

UNIVERSIDADE DE CAXIAS DO SUL

RODRIGO BRANDALISE

MÍNIMOS QUADRADOS APLICADOS À SUPER-RESOLUÇÃO DE VÍDEO

CAXIAS DO SUL

2014

RODRIGO BRANDALISE

MÍNIMOS QUADRADOS APLICADOS À SUPER-RESOLUÇÃO DE VÍDEO

Dissertação apresentada como requisito para a
obtenção do título de Mestre em Engenharia
Mecânica pela Universidade de Caxias do Sul.

Área de concentração: Projeto e Fabricação.

Orientador: Prof. Dr. Guilherme H. Costa.

CAXIAS DO SUL

2014

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
UCS - BICE - Processamento Técnico

B817m Brandalise, Rodrigo
Mínimos quadrados aplicados à super-resolução de vídeo / Rodrigo
Brandalise - 2014.
39 f. : il. ; 30 cm

Dissertação (Mestrado) – Universidade de Caxias do Sul, Programa de
Pós-Graduação em Engenharia Mecânica, 2014.
Apresenta bibliografia.
“Orientação: Prof. Dr. Guilherme H. Costa”.

1. Resolução (Óptica) . 2. Mínimos quadrados. I. Título.

CDU 2.ed.: 535.317.25

Índice para o catálogo sistemático:

1. Resolução (Óptica)	535.317.25
2. Mínimos quadrados	528.14

Catalogação na fonte elaborada pela bibliotecária
Carolina Meirelles Meroni – CRB 10/ 2187

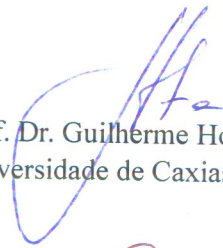
"Mínimos Quadrados Aplicados à Super-Resolução de Vídeo".

Rodrigo Brandalise


Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Mecânica da Universidade de Caxias do Sul, como parte dos requisitos necessários para a obtenção do título de Mestre em Engenharia Mecânica, Área de Concentração: Projeto e Fabricação.

Caxias do Sul, 20 de fevereiro de 2014.


Banca Examinadora:



Prof. Dr. Guilherme Holsbach Costa (orientador)
Universidade de Caxias do Sul




Prof. Dr. Carlos Roberto Altafini (coorientador)
Universidade de Caxias do Sul



Profa. Dra. Adriana Miorelli Adami
Universidade de Caxias do Sul



Prof. Dr. Alexandre Fassini Michels
Universidade de Caxias do Sul



Prof. Dr. Julio Cesar Ceballos Aya
Universidade de Caxias do Sul

RODRIGO BRANDALISE

MÍNIMOS QUADRADOS APLICADOS À SUPER-RESOLUÇÃO DE VÍDEO

Dissertação apresentada como requisito para a obtenção do título de Mestre em Engenharia Mecânica pela Universidade de Caxias do Sul.

Área de concentração: Projeto e Fabricação.

Aprovado em __/__/__

Banca Examinadora

Prof. Dr. Guilherme Holsbach Costa
Universidade de Caxias do Sul – UCS

Prof. Dr. Carlos Roberto Altafini
Universidade de Caxias do Sul – UCS

Prof. Dr. Alexandre Fassini Michels
Universidade de Caxias do Sul – UCS

Prof. Dr. Julio Cesar Ceballos Aya
Universidade de Caxias do Sul – UCS

Prof. Dra. Adriana Miorelli Adami
Universidade de Caxias do Sul – UCS

Dedico este trabalho à minha esposa Angélica e ao meu filho Leonardo. Muitas vezes abdiquei da companhia deles para realizar este projeto.

AGRADECIMENTOS

Agradeço primeiramente ao Professor Guilherme, que além proporcionar ensinamentos imensuráveis, teve muita paciência e persistência em nossas reuniões de orientação e acompanhamento do projeto.

Agradeço minha mãe Dione e a minha irmã Letícia, que nos momentos mais críticos estiveram ao nosso lado dando um suporte familiar.

Agradeço a minha esposa Angélica, que além do apoio dado em todos os momentos de dificuldade, teve muita resiliência comigo. Entendendo por resiliência a capacidade de o indivíduo lidar com problemas, superar obstáculos ou resistir à pressão de situações adversas - choque, estresse etc. Muito obrigado!

RESUMO

Neste trabalho é proposta uma adaptação do método de Mínimos Quadrados aplicado à reconstrução de imagens com Super-Resolução, visando a reconstrução de vídeo (sequências de imagens) em tempo real. Resultados demonstram que a implementação proposta pode apresentar um desempenho melhor que o algoritmo de menor custo computacional da literatura, com um pequeno incremento no número de operações. Por fim, a estrutura proposta sugere viabilidade de análise. De posse de um modelo teórico de comportamento, parâmetros ótimos de projeto podem ser obtidos melhorando ainda mais o desempenho do algoritmo.

Palavras-chaves: Super-Resolução, reconstrução de imagens, Mínimos Quadrados.

ABSTRACT

In this work an adaptation of the Least Squares method applied to Super-Resolution image reconstruction is proposed aiming real time video (image sequences) reconstruction. Results demonstrate that the proposed implementation presents a better performance than the algorithm with the lower computational cost presented in the literature, considering a small increase in the number of operations. Finally, the proposed structure suggests feasibility of analysis. A theoretical model for the algorithm behavior can lead to an optimal parameters design yielding further improvements in the algorithm performance.

Keywords: Super-Resolution, image reconstruction, Least Squares.

ÍNDICE DE FIGURAS

Figura 1- Sistema de som para análise de vibrações.	13
Figura 2 - Imagem das fontes de vibrações de um motor.	13
Figura 3: Diagrama de obtenção da SRR.	14
Figura 4: Diagrama de blocos do modelo de degradação.	19
Figura 5: Critérios de desempenho.....	20
Figura 6: Imagem de 225×225 pixels que gerou as imagens HR.	32
Figura 7: Imagens da Simulação: (a) imagem original. (b) imagem LR. (c) reconstrução via interpolação bicúbica. (d) reconstrução via LMS-SRR. (e) reconstrução via LS.	34
Figura 8: Evolução temporal do erro quadrático médio de reconstrução.....	34
Figura 9: Resultados de simulação: (a) imagem LR. (b) reconstrução via interpolação bicúbica. (c) reconstrução via LMS-SRR. (d) reconstrução via LS.	36
Figura 10: Resultados de simulação: (a) reconstrução via LMS-SRR (ampliação de 3x). (b) reconstrução via LS (ampliação de 3x).	36

LISTA DE SIGLAS

EQMR	Erro Quadrático Médio de Reconstrução
HR	<i>High Resolution</i>
LMS	<i>Least Mean Square</i>
LR	<i>Low Resolution</i>
LS	<i>Least Squares</i>
SDTV	<i>Standard Definition Television</i>
SRR	<i>Super Resolution Reconstruction</i>

LISTA DE SÍMBOLOS

D	Matriz de degradação.
$e(t)$	Vetor de ruído.
F	Matriz de registro entre $\mathbf{x}(t-1)$ e $\mathbf{x}(t)$.
$G(t)$	Matriz de registro entre $\mathbf{x}(t)$ e $\mathbf{x}(t-1)$.
I	Matriz identidade.
K	Número de observações.
k	Índice das observações.
$L(\hat{\mathbf{x}})$	Erro quadrático em função $\hat{\mathbf{x}}$.
L_1	Critério de erro absoluto.
L_2	Critério de erro quadrático.
l	Número de iterações do algoritmo por amostra (imagem) de entrada.
M	Número de linhas/colunas das imagens de alta resolução (HR).
N	Número de linhas/colunas das imagens de baixa resolução (LR).
$s(t)$	Vetor das inovações em $\mathbf{x}(t)$.
t	Tempo (discreto).
$\mathbf{X}(t)$	Representação matricial da imagem desejada (HR) no instante t .
$\mathbf{x}(t)$	Representação lexicográfica (vetorial) de $\mathbf{X}(t)$.
$\hat{\mathbf{X}}(t)$	Representação matricial da imagem reconstruída (HR) no instante t .
$\hat{\mathbf{x}}(t)$	Representação lexicográfica (vetorial) de $\hat{\mathbf{X}}(t)$.
$\mathbf{Y}(t)$	Representação matricial da imagem observada (LR) no instante t .
$\mathbf{y}(t)$	Representação lexicográfica (vetorial) de $\mathbf{Y}(t)$.
∂	Derivada.
$\delta(t)$	Vetor das inovações em $\mathbf{x}(t-1)$.
$\boldsymbol{\varepsilon}$	Vetor de ruído e inovações.
$\hat{\boldsymbol{\varepsilon}}$	Estimativa do vetor de ruído e inovações.
∇	Gradiente.
μ	Passo de adaptação do algoritmo.

