10^{-5}$ bit/bit.



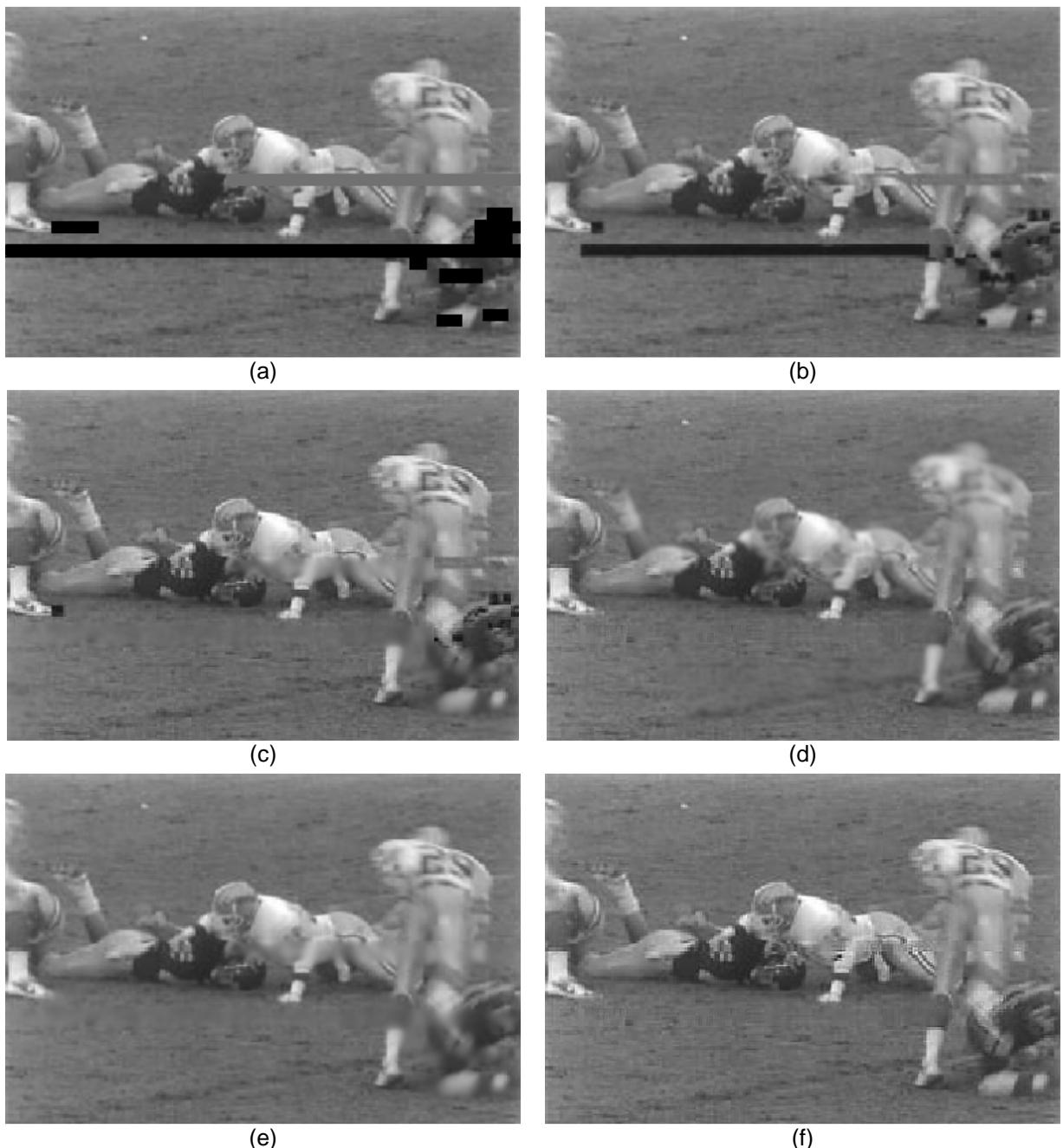
Fonte: O Autor.

Para a simulação de distorção por erros de *bits*, o algoritmo é aplicado à sequência ‘Football’, com 60 quadros codificados a 740×480 pixels em MPEG-2, a 30 quadros/s, no formato YUV 4:2:0, inserindo-se erros de *bits* em localizações aleatórias, a uma taxa de erros de *bits* (BER)¹⁰ que varia de 10^{-6} a 10^{-4} bits/bit. A

¹⁰ A BER (bit error rate) é a razão do número de *bits* incorretos recebidos pelo número de *bits* transmitidos. A medida clássica do BER é feita transmitindo-se um *bit* padrão conhecido e comparando este a um *bit* padrão recebido. Especificamente para transmissão de vídeo, altos valores de BER simbolizam quadros transmitidos de forma truncada. Para o cálculo da BER nesse caso, calcula-se, para um quadro transmitido, a razão entre o número de *bits* incorretos existentes e o número total de *bits* transmitidos. Por exemplo, para um quadro de 100×100 pixels e representação de 8 bits/pixel, um único pixel pode carregar um erro de 1 a 8 bits, do total de 80000 bits transmitidos na cena, conduzindo a um erro de $125 \cdot 10^{-7}$ a 10^{-4} bits/bit. Cada pixel com erro, por sua vez, contribui com uma parcela para o cálculo da BER, que, no pior caso, valerá 1 bit/bit.

