

Processamento de Imagens Aplicado à Monitoração de Processos

Jorge Leonardo Lima Barboza
Orientador: Antonio Carlos Gay Thomé

UFRJ

2004

Processamento de Imagens Aplicado à Monitoração de Processos

por

Jorge Leonardo Lima Barboza

Dissertação submetida ao Corpo Docente do Instituto de Matemática e Núcleo de Computação Eletrônica da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências em Informática.

Aprovada por

(Antônio Carlos Gay Thomé, Ph.D., DCC-IM/UFRJ - Orientador)

(Paulo Sérgio da Silva Borges, Ph.D., UFSC)

(Felipe Maia Galvão França, Ph.D., COPPE Sistemas/UFRJ)

(Adriano Joaquim de Oliveira Cruz, Ph.D., DCC-IM/UFRJ)

Biblioteca

“So we beat on, boats against the current, borne back ceaselessly into the past.”
F. Scott Fitzgerald, *The Great Gatsby*.

*“Good night, you princes of Maine,
You kings of New England.”*
John Irving, *The Cider House Rules*.

Para meus pais.

Agradecimentos

Ao professor Thomé, por sua assistência e idéias sempre construtivas, mas principalmente por sua confiança, demonstrada através da liberdade que me concedeu para a realização deste trabalho.

A meus pais, personagens principais no roteiro de tudo que já conquistei, e daquilo que venha a conquistar.

À minha namorada, Daniela Mauler, pelo apoio e companheirismo em todos os momentos desta jornada.

Aos amigos do Laboratório de Inteligência Computacional (LabIC) pelo enriquecimento intelectual e, principalmente, pelos momentos de descontração, que tornaram esta tarefa não só possível, mas agradável. São eles: Adriano Martins Moutinho, Airam (Mary Backwards) Carlos Paes Barreto Marques, Bruno Clemente Guingo, Cláudio Azevedo Passos, Eugênio (Sanders) Silva, Fernando Tamberlini (Thunder ... Thunder ...) Alves, Guilherme Matosinho Stiebler, Leticia Winkler, Marcelo de Oliveira Mello Marques (Álvaro) Alvim, Pedro (Petrovic) Demasi, Suelaine dos Santos Diniz, Thiago da Silva de Pontes.

Aos amigos Márcio Valle de Oliveira e Cruz, Ramiro Pereira de Magalhães, Raphael Gomes e Rodrigo Azevedo da Rocha.

A Maurício Carvalho dos Santos (CENPES/PETROBRAS), por confiar em nosso grupo e proporcionar a oportunidade da aplicação dos conhecimentos adquiridos ao longo do estudo a um problema prático e real.

A Diogo Fujio Takano e Nelson Eduardo Costa Cruz Lopes pela ajuda inestimável que ofereceram na montagem do experimento do primeiro estudo de caso.

À CAPES, pelo suporte que proporcionou através da concessão de uma bolsa de Mestrado.

Resumo

Este trabalho consiste do estudo de técnicas de Processamento de Imagens e seu emprego nos sistemas de monitoração de processos baseados em imagens, além da realização de três estudos de caso. O primeiro correspondente à detecção da queda da primeira gota em um processo de destilação de petróleo, o segundo ao controle do volume de petróleo acumulado em um frasco coletor no mesmo processo, e o terceiro a uma aplicação anti-furto de monitoração de um objeto específico em uma sala. Para cada estudo de caso, são apresentadas as soluções propostas e os resultados obtidos.

Abstract

This work consists of the study of Image Processing techniques and their employment in image-based process monitoring systems, in addition to the execution of three case-studies. The first one is the detection of the first drop in an oil distillation process, the second is volume control for accumulated oil in a collector recipient in the same distillation process, and the third is an anti-theft application for monitoring a specific object in a room. For each case-study, the proposed solutions and the obtained results are presented.

Sumário

1	Introdução	19
1.1	Considerações Iniciais	19
1.2	Sistemas de Monitoração de Processos	20
1.3	Sistemas de Tempo Real	21
1.4	Processamento de Imagens Digitais	22
1.5	Objetivo	25
1.6	Organização da Dissertação	25
2	Revisão Bibliográfica	27
2.1	Introdução	27
2.2	Histórico	27
2.3	Componentes de um Sistema de Monitoração	30
2.4	Abordagens Visuais	31
2.4.1	Olho Humano	32
2.4.2	Câmera de Vídeo	33
2.4.3	Raios X	36
2.4.4	Infra-Vermelho	37
2.4.5	Ressonância Magnética	39
2.4.6	Ultra-Som	42
2.5	Pesquisa no Brasil	44
2.6	Pesquisa no Exterior	44
3	Processamento de Imagens	45
3.1	Introdução	45
3.2	Modelo de Imagem	45
3.2.1	Amostragem e Quantização	46

3.3	Relações Entre Pixels	47
3.3.1	Vizinhança	47
3.3.2	Conectividade	48
3.3.3	Distância	49
3.4	Operações Lógico-Aritméticas	50
3.4.1	Transformações Geométricas	52
3.5	Transformações Radiométricas	56
3.5.1	Histograma de uma Imagem	57
3.5.2	Limiarização Simples	58
3.5.3	Expansão do Histograma	58
3.5.4	Equalização do Histograma	59
3.6	A Transformada de Fourier	60
3.6.1	Caso Contínuo	61
3.6.2	Caso Discreto	64
3.6.3	Compressão do Espectro	65
3.6.4	A Transformada Rápida de Fourier	66
3.7	Outras Ferramentas Matemáticas	67
3.7.1	Convolução	67
3.7.2	Correlação	70
3.8	Realce de Imagens	70
3.8.1	Tipos de Filtragem	72
3.8.2	Filtragem no Domínio do Espaço	72
3.8.3	Filtragem no Domínio da Frequência	76
3.9	Rotulação de Componentes Conexas	81
3.10	Detecção de Movimento	82
3.10.1	Comparação Simples	82
3.10.2	Diferenças Acumulativas	86
3.10.3	Atualização do Fundo	86
3.10.4	Cálculo de Velocidade	87
4	Estudo de Casos	89
4.1	1º Caso: Detecção da Queda da Primeira Gota em um Processo de Destilação de Petróleo em Batelada	89
4.1.1	Introdução	89

4.1.2	O Processo de Destilação	91
4.1.3	O Cenário	92
4.1.4	O Sistema	93
4.1.5	Localização do Bico	94
4.1.6	Detecção da Primeira Gota	97
4.1.7	Trepidações e Reflexos	98
4.1.8	Resultados	99
4.2	2º Caso: Controle do Nível de Petróleo em Frasco Destino de um Processo de Destilação em Batelada	101
4.2.1	Introdução	101
4.2.2	Categorias do Problema	102
4.2.3	Posicionamento da Câmera	103
4.2.4	Sub-tarefas do Sistema	104
4.2.5	Localização das Paredes do Frasco	104
4.2.6	Localização das Marcas	109
4.2.7	Determinação do Nível do Líquido	110
4.2.8	Tratamento de Distorções Projetivas	114
4.3	3º Caso: Controle Anti-Furto de Objetos	122
4.3.1	Abordagem Proposta	123
4.3.2	Amostragem do Objeto	123
4.3.3	Determinação dos Intervalos de Cor	124
4.3.4	Construção do Módulo de Reconhecimento	124
4.3.5	Detecção de Movimento no Recinto	128
4.3.6	Localização do Objeto	133
4.3.7	Exemplo Prático	134
5	Conclusões	139
5.1	Cumprimento de Objetivos	139
5.1.1	Estudo de Processamento de Imagens	139
5.1.2	Aplicação à Monitoração de Processos	140
5.1.3	Estudos de Caso	140
5.1.4	Análise de Adequação	143
5.1.5	Trabalhos Futuros	145
	Referências Bibliográficas	147

<i>SUMÁRIO</i>	12
Apêndice I	152
Apêndice II	153
Apêndice III	154

Lista de Figuras

1.1	Diagrama apresentando as entidades chave em um cenário de Sistema de Tempo Real.	22
1.2	Imagem transmitida com a utilização do sistema Bartlane.	23
1.3	Exemplo de aplicação de técnicas de desborramento em uma imagem, obtido em (PANASYUK, 2004).	24
2.1	Sistema baseado em relés.	29
2.2	Sistema baseado em circuitos integrados.	29
2.3	ENIAC, um dos primeiros computadores a serem desenvolvidos. . . .	29
2.4	PLC do tipo <i>rack</i>	29
2.5	Esquema de um sistema de monitoração de processos.	31
2.6	Visão externa do olho humano.	32
2.7	Visão interna do olho humano.	32
2.8	Tela do sistema de controle de tráfego aéreo do aeroporto de Sidney, Austrália.	34
2.9	Imagem de uma câmera da CET-RIO localizada na Av. Nossa Senhora de Copacabana, em Copacabana, Rio de Janeiro.	35
2.10	Primeira radiografia de Röntgen.	36
2.11	Espectro Eletromagnético.	38
2.12	Vítima em quarto repleto de fumaça.	39
2.13	Vítima de afogamento noturno.	39
2.14	Átomos de hidrogênio submetidos a um campo magnético.	41
2.15	Imagem da cabeça de uma pessoa gerada por ressonância magnética. .	42
2.16	Imagem gerada por ultra-som de um feto com 12 semanas de idade. .	43
3.1	Trecho de uma imagem.	48
3.2	Exemplo da aplicação da operação de soma para redução de ruído. . .	50

3.3	Exemplo da aplicação da operação de subtração para detecção de movimento.	51
3.4	Exemplo da aplicação da operação de multiplicação para destacar uma determinada região em uma imagem.	51
3.5	Exemplos de operações lógicas.	52
3.6	Região da imagem original de um escalamento bilinear.	53
3.7	Exemplos de operações de escalamento, utilizando os métodos de interpolação b) vizinho mais próximo, c) bilinear e d) bicúbico.	55
3.8	Exemplos de operações de rotação.	56
3.9	Imagem base para construção do histograma.	57
3.10	Histograma da imagem na Figura 3.9.	57
3.11	Limiarização com $\theta = 50$ aplicada à Figura 3.9.	58
3.12	Imagem com histograma expandido.	59
3.13	Histograma da imagem na Figura 3.12.	59
3.14	Imagem com histograma equalizada.	60
3.15	Histograma da imagem na Figura 3.14.	60
3.16	Função $f(x) = c$, definida no intervalo $[0, T]$	62
3.17	Espectro de Fourier da função $f(x) = c$, definida no intervalo $[0, T]$	63
3.18	Gráfico da função $f(x, y) = c$	63
3.19	Espectro de Fourier da função $f(x, y) = c$	63
3.20	Imagem base para visualização do <i>espectro</i>	65
3.21	<i>Espectro</i> da imagem na Figura 3.20 sem compressão de intervalo.	65
3.22	<i>Espectro</i> da imagem na Figura 3.20 com compressão de intervalo.	65
3.23	Funções discretas unidimensionais $f(x)$ e $g(x)$	68
3.24	Convolução entre as funções discretas $f(x)$ e $g(x)$	68
3.25	Exemplos de aplicações de filtros passa-baixa no domínio do espaço.	73
3.26	Exemplos de remoção de ruídos através da aplicação de filtragem passa-baixa por mediana.	74
3.27	Exemplos de filtragem passa-alta no domínio do espaço.	75
3.28	Exemplos de <i>espectros</i> de potência.	77
3.29	Gráfico de um filtro ideal passa-baixa.	78
3.30	Exemplo de filtragem ideal passa-baixa no domínio da frequência.	79
3.31	Gráfico de um filtro de Butterworth passa-baixa.	80
3.32	Exemplo de filtragem Butterworth passa-baixa no domínio da frequência.	80