SUMÁRIO

1	INTRODUÇÃO	12
1.1	CONTEXTUALIZAÇÃO	14
1.2	JUSTIFICATIVA.....	15
1.3	OBJETIVOS	17
1.3.1	Objetivo geral	17
1.3.2	Objetivos específicos.....	17
2	MATERIAL E MÉTODOS	18
2.1	MODELAGEM DOS SINAIS	18
2.2	ESTIMAÇÃO POR MÍNIMOS QUADRADOS	20
2.3	MÉTODO DO GRADIENTE	22
3	IMPLEMENTAÇÃO PROPOSTA.....	23
3.1	CUSTO COMPUTACIONAL	28
4	RESULTADOS	31
4.1	CONDIÇÕES GERAIS.....	31
4.2	AVALIAÇÃO QUANTITATIVA.....	32
4.3	AVALIAÇÃO QUALITATIVA.....	35
5	CONCLUSÕES	37
	REFERÊNCIAS BIBLIOGRÁFICAS.....	38

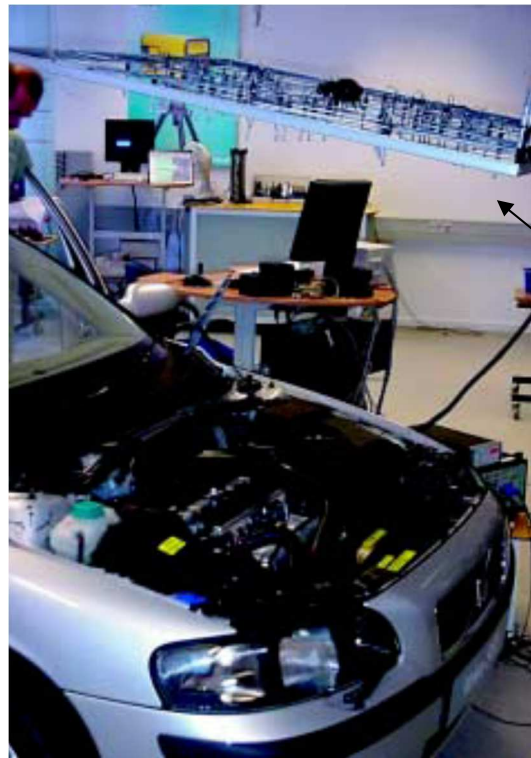
1 INTRODUÇÃO

A cada dia, com maior frequência, engenheiros mecânicos estão empregando sistemas de imagem para solucionar problemas de engenharia em seus mecanismos ou ainda para mensurar seus experimentos com mais eficácia e assertividade. Além do uso dos sensores tradicionais, imagens de peças e de máquinas podem ser geradas com a utilização de sensores baseados em ultrassom, som, infravermelho e outros. Em geral essas imagens são caracterizadas pela baixa definição, ainda que possam se beneficiar de técnicas de processamento digital de sinais para melhorar a sua qualidade.

Em (BATEL *et al.*, 2003) é apresentada uma técnica de análise de fontes de vibração que utiliza imagens geradas através do som. A Figura 1 ilustra um sistema de aquisição dessas imagens sendo aplicado em motores de combustão interna, em que é montado um arranjo de microfones para captar as fontes de vibrações através do ruído. Esse arranjo tem forma de uma matriz quadrada. Após a aquisição do som é construído um mapa no qual cada microfone corresponde a um pixel de uma imagem. A Figura 2 mostra a imagem resultante desse sistema de aquisição sobreposta à imagem do motor adquirida por um sensor convencional. A intensidade do som é representada por cores, indo do azul ao vermelho, conforme a amplitude do sinal. Cada microfone indica, portanto, a posição e a intensidade dos ruídos no equipamento facilitando a localização das fontes de vibração mais relevantes.

Como exemplo da resolução do arranjo apresentado na Figura 1, para que se tenha uma imagem com 20x20 pontos (pixels) são necessários 400 microfones. Já para uma imagem com 40x40 pixels, precisam-se 1600 microfones. Percebe-se que para dobrar as dimensões da imagem é necessário que se eleve ao quadrado o número de sensores, elevando-se sobremaneira os custos com equipamentos (sensores e condicionadores de sinais). Outra forma de melhorar a qualidade das imagens sem altos investimentos em *hardware* é o aumento das dimensões da imagem através de técnicas como a reconstrução com super resolução (PARK *et al.*, 2003) que, para esse aumento, utiliza apenas recursos computacionais, conforme será discutido na próxima seção.

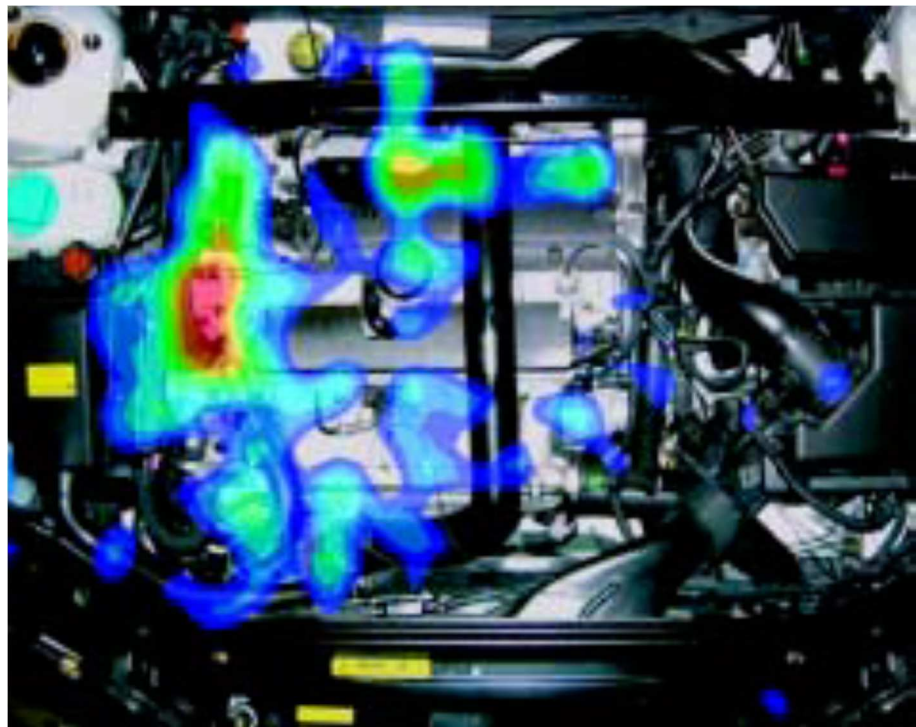
Figura 1- Sistema de som para análise de vibrações.



Arranjo de
microfones

Fonte: Batel at al (2003).

Figura 2 - Imagem das fontes de vibrações de um motor.

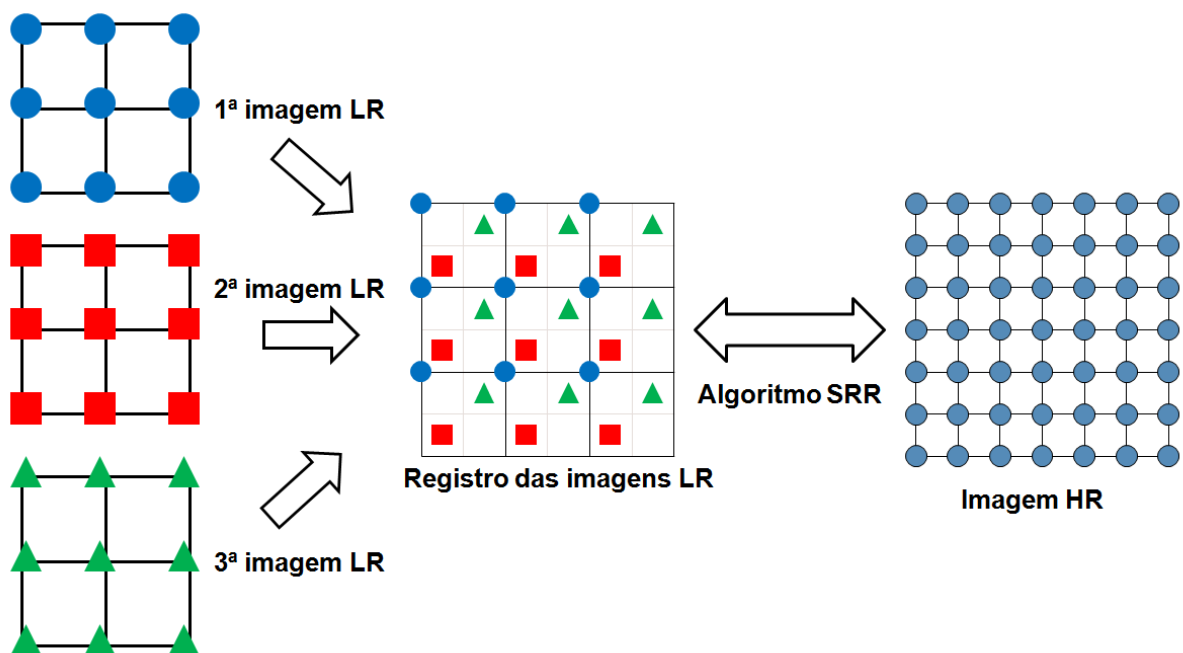


Fonte: Batel at al (2003).

1.1 CONTEXTUALIZAÇÃO

A reconstrução com super-resolução (SRR — *Super-Resolution Reconstruction*) consiste basicamente em combinar múltiplas imagens de baixa resolução (LR — *Low Resolution*) para formar uma imagem de alta (maior) resolução (HR — *High Resolution*). Na Figura 3 é ilustrado um processo simplificado de SRR em que três imagens de baixa resolução (3×3) pixels são reorganizadas através de uma etapa chamada registro, e fundidas em uma imagem com maior resolução. O processo de registro consiste na estimação da posição relativa entre imagens, considerando a existência de deslocamentos relativos entre os objetos presentes nas cenas (PARK *et al.*, 2003). A imagem de alta resolução pode ser construída se existirem inovações (novas informações) entre as imagens de baixa resolução. Essas novas informações manifestam-se quando ocorrem deslocamentos relativos com dimensão menor que um pixel, chamados deslocamentos de subpixel. Caso a posição relativa entre as imagens LR seja a mesma (sem deslocamento) ou caso o deslocamento relativo entre elas seja múltiplo da dimensão de um pixel, não se obtém a SRR, ainda que possa ocorrer, no segundo caso, um aumento da dimensão da imagem.

Figura 3: Diagrama de obtenção da SRR.