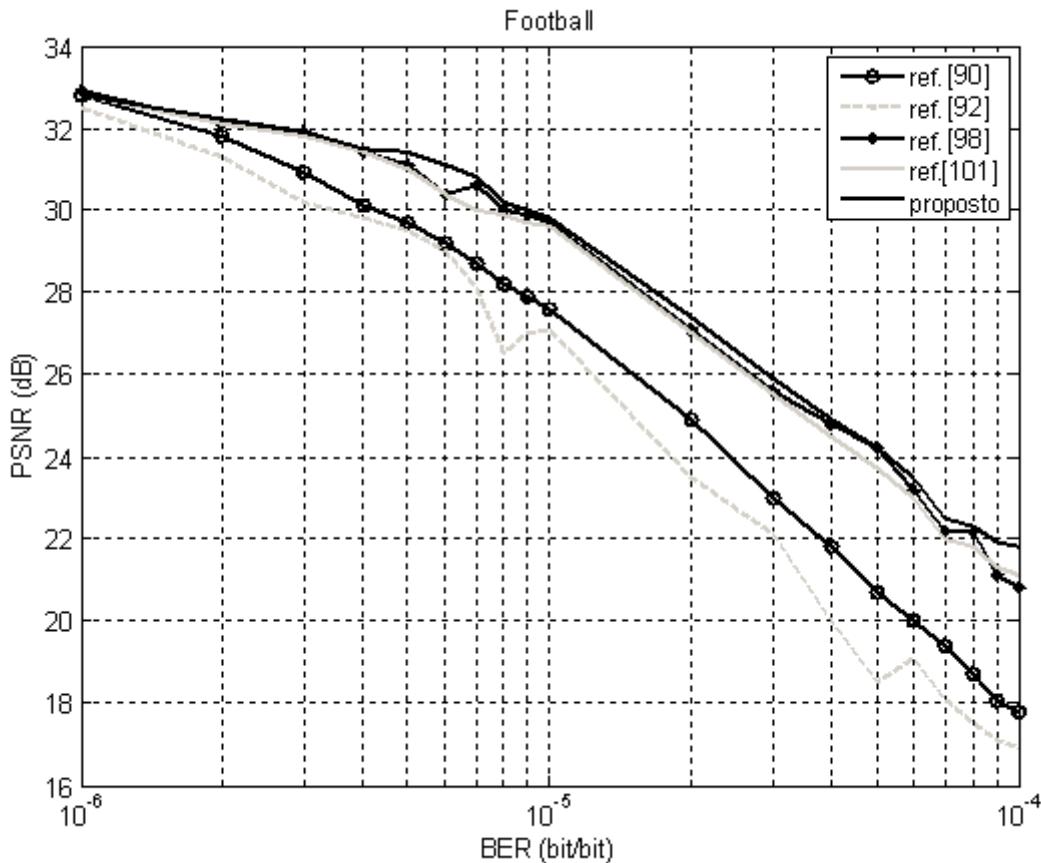
Figura 5.8 ilustra o quadro 40 da sequência, corrompido pelo efeito de erros de *bits*, bem como os quadros corrigidos pelos métodos de correção comparados [80, 82, 88, 91] e pelo método proposto. A Fig. 5.9 ilustra comparativamente o desempenho, em termos de PSNR, entre os métodos comparados.

Fig. 5.8 - Quadro da sequência ‘Football’. (a) Corrompido. (b) Corrigido pelo método apresentado em [90]. (c) Corrigido pelo método apresentado em [92]. (d) Corrigido pelo método apresentado em [98]. (e) Corrigido pelo método apresentado em [101]. (f) Corrigido pelo método aqui proposto.



Fonte: O Autor.

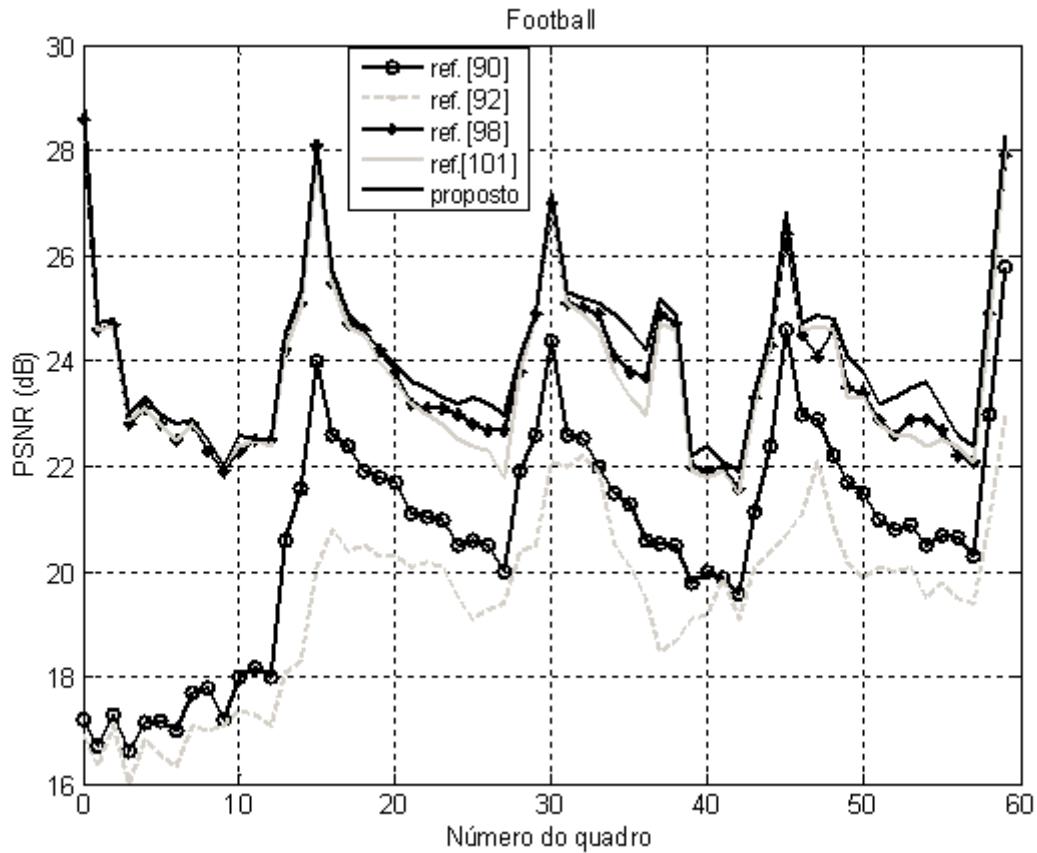
Fig. 5.9 - Curvas de PSNRxBER para o método proposto e para os apresentados em outras investigações [90, 92, 98, 101], aplicados sobre a sequência ‘Football’.