3.33	Gráfico de um filtro ideal passa-alta.	81
3.34	Exemplo de filtragem ideal passa-alta no domínio da frequência. . . .	82
3.35	Gráfico de um filtro de Butterworth passa-alta.	84
3.36	Exemplo de filtragem Butterworth passa-alta no domínio da frequência.	84
3.37	Exemplo de segmentação de imagens através de rotulação de compo- nentes conexas.	85
3.38	$f(x, y, t_i)$	85
3.39	$f(x, y, t_j)$	85
3.40	$d(x, y)$	85
3.41	Exemplo da aplicação da técnica de atualização do fundo em situações de detecção de movimento.	87
4.1	Unidade de destilação em batelada.	91
4.2	Cenário do sistema.	93
4.3	Esquema das <i>threads</i> do sistema.	94
4.4	Foto da interface do sistema de detecção da queda da primeira gota no processo de destilação de petróleo.	95
4.5	Imagem em detalhe do bico do aparelho de destilação.	96
4.6	Bico do aparelho de destilação iluminado.	96
4.7	Imagem limiarizada do bico do aparelho de destilação.	96
4.8	Região conexa correspondente ao bico encontrado.	96
4.9	Projeção vertical do bico separado.	97
4.10	Projeção horizontal do bico separado.	97
4.11	Localização do bico.	97
4.12	Marcação da região de interesse.	97
4.13	Quadros correspondentes à queda de massa líquida.	99
4.14	Resultados da aplicação da rotina de detecção de movimento nos qua- dros da Figura 4.13.	100
4.15	Imagem do frasco utilizado nos testes em laboratório para o sistema de monitoração da variação de volume no processo de destilação de petróleo em batelada.	102
4.16	Cenário básico do sistema de controle do nível de líquido no frasco da unidade de destilação.	103

4.17	Diferentes tipos de visão do líquido baseados no posicionamento da câmera.	104
4.18	Região do frasco destacada com o uso do <i>mouse</i> pelo usuário.	105
4.19	Resultado do processo de limiarização simples aplicado à região de interesse indicada na Figura 4.18.	105
4.20	Aplicação de um filtro passa-baixa para remoção de ruídos da Figura 4.19.	105
4.21	Imagem contendo o resultado da aplicação do algoritmo de rotulação de componentes conexas na Figura 4.20.	106
4.22	Figura 4.21 após processo de dilatação.	106
4.23	Imagem contendo os contornos do frasco dividida em partições de largura 10.	107
4.24	Gráfico que ilustra a relação entre cada uma das partições apresentadas na Figura 4.23 e o número de <i>pixels</i> brancos contidos em cada uma.	107
4.25	Partições candidatas eleitas pelo método 1.	108
4.26	Projeção vertical da partição de número 7.	109
4.27	Projeção vertical da partição de número 8.	109
4.28	Limites laterais do frasco de coleta de petróleo.	110
4.29	Marcas indicadoras dos limites inferior e superior do frasco.	111
4.30	Frasco segmentado do restante da imagem.	111
4.31	Imagens do processo de localização da marcas indicadoras dos limites inferior e superior do frasco.	111
4.32	Projeção vertical da Figura 4.31.c e a posterior remoção de regiões de pouca relevância para localização das marcas.	112
4.33	Imagens do processo de determinação do nível de líquido armazenado no frasco.	112
4.34	Projeção vertical do resultado da aplicação da técnica Comparação Simples na Figura 4.33.c.	113
4.35	a) Região contendo a superfície do líquido segmentada do restante do frasco e b) projeção vertical da mesma.	114
4.36	Localização da superfície do líquido determinada pela estratégia apresentada.	115

4.37	Cenário ilustrativo da primeira abordagem para tratamento das distorções projetivas da imagem.	116
4.38	Determinação do ponto de fuga através das marcações dos segmentos de reta \overline{PQ} e \overline{RS}	117
4.39	Imagem de um frasco na qual podem ser observadas as distorções projetivas.	119
4.40	Imagem do frasco mostrado na Figura 4.39 alterada pela abordagem corretiva.	121
4.41	Exemplo da aplicação do algoritmo de obtenção do contorno de um objeto.	127
4.42	Sala de testes para a monitoração anti-furto.	130
4.43	Exemplo de detecção de movimento com uma pessoa na sala.	130
4.44	Exemplo de detecção de movimento sem uma pessoa na sala.	131
4.45	Projeções verticais de situações com a ausência e presença de uma pessoa.	132
4.46	Sala na qual foi realizado o experimento da monitoração anti-furto de um objeto, estando o mesmo destacado por um retângulo que o envolve.	134
4.47	Imagens do objeto de interesse a serem utilizadas na construção do módulo de reconhecimento.	135
4.48	Pontos clicados pelo usuário para a determinação dos intervalos de cor do objeto.	135
4.49	Entidades conexas segmentadas de uma imagem posterior à saída da pessoa da sala e seus respectivos rótulos.	137

Lista de Tabelas

2.1	Pesquisa no Brasil.	44
2.2	Pesquisa no Exterior.	44
3.1	Conjuntos $N_4(E)$, $N_D(E)$ e $N_8(E)$	48
4.1	Relação de classificação entre dois quadros subseqüentes.	98
4.2	Resultados dos testes realizados para o sistema.	100
4.3	Resultados obtidos pela abordagem corretiva para o problema das distorções projetivas.	120
4.4	Relação distância - número de <i>pixels</i> para as marcações da Figura 4.39.121	
4.5	Resultado da rede neural para o problema das distorções projetivas. .	122
4.6	Resultados atingidos pelo módulo de reconhecimento.	138

Capítulo 1

Introdução

1.1 Considerações Iniciais

A automação de tarefas é uma tendência real e inegável na sociedade do século XXI. O debate sobre a substituição da mão-de-obra humana pela máquina repousa sobre o avanço tecnológico o véu da desconfiança, levando a humanidade a questionar sua função no planeta e a buscar o porquê de seus próprios passos.

É cada vez mais difícil fazer qualquer tipo de prognóstico confiável sobre a forma como a tecnologia afeterá a vida da população mundial nos anos por vir, e se fortalece de forma progressiva o fantasma do desenvolvimento tecnológico como provável catalisador de um colapso mundial de desemprego e pobreza.

Enquanto o discurso humanista prega a redução do ritmo de desenvolvimento da tecnologia, o discurso técnico propõe o controle populacional. Encurralados entre eles e, por que não, injustiçados, encontram-se os frutos essenciais do avanço tecnológico colhidos ao longo de séculos de estudo e trabalho. Com exceção de puristas ecológicos, quem contesta os benefícios da eletricidade, o conforto da Engenharia, a praticidade da Informática ou a capacidade da Medicina moderna?

A grande verdade é que à medida que a tecnologia avança e a população cresce, se torna cada vez mais difícil que qualquer um de nós tenha condições reais de fazer um julgamento coerente e imparcial no que diz respeito ao aspecto global do desenvolvimento. A teia social que se tece com o passar dos anos e a forma como a rapidez do movimento tecnológico e as próprias correntes culturais, políticas e econômicas mundiais a afetam nos colocam constantemente em posição de defesa e submissão.

Tal panorama não pode, porém, impedir que se vislumbrem as possíveis aplicações

do potencial tecnológico atingido, considerando-se que as mesmas têm abrangência de atuação indescritivelmente ampla, desde a simples facilitação de uma tarefa corriqueira e repetitiva ao salvamento de vidas.

Os sistemas de computação têm papel de importância singular neste cenário, já que desde o advento do computador de grande porte, o ritmo de conquista do conhecimento tecnológico teve sua aceleração exponencialmente catapultada, e tal fenômeno continua ocorrendo ainda nos dias de hoje.

Este estudo se concentra em uma das categorias de sistemas mais proeminentes do mundo da computação, tanto no meio acadêmico, quanto no âmbito comercial, a qual diz respeito aos sistemas que fazem uso da imagem, objeto fundamental da visão, para realizar monitoração de processos em geral, sejam eles biológicos, como a evolução de uma enfermidade ou doença, industriais, como procedimentos de controle de qualidade, ou situações cotidianas, como a vigilância de um determinado bem no interior de um recinto qualquer.

1.2 Sistemas de Monitoração de Processos

Em geral, sistemas de monitoração de processos têm como finalidade a observação contínua ou periódica do desenvolvimento de um procedimento qualquer, seja para controlar ou mesmo verificar determinadas características do mesmo, seja para identificar algum evento específico referente ao cenário em questão.

Existem inúmeras modalidades de sistemas de monitoração de processos, empregando sensores dos mais variados modelos e aplicados aos mais diversos tipos de ambiente. Neste trabalho, porém, serão considerados apenas os sistemas que utilizam imagens e técnicas de processamento das mesmas para atingir seus objetivos. Esses sistemas são, em geral, classificados como Sistemas de Tempo Real, uma vez que seu desempenho está normalmente, de certa forma, atrelado ao seu tempo de resposta, tendo em vista que os problemas que eles se propõem a resolver têm restrições temporais inerentes.

Esta característica confere a este tipo de sistema uma dificuldade considerável, dado que as técnicas de processamento de imagens são, em geral, muito custosas em termos computacionais. Portanto, os sistemas de monitoração de processos baseados em imagens exigem planejamento meticuloso e implementação otimizada, visando à redução dos efeitos que um código mal escrito pode provocar no tempo de resposta

do sistema.

As seções 1.3 e 1.4 têm como objetivo fazer uma apresentação introdutória de Sistema de Tempo Real e Processamento de Imagens Digitais, respectivamente. Não há nas seções mencionadas o interesse de se explorar os campos em questão de forma específica, mas sim posicionar o leitor quanto ao que representam, suas propostas e potencialidades.

1.3 Sistemas de Tempo Real

São ditos **reativos** os sistemas que interagem com um determinado ambiente e reagem a estímulos produzidos pelo mesmo. Diz-se então que Sistemas de Tempo Real são, na verdade, sistemas reativos cuja reação deve ocorrer dentro de prazos específicos, os quais podem apresentar tolerância variada (FARINES & FRAGA & OLIVEIRA, 2004).

Confere-se então aos Sistemas de Tempo Real, tendo em vista tal definição, requisitos temporais, o que transforma o conceito de correção desses sistemas em relação ao empregado em sistemas convencionais de computação, nos quais a partir de um certo conjunto de entradas deve ser atingido um conjunto específico de saídas correspondentes. Para Sistemas de Tempo Real a qualidade de uma resposta está atrelada tanto à correção na relação entrada \rightarrow saída, quanto ao cumprimento de um determinado prazo estipulado para a realização da tarefa em questão (FARINES & FRAGA & OLIVEIRA, 2004).

Grande parte dos cenários nos quais se enquadram os Sistemas de Tempo Real pode ser descrito conforme mostra a Figura 1.1, apresentando 5 entidades chave: o sistema a controlar, a interface de instrumentação, o sistema computacional de controle, a interface homem-máquina e o operador (FARINES & FRAGA & OLIVEIRA, 2004).

Não é incomum que requisitos temporais do sistema sejam estabelecidos em concordância com alguma exigência externa ao mesmo, especialmente em sistemas que oferecem determinados serviços a um usuário humano, como por exemplo a exibição de imagens (vídeo) a um observador em Sistemas Multimídia. Neste caso, os requisitos temporais do sistema são determinados pelas condições de contorno do ambiente final que será afetado pela atuação do operador (FARINES & FRAGA

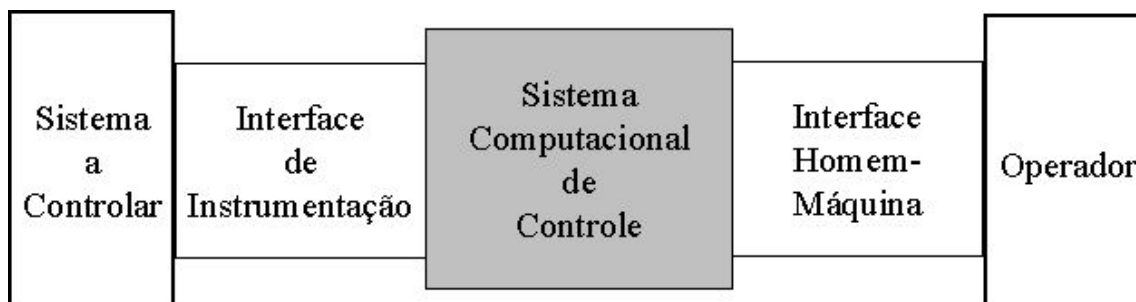


Figura 1.1: Diagrama apresentando as entidades chave em um cenário de Sistema de Tempo Real.