A SRR tem atraído considerável interesse nas últimas décadas. Em Park *et al.* (2003), são revisados importantes resultados em SRR disponíveis na literatura. Esses resultados têm sido aplicados com sucesso em áreas como fotografia digital (FARSIU *et al.*, 2006), sistemas de segurança (CHAUDHURI & TAUTR, 2005; FORESTI *et al.* 2005; HAMPAPUR *et al.*, 2005; CHEN *et al.* 2005), sistemas de visão computacional e sistemas biométricos (GUNTURK *et al.*, 2002), dentre outros.

1.2 JUSTIFICATIVA

Um dos principais problemas relacionados à SRR é o fato de os algoritmos serem fortemente dependentes de um modelo acurado das variáveis que compõem o problema (FARSIU *et al.*, 2004). Essa dependência é conhecida como sensibilidade a *outliers* (que podem ser definidos como dados cuja característica estatística não segue o modelo assumido). Por esse motivo, as abordagens mais tradicionais, que recaem na minimização da norma¹ L_2 do erro de reconstrução, vêm sendo gradativamente preteridas pela minimização da norma L_1 desse erro (FARSIU *et al.*, 2004; ZOOMET *et al.*, 2001). Neste caso, o estimador de média, tradicionalmente utilizado para calcular o gradiente nos algoritmos de SRR, é substituído por um estimador de mediana, mais robusto a *outliers*.

Outra característica da SRR é o alto custo computacional. Quando a aplicação de interesse requer processamento em tempo real, algoritmos rápidos devem ser empregados (COSTA & BERMUDEZ, 2007). Em (ELAD & FEUER, 1999), é proposto um algoritmo para SRR de sequências de imagens baseado no conceito do gradiente estocástico (HAYKIN, 2002). Esse algoritmo, chamado LMS-SRR (*Least Mean Square – Super Resolution Reconstruction*), apresenta um custo computacional consideravelmente baixo, quando comparado aos demais. Em contrapartida, o erro de reconstrução é acrescido de um erro em excesso causado pela estimativa instantânea do gradiente (HAYKIN, 2002). Em outras palavras, o custo computacional do algoritmo é reduzido em detrimento da qualidade de reconstrução.

¹ Norma é uma nomenclatura utilizada em processamento de sinais para mensurar a magnitude de uma matriz ou um vetor (MEYER, 2000). Neste trabalho são denominadas as norma L_1 e norma L_2 respectivamente para as normas absoluta e euclidiana.

Outro algoritmo para reconstrução de sequências de imagens, com boa robustez à influência de variáveis externas, foi proposto por Farsiu *et al.* (2004). Esse algoritmo admite uma implementação rápida, porém apenas para movimentos translacionais globais. Tanto o custo computacional quanto os requisitos de memória necessários para a implementação desse algoritmo superam os do LMS-SRR, o que caracteriza uma desvantagem. Outra desvantagem desse algoritmo é a formulação altamente não linear de suas equações, característica da maioria dos algoritmos baseados em norma L_1 , que praticamente inviabiliza a análise de seu comportamento.

Diferente do algoritmo proposto por Farsiu *et al.* (2004), a viabilidade de análise do algoritmo LMS-SRR tem permitido que seu desempenho seja maximizado frente a *outliers*, por ser baseado em norma L_2 . A formulação matemática desse algoritmo permite a aplicação de uma taxonomia de análise largamente empregada no estudo do comportamento de filtros adaptativos aplicados a sinais unidimensionais (COSTA & BERMUDEZ, 2007). Uma vez que se obtenha um modelo de comportamento para um determinado algoritmo, é possível interpretar de forma qualitativa a relação entre suas variáveis (COSTA & BERMUDEZ, 2008) e definir metodologias e parâmetros ótimos de projeto, bem como garantir determinados níveis de desempenho frente a qualquer tipo de aplicação (COSTA & BERMUDEZ, 2007). Esse tipo de análise teórica é bastante comum em outras áreas de aplicação de Processamento de Sinais, mas ainda é muito pouco difundida no campo de Processamento de Imagens e de Super-Resolução, provavelmente pela complexidade computacional oriunda da característica bidimensional dos sinais envolvidos, bem como da dificuldade de modelagem dos sinais de imagens. De outra forma, os algoritmos são geralmente avaliados apenas frente a casos específicos e testados através de um número relativamente pequeno exemplos, sem garantias de desempenho para qualquer outro conjunto de imagens de teste diferente do utilizado na avaliação.

Diversos algoritmos apresentados na literatura recaem na formulação de Mínimos Quadrados (LS - *Least Squares*) (PARK *et al.*, 2003). Neste trabalho é proposta uma implementação rápida dessa abordagem, favorecendo sua aplicação em sistemas tempo real. O algoritmo proposto apresenta uma estrutura muito similar à do LMS-SRR, indicando viabilidade de análise teórica, o que pode levá-lo a níveis desejáveis de robustez e a garantias de desempenho. Simulações demonstram que o algoritmo proposto pode apresentar resultados

consideravelmente superiores aos do LMS-SRR, com um pequeno acréscimo no custo computacional, sendo aplicável a qualquer modelo de movimento. Espera-se ainda que, de posse de um modelo teórico, um projeto ótimo possibilite resultados ainda melhores.

Na próxima seção são definidos os objetivos deste trabalho. No Capítulo 2 é brevemente apresentada a formulação LS, visto que a abordagem LS é tradicionalmente aplicada a imagens estáticas. No Capítulo 3 uma adequação da notação é apresentada de forma a aplicar essa solução à reconstrução de sequências de imagens (vídeos), como no caso do LMS-SRR, bem como é apresentada a implementação proposta. No Capítulo 4 são apresentados os resultados obtidos e, por fim, no último capítulo são apresentadas as conclusões do trabalho.

1.3 OBJETIVOS

1.3.1 Objetivo geral

Propor uma implementação para super-resolução de imagens baseada na formulação de mínimos quadrados (LS - *Least Squares*) adequada à reconstrução de sequências de imagens (vídeo), favorecendo aplicações em tempo real.

1.3.2 Objetivos específicos

- a) Obter resultados de super-resolução superiores aos do algoritmo LMS-SRR (o mais rápido encontrado na literatura atual) com custo computacional compatível;
- b) Obter um método de super-resolução de vídeo com custo computacional compatível ao do LMS-SRR frente a qualquer modelo de movimento;
- c) Obter um novo algoritmo ou nova implementação cuja estrutura matemática sugira viabilidade de análise teórica através das técnicas e taxonomias conhecidas.

2 MATERIAL E MÉTODOS

Neste capítulo a formulação de mínimos quadrados é brevemente revisada, de forma a definir a notação utilizada ao longo do trabalho. Deste ponto em diante, vetores serão representados por letras minúsculas em negrito, matrizes por letras maiúsculas em negrito e escalares por letras minúsculas itálicas. A variável t é assumida inteira e indexa amostras em tempo discreto de sinais e de operadores. A imagem de baixa resolução (observada) é chamada de imagem LR (*Low Resolution*) enquanto a imagem reconstruída (desejada) é chamada de imagem HR (*High Resolution*).

Na secção 2.1 é apresentada a modelagem dos sinais para a obtenção das imagens observadas, assim como a modelagem para estimativa do erro de reconstrução e seus critérios de desempenho. Na secção 2.2 é apresentada a modelagem matemática do método de estimação por mínimos quadrados e na secção 2.3 a modelagem pelo método de gradiente.

2.1 MODELAGEM DOS SINAIS

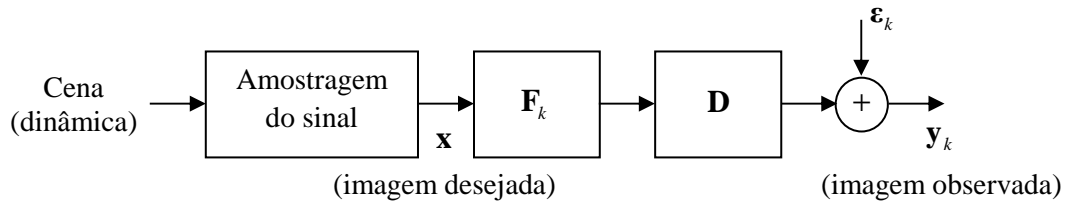
Dada uma representação matricial de uma imagem digital de alta resolução HR (desejada) \mathbf{X} , com dimensões $M \times M$, e a representação matricial da k -ésima observação desta mesma imagem com baixa resolução (LR), \mathbf{Y}_k , com dimensões $N \times N$ ($M > N$), \mathbf{y}_k ($N^2 \times 1$) e \mathbf{x} ($M^2 \times 1$) são respectivamente as representações lexicográficas das k observações (imagens observadas) e da imagem original (desejada), \mathbf{Y}_k e \mathbf{X} . A matriz \mathbf{D} de dimensões $N^2 \times M^2$ modela as degradações referentes à (sub) amostragem e às distorções ópticas inerentes ao processo de aquisição. A matriz \mathbf{F}_k é chamada de matriz de registro, e descreve o deslocamento relativo entre as diferentes observações \mathbf{y}_k e a imagem HR \mathbf{x} . O vetor $\boldsymbol{\varepsilon}_k$ ($N^2 \times 1$) modela o ruído (eletrônico) e as inovações presentes na cena, em cada observação. Essas inovações podem ser devidas a movimentos de objetos ou do sistema de aquisição.

De acordo com Park *et al.* (2003), o processo de obtenção da imagem observada é tradicionalmente modelado como:

$$\mathbf{y}_k = \mathbf{D}\mathbf{F}_k\mathbf{x} + \boldsymbol{\varepsilon}_k. \quad (1)$$

Esse modelo, representado na Figura 4 através de um diagrama de blocos, assume que a imagem \mathbf{x} é limitada em banda e adquirida de acordo com o teorema da amostragem de Nyquist (HAYKIN, 2007).