Fonte: O Autor.

Como pode ser observado na Figura 5.9, o algoritmo aqui proposto fornece valores melhores de PSNR em comparação aos outros quatro métodos testados. Nesse sentido, a Fig. 5.9 ilustra os valores de PSNR para os métodos comparados, para uma taxa de erro de *bits* fixada em 5×10^{-5} *bits/bit* (obtida para a inserção de 8 erros de *bits* em cada um dos 60 quadros da sequência), conforme a Fig. 5.10. Similarmente ao método da referência [96], no método proposto não há necessidade do quadro referência para correção do primeiro quadro, não existindo assim degradação de desempenho para o primeiro quadro.

Fig. 5.10 - Curvas de PSNR tomada para diferentes quadros da sequência de vídeo 'Football', com $\text{BER} = 5 \times 10^{-5}$ bits / bit , comparando o método proposto aos apresentados em outras investigações [90, 92, 98, 101].



Fonte: O Autor.

O algoritmo pode ser testado em aplicações práticas, como por exemplo, na remoção de legendas em filmes, tarefa que é espacialmente impossível de se realizar sem perda da informação visual de cada quadro. Isto se deve ao fato de que o processo de inserção de legendas consiste na sobreposição do texto, como imagem, sobre cada quadro da sequência de vídeo, perdendo-se a informação original do quadro. Para cada cena ou grupo de cenas com legendas sobrepostas, é sinalizado ao algoritmo interpretar tais legendas como distorções visuais que corrompem os quadros originais.

A Fig. 5.11 apresenta os resultados, para os métodos comparados, na remoção de legendas de um quadro da película 'Thank You for Smoking'.

Figura 5.11 - Quadro do filme ‘Thank You for Smoking’. (a) Quadro original. (b) Quadro com legendas. (c) Quadro reconstituído por meio do método apresentado em [90]. (d) Quadro reconstituído por meio do método apresentado em [92]. (e) Quadro reconstituído por meio do método apresentado em [98]. (f) Quadro reconstituído por meio do método apresentado em [101]. (g) Quadro reconstituído com a função de escalonamento definida pela equação (4.2). (h) Quadro reconstituído com a função de escalonamento corrigida, definida pela equação (4.9).



Fonte: O Autor.

Como observado pelos resultados da Fig. 5.11, o método aqui proposto não cancela de modo agressivo os componentes de características de altas frequências, enquanto reconstrói com alguma correção a informação original do quadro de vídeo. Entretanto, tais resultados podem ser refinados, quanto se ajuste a função de escalonamento corrigida que foi definida pela eq. (4.9).

Em termos de complexidade computacional, a carga exigida para implementação do algoritmo proposto é baixa, se comparada às demais. O método proposto realiza, para cada bloco do total de $M = P \times Q$ blocos de cada quadro, $O(M)$ operações aritméticas para cálculo da transformada *wavelet* dos blocos e detecção de quais blocos devem ser corrigidos. Para esses, são realizadas $O(N)$ operações aritméticas para correção dos blocos corrompidos. Isto totaliza uma carga computacional de $O(M + N)$ operações para o caso mais severo de erro (todos os blocos danificados e passíveis de correção). No pior caso, o algoritmo apresenta mesma complexidade computacional que o melhor caso dos métodos comparados, o que é vantajoso em termos de sua implementação em sistemas de transmissão de vídeo com poder de processamento limitado (bastante comuns em aplicações de vídeo). A tabela a seguir apresenta a função custo computacional para os métodos testados na correção de um quadro corrompido, que ressalta a vantagem de implementação da técnica proposta, em termos da menor quantidade de operações requerida para o *inpainting*. Isto assegura a aplicação do método sobre sequências de vídeo de longa duração ou agressivamente afetadas por distorções, sem perda significativa de qualidade na correção.

Tabela 5.2 – Complexidade computacional para os métodos de correção comparados.