& OLIVEIRA, 2004). Como exemplo, pode-se considerar uma esteira industrial de altíssima velocidade de processamento na qual são produzidas embalagens de um determinado produto alimentício, sendo o controle de qualidade das embalagens realizado por um operador que recebe imagens geradas por uma câmera em algum ponto da esteira. Supondo que falhas nas embalagens ocorram sempre em lote, pode-se considerar a análise amostral como a mais adequada para a identificação dessas falhas, desde que o sistema de vídeo (neste caso de Tempo Real) garanta que a amostragem é feita no tempo correto, de forma que a frequência de exibição das imagens das embalagens permita que o operador seja capaz de identificar defeitos em um determinado lote caso eles ocorram.

A monitoração de um processo, por mais simples que ele seja, implica, mesmo que implicitamente, a consideração do quesito tempo, seja na descrição do intervalo ao longo do qual o sistema opera, seja na representação dos momentos em que determinados eventos de interesse ocorrem. Conclui-se então que requisitos temporais são absolutamente inerentes a sistemas de monitoração de processos, alocando-os então na categoria de Sistemas de Tempo Real.

1.4 Processamento de Imagens Digitais

O objetivo desta seção é apresentar, de forma superficial e introdutória, o universo das aplicações do estudo de processamento de imagens digitais, mostrando, além de um pouco de sua História, sua crescente importância e aplicação nas mais diversas áreas do conhecimento humano. No Capítulo 3 abordar-se-á a área de forma mais específica e técnica, apresentando-se a teoria básica por trás da mesma.

Processamento de imagens digitais se refere, em geral, à manipulação, por um computador digital, de uma imagem bidimensional. De forma mais abrangente, envolve o processamento de qualquer conjunto bidimensional de dados (JAIN, 1989). Dentre os diversos campos de aplicação do estudo de processamento de imagens digitais pode-se citar como principais a melhoria da qualidade da informação visual da imagem, com a finalidade de facilitar a interpretação humana, e o processamento dos dados contidos em uma imagem visando possibilitar a percepção automática por máquinas (GONZALEZ & WOODS, 2000).

Estas duas categorias englobam grande parte do conjunto de técnicas existentes em processamento de imagens. A História, porém, conta que um outro campo foi de fundamental importância para o estudo destas técnicas: a compressão de dados, visando à redução do tempo de transmissão de imagens. No início dos anos 20, imagens de jornais eram transferidas de Londres para Nova York através de um cabo submarino que atravessava o oceano Atlântico. Desenvolvido por Harry G. Bartholomew e Maynard D. McFarlane, o sistema Bartlane de transmissão de imagens reduziu o tempo de transferência de uma imagem de pouco mais de uma semana para menos de três horas. O sistema era baseado em cartões perfurados e fita telegráfica, sendo as imagens codificadas na origem, transmitidas e decodificadas no destino. Inicialmente, a capacidade de representação do sistema limitava-se a cinco tons de brilho, os quais eram simulados com caracteres impressos por uma impressora telegráfica. Em 1929, esta capacidade foi aumentada para 15 tons de brilho. A Figura 1.2 mostra uma imagem transferida com a utilização do sistema Bartlane (GONZALEZ & WOODS, 2000; DIGICAMHISTORY, 2004).



Figura 1.2: Imagem transmitida com a utilização do sistema Bartlane.

Outras técnicas visando à transmissão de imagens digitais foram desenvolvidas com o passar dos anos. Porém, apenas a partir da combinação entre o surgimento dos computadores de grande porte e o programa espacial americano que o potencial da teoria de processamento de imagens foi explorado com maior atenção, empregado na correção de determinados tipos de distorções em imagens da Lua, geradas pela espaçonave Ranger 7, a qual correspondeu à primeira tentativa bem sucedida do governo americano de obter fotografias da superfície lunar (GONZALEZ & WOODS, 2000; NSSDC, 2004).



Figura 1.3: Exemplo de aplicação de técnicas de desborramento em uma imagem, obtido em (PANASYUK, 2004).

Até os dias de hoje, pesquisadores discordam sobre os limites entre processamento de imagens digitais e áreas correlacionadas, como análise de imagens e visão computacional. Além da divisão em categorias mencionada no primeiro parágrafo desta seção, pode-se separar as operações de processamento de imagens digitais em processos de baixo, médio e alto nível. Processos de nível baixo são operações de natureza mais primitiva, como pré-processamento para a redução de ruídos e aumento de contraste, sendo que tanto a entrada quanto a saída das operações são imagens. Processos de nível médio englobam segmentação e descrição de elemen-

tos da imagem para posterior processamento e classificação (reconhecimento). Para essas operações, a saída corresponde não a uma imagem em particular, mas sim a atributos extraídos da mesma, como contornos e a identidade de determinados objetos. Por fim, as operações de nível alto correspondem a tarefas mais relacionadas a interpretação e compreensão de imagens, nas quais se encontram vários objetos já reconhecidos (GONZALEZ, WOODS & EDDINS, 2004).

Conforme pode ser visto na Figura 1.3, o potencial da utilização de técnicas de processamento de imagens digitais é enorme, se aplicando a diversas áreas do conhecimento humano. Dentre elas, podem ser citadas Medicina (radiologia, ressonância magnética, tomografias computadorizadas e ultra-sonografia), Arqueologia (recuperação visual de artefatos danificados e simulação e visualização de estruturas desaparecidas ou extintas), Biologia (análise de organismos microscópicos), Criminalística (identificação de vítimas ou criminosos), Segurança (circuitos de vigilância por televisão), Defesa (radares e sonares utilizados na identificação de alvos ou inimigos), Geologia (estudo de relevo e formações rochosas em larga escala), Geografia (previsão de colheitas, desenvolvimento urbano, variações climáticas, controle de inundações e incêndios, etc), entre outras (JAIN, 1989; GONZALEZ & WOODS, 2000).

1.5 Objetivo

O objetivo deste trabalho consiste da realização de 3 tarefas: a) o estudo de técnicas de Processamento de Imagens e suas aplicações a sistemas de monitoração de processos baseados em imagens; b) a realização de três estudos de caso nos quais problemas práticos de monitoração são abordados e soluções propostas com o emprego das técnicas mencionadas e c) a análise da adequação de sistemas baseados em imagens à atividade de monitorar um processo qualquer, sendo sub-objetivos o levantamento de dificuldades inerentes ao uso de imagens para este fim e a identificação da possibilidade de melhorias e avanços naquilo que for implementado neste trabalho.

1.6 Organização da Dissertação

O Capítulo 2, correspondente à revisão bibliográfica deste trabalho, apresenta um apanhado da história e das diferentes modalidades de sistemas de monitoração de

processos baseados em imagem.

No Capítulo 3 é realizada uma introdução teórica ao Processamento de Imagens, sendo apresentadas técnicas fundamentais de manipulação de imagens, além de exemplos e aplicações.

O Capítulo 4 é dividido em três partes essenciais, sendo cada uma delas correspondente a um diferente estudo de caso realizado neste trabalho.

O primeiro deles corresponde a um sistema de monitoração de um processo de destilação de petróleo em batelada, parte integrante do projeto de Nacionalização da Unidade de Destilação em Batelada, uma parceria entre o CENPES/PETROBRAS e a Escola de Química/UFRJ. Nesse estudo de caso, o objetivo do sistema é detectar a queda da primeira gota de petróleo a partir do condensador lateral da Unidade de Destilação em Batelada.

No segundo estudo de caso, parte integrante do projeto mencionado no parágrafo anterior, o objetivo do sistema é acompanhar e determinar o volume de petróleo armazenado no frasco de coleta da unidade.

O terceiro estudo de caso é uma abordagem preliminar ao problema da monitoração anti-furto de um objeto em um recinto qualquer.

Em cada uma das subdivisões são expostos o problema, a solução concebida e os resultados obtidos.

No Capítulo 5 são exibidas as conclusões do trabalho e as propostas de trabalhos futuros.

Capítulo 2

Revisão Bibliográfica

2.1 Introdução

Embora o conceito atual de monitoração de processos apresente uma tendência a nos remeter mentalmente a ambientes modernos e de alta sofisticação tecnológica, tal procedimento ocorre na natureza desde o surgimento da vida no planeta. Um leão à espreita de uma gazela monitora seu processo de movimentação. Com isso, ele averigua a vulnerabilidade da presa e tenta identificar o momento perfeito para o ataque. Uma leoa que protege seus filhotes de uma alcatéia de hienas monitora o posicionamento do grupo de predadores com o intuito de ocupar a posição mais adequada para defender sua prole.

A monitoração de um processo consiste da observação de um determinado procedimento ou ambiente com a finalidade de controlar suas características ou identificar eventos relacionados ao mesmo. A variabilidade da complexidade desta prática é extremamente ampla, abrangendo do cozinheiro que, baseado no cheiro, textura e outros aspectos, decide se o alimento sendo preparado está pronto ou não, até o controle de segurança de uma usina nuclear na ocorrência de algum acidente que pode causar a morte de milhares de pessoas.

2.2 Histórico

Comercialmente, a monitoração de processos é empregada desde a Idade Média, quando a economia ainda girava em torno do Mercantilismo. Uma operação delicada naqueles dias correspondia ao desembarque de especiarias, metais preciosos e outras riquezas trazidas de mercados longínquos ou colônias das grandes potências

da época.

Como não poderia deixar de ser, foi durante a Revolução Industrial (Inglaterra, séculos XVIII-XIX) que a monitoração de processos passou a receber maior atenção de comerciantes e estudiosos, especialmente à medida que os procedimentos de trabalho e produção ganhavam em complexidade e magnitude. Porém, com as limitações tecnológicas da época, os sistemas de controle construídos eram dispositivos mecânicos altamente engenhosos de proporções físicas pouquíssimo práticas. Tais sistemas, empregados na realização de algumas tarefas críticas e repetitivas do processo de produção, apresentavam, além do tamanho, a desvantagem de serem absolutamente inflexíveis e apresentarem tempo de vida muito curto (FILHO, 2002).

No início do século XX, com o surgimento dos relés elétricos, tornou-se possível a implementação de funções de controle mais complexas e poderosas. Porém, considerando que cada relé tem aproximadamente o tamanho de uma caixa de fósforos comum, o mesmo se mostrou como uma opção viável, em termos de custo, para a automação de procedimentos de menor escala. Já em sistemas mais complexos, o número de relés exigido seria muito elevado, comprometendo a eficácia espacial da solução (FILHO, 2002).

Mesmo assim, muitos sistemas de controle, sendo eles complexos ou não, foram implementados com o uso de relés. E embora os relés não sejam mais adotados para projetos deste tipo, ainda existe um número considerável de sistemas legados que utilizam tais dispositivos mundo afora.

Na segunda metade do século XX, com o surgimento dos circuitos integrados, deu-se mais um passo no desenvolvimento e evolução dos sistemas em geral. As maiores vantagens dos circuitos integrados eram sua precisão e principalmente seu tamanho. Justamente por causa deste fator, foi possível miniaturizar sistemas outrora inviáveis devido a questões espaciais (FILHO, 2002).

A partir da década de 70, o computador comercial foi incorporado ao controle de processos industriais. Sua maior vantagem era a possibilidade de se programar e modificar a lógica do sistema com facilidade, o que não era possível em sistemas baseados em relés ou circuitos integrados. Porém, o tamanho do computador, seu custo e sua pouca resistência a ambientes hostis (cenário comum em centros de produção em série) representavam dificuldades consideráveis para sua aplicação à monitoração de processos (FILHO, 2002).

Baseado nestas dificuldades, e em uma demanda muito grande da indústria au-

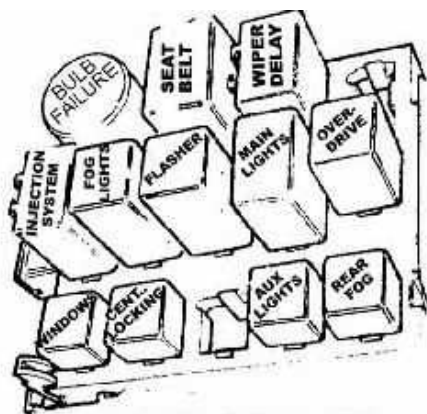


Figura 2.1: Sistema baseado em relés.