Figura 4: Diagrama de blocos do modelo de degradação.



Na SRR deseja-se buscar uma estimativa $\hat{\mathbf{x}}$ de \mathbf{x} . Uma avaliação dessa estimativa poderia ser obtida a partir da diferença entre $\hat{\mathbf{x}}$ e \mathbf{x} . Entretanto, em aplicações reais, \mathbf{x} não está disponível. Uma forma de avaliação da estimativa, então, pode ser obtida diretamente de (1), conforme:

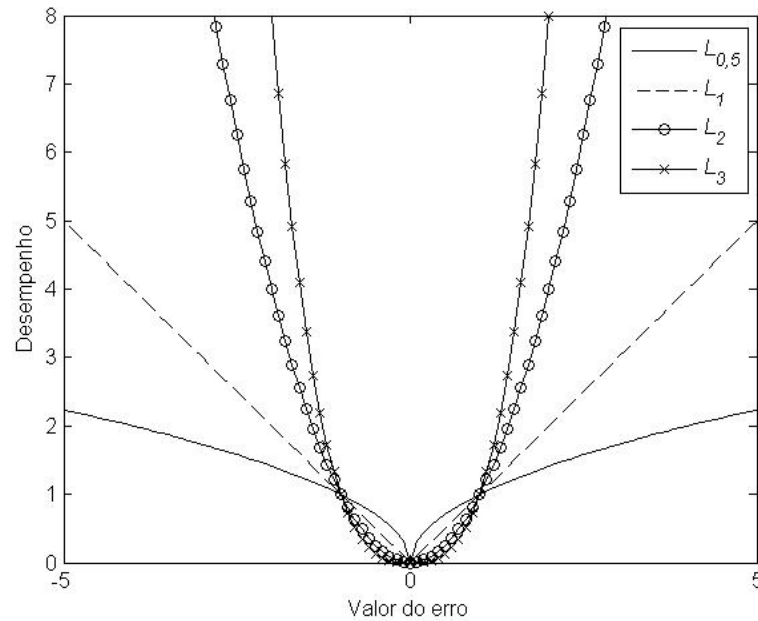
$$\boldsymbol{\varepsilon}_k = \mathbf{y}_k - \mathbf{D}\mathbf{F}_k\mathbf{x}. \quad (2)$$

Substituindo-se \mathbf{x} por sua estimativa $\hat{\mathbf{x}}$ em (2), pode-se avaliar o erro de reconstrução conforme

$$\hat{\boldsymbol{\varepsilon}}_k = \mathbf{y}_k - \mathbf{D}\mathbf{F}_k\hat{\mathbf{x}}. \quad (3)$$

Uma vez que, na maioria das aplicações, erros negativos e positivos são igualmente prejudiciais ao sistema, deve ser escolhido um critério de desempenho que considera essa ambiguidade. A forma mais comum é escolher uma figura de mérito de valor absoluto. Existem diversas possibilidades para essa escolha, estando entre elas a norma L_1 , $\|\hat{\boldsymbol{\varepsilon}}_k\|$, a norma L_2 (ou erro quadrático), $\|\hat{\boldsymbol{\varepsilon}}_k\|_2^2$, ou ainda algum outro critério a exemplo dos ilustrados na Figura 5.

Figura 5: Critérios de desempenho.



O critério de desempenho pode ser decisivo no comportamento do algoritmo. Por exemplo, o critério de erro quadrático L_2 realça o efeito dos grandes erros muito mais que o critério de erro absoluto L_1 , e é também mais sensível a *outliers* (MANOLAKIS *et al.*, 2000). Por esses motivos é que o método de estimação por mínimos quadrados requer sua aplicação baseado na norma L_2 conforme descrito a seguir na secção 2.2.

2.2 ESTIMAÇÃO POR MÍNIMOS QUADRADOS

A abordagem de mínimos quadrados, também conhecida como regressão linear, leva à estimativa $\hat{\mathbf{x}}$ pela minimização da soma dos erros quadráticos (MANOLAKIS *et al.*, 2000),

$$\begin{aligned}
 L(\hat{\mathbf{x}}) &= \sum_{k=0}^{K-1} \|\hat{\mathbf{e}}_k\|_2^2 \\
 &= \sum_{k=0}^{K-1} \|\mathbf{y}_k - \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}}\|_2^2,
 \end{aligned} \tag{4}$$

em que K é o número total de observações da imagem de baixa resolução.

A solução $\hat{\mathbf{x}}$ que minimiza (4) pode ser obtida como se segue. Expandindo-se (4), tem-se que

$$\begin{aligned} L(\hat{\mathbf{x}}) &= \sum_{k=0}^{K-1} [\mathbf{y}_k - \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}}]^T [\mathbf{y}_k - \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}}] \\ &= \sum_{k=0}^{K-1} \mathbf{y}_k^T \mathbf{y}_k - \mathbf{y}_k^T \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}} - \hat{\mathbf{x}}^T \mathbf{F}_k^T \mathbf{D}^T \mathbf{y}_k + \hat{\mathbf{x}}^T \mathbf{F}_k^T \mathbf{D}^T \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}}. \end{aligned} \quad (5)$$

O ponto de mínimo dessa função pode ser encontrado igualando-se a zero a primeira derivada de (5), ou seja,

$$\frac{\partial L(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}} = \sum_{k=0}^{K-1} -2\mathbf{F}_k^T \mathbf{D}^T \mathbf{y}_k + 2\mathbf{F}_k^T \mathbf{D}^T \mathbf{D}\mathbf{F}_k \hat{\mathbf{x}} = \mathbf{0}. \quad (6)$$

De fato, para que este seja um ponto de mínimo, a Hessiana de (5) deve ser verificada como definida positiva em $\hat{\mathbf{x}}$, ou seja,

$$\frac{\partial^2 L(\hat{\mathbf{x}})}{\partial \hat{\mathbf{x}}^2} = \sum_{k=0}^{K-1} \mathbf{F}_k^T \mathbf{D}^T \mathbf{D}\mathbf{F}_k. \quad (7)$$

Assim sendo, isolando-se $\hat{\mathbf{x}}$ em (6), tem-se:

$$\left[\sum_{k=0}^{K-1} \mathbf{F}_k^T \mathbf{D}^T \mathbf{D}\mathbf{F}_k \right] \hat{\mathbf{x}} = \sum_{k=0}^{K-1} \mathbf{F}_k^T \mathbf{D}^T \mathbf{y}_k. \quad (8)$$

A matriz de degradação \mathbf{D} possui dimensões $N^2 \times M^2$ e, portanto, $\mathbf{D}^T \mathbf{D}$ não possui posto completo. Sendo assim, a soma de termos $\mathbf{F}_k^T \mathbf{D}^T \mathbf{D}\mathbf{F}_k$ não possui inversa. A estimativa $\hat{\mathbf{x}}$ pode ser determinada a partir da equação (8) pela pseudo-inversa do termo que pré-multiplica $\hat{\mathbf{x}}$. Porém, dadas as grandes dimensões das matrizes envolvidas, esse cálculo é geralmente impraticável na maioria das aplicações em tempo real. Uma solução comumente empregada é a aplicação do método do gradiente, discutido na seção seguinte.

Cabe ainda ressaltar que, assim como pela abordagem de mínimos quadrados, diversas outras soluções recaem na minimização da função em (4), que pode ou não incluir restrições (PARK *et al.*, 2003). Isso torna o método proposto ainda mais abrangente.

2.3 MÉTODO DO GRADIENTE

A solução $\hat{\mathbf{x}}$ que minimiza a função em (4) pode ser obtida de forma iterativa pelo método do gradiente, de acordo com a seguinte equação recursiva (HAYKIN, 2007):

$$\hat{\mathbf{x}}^{l+1} = \hat{\mathbf{x}}^l + \frac{\mu}{2} \nabla L(\hat{\mathbf{x}}^l), \quad (9)$$

em que μ é chamado passo do algoritmo, e controla a estabilidade e velocidade de convergência.

Substituindo-se em (9) o gradiente determinado em (6) tem-se:

$$\begin{aligned} \hat{\mathbf{x}}^{l+1} &= \hat{\mathbf{x}}^l + \frac{\mu}{2} \sum_{k=0}^{K-1} -2\mathbf{F}_k^T \mathbf{D}^T \mathbf{y}_k + 2\mathbf{F}_k^T \mathbf{D}^T \mathbf{D} \mathbf{F}_k \hat{\mathbf{x}}^l \\ &= \hat{\mathbf{x}}^l + \mu \sum_{k=0}^{K-1} \mathbf{F}_k^T \mathbf{D}^T \left[\mathbf{D} \mathbf{F}_k \hat{\mathbf{x}}^l - \mathbf{y}_k \right]. \end{aligned} \quad (10)$$

A equação (10) corresponde a uma solução iterativa que leva à imagem de alta resolução estimada, $\hat{\mathbf{x}}$, pela minimização do erro quadrático através do método de gradiente. No entanto, precisa-se primeiro adquirir as K observações da imagem de baixa resolução para que o algoritmo seja executado, em outras palavras, o processamento deve ocorrer *off-line*. No Capítulo 3 é proposta uma nova formulação que viabiliza a execução em tempo real desse algoritmo, ou seja, em sua nova forma o processamento pode ser iniciado mesmo antes que se tenham todas as observações disponíveis.