Método	Custo computacional
Imagen reconstruída pelo método de [90]	$O(N^2)$
Imagen reconstruída pelo método de [92]	$O(M + N^2)$
Imagen reconstruída pelo método de [98]	$O(M + N)$
Imagen reconstruída pelo método de [101]	$O(M + N^2)$
Imagen reconstruída pelo método proposto	$O(M + N)$

5.2 CONCLUSÕES

Por meio de simulações, utilizando imagens estáticas e quadros de sequências de vídeo existentes na literatura, assim como utilizando películas de longa-metragem, esse capítulo apresentou a análise de desempenho, em termos subjetivos e de PSNR, para o algoritmo de correção proposto, considerando-se o efeito de blocos danificados por ruído aditivo e por erros de *bits*. Os resultados obtidos validam, qualitativamente e para uma ampla faixa de taxa de erros de *bits* (BER), o algoritmo proposto, qualificando-o como superior aos algoritmos de correção de erro utilizados para comparação. A utilização da propriedade de regularidade das *wavelets* é útil para tal propósito, comparativamente a técnicas espaciais ou frequências descritas no capítulo, no sentido de que as operações realizadas se mostram matematicamente mais bem definidas e simples, requerendo menor custo computacional para sua implementação, obtendo também melhores resultados em termos objetivos (PSNR) e subjetivos (perceptuais), aliados a uma robusta caracterização matemática dos sinais analisados – que assegura de forma bastante explícita a associação entre elementos indesejados nas sequências de vídeo e componentes *wavelet* passíveis de descarte/ponderação. A simplicidade de implementação e a rapidez de execução apresentadas pela referida abordagem permite, assim, o emprego do método em diferentes interfaces e aplicações que envolvam manipulação e transmissão de imagens.

6 CONCLUSÕES GERAIS E PERSPECTIVAS DO TRABALHO

Este trabalho descreve uma técnica de *inpainting* em imagens baseada na preservação da regularidade dos coeficientes *wavelet* do sinal corrompido por ruído. O método de *inpainting* proposto explora a preservação da propriedade da regularidade da transformada *wavelet*. Por meio dessa propriedade, estima-se a característica de decaimento dos coeficientes relativos a altas frequências, com o qual se constrói um critério de detecção e correção de erro, segundo a análise da função de correlação espacial entre blocos da imagem corrompida. Os resultados preliminares obtidos mostram o bom desempenho da proposta tanto em termos de PSNR como subjetivos, o que evidencia a boa escolha do critério para as etapas de detecção e correção de quadros corrompidos por ruído, com resultados superiores em relação a outros métodos do estado-da-arte. Buscou-se validar a eficiência do algoritmo para situações realísticas, fazendo-se, para tal, testes que simulassem a ocorrência de erros de *bits* e de blocos transmitidos, obtendo-se perceptível vantagem da aplicação do método em termos de velocidade de processamento.

Deve-se ressaltar que o algoritmo proposto não altera as características de codificação do sinal, uma vez que opera sobre os valores de *pixels* codificados. A funcionalidade do algoritmo como etapa de pós-processamento assegura que a tarefa de codificação não seja sobrecarregada em termos de complexidade computacional, bem como não exige da fonte, para melhor desempenho de correção, a adição de redundância à informação, já que a exploração de similaridade dos coeficientes através de múltiplas escalas manipula a redundância da informação no domínio transformado.

Vislumbrando-se a continuação do trabalho, outros tópicos podem ser explorados, de forma a aprimorar o algoritmo desenvolvido. Algumas propostas de continuação do trabalho são descritas a seguir.

- Desenvolvimento de um algoritmo de *Inpainting* baseado na obtenção bidimensional de expoentes de decaimento.

Apesar de a regularidade ser explorada em seus valores médios sobre blocos de *pixels*, a obtenção dos expoentes de decaimento ainda é realizada em uma única dimensão. A implementação dessa tarefa em blocos, de tamanho adequado para estabelecer um bom compromisso complexidade computacional *versus* precisão, tornaria o algoritmo mais rápido, sem perda de sua eficiência.

- Estudo da robustez do algoritmo a degradações mais agressivas provocadas por ruído.

Situações mais extremas no que se refere ao efeito do ruído sobre os quadros de vídeo, como perda total de quadros ou movimento brusco de câmera, não foram ainda consideradas, devendo ser testadas tanto em termos de avaliação subjetiva quanto em termos de complexidade computacional do algoritmo proposto para esse caso particular.

- Utilização do algoritmo como modelo de teste da recomendação H.264.

O algoritmo aqui apresentado se destina à operação em uma etapa de pós-processamento. Todavia, pode ser vislumbrada a possibilidade de utilizá-lo como modelo de teste na recomendação H.264 do ITU-T, que padroniza a codificação avançada de vídeo para serviços audiovisuais genéricos. Para isso, podem-se propor modificações na estrutura de correção do algoritmo proposto, de forma a se adaptar aos requisitos da recomendação H.264, o que se aplicaria simultaneamente ao processo de codificação de fonte.

- Simulações em tempo real.

Apesar de terem sido elaboradas formas alternativas de se avaliar, de maneira *offline*, o desempenho do algoritmo como se operasse em tempo real, através da inserção de erros de *bits*, sua avaliação em um ambiente de transmissão

que opere verdadeiramente em tempo real garantiria maior consistência dos resultados. Para tal, podem-se adicionar, ao mecanismo de detecção proposto nesse trabalho, outros algoritmos que se prestem a detectar particularmente erros do meio de transmissão.