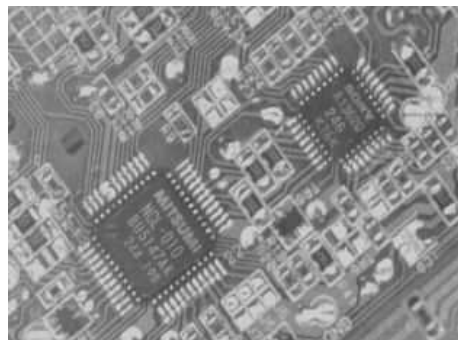


Figura 2.2: Sistema baseado em circuitos integrados.

tomobilística americana, foi criado o PLC (*Programmable Logic Controller*), desenvolvido especialmente para o ambiente industrial. Inicialmente, os PLCs admitiam apenas entradas digitais. A capacidade de lidar com entradas analógicas foi incorporada alguns anos mais tarde (FILHO, 2002).

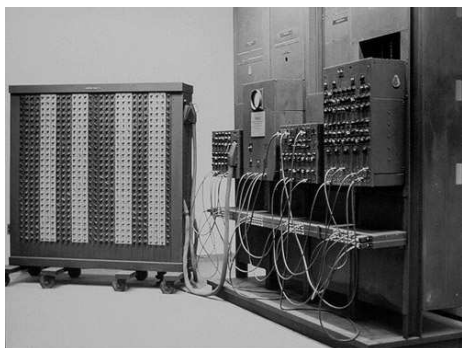


Figura 2.3: ENIAC, um dos primeiros computadores a serem desenvolvidos.

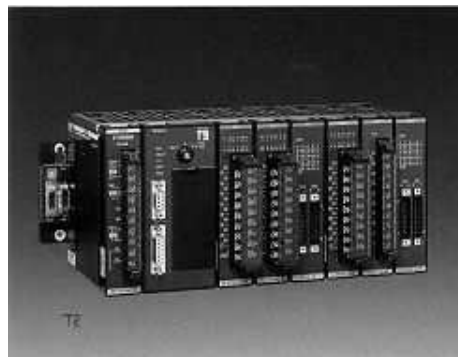


Figura 2.4: PLC do tipo rack.

À medida que os anos passaram, os sistemas de controle ficaram cada vez mais sofisticados. Nos dias atuais, sistemas embutidos (*embedded systems*) e microcontroladores representam uma vertente de pesquisa bastante proeminente neste enfoque de monitoração de processos (FILHO, 2002).

Por último, não poderia se deixar de fora o PC (*Personal Computer*), que embora por nome não se enquadre no assunto, é amplamente utilizado para hospedar

e executar sistemas de monitoração de processos. O avanço na capacidade de processamento, armazenamento e compactação permitiram ao PC figurar como uma opção eficiente e econômica ao controle de procedimentos.

2.3 Componentes de um Sistema de Monitoração

Os sistemas de monitoração de processos tem como função a observação de um determinado procedimento ou ambiente, identificando eventos ou controlando parâmetros de funcionamento ou estabilidade do mesmo. Para tanto, estes sistemas são divididos em três partes essenciais: os transdutores, os controladores e os atuadores (FILHO, 2002).

Os transdutores representam o elemento de sensibilidade dos sistemas de monitoração de processos, pois são eles os responsáveis por converter estímulos do meio no qual estão inseridos para sinais elétricos que possam ser interpretados pelo controlador. Transdutores digitais são capazes de monitorar variáveis discretas de um processo, como por exemplo o estado de um determinado aparelho, podendo o mesmo estar ligado ou desligado. Já para variáveis com faixas contínuas de valoração, como pressão, temperatura, nível e velocidade, são empregados transdutores analógicos (FILHO, 2002).

Os controladores são o cérebro do sistema de monitoração, sendo eles os responsáveis por interpretar os sinais enviados pelos transdutores e identificar o estado das componentes de interesse do procedimento ou ambiente monitorado. É também responsabilidade dos controladores identificar que ação deve ser executada, quando necessário, e enviar a ordem de execução da mesma aos atuadores (FILHO, 2002).

Os atuadores correspondem à ponte de comunicação entre os controladores e o ambiente ou procedimento monitorado. Através dos atuadores, são enviadas ordens para execução de alguma tarefa necessária para o sucesso do processo em questão (FILHO, 2002).

Voltando ao exemplo do leão que monitora a gazela, pode-se dizer que os olhos, ouvidos e nariz do leão agem como transdutores, através da visão, audição e olfato, para seu cérebro. Este representa o controlador, que processa os estímulos captados pelos transdutores e identifica o momento correto para atacar. Através do restante de seu sistema nervoso, que corresponde ao atuador, o cérebro dá a ordem ao corpo para que avance em direção à presa. A Figura 2.5 apresenta um esquema de um

sistema de monitoração.

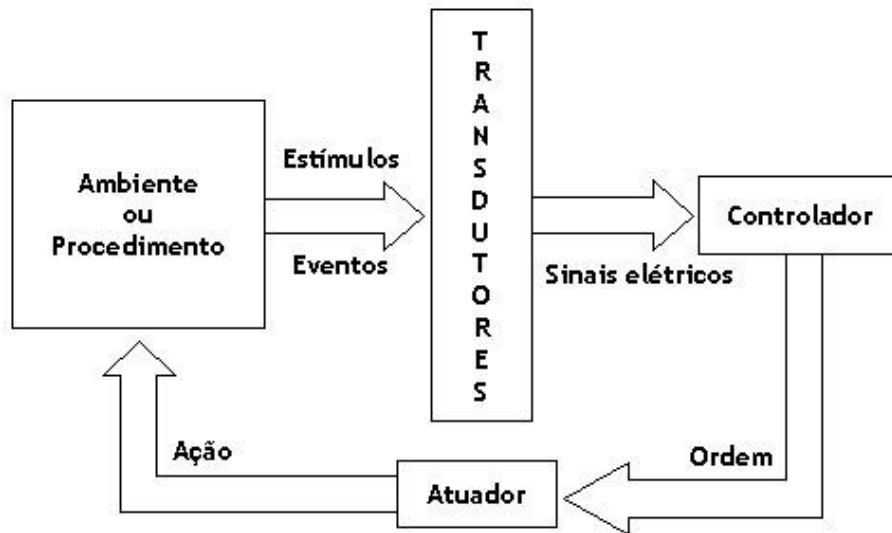


Figura 2.5: Esquema de um sistema de monitoração de processos.

2.4 Abordagens Visuais

Conforme foi mencionado, o campo da monitoração de processos é bastante vasto. Por este motivo, são inúmeras as maneiras pelas quais podemos abordar um problema e obter resultados satisfatórios. A título de exemplificação, considere um sistema modular físico no qual se deseja controlar a diferença de temperatura entre dois de seus módulos. Em uma primeira abordagem, poderiam ser usados sensores termo-sensíveis acoplados a diferentes regiões dos módulos de interesse, realizando-se algum tipo de interpolação entre as diferentes temperaturas computadas em cada sensor para se obter as medidas desejadas para cada módulo. Em uma segunda abordagem, poderia ser aproveitada a emissão de radiação infra-vermelha a partir dos módulos em questão, obtendo-se uma descrição térmica (representada através de uma imagem 2D) dos mesmos. E por que não uma solução que combine as duas abordagens?

A seguir, serão apresentadas diferentes abordagens adotadas para o controle de processos, porém, como este trabalho está basicamente direcionado ao uso de imagens com tal intuito, serão discutidas apenas técnicas que impliquem o emprego da visão, seja ela computacional ou não, na realização da monitoração em si.

2.4.1 Olho Humano

O olho humano (Figura 2.6) é o dispositivo de monitoração de processos mais utilizado em nosso dia a dia. Tal afirmativa se baseia no fato de que qualquer pessoa com nível mínimo de consciência realiza esta atividade diversas vezes durante o curso de um dia. Seja esta pessoa uma dona de casa que monitora a comida sendo preparada, ou um controlador de tráfego aéreo, que observa e orienta as chegadas e partidas de vôos em um aeroporto.

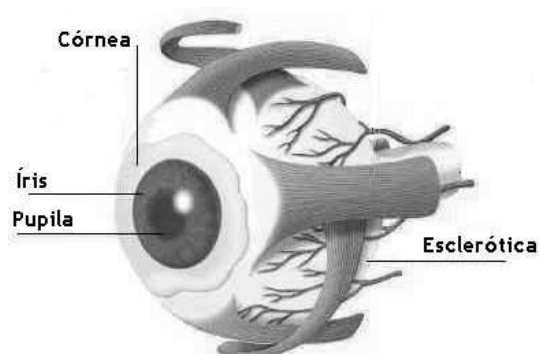


Figura 2.6: Visão externa do olho humano.

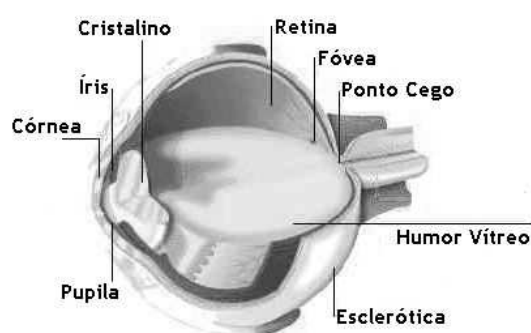


Figura 2.7: Visão interna do olho humano.

Nossos olhos agem como transdutores para que nosso cérebro nos conceda o privilégio da visão. Através da visão, interpretamos os estímulos luminosos do mundo em que vivemos e desfrutamos da capacidade de orientação, controle e proteção que ela nos fornece. E embora a utilização de nossos olhos seja simples e imediata, o funcionamento dos mesmos não tem nada de simples. Compostos por estruturas complexas e sofisticadas, os olhos são capazes de atuar sob diferentes condições de iluminação, umidade, temperatura, pressão, velocidade e inclinação.

O olho tem a forma aproximada de uma esfera, com diâmetro médio de 2 cm, e é coberto frontal e externamente por uma fina película transparente chamada córnea. Quando a luz de um objeto atinge nossos olhos, tem na córnea seu primeiro nível de contato com nosso organismo. Ao passar pela córnea, a luz incidente encontra a íris, que regula a quantidade de luz que penetra o olho. Isto é feito através da abertura ou fechamento da pupila, permitindo-se então que a quantidade de luz adequada atinja o cristalino. Este último funciona como uma lente, focando a imagem recebida sobre a retina, provocando inclusive a movimentação do olho de forma que a fóvea, região

central da retina, esteja alinhada com o centro da imagem (GONZALEZ & WOODS, 2000).

A retina é coberta por dois tipos de elementos foto-receptores, chamados cones e bastonetes. O olho contém cerca de 6 a 7 milhões de cones, estando a maioria deles concentrada na fóvea. Os cones são altamente sensíveis a cores, e cada um deles está diretamente ligado a uma fibra nervosa própria. Já os bastonetes, cuja quantidade varia entre 75 e 150 milhões, estão conectados a uma única fibra nervosa. Enquanto os cones se encarregam de possibilitar a visão de cores e detalhes finos de um objeto qualquer, os bastonetes operam sobre nossa capacidade de identificar a intensidade (luminância) das diferentes regiões do ambiente. (GONZALEZ & WOODS, 2000).

Vale ainda observar que a imagem projetada sobre a retina e recebida pelos cones e bastonetes é uma imagem invertida. Porém, nosso cérebro se encarrega de corrigir este erro e impedir que percebamos tal fato (GONZALEZ & WOODS, 2000).

Embora a modernidade e a tecnologia tenham se infiltrado em boa parte dos segmentos da vida humana, ainda existem muitas tarefas comerciais que empregam a monitoração de processos através da utilização do olho humano. Tome como exemplo o controle de tráfego aéreo. Embora existam softwares confiáveis que realizam a tarefa com eficiência, há uma resistência em se delegar tamanha responsabilidade a um programa. Por isso, nas torres de controle dos aeroportos ainda são humanos os responsáveis principais pelo controle dos vôos (SATCT, 2003). A Figura 2.8 apresenta a tela do sistema de controle de tráfego aéreo do aeroporto de Sidney, na Austrália. Em frente a este monitor, senta-se um operador que, através da utilização da visão, acompanha a movimentação das aeronaves localizadas nas proximidades do aeroporto e as orienta visando à segurança de pousos e decolagens das mesmas.

