

À minha querida esposa Claudia,
com amor e gratidão.

Agradecimentos

Gostaria de manifestar os meus agradecimentos a todos os amigos e amigas que tornaram possível a realização desta tese. Em especial, agradeço:

- A todos os meus orientandos e ex-orientandos, em especial àqueles que contribuíram nas pesquisas relatadas nesta tese: Paulo Barreto, Harold Bustos, Marco Antonio de Melo e Amir Afif.
- A todos os professores que colaboraram nas pesquisas relatadas nesta tese, especialmente ao prof. Zang Hee Cho, que me convidou como pesquisador visitante da University of California at Irvine, e ao prof. Ricardo de Queiroz que colaborou nas pesquisas de marca d'água para imagens binárias.
- Aos amigos(as) e colegas do Laboratório de Processamento de Sinais, professores(as) Denise Consoni, Flávio Cipparrone, Miguel Ramirez, Vitor Nascimento e Wagner Zucchi.
- Ao prof. Francisco Javier que me estimulou a fazer a livre docência.
- À FAPESP e ao CNPq, pelos auxílios financeiros que possibilitaram as nossas pesquisas.
- Aos meus pais e à minha irmã.

Resumo

Esta tese descreve as principais contribuições científicas do meu grupo de pesquisa após o meu doutoramento. Estas contribuições estão agrupadas em três capítulos:

1. Projeto de operadores pela aprendizagem: Tradicionalmente, um operador restrito à janela (W-operador), que desempenha uma determinada função no Processamento e Análise de Imagens, é projetada manualmente e esta tarefa pode ser tediosa. Um W-operador pode ser projetado automaticamente a partir das imagens amostras de entrada-saída por um processo de aprendizagem de máquina. Nesta tese, descrevemos o projeto automático de W-operadores “binária para binária” e “binária para níveis de cinza”. Primeiro, descrevemos a aprendizagem provavelmente aproximadamente correta e a estimação estatística que constituem o embasamento teórico do projeto automático de W-operadores. Depois, analisamos os diferentes algoritmos de aprendizagem e propomos as adaptações neles para aumentar os seus desempenhos ao resolver os problemas tratados. Por fim, utilizamos as teorias e os algoritmos desenvolvidos para aumentar a resolução espacial das imagens binárias e meio-tom, e para efetuar o meio-tom inverso.

2. Difusão anisotrópica: A difusão anisotrópica é frequentemente utilizada na segmentação de imagens, atenuação de ruídos e detecção de arestas. Esta tese descreve o uso da difusão anisotrópica em várias aplicações de Processamento e Análise de Imagens. Primeiro, descrevemos a teoria do espaço de escala linear (de onde se originou a difusão anisotrópica). Depois, descrevemos a difusão anisotrópica, incluindo uma versão baseada na estatística robusta. Mostramos, através de algumas aplica-

ções, que a difusão anisotrópica robusta é superior ao tradicional em termos da qualidade da imagem filtrada. Descrevemos o melhoramento do algoritmo de reconstrução tomográfica máxima entropia usando a difusão anisotrópica robusta. Por fim, descrevemos o aperfeiçoamento do modelo linear geral (um processo para detectar as áreas ativadas do cérebro em imagens de ressonância magnética funcional) usando a difusão anisotrópica robusta.

3. Marcas d'água de autenticação: Uma marca d'água é um sinal portador de informação embutido numa imagem digital que pode ser extraída mais tarde para fazer alguma asserção sobre a imagem hospedeira. As marcas d'água digitais são normalmente classificadas em robustas e frágeis. Esta tese trata somente das marcas d'água frágeis, também chamadas de autenticação. Primeiro, descrevemos a assinatura digital, um conceito amplamente utilizado nas marcas de autenticação de chave pública. Em segundo lugar, descrevemos as principais marcas de autenticação para as imagens estáticas de tonalidade contínua: Yeung-Mintzer e Wong. Descrevemos os principais ataques contra estas marcas e os meios para se defender contra eles. Em terceiro lugar, descrevemos as marcas d'água de autenticação para as imagens binárias e meio-tom.