3 IMPLEMENTAÇÃO PROPOSTA

Quando a reconstrução com super-resolução trata de sequências temporais de imagens, a modelagem dos sinais é definida de maneira ligeiramente diferente de (1), geralmente através de um modelo de aquisição e de um modelo de dinâmica da cena. O modelo de aquisição pode ser definido como (ELAD & FEUER, 1999):

$$\mathbf{y}(t) = \mathbf{D}\mathbf{x}(t) + \mathbf{e}(t), \quad (11)$$

em que agora as imagens $\mathbf{y}(t)$ e $\mathbf{x}(t)$ são funções do tempo (discreto), bem como $\mathbf{e}(t)$, que neste caso modela apenas o ruído de aquisição (eletrônico). Note-se que, neste modelo, a cena original $\mathbf{x}(t)$, que geralmente representa uma sequência de vídeo, possui uma dinâmica temporal (não é estática como no caso anterior). Em aplicações práticas de SRR de vídeo, o efeito de subamostragem é geralmente invariante no tempo, visto que, por via de regra, é oriundo de um único sensor de imagem. Por esse motivo, a matriz de degradação \mathbf{D} não é modelada como uma função do tempo.

A imagem em alta resolução $\mathbf{x}(t)$ quando observada em diferentes instantes de tempo, sofre uma transformação modelada conforme

$$\mathbf{x}(t) = \mathbf{G}(t)\mathbf{x}(t-1) + \mathbf{s}(t), \quad (12)$$

em que $\mathbf{G}(t)$ é a matriz de registro que modela o movimento relativo de $\mathbf{x}(t)$ em relação a imagem anterior $\mathbf{x}(t-1)$. A equação (12) é denominada modelo de dinâmica do sinal e o detalhamento deste processo é apresentado (COSTA & BERMUDEZ, 2007), que aborda a construção da matriz de registro a partir do conhecimento dos deslocamentos referentes a cada pixel da imagem. O vetor $\mathbf{s}(t)$ modela as inovações em $\mathbf{x}(t)$, também em relação a $\mathbf{x}(t-1)$. Essas inovações ocorrem muitas vezes devido a oclusões de objetos, causadas por sobreposições em apenas uma das imagens, ou por causa do movimento relativo entre a câmera e a cena, que altera a região da cena adquirida pelo sensor de um quadro em relação ao anterior.

Costa e Bermudez (2007) também descrevem a importância das condições de fronteira na construção da matriz de registro, considerando as diversas hipóteses para o preenchimento dos pixels referentes às inovações. Assumir que esses pixels têm valores

nulos, por exemplo, equivale à hipótese de que todos os pixels externos à cena observada no instante t são pretos. Esse tipo de implementação é chamado de preenchimento com zeros (*zero padding*), também conhecido como condição de fronteira de Dirichlet. Essa condição de fronteira geralmente produz um efeito indesejável nas regiões próximas às bordas da imagem reconstruída (COSTA & BERMUDEZ, 2007). Portanto, a hipótese assumida para o preenchimento dos pixels referentes às inovações deve ser incluída na construção da matriz $\mathbf{G}(t)$.

De forma a permitir a futura relação entre (1) e (11), inicialmente é possível dizer que, por analogia com (12), a dinâmica do sinal pode ser também descrita por

$$\mathbf{x}(t-1) = \mathbf{F}(t)\mathbf{x}(t) + \boldsymbol{\delta}(t), \quad (13)$$

em que, no sentido (temporal) inverso de $\mathbf{G}(t)$, $\mathbf{F}(t)$ modela o movimento relativo da imagem no instante de tempo anterior, $\mathbf{x}(t-1)$, em relação a imagem no instante atual t . O vetor $\boldsymbol{\delta}(t)$ contém as inovações em $\mathbf{x}(t-1)$, em relação a $\mathbf{x}(t)$.

Substituindo-se t por $t-1$ em (13), tem-se

$$\mathbf{x}(t-2) = \mathbf{F}(t-1)\mathbf{x}(t-1) + \boldsymbol{\delta}(t-1). \quad (14)$$

Substituindo-se (13) em (14), chega-se em

$$\begin{aligned} \mathbf{x}(t-2) &= \mathbf{F}(t-1)[\mathbf{F}(t)\mathbf{x}(t) + \boldsymbol{\delta}(t)] + \boldsymbol{\delta}(t-1) \\ &= \mathbf{F}(t-1)\mathbf{F}(t)\mathbf{x}(t) + \mathbf{F}(t-1)\boldsymbol{\delta}(t) + \boldsymbol{\delta}(t-1). \end{aligned} \quad (15)$$

Substituindo-se t por $t-1$ em (15), expandindo-se o resultado, e assim sucessivamente, tem-se

$$\begin{aligned} \mathbf{x}(t-3) &= \mathbf{F}(t-2)\mathbf{F}(t-1)\mathbf{F}(t)\mathbf{x}(t) + \\ &\quad + \mathbf{F}(t-2)\mathbf{F}(t-1)\boldsymbol{\delta}(t) + \\ &\quad + \mathbf{F}(t-2)\boldsymbol{\delta}(t-1) + \\ &\quad + \boldsymbol{\delta}(t-2), \end{aligned} \quad (16)$$

$$\begin{aligned}
\mathbf{x}(t-4) &= \mathbf{F}(t-3)\mathbf{F}(t-2)\mathbf{F}(t-1)\mathbf{F}(t)\mathbf{x}(t) + \\
&\quad + \mathbf{F}(t-3)\mathbf{F}(t-2)\mathbf{F}(t-1)\boldsymbol{\delta}(t) + \\
&\quad + \mathbf{F}(t-3)\mathbf{F}(t-2)\boldsymbol{\delta}(t-1) + \\
&\quad + \mathbf{F}(t-3)\boldsymbol{\delta}(t-2) + \\
&\quad + \boldsymbol{\delta}(t-3)
\end{aligned} \tag{17}$$

e assim consecutivamente. Por inspeção de (13) a (17), é possível descrever a dinâmica do sinal em termos de uma imagem atrasada em k amostras, $\mathbf{x}(t-k)$, conforme

$$\begin{aligned}
\mathbf{x}(t-k) &= \mathbf{F}(t-k+1)\mathbf{F}(t-k+2)\cdots\mathbf{F}(t)\mathbf{x}(t) + \\
&\quad + \mathbf{F}(t-k+1)\mathbf{F}(t-k+2)\cdots\mathbf{F}(t-1)\boldsymbol{\delta}(t) + \\
&\quad + \mathbf{F}(t-k+1)\mathbf{F}(t-k+2)\cdots\mathbf{F}(t-2)\boldsymbol{\delta}(t-1) + \\
&\quad + \mathbf{F}(t-k+1)\mathbf{F}(t-k+2)\cdots\mathbf{F}(t-3)\boldsymbol{\delta}(t-2) + \\
&\quad \vdots \\
&\quad + \mathbf{F}(t-k+1)\boldsymbol{\delta}(t-k+2) + \\
&\quad + \boldsymbol{\delta}(t-k+1).
\end{aligned} \tag{18}$$

Segue da propriedade da transposição de um produto de matrizes que (18) pode ser reescrita como

$$\begin{aligned}
\mathbf{x}(t-k) &= \left[\mathbf{F}^T(t)\cdots\mathbf{F}^T(t-k+2)\mathbf{F}^T(t-k+1) \right]^T \mathbf{x}(t) + \\
&\quad + \left[\mathbf{F}^T(t-1)\cdots\mathbf{F}^T(t-k+2)\mathbf{F}^T(t-k+1) \right]^T \boldsymbol{\delta}(t) + \\
&\quad + \left[\mathbf{F}^T(t-2)\cdots\mathbf{F}^T(t-k+2)\mathbf{F}^T(t-k+1) \right]^T \boldsymbol{\delta}(t-1) + \\
&\quad + \left[\mathbf{F}^T(t-3)\cdots\mathbf{F}^T(t-k+2)\mathbf{F}^T(t-k+1) \right]^T \boldsymbol{\delta}(t-2) + \\
&\quad \vdots \\
&\quad + \left[\mathbf{F}^T(t-k+1) \right]^T \boldsymbol{\delta}(t-k+2) + \\
&\quad + \boldsymbol{\delta}(t-k+1).
\end{aligned} \tag{19}$$

Assumindo-se $k > 1$, é possível agrupar os produtos matriciais em (19) conforme:

$$\begin{aligned}
\mathbf{x}(t-k)|_{k>1} &= \left[\prod_{m=0}^{k-1} \mathbf{F}^T(t-m) \right]^T \mathbf{x}(t) + \\
&+ \left[\prod_{m=1}^{k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t) + \\
&+ \left[\prod_{m=2}^{k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t-1) + \\
&+ \left[\prod_{m=3}^{k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t-2) + \\
&\vdots \\
&+ \left[\prod_{m=k-1}^{k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t-k+2) + \\
&+ \boldsymbol{\delta}(t-k+1)
\end{aligned} \tag{20}$$

Por fim, introduzindo um somatório dos termos em função de $\boldsymbol{\delta}$ (exceto o último deles) tem-se que

$$\begin{aligned}
\mathbf{x}(t-k)|_{k>1} &= \left[\prod_{m=0}^{k-1} \mathbf{F}^T(t-m) \right]^T \mathbf{x}(t) + \\
&+ \sum_{n=1}^{k-1} \left[\prod_{m=n}^{k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t-n+1) \\
&+ \boldsymbol{\delta}(t-k+1).
\end{aligned} \tag{21}$$