Outras atividades controladas por humanos apenas com o emprego da visão são:

- Controle de qualidade em processos industriais;
- Controle de entrada e saída de pessoal em portarias;
- Vigilância;

2.4.2 Câmera de Vídeo

Dentre os artificiais, a câmera de vídeo é o mais comum dispositivo visual de monitoração de processos, e seu emprego com tal finalidade se dá de duas formas:

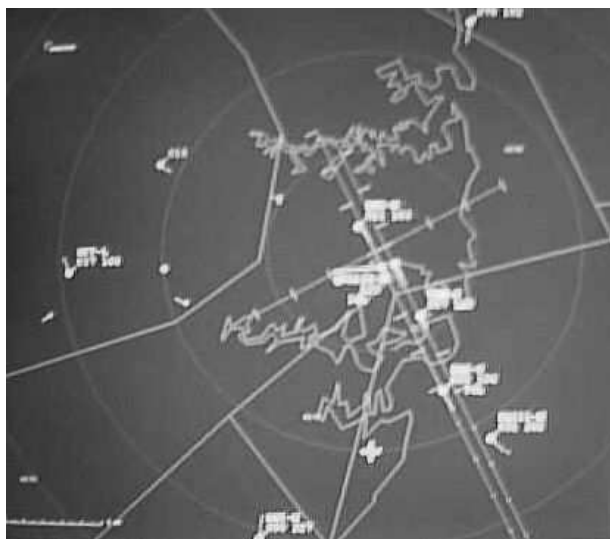


Figura 2.8: Tela do sistema de controle de tráfego aéreo do aeroporto de Sidney, Austrália.

1. As imagens geradas pela câmera são exibidas para um observador humano que aciona serviços e controles, ou toma as medidas que achar necessárias ou cabíveis. Esta prática é mais comum em atividades como segurança, monitoração de tráfego de veículos (Figura 2.9) e controle de entrada e saída de pessoal;
2. Através de uma placa de captura, os quadros gerados pela câmera são enviados para um computador, no qual os mesmos são analisados por um sistema que utiliza técnicas de processamento de imagens. Este tipo de estrutura é encontrado com mais frequência em ambientes de produção, como mecanismos de controle de qualidade ou de detecção de eventos.

O funcionamento de uma câmera de vídeo é, de certa forma, bem parecido com o do olho humano. Quando a luz atinge a lente da câmera, ela é direcionada a um prisma (ou a filtros de cores) e separada nas cores vermelho, verde e azul. Os raios de cada cor incidem sobre um dispositivo repleto de receptores foto-sensíveis, chamados de CCD (*Charged-Coupled Device*). Cada receptor atingido adquire uma carga elétrica que é proporcional a intensidade do brilho detectado. Posteriormente, um feixe de elétrons é direcionado à cada linha dos CCDs, de forma que, para cada receptor, gere-se uma corrente elétrica de intensidade proporcional à carga



Figura 2.9: Imagem de uma câmera da CET-RIO localizada na Av. Nossa Senhora de Copacabana, em Copacabana, Rio de Janeiro.

adquirida pelo receptor no momento de incidência da luz (AW, 2003; GOMES & VELHO, 1998).

As correntes geradas a partir dos receptores dos três CCDs são decodificadas e utilizadas para determinar o valor das componentes de vermelho, verde e azul de cada *pixel* da imagem. Esta operação é executada de forma repetitiva, sendo cada realização da mesma responsável pela geração de um quadro (AW, 2003).

As aplicações da câmera de vídeo são bastante variadas, especialmente com o crescimento acelerado do interesse de pesquisadores no campo da visão computacional e com o desenvolvimento da microeletrônica, o que tornou possível a construção de câmeras de maior portabilidade. Em monitoração de processos, a câmera de vídeo é empregada em diversas atividades, estando algumas delas citadas a seguir:

- Controle de entrada e saída de pessoal e materiais;
- Controle de eventos a grandes profundidades no mar;
- Detecção de falhas ou rachaduras em embalagens;
- Detecção de eventos específicos em processos industriais;
- Monitoração de alteração gradativa de cor em líquidos específicos;
- Vigilância;

2.4.3 Raios X

Data de novembro de 1895 a descoberta dos raios X, pelo físico Wilhelm Röntgen, então reitor da Universidade de Würzburg na Alemanha. Conta a História que Röntgen realizava testes com o tubo de raios catódicos, inventado pelo cientista inglês William Crookes, que nada mais era que um condutor metálico aquecido que emitia elétrons (então chamados raios catódicos) em direção a outro condutor, quando percebeu algo extremamente peculiar. Uma placa de material fluorescente próxima ao aparelho apresentava um brilho que Röntgen não conseguia explicar. Mesmo depois de colocar um livro e uma folha de alumínio entre a placa e o tubo, o brilho persistiu. Röntgen concluiu que partículas emitidas pelo condutor atravessavam as barreiras físicas e atingiam a placa (UCB, 2003).

Dias depois, após ter estudado o fenômeno intensamente, Röntgen emitiu radiação através da mão de uma mulher contraposta a uma chapa fotográfica por aproximadamente 15 minutos. Quando a chapa foi revelada, era possível ver os contornos e sombras dos ossos da mão da mulher, sendo esta a primeira radiografia da qual se tem notícia (Figura 2.10). Röntgen decidiu chamar os raios de X pelo que o símbolo representava matematicamente, ou seja, o desconhecido (UCB, 2003).



Figura 2.10: Primeira radiografia de Röntgen.

O aparelho de raios X dos dias de hoje consiste de um tubo de onde são disparados elétrons em direção a um alvo sólido, composto de átomos pesados, como o tungstênio. Os raios X são produzidos em parcelas de energia chamadas fótons, assim como a luz visível. Quando atingem o material, os elétrons perdem velocidade e liberam energia, emitindo fótons de radiação X, que, por seu elevado nível

de energia, é capaz de atravessar objetos (UCB, 2003).

Coloca-se então o objeto de interesse entre um anteparo sensível a radiação X e o aparelho de emissão da mesma. De acordo com a densidade e disposição dos materiais que compõe o objeto, o nível de absorção da radiação emitida varia ao longo do objeto, fazendo com que o anteparo registre as formas das partes mais densas do interior do objeto. Desta forma, tem-se a radiografia (UCB, 2003).

Em sistemas de monitoração, a radiação X é bastante utilizada pela indústria de alimentos, especialmente com o intuito de localizar corpos estranhos ou indesejados nos produtos, como caroços em tortas ou doces feitos a partir de alguma fruta. É comum, por exemplo, na atividade pecuária, que se apliquem injeções no gado para evitar doenças que prejudiquem a qualidade da carne. Muitas vezes o animal se move no momento da aplicação, causando a quebra da agulha, que permanece no interior do seu corpo até o abate. A radiação X é então empregada na tentativa de identificar a presença destas agulhas na carne, antes da distribuição da mesma (SWRI I, 2003).

2.4.4 Infra-Vermelho

Em 1800, interessado em estudar a quantidade de calor que passava pelos filtros coloridos que usava para observar o sol, o astrônomo alemão Frederick William Herschel realizou um experimento peculiar que o levou a descoberta dos raios infra-vermelhos (IPAC, 2003).

A experiência consistia da incidência direta de raios do sol em um prisma de vidro, com o objetivo de criar o espectro de cores, e então, com a utilização de termômetros, medir a temperatura de cada faixa de cor, utilizando termômetros de controle para comparação. Ao medir a temperatura das faixas violeta, azul, verde, amarelo, laranja e vermelha, Herschel constatou que não só suas temperaturas eram mais altas do que as registradas pelos termômetros de controle, mas que essas temperaturas aumentavam da faixa violeta para a vermelha. Por curiosidade, Herschel mediu a temperatura de uma faixa imaginária localizada após o vermelho. Para sua surpresa, foi nesta faixa que ele encontrou a mais elevada temperatura dentre todas as medidas anteriormente (IPAC, 2003).

Herschel deu continuidade aos estudos do que ele chamava de raios caloríficos, e constatou que eles eram refletidos, refratados, absorvidos e transmitidos exatamente como a luz visível. Estes raios são o que hoje conhecemos como raios infra-vermelhos. A experiência do astrônomo alemão foi importantíssima, não só pela descoberta dos

raios infra-vermelhos, mas também pelo fato de que, pela primeira vez, alguém havia mostrado que existiam formas de luz que não podiam ser vistas por nossos olhos (IPAC, 2003).

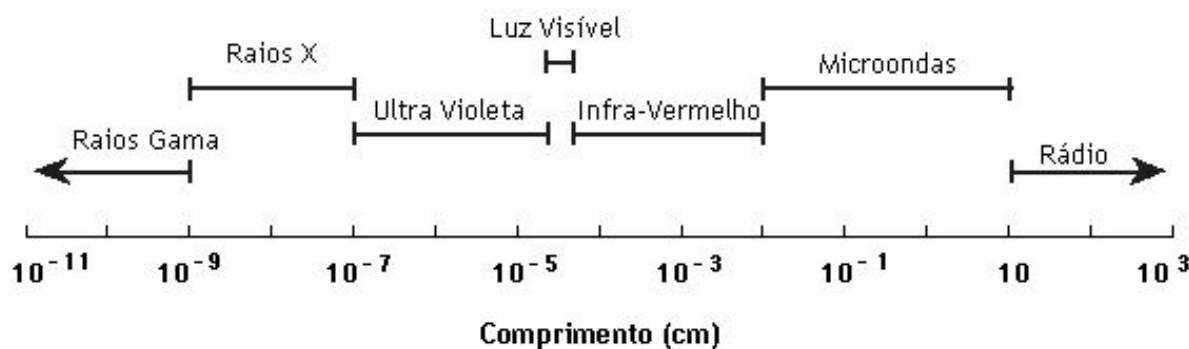


Figura 2.11: Espectro Eletromagnético.

Conforme podemos ver na Figura 2.11, o espectro eletromagnético apresenta diferentes tipos de ondas, cujas classificações são baseadas em seus comprimentos (ou frequências). As ondas do tipo infra-vermelho estão localizadas no espectro entre a luz visível e as microondas. Sua principal fonte é o calor (radiação térmica), produzida pelo movimento de moléculas e átomos de qualquer matéria cuja temperatura não corresponda ao zero absoluto (-273°C), na qual o movimento molecular e atômico cessa. Quanto maior a temperatura, maior será o movimento molecular e atômico, e por consequência, maior será a emissão de infra-vermelho (IPAC, 2003).

Tudo na natureza emite raios infra-vermelhos, e embora não possamos vê-los, podemos senti-los na forma de calor. Para que possamos interpretá-los, é necessário que a captação de infra-vermelhos a partir de um objeto seja transformada em imagem, de forma que as regiões de temperaturas diferentes sejam preenchidas com cores diferentes. E como a emissão de raios infra-vermelhos está relacionada a radiação térmica, se torna possível identificar objetos até sob completa e total escuridão (IPAC, 2003). Nas Figuras 2.12 e 2.13 são exibidos exemplos de operações de resgate com o emprego de imagens obtidas a partir da emissão de raios infra-vermelhos, sendo a primeira referente à identificação de uma pessoa em meio a um quarto repleto de fumaça e o segundo à localização de uma vítima de afogamento noturno.

Algumas aplicações de sensores de raios infra-vermelhos no âmbito de sistemas

de monitoração de processos ou ambientes são listados abaixo:

- Análise e monitoração de condições climáticas e fenômenos meteorológicos;
- Astronomia;
- Inspeções elétricas e mecânicas;
- Inspeção de alimentos;
- Localização e combate a incêndios (EL-SHEIMY & WRIGHT, 2003);
- Militares;
- Vigilância;



Figura 2.12: Vítima em quarto repleto de fumaça.



Figura 2.13: Vítima de afogamento noturno.

2.4.5 Ressonância Magnética

A ressonância magnética é provavelmente o que há de mais moderno hoje em dia na obtenção de imagens do interior de um objeto. Descoberta paralela e independentemente em 1946 pelos cientistas Felix Bloch e Edward Purcell (Universidade de Stanford e Universidade de Harvard respectivamente, ambas nos Estados Unidos), a ressonância magnética multiplicou a capacidade do ser humano de enxergar o recheio volumétrico de objetos com mais acurácia e segurança do que métodos já existentes, como a radiografia e a tomografia computadorizada, eram capazes.