Abstract

This thesis describes the main scientific contributions of my research group after my doctorate. These contributions are grouped in three chapters:

1. Operator design by machine learning: Traditionally, a windowed operator (W-operator) that plays a certain role in the Image Processing and Analysis is designed manually, and this task can be tedious. A W-operator can be designed automatically from sample in-out images by a machine learning process. In this thesis, we describe the automatic design of binary-to-binary and binary-to-grayscale W-operators. First, we describe the probably approximately correct learning and the statistical estimation that constitute the theoretic framework of the automatic W-operator design. Afterwards, we analyze the various learning algorithms and propose their adaptations to increase their performances in solving the applications addressed. We use the developed theories and algorithms to increase the spatial resolution of binary and halftone images, and to perform the inverse halftoning.

2. Anisotropic diffusion: Anisotropic diffusion is used frequently in image segmentation, noise attenuation and edge detection. This thesis describes the use of the anisotropic diffusion in various Image Processing and Analysis applications. First, we describe the linear scale space theory (from where the anisotropic diffusion has been derived). Afterwards, we expound the anisotropic diffusion theory, including a version based on the robust statistics. We show, through some applications, that the robust anisotropic diffusion is superior to the traditional in terms of the quality of the filtered image. We describe the improvement of the maximum entropy tomography

algorithm using the robust anisotropic diffusion. Finally, we describe the enhancement of the general linear model (a process to detect activated regions of brain in the functional magnetic resonance images) using the robust anisotropic diffusion.

3. Authentication watermarks: A watermarking is an information-bearing signal embedded in a digital image that can be extracted later to make some assertion on the host image. Digital watermarks are usually classified in robust and fragile. This thesis is concerned only with the fragile watermarks, also called authentication watermarks. First, we explain the digital signature, a concept widely used for the public-key authentication watermarking. Second, we describe the main authentication watermarks for static continuous-tone images: Yeung-Mintzer and Wong. We describe the principal attacks against these watermarks and the means to defend against them. Third, we describe the authentication watermarking techniques for binary and half-tone images.

Sumário

| | | |
|----------|---------------------------------------------------------------|------------|
| 1 | Introdução | 1 |
| 2 | Projeto de Operadores pela Aprendizagem | 3 |
| 2.1 | Introdução | 6 |
| 2.2 | Aprendizagem de W-Operadores Binários | 14 |
| 2.3 | Aumento de Resolução de Imagens Binárias | 32 |
| 2.4 | Aumento de Resolução de Imagens Meio-Tom | 44 |
| 2.5 | Meio-Tom Inverso pela Aprendizagem | 62 |
| 2.6 | Conclusões | 72 |
| 3 | Difusão Anisotrópica | 73 |
| 3.1 | Introdução | 75 |
| 3.2 | Espaço de Escala Linear | 79 |
| 3.3 | Difusão Anisotrópica | 87 |
| 3.4 | Melhoramento da Tomografia pela RAD | 103 |
| 3.5 | Melhoramento da fMRI pela RAD | 112 |
| 3.6 | Conclusões | 126 |
| 4 | Marcas d'Água de Autenticação | 127 |
| 4.1 | Introdução | 129 |
| 4.2 | Assinatura Digital | 136 |
| 4.3 | Marcas de Autenticação para Imagens Contone | 139 |
| 4.3.1 | Marca de Autenticação de Yeung-Mintzer | 139 |
| 4.3.2 | Marca de Wong e Hash Block Chaining | 143 |
| 4.4 | Marcas de Autenticação para Imagens Binárias e Meio-Tom | 158 |

| | | |
|----------|-----------------------------------------|------------|
| 4.4.1 | Introdução | 158 |
| 4.4.2 | Marca de Autenticação AWST | 161 |
| 4.4.3 | Marca de Autenticação AWSF | 167 |
| 4.5 | Conclusões | 175 |
| 5 | Referências Bibliográficas | 176 |
| 5.1 | Publicações do Autor | 176 |
| 5.2 | Referências da Literatura | 180 |