REFERÊNCIAS

- [1] Don R. House. "A Synopsis of Teletype Corporation History". NADCOOM, 2001.
- [2] "Video coding for low bitrate communication," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.263, 1998.
- [3] Liu Xuedong and Wang Hong, "A fast motion segmentation algorithm based on hypothesis test for surveillance video coding," *2010 International Conference on Image Analysis and Signal Processing (IASP)*, vol. 1, no. 1, pp. 653-655, Apr. 2010.
- [4] A. Abou-Elailah *et al.*, "Fusion of Global and Local Motion Estimation for Distributed Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.23, no.1, pp. 158-172, Jan. 2013.
- [5] H. Magal *et al.*, "A Robust Error Resilient Video Compression Algorithm," *IEEE Military Communications Conference*, Conf. Record, Vol. 1, pp. 247-251, Oct. 1994.
- [6] Xiaotao Hou and Yonglin Xue, "Fast coding unit partitioning algorithm for HEVC," *2014 IEEE International Conference on Consumer Electronics (ICCE)*, vol. 1, no. 1, pp. 7-10, Jan. 2014.
- [7] M. Flierl and B. Girod, "Video Coding With Motion Compensation For Groups of Pictures," *IEEE Int. Conf. on Image Processing*, vol. 1, pp. 69-72, 2002.
- [8] Li Yu *et al.*, "A novel motion compensated prediction framework using weighted AMVP prediction for HEVC," *Visual Communications and Image Processing (VCIP)*, vol. 1, no. 1, pp. 1-6, Nov. 2013.
- [9] Ka-Ho Ng *et al.*, "A new motion compensation method using superimposed inter-frame signals," *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 1, no. 1, pp. 1213-1216, Mar. 2012.
- [10] Jia Su *et al.*, "Coarse to fine adaptive interpolation filter for high resolution video coding," *2011 IEEE International Conference on Multimedia and Expo (ICME)*, vol. 1, no. 1, pp. 1-6, Jul. 2011.
- [11] Tianmi Chen *et al.*, "Adaptive patch matching for motion compensated prediction," *2011 IEEE International Symposium on Circuits and Systems (ISCAS)*, vol. 1, no. 1, pp. 2621-2624, May 2011.

- [12] H. Lakshman *et al.*, "Generalized interpolation for motion compensated prediction," *2011 18th IEEE International Conference on Image Processing (ICIP)*, vol. 1, no. 1, pp. 1213-1216, Sep. 2011.
- [13] V. Ramanathan, S. Mishra and P. Mitra, "Quadtree decomposition based extended vector space model for image retrieval," *2011 IEEE Workshop on Applications of Computer Vision (WACV)*, vol. 1, no. 1, pp. 139-144, Jan. 2011.
- [14] T. K. Shih, N. C. Tang and Hwang Jenq-Neng, "Exemplar-Based Video Inpainting Without Ghost Shadow Artifacts by Maintaining Temporal Continuity," *IEEE Transactions on Circuits and Systems for Video Technology*, vol.19, no.3, pp. 347-360, Mar. 2009.
- [15] B. Vidhya and S. Valarmathy, "Novel video inpainting using patch sparsity," *International Conference on Recent Trends in Information Technology (ICRTIT)*, vol.1, no.1, pp. 810-814, Jun. 2011.
- [16] Faniu Wang *et al.*, "An new method for image inpainting using wavelets," *International Conference on Multimedia Technology (ICMT)*, vol.1, no.1, pp. 201-204, Jul. 2011.
- [17] D. Bhatnagar and S. Budhiraja, "Image compression using deterministic Compressive Sensing and Vector Quantization," *2014 Recent Advances in Engineering and Computational Sciences (RAECS)*, vol. 1, no. 1, pp. 1-5, Mar. 2014.
- [18] E. M. Eksioglu and O. Bayir, "K-SVD Meets Transform Learning: Transform K-SVD," *Signal Processing Letters*, IEEE, vol.21, no.3, pp. 347-351, Mar. 2014.
- [19] L. Jimenez-Rodriguez, F. Auli-Llinas and M. W. Marcellin, "Visually Lossless Strategies to Decode and Transmit JPEG2000 Imagery," *Signal Processing Letters*, IEEE, vol.21, no.1, pp. 35-38, Jan. 2014.
- [20] A. Huszak and S. Imre, "Analysing GOP structure and packet loss effects on error propagation in MPEG-4 video streams," *4th Int. Symposium on Communications, Control and Signal Processing (ISCCSP)*, vol.1, no.1, pp. 1-5, Mar. 2010.
- [21] J. Bhadviya *et al.*, "Exploitation of temporal redundancy for lossless video coding," *2012 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, vol. 1, no. 1, pp. 663-666, Dec. 2012.
- [22] Zhihong Ding, Lizhu Liu and Meijun Fan, "Correction of random and outburst error in MPEG-2 code," *2010 3rd IEEE International Conference on*