Por esta descoberta, Bloch e Purcell receberam o prêmio Nobel de Física em 1952 (HORNAK, 2003).

A evolução da ressonância magnética, que era inicialmente chamada de ressonância magnética nuclear, o que deixou de ocorrer pela conotação negativa da palavra nuclear na época, é repleta de marcos e inovações que a levaram a seu estado atual. Entre as décadas de 1950 e 1970, a ressonância magnética era aplicada à análise química e física de moléculas. Em 1971, o médico e cientista americano Dr. Raymond Damadian mostrou que existia uma diferença significativa entre os tempos de relaxamento magnético de tecidos normais e tumores, o que motivou a comunidade científica a direcionar esforços à ressonância magnética com o intuito de diagnosticar doenças, reforçando e acelerando pesquisas na área (HORNAK, 2003).

O químico americano Paul Lauterbur foi responsável, em 1973, pela primeira imagem gerada a partir da ressonância magnética, utilizando uma técnica de projeção similar a da radiografia. Em 1975, o químico suíço Richard Ernst propôs um método para geração da imagem baseado na transformada de Fourier, que se tornou base para os métodos utilizados hoje em dia. Em 1977, o Dr. Damadian apresentou ao mundo a primeira imagem obtida através da ressonância magnética do corpo humano inteiro. Damadian e sua equipe levaram 5 horas para obter a primeira imagem. Diversos outros eventos ocorreram no curso dos anos até os dias atuais, e tornaram a ressonância magnética uma das técnicas mais eficientes de obtenção de imagens internas de corpos fechados (HORNAK, 2003).

Antes da apresentação da técnica, é importante observar que a ressonância magnética foi inicialmente desenvolvida com base no átomo de hidrogênio, por ter apenas um próton e um alto momento magnético. Por este motivo, quando submetido a um campo magnético, um átomo de hidrogênio apresenta uma forte tendência a se alinhar com a direção do campo. Fora o hidrogênio, átomos de flúor, alumínio, fósforo e sódio também oferecem bons resultados (HOWSTUFFWORKS I, 2003).

Considere então uma pessoa sendo submetida a um exame de ressonância magnética. Seu corpo é posicionado ao longo de um tubo e submetido a um campo magnético direcionado horizontalmente, partindo do fundo do tubo para a abertura do mesmo. Desta forma, os átomos de hidrogênio das moléculas que compõe os tecidos do paciente serão alinhados na horizontal, seja na direção da cabeça, ou dos pés, conforme indica a Figura 2.14 (HOWSTUFFWORKS I, 2003).

A maioria dos átomos de hidrogênio serão anulados por átomos alinhados em

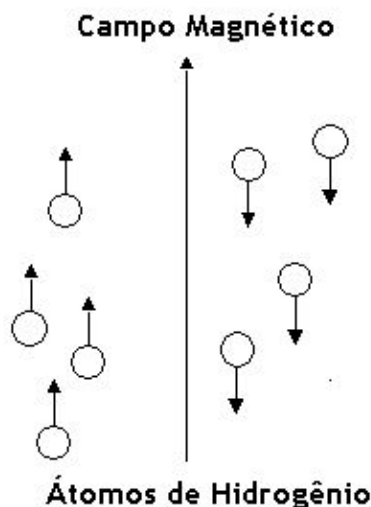


Figura 2.14: Átomos de hidrogênio submetidos a um campo magnético.

sentidos opostos. Porém, isso não ocorre com todos os átomos. Estima-se que a cada um milhão de átomos, apenas dois ou três não sejam cancelados. E embora isso não pareça muito, estes poucos não cancelados são suficientes para que seja possível obter informações extremamente precisas (HOWSTUFFWORKS I, 2003).

O passo seguinte corresponde à aplicação de ondas de rádio específicas para o tipo de átomo em questão (no caso, hidrogênio) sobre a região de interesse do corpo do paciente. Com isso, os átomos não cancelados são forçados a alterar seu movimento de precessão, que corresponde a um percurso orbital do núcleo do átomo em torno de um eixo imaginário. Para tal, estes átomos absorvem energia (HOWSTUFFWORKS I, 2003).

Quando a aplicação de ondas de rádio é interrompida, os átomos não cancelados retornam ao seu estado de alinhamento natural com o campo magnético. Com isso, eles liberam energia, gerando sinais que são enviados para um sistema que, através da transformada de Fourier, os transforma em imagem. A Figura 2.15 apresenta um exemplo de uma imagem da cabeça de uma pessoa gerada por ressonância magnética (HOWSTUFFWORKS I, 2003).

As possibilidades de aplicação da ressonância magnética na monitoração de processos são bastante numerosas. Abaixo segue uma lista de algumas destas possíveis aplicações (SWRI I, 2003):

- Detecção de deterioração de materiais, como óleos, asfalto, alimentos;

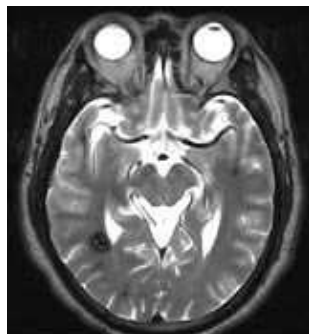


Figura 2.15: Imagem da cabeça de uma pessoa gerada por ressonância magnética.

- Detecção de explosivos e contrabando;
- Determinação de umidade, viscosidade e composição de produtos químicos, alimentícios e outros materiais;
- Determinação de pontos de congelamento, derretimento e evaporação de líquidos específicos;
- Determinação de pressão em gases;
- Monitoração da evolução de doenças;

2.4.6 Ultra-Som

Embora o ultra-som seja uma técnica empregada, em grande parte, apenas em procedimentos médicos, por ser a principal forma de monitoração do processo mais importante de todos, o da gestação, tal abordagem não poderia ser esquecida. Muito similar a ao mecanismo de eco-localização de morcegos e golfinhos e ao SONAR de um submarino, o ultra-som se baseia na emissão e reflexão de ondas sonoras de alta frequência para gerar uma imagem (Figura 2.16) (HOWSTUFFWORKS II, 2003).

A emissão de ondas sonoras no exame de ultra-som é responsabilidade de uma sonda transdutora, que funciona com base no efeito *piezo-elétrico*, descoberto por Pierre e Jacques Curie em 1880. Na sonda, encontram-se um ou mais cristais *piezo-elétricos*, que quando submetidos a uma corrente elétrica, mudam de forma rapidamente. Esta mudança de forma gera ondas sonoras que são propagadas para o exterior do cristal. E estes mesmos cristais, quando atingidos por ondas sonoras,



Figura 2.16: Imagem gerada por ultra-som de um feto com 12 semanas de idade.

emitem corrente elétrica. Desta forma, a sonda transdutora atua tanto como emissora de ondas de som, como receptora de ondas refletidas (HOWSTUFFWORKS II, 2003).

Os processos tradicionalmente monitorados através do ultra-som são a gestação e o fluxo sanguíneo através de órgãos, como o coração e os rins (HOWSTUFFWORKS II, 2003).

O exame de ultra-som consiste dos seguintes passos (HOWSTUFFWORKS II, 2003):

1. Através de uma sonda transdutora, a máquina de ultra-som controla a emissão de ondas de som de alta frequência em direção à região de interesse do exame.
2. As ondas viajam pelo interior do corpo até atingirem regiões limítrofes entre diferentes tipos de tecidos, onde são refletidas.
3. As ondas refletidas são detectadas pela mesma sonda que as emitiu, cuja função agora é passar os dados da reflexão à máquina de ultra-som, que calcula, baseando-se no tempo de retorno e velocidade das ondas, as distâncias das regiões limítrofes entre tecidos.
4. A partir destas distâncias, é gerada uma imagem.

2.5 Pesquisa no Brasil

A monitoração de processos visual em si não é, de forma direta, assunto de grande interesse de pesquisa no Brasil. Geralmente, projetos em andamento relacionados a tal atividade são encontrados sob a tutela de grupos de pesquisa das áreas de Computação Gráfica, Visão Computacional, Robótica e Inteligência Artificial em geral. A Tabela 2.1 apresenta alguns destes grupos, suas respectivas universidades e endereço para contato através da Internet.

Universidade	Grupo	URL
UFRJ (NCE)	LABIC	www.labic.nce.ufrj.br
UFRJ (COPPE)	LCG	www.lcg.ufrj.br
UFRGS	GCGPI	www.inf.ufrgs.br/cg
UFRGS	LRI	www.inf.ufrgs.br/lri
UNICAMP	-	www.ic.unicamp.br/ic-main/research-areas-p.html
PUC-RIO	TECGRAF	www.tecgraf.puc-rio.br

Tabela 2.1: Pesquisa no Brasil.

2.6 Pesquisa no Exterior

No exterior, a monitoração de processos é atividade de pesquisa bastante proeminente em diversas instituições, e de forma mais direta que no Brasil. A Tabela 2.2 apresenta algumas destas instituições.

Instituição	País	URL
SW Research Inst.	EUA	www.swri.edu
Fraunhofer IPM	ALE	www.ipm.fhg.de
Univ. of Kaiserslautern	ALE	ag-vp-www.informatik.uni-kl.de
KTH	SUE	www.kth.se
Univ. Karlsruhe	ALE	wwwipr.ira.uka.de

Tabela 2.2: Pesquisa no Exterior.

Capítulo 3

Processamento de Imagens

3.1 Introdução

Conforme mencionado no Capítulo 1, o processamento de imagens digitais se enquadra em duas grandes categorias: a da melhoria visual da imagem visando à interpretação humana, e a da manipulação de processamento de dados para a análise automática realizada por máquinas.

Tendo sido feitas, no capítulo acima mencionado, todas as apresentações referentes à importância e à contextualização do processamento de imagens digitais no mundo de hoje, este capítulo se concentrará apenas nos aspectos técnicos do assunto. O objetivo é abordar, de forma introdutória e apresentativa, o universo das técnicas de manipulação de imagens, de forma a facilitar a compreensão dos estudos de caso realizados neste trabalho, a serem abordados no Capítulo 4.

3.2 Modelo de Imagem

Conforme foi dito na seção anterior, neste capítulo serão apresentadas diversas técnicas de processamento de imagens digitais, o que implica a definição de um modelo de imagem a ser utilizado no texto. É importante observar que, neste trabalho, não serão consideradas imagens coloridas, ou seja, a abordagem aqui apresentada se restringe a imagens em tons de cinza.

O modelo de imagem empregado corresponde a uma função bidimensional de intensidade luminosa e será denotado por $f(x, y)$. Denota-se por $f(a, b)$, sendo a e b dois valores particulares de x e y , a intensidade (luminância) da imagem no ponto (a, b) . Partindo do princípio que a luz é uma forma de energia (GONZALEZ &

WOODS, 2000), temos

$$0 < f(x, y) < \infty, \forall x, y \quad (3.1)$$

As imagens percebidas por nossos olhos em atividades diárias correspondem a quantidade de luz refletida a partir de objetos e matéria em geral. Por este motivo, $f(x, y)$ pode ser definida em função de dois aspectos chave: a) a luminância (denotada por $i(x, y)$), correspondente à quantidade de luz incidente sobre a cena em questão, e b) a reflectância (denotada por $r(x, y)$), que corresponde à quantidade de luz refletida por seus objetos (GONZALEZ & WOODS, 2000). Define-se $f(x, y)$ da seguinte forma:

$$f(x, y) = i(x, y) \cdot r(x, y) \quad (3.2)$$

Obviamente, os limites apresentados na Equação 3.1 são apenas teóricos. Geralmente utiliza-se 0 (zero) como valor mínimo para $f(x, y)$ e um L qualquer como valor máximo, dependendo da forma de representação da imagem.

3.2.1 Amostragem e Quantização

Dois processos específicos devem ser abordados antes das técnicas de processamento de imagens propriamente ditas, sendo eles a amostragem e a quantização, necessários para adaptação do modelo de imagem definido acima ao paradigma computacional de análise.