Lista de Figuras

| | | |
|-------------|--------------------------------------------------------------------|-----|
| Figura 2.1 | W-operador | 12 |
| Figura 2.2 | Aprendizagem de W-operador em ambiente sem ruído | 20 |
| Figura 2.3 | Aprendizagem de W-operador em ambiente ruidoso | 23 |
| Figura 2.4 | Janelas sem e com pesos, com 17 furos-de-espiar | 23 |
| Figura 2.5 | Operador de aumento de resolução restrito à janela | 37 |
| Figura 2.6 | Aumento de resolução de caracteres impressos | 40 |
| Figura 2.7 | Aumento de resolução de um documento manuscrito | 43 |
| Figura 2.8 | Ampliação das imagens meio-tom “HP driver pontos grandes” | 51 |
| Figura 2.9 | Ampliação das imagens meio-tom “HP driver pontos pequenos” | 60 |
| Figura 2.10 | Ampliação das imagens meio-tom excitação ordenada | 61 |
| Figura 2.11 | Meio-tom inverso das imagens obtidas por difusão de erro | 70 |
| Figura 2.12 | Meio-tom inverso 10-ID3 em diferentes tipos de imagens | 71 |
| Figura 3.1 | Funções gaussianas 1-D, 2-D e suas derivadas | 84 |
| Figura 3.2 | Detecção de arestas no espaço de escala linear | 86 |
| Figura 3.3 | Funções parada-na-aresta e de influência | 91 |
| Figura 3.4 | Filtragem de um sinal sintetizado pela difusão anisotrópica | 97 |
| Figura 3.5 | Filtragem do sinal do sensor de aceleração ADLX202E | 98 |
| Figura 3.6 | Detecção de arestas usando a difusão anisotrópica | 101 |
| Figura 3.7 | Difusão anisotrópica com grande número de iterações | 102 |
| Figura 3.8 | Comparação entre o MENT-estendido e a reconstrução-difusão | 109 |
| Figura 3.9 | Diferença entre as projeções originais e reconstruídas | 110 |
| Figura 3.10 | Comparação entre MENT, reconstrução-difusão e retro-projeção | 111 |

| | | |
|-------------|--------------------------------------------------------------|-----|
| Figura 3.11 | Um fantom fMRI simulado | 121 |
| Figura 3.12 | SPM{t} obtido da fMRI da figura 3.11 | 122 |
| Figura 3.13 | Imagens fMRI reais com áreas artificialmente ativadas | 123 |
| Figura 3.14 | SPM{t} obtido da figura 3.13 sem filtragem | 124 |
| Figura 3.15 | SPM{t} obtido da figura 3.13 usando a técnica proposta | 125 |
| Figura 4.1 | Uso da informação contextual | 149 |
| Figura 4.2 | Impedindo o ataque “recortar-e-colar” com HBC2 | 150 |
| Figura 4.3 | Ilustração da AWST chave pública | 165 |
| Figura 4.4 | Qualidade dos documentos marcados com AWST | 166 |
| Figura 4.5 | Qualidade visual de um documento marcado com AWSF | 170 |
| Figura 4.6 | Falsificação “ataque de paridade” | 172 |

Lista de Tabelas

| | | |
|------------|-------------------------------------------------------------------|----|
| Tabela 2.1 | Erros obtidos usando e1-NN ao ampliar documentos impressos | 37 |
| Tabela 2.2 | Erros usados para comparar os diferentes vieses indutivos | 40 |
| Tabela 2.3 | Erros obtidos usando WZDT com janelas de diferentes tamanhos | 54 |
| Tabela 2.4 | Erros dos diferentes algoritmos de aprendizagem | 54 |
| Tabela 2.5 | Os erros diminuem quando os tamanhos das amostras crescem | 54 |
| Tabela 2.6 | Erros observados em ampliação pelo meio-tom inverso | 57 |
| Tabela 2.7 | PSNRs obtidas usando meio-tom inverso pela aprendizagem | 69 |
| Tabela 3.1 | Erros obtidos pela difusão anisotrópica com 50 iterações | 96 |
| Tabela 3.2 | Erros obtidos com 100 iterações | 96 |