Definindo-se os produtórios transpostos como sendo

$$\mathbf{F}_k(t)|_{k \geq 1} = \left[\prod_{m=0}^{m=k-1} \mathbf{F}^T(t-m) \right]^T \tag{22}$$

e a soma dos termos em $\boldsymbol{\delta}$ como sendo

$$\boldsymbol{\sigma}_k(t)|_{k \geq 1} = \sum_{n=0}^{m=k-2} \left\{ \left[\prod_{m=n+1}^{m=k-1} \mathbf{F}^T(t-m) \right]^T \boldsymbol{\delta}(t-n) \right\} + \boldsymbol{\delta}(t-k+1), \tag{23}$$

e fazendo-se $\mathbf{F}_0(t) = \mathbf{I}$, $\boldsymbol{\sigma}_0(t) = 0$ e $\boldsymbol{\sigma}_1(t) = \boldsymbol{\delta}(t)$, tem-se:

$$\mathbf{x}(t-k) = \mathbf{F}_k(t)\mathbf{x}(t) + \boldsymbol{\sigma}_k(t) . \quad (24)$$

Substituindo-se t por $t-k$ em (11) tem-se:

$$\mathbf{y}(t-k) = \mathbf{D}\mathbf{x}(t-k) + \mathbf{e}(t-k) . \quad (25)$$

Substituindo (24) em (25), chega-se em:

$$\begin{aligned} \mathbf{y}(t-k) &= \mathbf{D}[\mathbf{F}_k(t)\mathbf{x}(t) + \boldsymbol{\sigma}_k(t)] + \mathbf{e}(t-k) \\ &= \mathbf{D}\mathbf{F}_k(t)\mathbf{x}(t) + \mathbf{D}\boldsymbol{\sigma}_k(t) + \mathbf{e}(t-k). \end{aligned} \quad (26)$$

Fazendo-se

$$\mathbf{D}\boldsymbol{\sigma}_k(t) + \mathbf{e}(t-k) = \boldsymbol{\varepsilon}_k(t) , \quad (27)$$

onde $\boldsymbol{\varepsilon}_k(t)$ corresponde à contribuição do ruído de aquisição e das inovações, na baixa resolução, segue de (26) e (27) que:

$$\mathbf{y}(t-k) = \mathbf{D}\mathbf{F}_k(t)\mathbf{x}(t) + \boldsymbol{\varepsilon}_k(t) \quad (28)$$

A equação (28), função de k , pode ser, por comparação, facilmente adequada a formulações que utilizam o modelo descrito em (1).

Com base em (28), seja por inspeção, seja reformulando o problema de otimização e seguindo-se os mesmos passos que levam a (10), (10) pode ser reescrita como uma função do tempo:

$$\hat{\mathbf{x}}^{l+1}(t) = \hat{\mathbf{x}}^l(t) + \mu \sum_{k=0}^{K-1} \mathbf{F}_k^T(t) \mathbf{D}^T [\mathbf{D}\mathbf{F}_k(t)\hat{\mathbf{x}}^l(t) - \mathbf{y}(t-k)] . \quad (29)$$

Esta equação mostra claramente que, aproximando-se o gradiente por sua estimativa instantânea ($K=1$) e considerando que $\mathbf{F}_0(t) = \mathbf{I}$, (29) torna-se idêntica à equação de atualização do gradiente estocástico derivada em (COSTA & BERMUDEZ, 2007).

Geralmente, os algoritmos de super-resolução são inicializados (estimativa inicial da imagem HR) com uma imagem LR interpolada para a alta resolução. A partir daí, diversas

observações são requeridas para que a imagem desejada seja super-resolvida, elevando K e, portanto, o custo computacional. Utilizando-se a última reconstrução (instante $t-1$) devidamente registrada como inicialização do algoritmo (no instante t), como feito em (ELAD & FEUER, 1999), é possível que o valor de K seja reduzido substancialmente. No limite, quando $K=1$, chega-se ao algoritmo LMS. Resultados ilustrados no Capítulo 4 mostram que, quando consideradas situações em que os movimentos possam ser modelados por um processo de *random walking*, a exemplo de movimentos involuntários do operador de uma câmera fotográfica ou de vídeo (*câmera shaking*), o uso de apenas duas observações ($K=2$) leva a resultados substancialmente melhores do que o LMS ($K=1$) com um pequeno acréscimo computacional.

3.1 CUSTO COMPUTACIONAL

A determinação do custo computacional de algoritmos de super-resolução depende de hipóteses acerca do movimento. Assumindo-se movimentos fracionários (de subpixel) na grade de alta resolução, por exemplo, é necessário o uso de interpolações, o que pode aumentar o número de operações do algoritmo em relação aos casos em que os movimentos são inteiros na grade de alta resolução. Entretanto, essas hipóteses tendem a influenciar da mesma forma o custo de todos os algoritmos. Nesta seção, são assumidos apenas movimentos translacionais e globais, com valores inteiros na grade de alta resolução. Essa hipótese tende a reduzir o custo computacional dos algoritmos e, portanto, é bastante empregada em sistemas tempo-real. Além disso, geralmente, bons resultados são obtidos mesmo que o movimento presente na cena não seja, de fato, translacional e global (COSTA & BERMUDEZ, 2008). Sendo assim, as matrizes de registro, $\mathbf{G}(t)$ e $\mathbf{F}(t)$, não implicam em operações de ponto flutuante, FLOP(s) - *float point operation* (s), resumindo-se a deslocamentos (rearranjos) em memória.

A implementação proposta, resumida na Tabela 01, tem seu custo computacional determinado por

$$\hat{\mathbf{x}}^{l+1}(t) = \hat{\mathbf{x}}^l(t) + \mu \sum_{k=0}^{K-1} \mathbf{F}_k^T(t) \mathbf{D}^T \left[\underbrace{\underbrace{\underbrace{\mathbf{D}\mathbf{F}_k(t)\hat{\mathbf{x}}^l(t) - \mathbf{y}(t-k)}_{1 \text{ convolução}}}_{N^2 \text{ somas}}}_{1 \text{ convolução}} \right]. \quad (30)$$

$\underbrace{\hspace{10em}}_{K \text{ vezes}}$
 $\underbrace{\hspace{10em}}_{M^2 \text{ produtos}}$
 $\underbrace{\hspace{10em}}_{M^2 \text{ somas}}$

As convoluções consideram imagens HR e uma máscara com dimensões $(H \times H)$, que depende do modelo do sistema óptico. O custo computacional das convoluções, caso executadas no domínio da imagem, é $M^2 H^2$ produtos e $M^2(H^2 - 1)$ somas, ou $M^2(2H^2 - 1)$ FLOPs.

Tabela 01: Implementação proposta

Inicialização:

$L = 1$ (número de iterações para cada instante t)

$K = 2$ (número de observações)

$\mu = 8$ (passo de convergência)

$\hat{x}^0(0) =$ interpolação bicubica $y(0)$

Algoritmo:

$$\left\{ \begin{array}{l} \text{Laço em } l = 0, 1, 2, \dots, L \\ \text{Laço em } t = 0, 1, 2, \dots, K \\ \mathbf{y}(t-k) = \mathbf{D}\mathbf{F}_k(t)\mathbf{x}(t) + \boldsymbol{\varepsilon}_k(t) \\ \hat{\mathbf{x}}^{l+1}(t) = \hat{\mathbf{x}}^l(t) + \mu \sum_{k=0}^{K-1} \mathbf{F}_k^T(t) \mathbf{D}^T \left[\mathbf{D}\mathbf{F}_k(t)\hat{\mathbf{x}}^l(t) - \mathbf{y}(t-k) \right] \end{array} \right.$$

Sendo assim, o custo total do algoritmo corresponde a

$$\text{custo} = K(N^2 \text{ somas} + 2 \text{ convoluções}) + M^2 \text{ produtos} + M^2 \text{ somas}. \quad (31)$$

Considerando as somas e produtos como 1 FLOP, (31) se reduz a

$$\text{custo} = K(N^2 + 2 \text{ convoluções}) + 2M^2 \text{ [FLOPs]} \quad (32)$$

ou, considerando-se uma convolução como demandando $M^2(H^2 - 1)$ FLOPS,

$$custo = K[N^2 + 2M^2(2H^2 - 1)] + 2M^2 \text{ [FLOPs]}, \quad (33)$$

Dividindo-se o custo em (33) por N^2 , resulta no custo computacional por pixel da imagem LR

$$custo = K[1 + 2A^2(2H^2 - 1)] + 2A^2 \text{ [FLOPs por pixel LR]}, \quad (34)$$

em que A corresponde à ampliação desejada com a super-resolução ($A^2 = M^2 / N^2$).

Assumindo a utilização de apenas uma observação $K=1$ para cada iteração no instante t , tem-se que o algoritmo baseado na equação (29) é igual ao LMS-SRR. Baseado na premissa de que o algoritmo proposto necessita de no mínimo duas observações para sua implementação, $K=2$, e utilizando dos critérios de operações de ponto flutuante, conclui-se que a implementação proposta apresenta um custo computacional menor que duas vezes o custo do LMS-SRR.

4 RESULTADOS

Neste capítulo são apresentados resultados de simulações utilizando a implementação proposta no Capítulo 3. O objetivo de tais simulações é comparar o desempenho dessa estrutura com o do LMS-SRR, proposto em (ELAD & FEUER, 1999). Inicialmente serão descritas algumas condições gerais de simulação, e serão fornecidos os detalhes de modelagem comuns aos exemplos a serem apresentados. As especificidades de cada simulação e os resultados obtidos serão então descritos em duas seções distintas.

4.1 CONDIÇÕES GERAIS

A matriz de degradação $\mathbf{D}(t)$ foi assumida determinística e invariante no tempo, e, portanto, $\mathbf{D}(t) = \mathbf{D}$. Essa matriz foi modelada como um processo de filtragem média-móvel por uma máscara 2×2 , seguido de uma operação de dizimação, que consiste em reduzir a taxa de amostragem de um sinal discreto (OPPENHEIM & SHAFER, 2013).