A amostragem corresponde a restrições espaciais impostas sobre a imagem, de forma que

$$0 \leq x < H, \text{ onde } H \text{ representa a largura da imagem e} \quad (3.3)$$

$$0 \leq y < W, \text{ onde } W \text{ representa a altura da imagem.} \quad (3.4)$$

Desta forma, temos a imagem representada por uma matriz com o seguinte formato:

$$\begin{bmatrix} f(0, 0) & f(0, 1) & \cdots & f(0, H-1) \\ f(1, 0) & f(1, 1) & \cdots & f(1, H-1) \\ \vdots & \vdots & & \vdots \\ f(W-1, 0) & f(W-1, 1) & \cdots & f(W-1, H-1) \end{bmatrix}$$

A quantização corresponde à discretização dos níveis de cinza da imagem, e está diretamente ligada ao número de *bits* utilizados para representar cada *pixel* da imagem. Para cada *pixel* representado por n *bits*, são possíveis 2^n tons de cinza diferentes (Equação 3.5). Em todos os experimentos deste trabalho, foram usadas imagens de 8 *bits*.

$$f(x, y) \in \{0, 1, 2, 3, 4, \dots, 2^n\} \quad \forall x, y \mid 0 \leq x < H \text{ e } y < W. \quad (3.5)$$

A amostragem define a imagem em um número finito e definido de pontos, enquanto a quantização faz o mesmo para o conjunto de cores.

3.3 Relações Entre Pixels

Nesta seção, serão apresentados alguns importantes tipos de relações entre *pixels* de uma mesma imagem. São eles: vizinhança, conectividade e distância. Serão também estudadas algumas operações lógico-aritméticas aplicadas a imagens.

3.3.1 Vizinhança

A relação de vizinhança entre *pixels* é determinada pelo posicionamento dos *pixels* de uma imagem. São 3 os principais tipos de vizinhança entre *pixels* de uma mesma imagem. São elas: vizinhança-de-4, vizinhança-diagonal, e vizinhança-de-8 (GONZALEZ & WOODS, 2000).

- **Vizinhança-de-4:** os vizinhos de 4 de um *pixel* p qualquer correspondem aos *pixels* vertical e horizontalmente adjacentes a p . Considerando então $p = (x, y)$, o conjunto dos vizinhos de 4 de p , denotado por $N_4(p)$, seria formado por $(x - 1, y)$, $(x + 1, y)$, $(x, y - 1)$ e $(x, y + 1)$.
- **Vizinhança-diagonal:** analogamente, os vizinhos diagonais de p são aqueles diagonalmente adjacentes ao mesmo. Sendo assim, $N_D(p)$ consistiria de $(x - 1, y - 1)$, $(x - 1, y + 1)$, $(x + 1, y - 1)$ e $(x + 1, y + 1)$.
- **Vizinhança-de-8:** o conjunto de vizinhos de 8 de p corresponde a união entre os conjuntos $N_4(p)$ e $N_D(p)$. Sendo assim, $N_8(p) = N_4(p) \cup N_D(p)$.

Considere as matrizes apresentadas na Figura 3.1 como o mesmo trecho de uma imagem qualquer. À esquerda, vemos os valores dos *pixels*, enquanto, à direita, temos A, B, \dots, H, I como identificadores dos mesmos em relação a suas posições na imagem. A Tabela 3.1 mostra os conjuntos N_4, N_D e N_8 do *pixel* representado pela letra E .

1	0	0	0	1	0	0	0
0	0	1	0	A	B	C	0
1	1	0	0	D	E	F	0
0	1	0	0	G	H	I	0

Figura 3.1: Trecho de uma imagem.

$N_4(E)$	$\{B, D, F, H\}$
$N_D(E)$	$\{A, C, G, I\}$
$N_8(E)$	$\{A, B, C, D, F, G, H, I\}$

Tabela 3.1: Conjuntos $N_4(E), N_D(E)$ e $N_8(E)$.

3.3.2 Conectividade

A conectividade entre *pixels* é um conceito bastante útil, aplicado especialmente na detecção de bordas ou regiões homogêneas em uma imagem (GONZALEZ & WOODS, 2000), e está diretamente relacionada a dois fatores: a) a relação de vizinhança entre dois *pixels* e b) algum critério de similaridade S entre os valores (níveis de cinza) dos mesmos. São 4 os tipos de conectividade entre *pixels*: conectividade-de-4, conectividade-diagonal, conectividade-de-8 e conectividade-de- m .

- **Conectividade-de-4:** considerando p e q dois *pixels*, os mesmos são ditos conectados-de-4 se atendem a S e $q \in N_4(p)$.
- **Conectividade-diagonal:** análogo à conectividade-de-4, sendo que q deve pertencer a $N_D(p)$.
- **Conectividade-de-8:** idem ao anterior, trocando-se $N_D(p)$ por $N_8(p)$.

- **Conectividade-de- m :** sendo p e q dois *pixels* de uma imagem, eles são conectados-de- m se $q \in N_4(p)$ ou se $q \in N_D(p)$ e $\forall v \in N_4(q)$, v não está conectado-de-4 com p .

Aproveitando o exemplo da Figura 3.1, e sendo $S = \{p = q\}$, temos que E está conectado-de-4 com D e H , conectado-diagonalmente com C , conectado-de-8 com C , D e H e conectado-de- m com C , D e H . Repare como B e F são vizinhos-de-4 de C , porém não estão conectados-de- m com E .

3.3.3 Distância

Outra relação interessante entre *pixels* de uma imagem é a distância. Sejam p , q e t pixels com coordenadas (x_1, y_1) , (x_2, y_2) e (x_3, y_3) respectivamente. Diz-se que M é uma medida de distância se:

1. $M(p, q) \geq 0$.
2. $M(p, q) = 0$ se e somente se $p = q$.
3. $M(p, q) = M(q, p)$.
4. $M(p, t) \leq M(p, q) + M(q, t)$.

Alguns tipos de medidas de distância são a Euclidiana, a D_4 , também conhecida como distância de *Manhattan*, e a D_8 , conhecida como distância de xadrez.

- **Distância Euclidiana:** $D_E(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$
- **Distância de *Manhattan*:** $D_4(p, q) = |(x_1 - x_2)| + |(y_1 - y_2)|$
- **Distância de Xadrez:** $D_8(p, q) = \max(|(x_1 - x_2)|, |(y_1 - y_2)|)$,
onde $\max(a, b)$ corresponde ao maior valor entre a e b .

Repare que a distância de *Manhattan* entre p e q corresponde ao comprimento do menor caminho possível entre os dois *pixels* não sendo permitida a transição diagonal entre *pixels*. Já a distância de xadrez é o comprimento do menor caminho permitindo-se inclusive a transição diagonal.

3.4 Operações Lógico-Aritméticas

As operações lógico-aritméticas são extensivamente aplicadas em diversos ramos de processamento de imagens. As mesmas são, em geral, aplicadas *pixel a pixel* ou entre *pixels* de imagens distintas.

Operações Aritméticas

As operações aritméticas aplicadas a imagens são a soma, a subtração, a multiplicação e a divisão.

- **Soma:** geralmente utilizada entre duas ou mais imagens para, através da média, reduzir ruídos gerados pelo equipamento de captura de imagens. Observe as Figuras 3.2.a e 3.2.b e repare nos ruídos em linhas de cor cinza. Calculando-se a média entre estas imagens é possível notar que os ruídos foram bastante atenuados, conforme é mostrado na Figura 3.2.c.

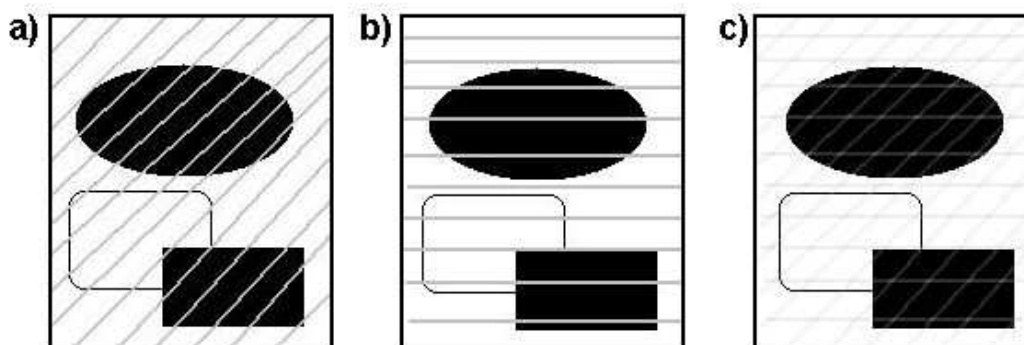


Figura 3.2: Exemplo da aplicação da operação de soma para redução de ruído.

- **Subtração:** muito utilizada em aplicações de detecção de movimento. A Figura 3.3.a apresenta uma região em repouso. Na Figura 3.3.b vemos um veículo atravessando a mesma região. Na Figura 3.3.c vemos o objeto em movimento destacado da região em questão.
- **Multiplicação e Divisão:** geralmente utilizadas para correção de sombras nos níveis de cinza de uma imagem, provenientes de não-uniformidades no sensor de aquisição ou na iluminação local (GONZALEZ & WOODS, 2000). Na Figura 3.4.c, vê-se um exemplo de aplicação da multiplicação ponto a ponto



Figura 3.3: Exemplo da aplicação da operação de subtração para detecção de movimento.

de matrizes entre as Figuras 3.4.a e 3.4.b, sendo o objetivo atenuar regiões da imagem que forneçam pouca informação.

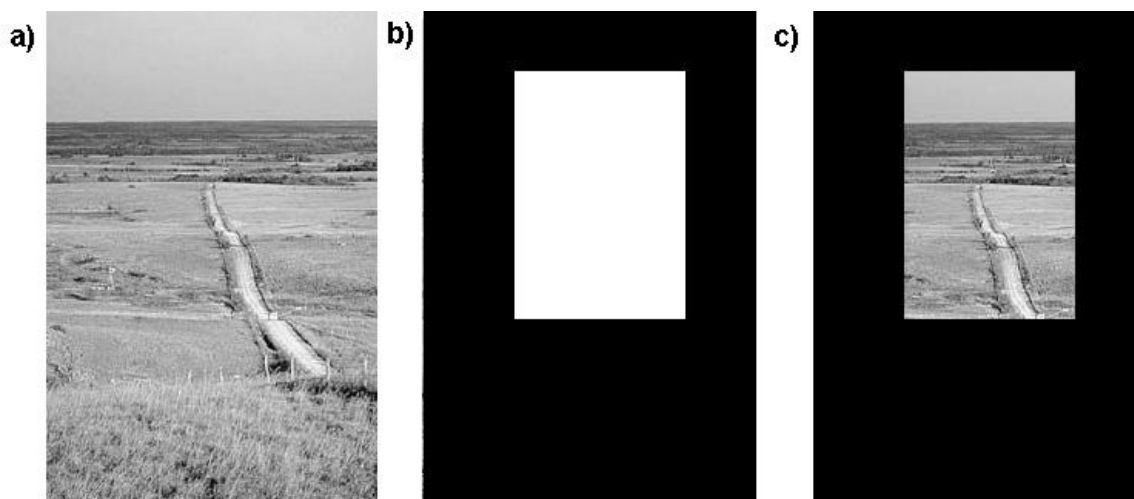


Figura 3.4: Exemplo da aplicação da operação de multiplicação para destacar uma determinada região em uma imagem.

Operações Lógicas

Operações lógicas são empregadas no tratamento de imagens binárias, de forma que o valor de cada *pixel* corresponde a 0 (zero) ou 1 (um). São utilizadas em aplicações de detecção de características e análise de forma de corpos. Dentre as possíveis operações lógicas aplicadas a imagens estão o **OR**, o **AND**, o **XOR**, o **NOT** e outras mais. As Figuras 3.5.c e 3.5.d apresentam resultados da aplicação de algumas operações lógicas nas Figuras 3.5.a e 3.5.b.