Lista de Abreviaturas

| | |
|---------|-------------------------------------------------------------------------------------------------------------------------|
| Alice | Nome fictício que da pessoa que assina um documento digital usando a sua chave privada. |
| AWSF | Authentication watermarking by shuffling and flipping (marca d'água de autenticação pelo embaralhamento e reviramento). |
| AWST | Authentication watermarking by self toggling (marca d'água de autenticação pelo auto-reviramento). |
| A^x | Imagem amostra de entrada. |
| A^y | Imagem amostra de saída. |
| Bob | Nome fictício da pessoa que verifica a assinatura de um documento digital usando a chave pública. |
| BOLD | Blood oxygen level dependent (dependente do nível de oxigenação do sangue). |
| Contone | Continuous tone (tonalidade contínua). |
| DHPT | Data Hiding by Pair-Toggling (embutimento de dados por reviramento aos pares). |
| DHSPT | Data Hiding by Smart Pair Toggling (embutimento de dados por reviramento inteligente aos pares). |
| DHST | Data hiding by self toggling (embutimento de dados por auto-reviramento). |
| DS | Digital signature (assinatura digital). |
| DSA | Digital signature algorithm (algoritmo de assinatura digital). |
| DT | Decision tree (árvore de decisão). |

| | |
|-------------|--------------------------------------------------------------------------------------------|
| EPM | Estimated parameters map (mapa dos parâmetros estimados). |
| e_error | Empirical error (erro empírico ou observado, escrito como “e-erro” no texto em português). |
| e-ótimo | Empiricamente ótimo, isto é, o melhor sobre os dados observados. |
| ek-NN | Aprendizagem k -NN empiricamente ótima. |
| fMRI | Functional magnetic resonance imaging (imageamento por ressonância magnética funcional). |
| Furo | Abreviação de “furo de espiar” (peephole). |
| HBC | Hash block chaining (encadeamento dos blocos de hash). |
| HSI | Hue, saturation, intensity (tonalidade, saturação, intensidade). |
| IH | Inverse halftoning (meio-tom inverso). |
| kd-árvore | Árvore k dimensional (kd-tree). |
| k -NN | k nearest neighbors (k vizinhos mais próximos). |
| LSB | Least significant bit (bit menos significativo). |
| LUT | Look-up-table (tabela de busca). |
| MAC | Message authentication code (código de autenticação de mensagem). |
| Mallory | Nome fictício de um “hacker” malicioso. |
| Marca | Abreviação de “marca d’água” (watermark). |
| MENT | Maximum entropy (máxima entropia). |
| mod | Módulo, resto de divisão inteira. |
| NN | Nearest neighbor (vizinho mais próximo). |
| OCR | Optical character recognition (reconhecimento de caracteres ópticos). |
| OD | Ordered dithering (excitação ordenada). |
| PAC | Probably approximately correct (provavelmente aproximadamente correto). |
| PET | Positron emission tomography (tomografia de emissão de pósitrons). |
| PSNR | Peak signal-to-noise ratio (razão entre o pico do sinal e o erro). |
| Q^x | Imagem a-ser-processada. |
| Q^y | Imagem de saída ideal. |
| \hat{Q}^y | Imagem processada. |
| RAD | Robust anisotropic diffusion (difusão anisotrópica robusta). |

| | |
|-------------|------------------------------------------------------------------------------------|
| RGB | Red, green, blue (vermelho, verde, azul). |
| RSA | Esquema de criptografia de chave pública de Rivest, Shamir e Adleman. |
| RMS | Root mean square (raiz da média quadrática). |
| SPM | Statistical parametric map (mapa estatístico paramétrico). |
| SPM{t} | SPM das estatísticas t de Student. |
| t_error | True error (erro real ou verdadeiro, escrito como “t-erro” no texto em português). |
| t-ótimo | Verdadeiramente ótimo (isto é, o melhor sobre a distribuição de probabilidade). |
| VIS | Visual impact score (nota de impacto visual). |
| W-operador | “Window operator” ou “windowed operator” (operador restrito à janela). |
| WZ-operador | “Windowed zoom operator” (operador restrito à janela para ampliação). |
| WZDT | Windowed zoom decision tree (árvore de decisão restrita à janela para ampliação). |