- Broadband Network and Multimedia Technology* (IC-BNMT), vol. 1, no. 1, pp. 771-774, Oct. 2010.
- [23] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. on Information Theory*, Vol. 13, no. 02, pp. 260-269, 1967.
- [24] P. Kannan, "PRATM - Derived PRASAN Error Concealment Model for Efficient Transmission of MPEG Video Streams Over ATM Networks," *2006 International Conference on Wireless Communications, Networking and Mobile Computing* (WiCOM) , vol. 1, no. 1, pp. 1-4, Sep. 2006.
- [25] R. Bernardini *et al.*, "Error concealment of intra coded video frames," *2010 11th International Workshop on Image Analysis for Multimedia Interactive Services* (WIAMIS), vol. 1, no.1, pp. 1-4, Apr. 2010.
- [26] H. Sun *et al.*, "Error Concealment in Digital Simulcast AD-HDTV Decoder," *IEEE Trans. on Consumer Electronics*, Vol. 38, no. 3, pp. 108-118, Aug. 1992.
- [27] Zhe Wu and Shaoshuai Gao, "An adaptive multi-level matching interpolation algorithm for spatial error concealment," *IET International Conference on Information and Communications Technologies* (IETICT 2013), vol. 1, no. 1, pp. 539-543, Apr. 2013.
- [28] S. Aign, "Error Concealment, Early Re-synchronization, and Iterative Decoding for MPEG-2," *IEEE International Conference on Communications*, Vol. 3, pp. 1654 -1658, Jun. 1997.
- [29] Chi-Cheng Ju *et al.*, "A full-HD 60fps AVS/H.264/VC-1/MPEG-2 video decoder for digital home applications," *2011 International Symposium on VLSI Design, Automation and Test* (VLSI-DAT), vol. 1, no. 1, pp. 1-4, Apr. 2011.
- [30] S. Tsekeridou *et al.*, "An Error Concealment Scheme for MPEG-2 Coded Video Sequences," *Proc. of IEEE International Symposium on Circuits and Systems*, Vol. 2, pp. 1289-1292, Jun. 1997.
- [31] H. R. Rabiee *et al.*, "Error Concealment of Still Image and Video Streams With Multi-directional Recursive Nonlinear Filters," *International Conference on Image Processing*, Vol. 2, pp. 37-40. Sep. 1996.
- [32] Guangtao Zhai *et al.*, "Bayesian error concealment with DCT pyramid," *2010 IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), vol. 1, no. 1, pp. 1366-1369, Mar. 2010.
- [33] B. Langari, S. Vaseghi and S. K. Pedram, "Multi-resolution edge-guided image gap restoration," *2013 IEEE International Symposium on Signal Processing and Information Technology* (ISSPIT), vol. 1, no. 1, pp. 374-379, Dec. 2013.

-
- [34] Z. Ding, L. Liu and M. Fan, "Correction of random and outburst error in MPEG-2 code," *2010 3rd IEEE International Conference on Broadband Network and Multimedia Technology* (IC-BNMT), vol. 1, no. 1, pp. 771-774, Oct. 2010.
 - [35] F. Sohel, M. Bennamoun and M. Hahn, "Spatial shape error concealment utilising image texture," *2011 6th IEEE Conference on Industrial Electronics and Applications* (ICIEA), vol. 1, no. 1, pp. 265-270, Jun. 2011.
 - [36] P. Cuenca *et al.*, "A Survey of Error Concealment Schemes for MPEG-2 Video Communications Over ATM Networks," *IEEE Canadian Conference on Electrical and Computer Engineering*, Vol. 1, pp. 118-121, May 1997.
 - [37] B. Yun, M. Park and H. Hong, "POCS-based error concealment using interlayer correlation and features of neighbor blocks in multi-layer video coding," *2010 Digest of Technical Papers International Conference on Consumer Electronics* (ICCE), vol. 1, no. 1, pp. 385-386, Jan. 2010.
 - [38] M. Elad, M. A. T. Figueiredo and Yi Ma, "On the Role of Sparse and Redundant Representations in Image Processing," Proceedings of the IEEE, vol.98, no.6, pp. 972-982, Jun. 2010.
 - [39] M. V. Afonso, J. M. Bioucas-Dias, M. A. T. Figueiredo, "An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems," *Image Processing, IEEE Transactions on*, vol.20, no.3, pp. 681-695, Mar. 2011.
 - [40] F.-C. Jeng and S. H. Lee, "Concealment of Bit Error and Cell Loss in Inter-frame Coded Video Transmission," *IEEE Trans. on Communications*, Vol. 40, pp. 496-500, 1991.
 - [41] MPEG-1 Video Coding Standard, ISO/IEC 11172-2, Information Technology - coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s-Part2: Video. ISO, 1993.
 - [42] International Telecommunication Union, Telecommunication Standardization Sector – "Line Transmission of non-Telephone Signals - Video Codec for Audiovisual Services at $p \times 64$ Kbit/s" - ITU-T Recommendation H.261, March 1993.
 - [43] T. Na *et al.*, "A Hybrid Stereoscopic Video Coding Scheme Based on MPEG-2 and HEVC for 3DTV Services," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 9, pp. 1542-1554, Sep. 2013.
 - [44] Wen-Nung Lie *et al.*, "Motion Vector Recovery for Video Error Concealment by Using Iterative Dynamic-Programming Optimization," *IEEE Transactions on Multimedia*, vol.16, no.1, pp. 216-227, Jan. 2014.

-
- [45] B. Macchiavello *et al.* "Loss-Resilient Coding of Texture and Depth for Free-Viewpoint Video Conferencing," *IEEE Transactions on Multimedia*, vol.16, no.3, pp. 711-725, Apr. 2014.
 - [46] J. Chakareski, "Transmission Policy Selection for Multi-View Content Delivery Over Bandwidth Constrained Channels," *IEEE Transactions on Image Processing*, vol.23, no.2, pp. 931-942, Feb. 2014.
 - [47] W. El-Shafai, "Optimized Adaptive Space-Time-View Multi-Dimensional Error Concealment for 3D Multi-View Video transmission," *2013 Saudi International Electronics, Communications and Photonics Conference (SIECPC)*, vol. 1, no. 1, pp. 1-6, Apr. 2013.
 - [48] J. Seiler, M. Schoberi, A. Kaup, "Spatio-temporal error concealment in video by denoised temporal extrapolation refinement," *20th IEEE International Conference on Image Processing (ICIP)*, vol. 1, no. 1, pp. 1613-1616, Sept. 2013.
 - [49] B. Yan, J. Zhou, "Efficient Frame Concealment for Depth Image-Based 3-D Video Transmission," *IEEE Transactions on Multimedia*, vol.14, no.3, pp. 941-945, Jun. 2012.
 - [50] Dong Nguyen, T. Tran, Thinh Nguyen and B. Bose, "Wireless Broadcast Using Network Coding," *IEEE Transactions on Vehicular Technology*, vol.58, no.2, pp. 914-925, Feb. 2009.
 - [51] Bo Yan and Jie Zhou, "Efficient Frame Concealment for Depth Image-Based 3-D Video Transmission," *IEEE Transactions on Multimedia*, vol.14, no.3, pp. 936-941, Jun. 2012.
 - [52] Y. Wang *et al.*, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol.17, no.4, pp. 61-82, Jul. 2000.
 - [53] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - Video Coding for Low Bitrate Communications" - Draft ITU-T Recommendation H.263, July 1995.
 - [54] Y. Wang and Q.-F. Zhu, "Error Control and Concealment for Video Communication: A Review," *Proc. of the IEEE*, Vol. 86, Issue 5, pp. 974-997, May 1998.
 - [55] "Information technology - Generic coding of moving pictures and associated audio information - Part 2: Video," ISO/IEC DIS 13818-2, 1994.
 - [56] "Frame structure for a 64 to 1920 kbit/s channel in audiovisual teleservices," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.221, 1993.