Para o algoritmo LMS-SRR foi utilizado passo de convergência $\mu = 4$ e a execução de uma iteração por amostra temporal, $l = 1$. Já foi mostrado em (COSTA & BERMUDEZ, 2007) que, dado o modelo de degradação adotado, esses parâmetros levam o algoritmo ao seu desempenho ótimo.

Os valores dos parâmetros utilizados no algoritmo proposto foram escolhidos empiricamente, uma vez que não se têm uma análise teórica para determinação dos parâmetros ótimos. Simulações com diversos valores do passo de adaptação levaram a um melhor desempenho com um passo de convergência $\mu = 8$. O *buffer* foi construído com duas observações LR por iteração, $K = 2$, visto ser este o valor mínimo para a implementação da solução pelo método de estimação de mínimos quadrados, conforme a equação (23). O número de iterações do algoritmo por amostra temporal foi definido com o mesmo valor do LMS-SRR, $l = 1$, com o objetivo de minimizar o custo computacional e facilitar a comparação entre os algoritmos.

4.2 AVALIAÇÃO QUANTITATIVA

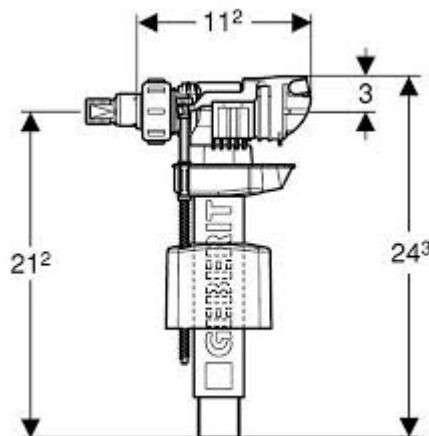
Esta seção tem como objetivo comparar os desempenhos entre a implementação proposta no Capítulo 3, baseada em LS, e o algoritmo LMS-SRR proposto em (ELAD & FEUER, 1999). A comparação proposta, quantitativa, baseia-se no erro quadrático médio² de reconstrução, $EQMR$, ou seja, em

$$EQMR = \frac{1}{M^2} \|\mathbf{x}(t) - \hat{\mathbf{x}}(t)\|_2^2, \quad (35)$$

em que M^2 é o número de pixels da imagem de alta resolução. Para tanto, é necessário que se tenha a imagem desejada $\mathbf{x}(t)$ disponível. Isto apenas é possível quando a simulação baseia-se em sequências sintéticas.

A sequência de alta resolução foi gerada a partir de uma imagem escolhida aleatoriamente através de um painel com diversas imagens disponíveis na internet e é ilustrada da Figura 6. Essa imagem tem tamanho de 225×225 pixels e a partir dela foram gerados 50 quadros, $\mathbf{X}(t)$, assumidos de alta resolução, com tamanho 140×140 pixels, simulando-se uma cena estática adquirida por uma câmera em movimento global de translação. Esse movimento foi gerado a partir de um processo aleatório bidimensional com passos independentes e identicamente distribuídos, de magnitude zero ou ± 1 , nas direções das linhas e/ou das colunas da imagem simulando o movimento de uma câmera trêmula (*camera shaking*).

Figura 6: Imagem de 225×225 pixels que gerou as imagens HR.



² Considera-se, neste caso, média espacial.

As imagens HR foram degradadas por um sistema de subamostragem com fator de 2:1, resultando em imagens de baixa resolução (LR), $\mathbf{Y}(t)$, com dimensões 70×70 . O movimento entre as imagens foi assumido conhecido (sem erro de registro). Não foi considerado ruído aditivo e em nenhum dos algoritmos foi considerada regularização.

A partir das observações $\mathbf{Y}(t)$ foi realizado o processo de reconstrução das imagens $\hat{\mathbf{X}}(t)$ com o uso de ambos algoritmos. Nessa simulação também foram geradas imagens com as mesmas dimensões de $\hat{\mathbf{X}}(t)$, baseadas apenas na interpolação bicúbica³ das observações, para facilitar a avaliação perceptual dos resultados da SRR. Na Figura 7 são apresentados resultados obtidos com o uso dos algoritmos avaliados. Na Figura 7(a) é apresentado um dos quadros da imagem original (desejada). Na Figura 7(b) é apresentada a imagem observada (LR) do quadro em (a). Na Figura 7(c) é apresentada a interpolação bicúbica de (b). Na Figura 7(d) é apresentado o resultado da reconstrução do quadro em (b) através do LMS-SRR. Por fim, na Figura 5(e), é apresentado o resultado da reconstrução do quadro em (b) através da implementação proposta.

Através dos resultados apresentados na Figura 7 é possível perceber que ambos os métodos de SRR levam a resultados perceptuais substancialmente melhores que a imagem interpolada. Comparando-se as figuras 7(d) e 7(e), é possível verificar que o método LS leva a uma preservação de bordas ligeiramente maior que o LMS-SRR.

Na Figura 8 é ilustrada uma avaliação quantitativa da reconstrução no decorrer do tempo. Como pode ser observado, o LS leva a um erro quadrático médio inferior ao obtido com o LMS-SRR. Diferente do LMS-SRR, em que, nas configurações utilizadas, o aumento do número de iterações por amostra de entrada não leva a um erro de reconstrução menor, o aumento do número de iterações pode levar o LS a um melhor desempenho, conforme ilustrado para o caso de $l = 3$. Naturalmente, o aumento do número de iterações implica em um aumento do custo computacional em três vezes, o que pode não ser interessante para processamento tempo real.

³ Considera-se a interpolação bicúbica como um padrão de comparação de qualidade, por ser amplamente utilizada como *zoom* em *softwares* computacionais.

Figura 7: Imagens da Simulação: (a) imagem original. (b) imagem LR. (c) reconstrução via interpolação bicúbica. (d) reconstrução via LMS-SRR. (e) reconstrução via LS.

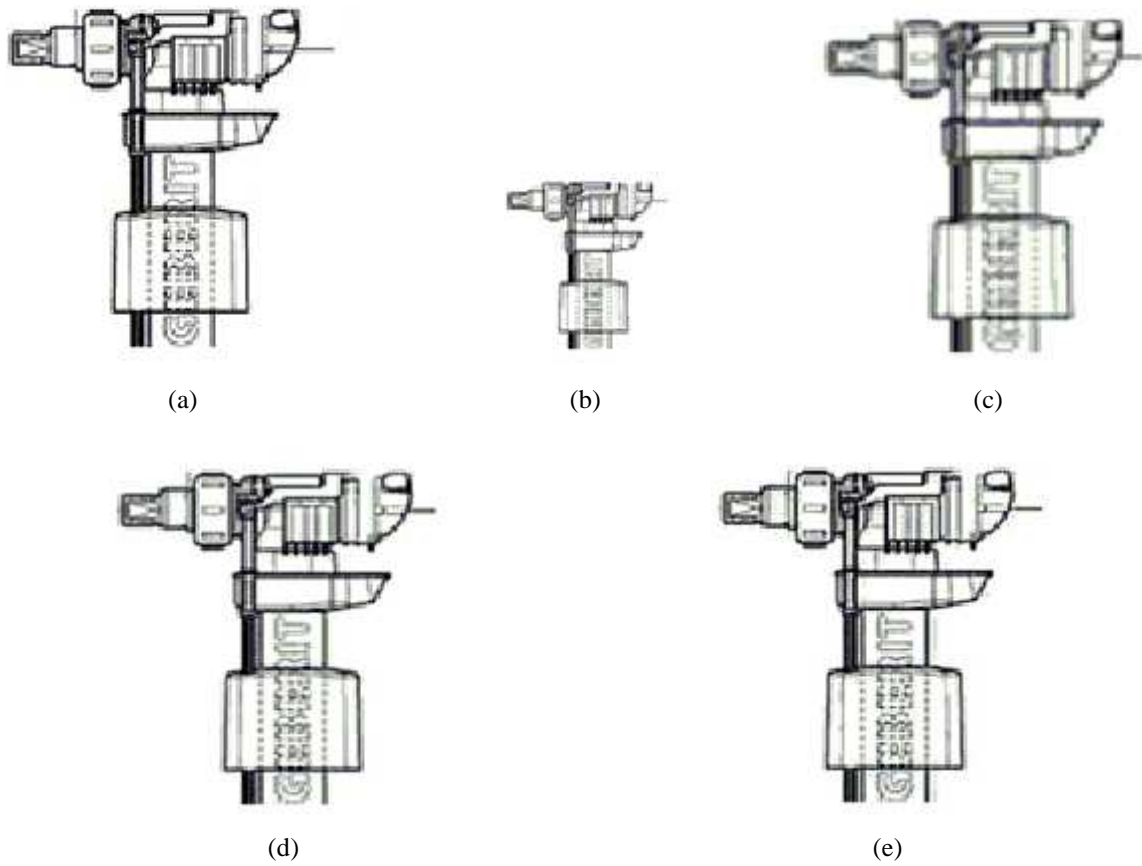
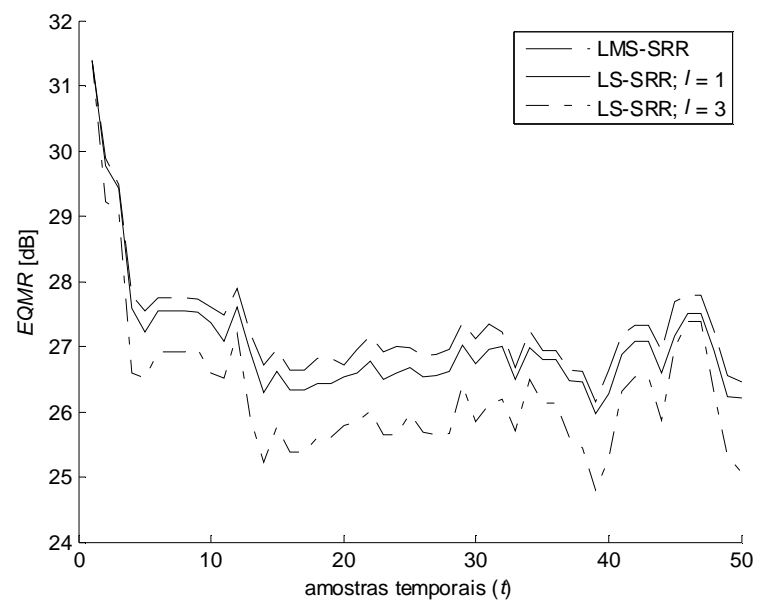


Figura 8: Evolução temporal do erro quadrático médio de reconstrução.