- [57] "Narrow-band visual telephone systems and terminal equipment," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.320, 1993.
- [58] T. Wiegand *et al.*, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, 2003.
- [59] V.N.G.J. Soares, P.A.C.S. Neves and J.J.P.C. Rodrigues, "Past, Present and Future of IP Telephony," *International Conference on Communication Theory, Reliability, and Quality of Service*, vol.1, no.1, pp. 19-24, Jun. 2008.
- [60] "Terminal for low bitrate multimedia communication," International Telecommunications Union, Geneva, Switzerland, ITU-T Draft Recommendation H.324, Dec. 1997.
- [61] "Media stream packetization and synchronization on nonguaranteed quality of services LANs," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.225.0, 1996.
- [62] "Telephone systems and equipment for local area networks which provide a non-guaranteed quality of service," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.323, Visual 1996.
- [63] Y. Wu *et al.*, "Cross-Layer Forward Error Correction Scheme Using Raptor and RCPC Codes for Prioritized Video Transmission Over Wireless Channels," *IEEE Trans. on Circuits and Systems for Video Technology*, vol.24, no.6, pp. 1047-1060, Jun. 2014.
- [64] S. Mishra and S. Sawarkar, "Various techniques for low bit rate video coding," *2012 International Conference on Communication, Information & Computing Technology (ICCICT)*, vol. 1, no. 1, pp. 1-6, Oct. 2012.
- [65] L. Wang and Y. Fan, "A Kind of Dynamic Bandwidth Allocation Scheme Suitable for MPEG Video in Broadband Network," *'08 International Symposium on Computational Intelligence and Design (ISCID)*, vol.2, no. 1, pp. 79-82, Oct. 2008.
- [66] "Multiplexing protocol for low bitrate multimedia communication," International Telecommunications Union, Geneva, Switzerland, ITU-T Recommendation H.223, 1995.
- [67] P. Farkas, M. Rakús and T. Páleník, "One error control scheme for ATM header," *2010 IEEE Region 8 International Conference on Computational Technologies in Electrical and Electronics Engineering (SIBIRCON)*, vol. 1, no. 1, pp. 55-58, Jul. 2010.

-
- [68] Y. Fang, H. Li and X. Li, "Lifetime Enhancement Techniques for PCM-Based Image Buffer in Multimedia Applications," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no.6, pp. 1450-1455, Jun. 2014.
 - [69] A. Pushpakar *et al.* "Resourcesat-2 image restoration for differential pulse code modulation compression artifacts," *2013 IEEE Second International Conference on Image Information Processing (ICIIP)*, vol. 1, no. 1, pp. 545-549, Dec. 2013.
 - [70] O. R. Mitchell and A. J. Tabatabai, "Channel Error Recovery for Transform Image Coding," *IEEE Trans. on Communications*, Vol. COM-29, pp. 1754-1762, Dec. 1981.
 - [71] S. Narayanan and A. Makur, "Compressive coded video compression using measurement domain motion estimation," *2014 IEEE International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCCT)*, vol. 1, no. 1, pp. 1-6, Jan. 2014.
 - [72] S. Benchikh and M. Corinthios, "Efficiency evaluation of different wavelets for image compression," *11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, vol. 1, no.1, pp. 1420-1421, 2-5 July 2012.
 - [73] A. Eleyan, H. Ozkaramanli and H. Demirel, "Dual-tree and single-tree complex wavelet transform based face recognition," *IEEE 17th Signal Processing and Communications Applications Conference, SIU 2009*, vol.1, no.1, pp. 536-539, 9-11 April 2009.
 - [74] R. A. Vora, V.A. Bharadi and H. B. Kekre, "Retinal scan recognition using wavelet energy entropy," *International Conference on Communication, Information & Computing Technology (ICCICT)*, vol.1, no.1, pp. 1-6, 19-20 Oct. 2012.
 - [75] C. Naftornita and A. Isar, "Estimating directional smoothness of images with the aid of the hyperanalytic wavelet packet transform," *2013 International Symposium on Signals, Circuits and Systems (ISSCS)*, vol. 1, no. 1, pp. 1-4, Jul. 2013.
 - [76] Y. Meyer, "Wavelets and Operators," *Cambridge University Press*, Cambridge, 1992.
 - [77] I. Daubechies, "Ten Lectures on Wavelets," CBMS-NSF Reg. Conf. Series in Appl. Mathematics (SIAM), Vol. 61, Philadelphia, 1992.
 - [78] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, pp. 674-693, Jul. 1989.

- [79] G. Strang, "Creating and Comparing Wavelets," Department of Mathematics, MIT, Cambridge, Massachusetts, 1994.
- [80] M. F. Cohen *et al.* "Wavelets and their Applications in Computer Graphics," Siggraph '95 Notes, 1995.
- [81] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 7, pp. 710-732, 1992.
- [82] W. K. Carey, D. B. Chuang and S. S. Hemami, "Regularity-Preserving Image Interpolation," *IEEE Transactions on Image Processing*, vol.8, no.9, pp. 1293-1297, Sep. 1999.
- [83] S. R. M. Penedo, "Interpolação de Imagens Através da Análise de Regularidade Baseada em Decomposição Wavelet." Florianópolis, 2000. Dissertação (Mestrado em Engenharia Elétrica) - LINSE – UFSC.
- [84] K. Berkner, "Enhancement of scanned documents in Besov spaces with applications to image enhancement using wavelet domain representations," Proceedings of SPIE, vol. 4670, pp. 143-154, San Jose, 2002.
- [85] K. Berkner, E.L. Schwartz and M.J. Gormish, "Multiscale sharpening and smoothing in Besov spaces with applications to image enhancement," Applied and Computational Harmonic Analysis, special issue on wavelet applications, vol. 11, no. 1, pp. 2-31, July 2001.
- [86] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projections into convex sets," *IEEE Trans. on Image Processing*, vol. 4, pp. 470–477, Apr. 1995.
- [87] X. Lee *et al.*, "Information Loss Recovery for Block-based Image Coding Techniques—A Fuzzy Logic Approach," *IEEE Trans. Image Processing*, vol. 4, pp. 259–273, Mar. 1995.
- [88] Y. Wang and Q. Zhu, "Signal Loss Recovery in DCT-based Image and Video Codecs," in *Proc. SPIE Visual Communications and Image Processing '91*, Boston, MA, pp. 667–678, Nov. 1991.
- [89] I. Wang *et al.*, "Coding and cell-loss recovery in DCT-based packet video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 248-258, Jun. 1993.
- [90] Y. Wang, Q. Zhu, and L. Shaw, "Maximally Smooth Image Recovery in Transform Coding," *IEEE Trans. on Communications*, vol. 41, pp. 1544–1551, Oct. 1993.

-
- [91] S. S. Hemami and T. H. -Y. Meng, "Transform coded image reconstruction exploiting interblock correlation," *IEEE Trans. Image Processing*, vol. 4, no. 7, pp. 1023–1027, 1995.
 - [92] W. Zhu and Y. Wang, "Second-order derivative-based smoothness measure for error concealment in DCT-based codecs," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.8, Iss. 6, pp. 713-718, Oct. 1998.
 - [93] S. S. Hemami and T. H. -Y. Meng, "Spatial and temporal video reconstruction for nonlayered transmission," in *Proc. VISCOM '93, Fifth Int. Workshop on Packet Video*, Berlin, Germany, Mar. 1993.
 - [94] S. S. Hemami and R. M. Gray, "Image reconstruction using vector quantized linear interpolation," in *Proc. ICASSP '94*, Adelaide, Australia, pp. 629–631, Apr. 1994.
 - [95] P. Kannan and R. Pravin Kumar, "Adaptive Localized Curve Tracing Methodology for Efficient Transmission of Audio Streams over Noisy ATM Networks," *2006 Annual IEEE India Conference*, vol. 1, no. 1, pp. 1-5, Sep. 2006.
 - [96] R. Zhang, Y. Chen and M. Karczewicz, "Adaptive Depth edge sharpening for 3D video depth coding," *2012 IEEE Visual Communications and Image Processing (VCIP)*, vol. 1, no. 1, pp. 1-6, Nov. 2012.
 - [97] M. Ghanbari and V. Seferidis, "Cell-loss concealment in ATM video codecs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, pp. 238–247, June 1993.
 - [98] C.-T. Chen, "Error detection and concealment with an unsupervised MPEG2 video decoder," *J. Visual Commun. Image Represent.*, vol. 6, pp. 265–279, Sep. 1995.
 - [99] O. Hadar *et al.*, "New Hybrid Error Concealment for Digital Compressed Video," *EURASIP Journal on Applied Signal Processing*, vol. 12, pp. 1821-1833, Dec. 2005.
 - [100] W. Lie *et al.*, "Motion Vector Recovery for Video Error Concealment by Using Iterative Dynamic-Programming Optimization," *IEEE Transactions on Multimedia*, vol. 16, no. 1, pp. 216-227, Jan. 2014.
 - [101] V. V. Voronin *et al.*, "Image inpainting using cubic spline-based edge reconstruction," Proc. SPIE8295, Image Processing: Algorithms and Systems X; and Parallel Processing for Imaging Applications II, 82950I, Feb. 2012.
 - [102] P. Nasiopoulos and R. K. Ward, "Improving the HDTV picture performance under noisy transmission conditions," in *Proc. ICASSP '93*, Minneapolis, MN, pp. 277-280. Apr. 1993.

- [103] Wen-Nung Lie *et al.* “Motion Vector Recovery for Video Error Concealment by Using Iterative Dynamic-Programming Optimization,” *IEEE Transactions on Multimedia*, vol.16, no.1, pp. 216-227, Jan. 2014.
- [104] A. B. Ibrahim and A. H. Sadka, “Implementation of error resilience technique for Multiview Video Coding,” *2014 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, vol. 1, no. 1, pp. 25-28, Apr. 2014.