As oscilações bruscas no gráfico da Figura 8 são resultantes das inovações nas regiões de fronteira das imagens observadas. Como foi visto no Capítulo 3 se faz necessário assumir uma hipótese para o preenchimento dos pixels na região de fronteira da imagem observada em relação à imagem anterior, $\mathbf{x}(t-1)$, portanto, não sendo super-resolvida, a região de fronteira pode gerar um *EQMR* (erro quadrático médio de reconstrução) maior em função da relevância das inovações que surgem em cada nova observação k .

Considerando o modelo de degradação e os parâmetros definidos, o custo computacional do algoritmo LMS-SRR é equivalente a 65 FLOPS por pixel, enquanto a implementação proposta resulta em 122 FLOPs por pixel.

4.3 AVALIAÇÃO QUALITATIVA

Assim como na secção 4.2, esta secção também tem como objetivo comparar os desempenhos entre a implementação proposta no Capítulo 3, baseada em LS, e o algoritmo LMS-SRR proposto em (ELAD & FEUER, 1999). Porém, nesta secção é considerado um caso real de aplicação. Sendo assim, as imagens desejadas $\mathbf{X}(t)$ não estão disponíveis e a comparação passa a ser apenas qualitativa, visto que não é possível determinar o erro de reconstrução.

Como observações (LR), para as imagens $\mathbf{Y}(t)$ foram utilizados os primeiros 20 quadros da sequência *Disk*. Desses quadros, foram considerados apenas os primeiros 48×48 pixels, por questões de implementação (imagens/matriz quadradas).

O processo de reconstrução das imagens (HR) $\hat{\mathbf{X}}(t)$ através dos algoritmos LMS-SRR e LS foi inicializado a partir da interpolação bicúbica do primeiro quadro da sequência natural das observações em baixa resolução (LR) $\mathbf{Y}(t)$. As imagens $\hat{\mathbf{X}}(t)$ foram reconstruídas com as dimensões de 96×96 pixels. Nesta simulação, também foram geradas imagens com 96×96 pixels baseadas apenas na interpolação bicúbica das observações, para facilitar a avaliação dos resultados de Super-Resolução. Na Figura 9 são apresentados resultados obtidos com o uso dos algoritmos avaliados. Na Figura 9(a) é apresentado o

vigésimo quadro da sequência de baixa resolução (LR) $Y(t)$. Na Figura 9(b) é apresentada a interpolação bicúbica de (a). Na Figura 9(c) é apresentado o resultado da reconstrução do quadro em (a) através do LMS-SRR. Por fim, na Figura 9(d), é apresentado o resultado da reconstrução do quadro em (a) através da implementação proposta.

Figura 9: Resultados de simulação: (a) imagem LR. (b) reconstrução via interpolação bicúbica. (c) reconstrução via LMS-SRR. (d) reconstrução via LS.



Através dos resultados apresentados na Figura 9 é possível perceber que ambos os métodos de Super-Resolução levam a resultados perceptuais substancialmente melhores que a imagem interpolada. Comparando-se as figuras 9(c) e 9(d), é possível verificar que o método LS leva a uma imagem perceptualmente mais agradável, visto que na Figura 9(c) aparecem alguns artefatos de alta frequência (variação abrupta no tom de cinza). Ampliações desses resultados são apresentadas na Figura 10, para uma melhor visualização.

Figura 10: Resultados de simulação: (a) reconstrução via LMS-SRR (ampliação de 3x). (b) reconstrução via LS (ampliação de 3x).



5 CONCLUSÕES

Neste trabalho foi proposta uma adaptação do método de mínimos quadrados (LS) aplicado à reconstrução de imagens com super-resolução, visando a reconstrução de vídeo (sequências de imagens) em tempo real. Resultados do método proposto foram comparados com resultados do algoritmo LMS-SRR (ELAD & FEUER, 1999), dado ser o algoritmo de menor custo computacional na literatura. Frente aos resultados obtidos, conclui-se que:

1. A implementação proposta para o LS mostra-se capaz de atingir resultados superiores ao LMS-SRR com um pequeno acréscimo no custo computacional;
2. Os resultados apresentados foram obtidos mediante o desempenho ótimo do LMS-SRR e mediante uma configuração do LS realizada com base empírica. Isto sugere que, de posse de um modelo analítico do comportamento do LS, possam ser obtidos resultados ainda melhores.

Como sugestão para um trabalho futuro recomenda-se o desenvolvimento de uma análise teórica do comportamento do algoritmo de forma a obter-se o passo de adaptação ótimo.

REFERÊNCIAS BIBLIOGRÁFICAS

- BATEL, M., MARROQUIN, M., HALD, J., CHRISTENSEN, J. J., SCHUHMACHER, A. P. E NIELSEN, T. G. Noise Source Location Techniques - Simple to Advanced Applications. **Sound and Vibration**, Brüel & Kjør, Mar. 2003.
- CHAUDHURI, S. E TAUR, D. R. High-resolution slow-motion sequencing. **IEEE Signal Processing Magazine**, vol. 22, no. 2, pp. 16–24, Mar. 2005.
- CHEN, H.M., LEE, S., RAO, R. M., SLAMANI, M.A. E VARSHNEY, P. K. Imaging for canceled weapon detection. **IEEE Signal Processing Magazine**, vol. 22, no. 2, pp. 52–61, Mar. 2005.
- COSTA, G. H. E BERMUDEZ, J. C. M. Informed choice of the LMS parameters in super-resolution video reconstruction applications. **IEEE Transactions on Signal Processing**, vol. 56, no. 2, pp. 555–564, Feb. 2008.
- COSTA, G. H. E BERMUDEZ, J. C. M. Statistical analysis of the LMS algorithm applied to super-resolution image reconstruction. **IEEE Transactions on Signal Processing**, vol. 55, no. 5, pp. 2084–2095, May 2007.
- ELAD, M. E FEUER, A. Superresolution restoration of an image sequence: Adaptive filtering approach. **IEEE Transactions on Signal Processing**, vol. 8, no. 3, pp. 387–395, Mar. 1999.
- FARSIU, S., ROBINSON, D., ELAD, M. E MILANFAR, P. Fast and robust multiframe super resolution. **IEEE Transactions on Image Processing**, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- FARSIU, S., ROBINSON, D., ELAD, M., E MILANFAR, P. Advances and challenges in super-resolution. **Int'l Journal of Imaging Systems and Technology**, vol. 14, no. 2, pp. 47–57, Aug. 2004.
- FARSIU, S., ELAD, M., E MILANFAR, P. Multi-frame demosaicing and super-resolution of color images. **IEEE Transactions on Signal Processing**, vol. 15, no. 1, pp. 141–159, Jan. 2006.
- FORESTI, G. L., MICHELONI, C., SNIDARO, L., REMAGNINO, P. E ELLIS, T. Active video-based surveillance system. **IEEE Signal Processing Magazine**, vol. 22, no. 2, pp. 25–37, Mar. 2005.
- GUNTURK, B., BATUR, A., M. H. III E MERSEREAU, R. M. Eigenfacebased super-resolution for face recognition. **IEEE International Conference on Image Processing**, vol. 2, pp. 845–848, Sept. 2002.
- HAMPAPUR, A., BROWN, L., CONNEL, J., EKIN, A., HAAS, N., LU, M., MERKL, H., PANKANTI, S., SENIOR, A., SHU, C.F. E TIAN, Y. L. Smart video surveillance. **IEEE Signal Processing Magazine**, vol. 22, no. 2, pp. 38–51, Mar. 2005.

HAYKIN, S. **Sistemas de Comunicação Analógicos e Digitais**. Artmed Editora S.A., 4ª edição, pp 285-288, 2004 reimpressão 2007.

HAYKIN, S. **Adaptive Filter Theory**. 4th Edition, Prentice Hall, pp 863-865, 2002.

MANOLAKIS, D. G., INGLE, V. K., KOGON, S. M. **Statistical and Adaptive Signal Processing**. McGraw-Hill, 796 pp, 2000.

MEYER, C. D., **Matrix Analysis and Applied Linear Algebra**. Society for Industrial and Applied Mathematics, pp 269-293, 2000.

OPPENHEIM, A. V., SHAFER, R. W. **Processamento em tempo discreto de sinais**. Editora PEARSON, 3ª Edição, 2013.

PARK, S. C., PARK, M. K., E KANG, M. G. Super-resolution image reconstruction: A technical overview. **IEEE Signal Processing Magazine**, vol. 20, n° 3, pp. 21–36, Mai. 2003.

ZOOMET, A., RAV-ACHA, A. E PELEG, S. Robust super-resolution. **Int'l Conf. Computer Vision and Pattern Recognition (CVPR)**, vol. 1, pp. 645–650, Dec. 2001.