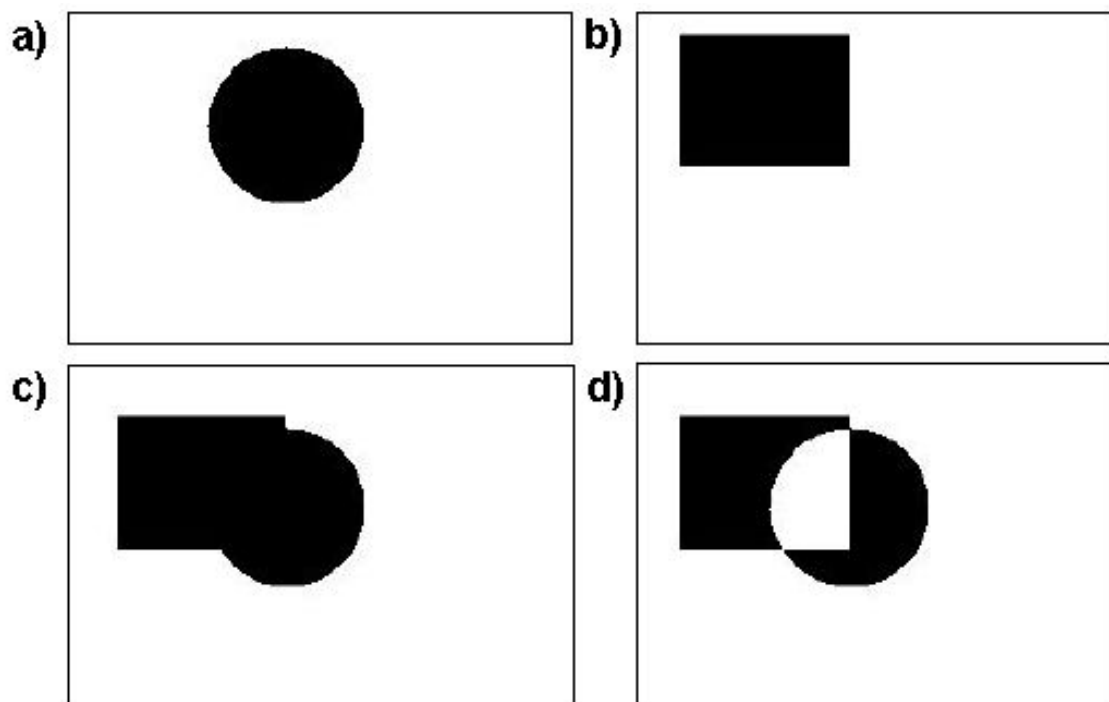


Figura 3.5: Exemplos de operações lógicas.

- a) Conjunto A.
- b) Conjunto B.
- c) A **OR** B.
- d) A **XOR** B.

3.4.1 Transformações Geométricas

As transformações geométricas atuam sobre a distribuição dos *pixels* de uma imagem, modificando-a de acordo com a operação em questão. Dos diversos tipos de transformações conhecidas, serão abordadas, neste trabalho, o escalamento e a rotação de imagens.

Escalamento

O escalamento corresponde aos processos de ampliação ou redução das dimensões de uma imagem. Nos métodos aqui abordados, o valor da luminância de um *pixel* na imagem destino é calculado através de uma interpolação de *pixels* da imagem origem. Existem diversos métodos de interpolação, sendo três dos mais utilizados o do vizinho mais próximo (*nearest neighbor*), o bilinear e o bicúbico.

Vizinho Mais Próximo Dos três métodos a serem abordados, este é o único que não altera o valor da luminância dos *pixels*. Sua implementação é simples, dado que o mesmo se baseia na relação espacial direta entre a imagem origem e a imagem destino, ou seja, aplica-se a transformação de escalamento em cada *pixel* da imagem original e calcula-se a posição correspondente do mesmo na imagem destino. Copia-se então o valor da luminância do *pixel* da imagem original para seu correspondente na imagem destino (WINKLER & THOMÉ I, 2003). É importante observar que ao aplicar a transformação de escalamento em um *pixel* torna-se possível que as coordenadas da posição calculada na imagem destino sejam dadas por números não inteiros. Quando isto ocorre, arredonda-se os mesmos para os valores inteiros mais próximos. É importante ressaltar que as operações de escalamento não são, em geral, sobrejetoras, e, por este motivo, podem deixar descontinuidades na imagem destino. Sendo assim, realiza-se esta operação na direção inversa, ou seja, da imagem destino para a imagem origem, o que soluciona por completo o problema mencionado. A Figura 3.7.b mostra o resultado da aplicação deste método no escalamento em 150 % da Figura 3.7.a.

Bilinear Seja p um *pixel* da imagem destino. Utilizando a interpolação bilinear, o valor da luminância de p é calculado da seguinte forma: aplica-se a transformada de escalamento inversa a p e calcula-se uma média ponderada das luminâncias dos 4 *pixels* mais próximos (ver Figura 3.6) (WINKLER & THOMÉ I, 2003) (JIANG, 2004).

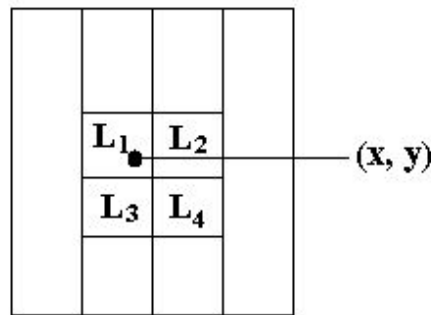


Figura 3.6: Região da imagem original de um escalamento bilinear.

Considere (x, y) como as coordenadas da transformação inversa aplicada ao *pixel* p de coordenadas (z, t) na imagem destino e L_1 , L_2 , L_3 e L_4 os valores das

luminâncias dos *pixels* vizinhos de coordenadas (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) , respectivamente. Sendo assim, para calcular $f(z, t)$ (JAIN, 1989), ter-se-ia

$$dx_1 = \frac{|y - y_1|}{|y_3 - y_1|} \text{ e } dy_3 = \frac{|y - y_3|}{|y_3 - y_1|}$$

$$l_1 = dy_1 L_3 + dy_3 L_1$$

$$dy_2 = \frac{|y - y_2|}{|y_4 - y_2|} \text{ e } dy_4 = \frac{|y - y_4|}{|y_4 - y_2|}$$

$$l_2 = dy_2 L_4 + dy_4 L_2$$

$$dx_1 = \frac{|x - x_1|}{|x_2 - x_1|} \text{ e } dx_2 = \frac{|x - x_2|}{|x_2 - x_1|}$$

$$f(z, t) = l_1 \, dy_2 + l_2 \, dy_1$$

A Figura 3.7.c mostra o resultado da aplicação deste método no escalamento em 150 % da Figura 3.7.a.

Bicúbico Considerando (x', y') como as coordenadas não necessariamente inteiras provenientes da transformação inversa ao escalamento desejado e (x, y) o ponto mais próximo a (x', y') , o método bicúbico calcula o valor da luminância de (z, t) através de uma média ponderada entre os 16 *pixels* mais próximos a (x, y) . Tem-se que $d_x = x' - x$ e $d_y = y' - y$. O valor de $f(z, t)$ é calculado conforme indica a Equação 3.6 (BOURKE, 2004).

$$f(z, t) = \sum_{i=-1}^2 \sum_{j=-1}^2 f(x + i, y + j) R(i - d_x) R(j - d_y) \quad (3.6)$$

$$R(x) = \frac{1}{6} [P(x + 2)^3 - 4.P(x + 1)^3 + 6.P(x)^3 - 4.P(x - 1)^3]$$

$$P(x) = \begin{cases} x & , \text{ se } x > 0 \\ 0 & , \text{ se } x \leq 0 \end{cases}$$

O método de escalamento que emprega a interpolação bicúbica suaviza as bordas dos objetos na imagem. Por este motivo, é indicado para aplicações de interpretação visual e ampliação de produtos fotográficos (WINKLER & THOMÉ I, 2003). A Figura 3.7.d mostra o resultado da aplicação deste método no escalamento em 150 % da Figura 3.7.a.

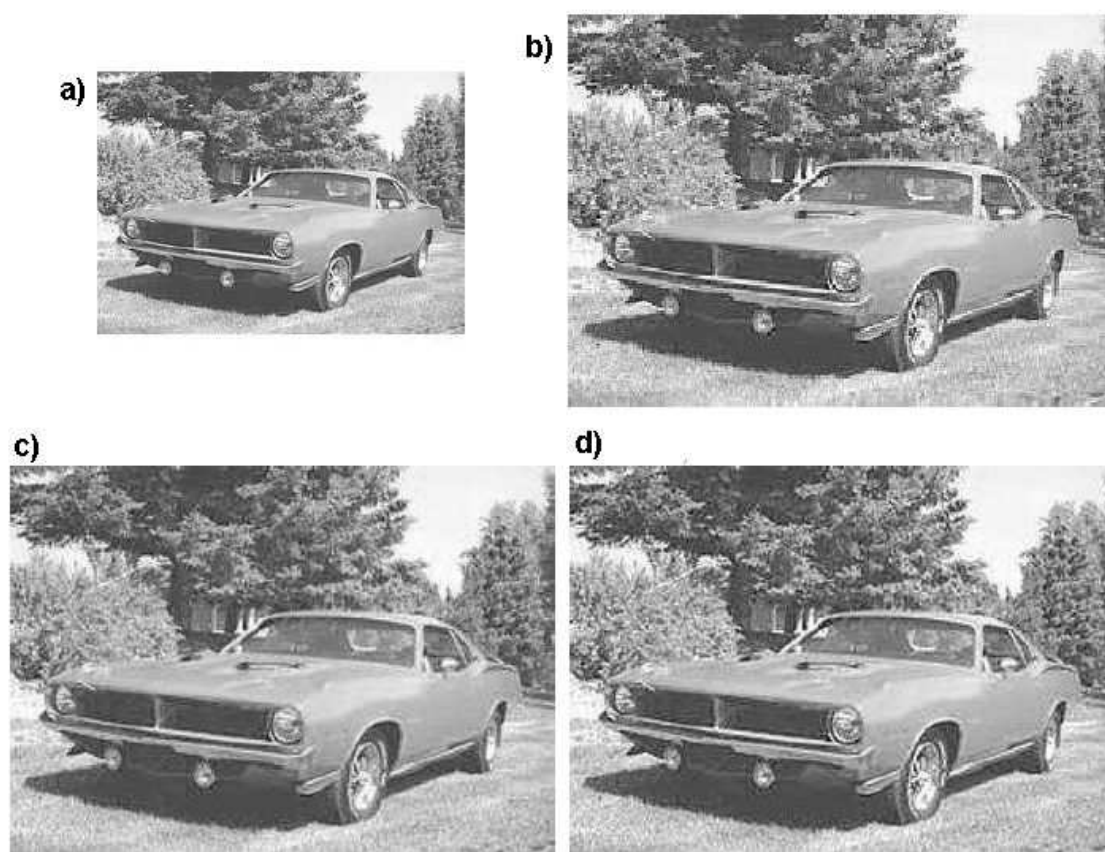


Figura 3.7: Exemplos de operações de escalamento, utilizando os métodos de interpolação b) vizinho mais próximo, c) bilinear e d) bicúbico.

Rotação

A rotação, que embora tenha maior aplicabilidade na computação gráfica, devido a forma de representação do movimento de objetos, dada pela variação da posição e orientação dos mesmos (GOMES & VELHO, 1998), representa, em processamento

de imagens, uma das mais importantes operações geométricas, especialmente em aplicações de processamento eletrônico de formulários e visão computacional.

Considere (x, y) um *pixel* de uma imagem I . Sejam I_θ o resultado da rotação de I por θ radianos e (x', y') as coordenadas correspondentes a (x, y) em I_θ . Desta forma, tem-se:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3.7)$$

Aplicando-se esta operação *pixel* a *pixel* sobre I e copiando-se o valor da luminância dos mesmos em seus correspondentes em I_θ , obtém-se a imagem rotacionada. A desvantagem de se utilizar este procedimento é que a imagem rotacionada pode apresentar descontinuidades, conforme mostra a Figura 3.8.b, que apresenta o resultado da rotação anti-horária por $\frac{\pi}{20}$ radianos da Figura 3.8.a. Sendo assim, aplica-se o procedimento inverso, varrendo-se a imagem rotacionada e calculando-se os correspondentes de seus *pixels* na imagem original através da matriz inversa àquela apresentada na Equação 3.7. Com o emprego deste método, torna-se necessário recorrer a alguma estratégia de interpolação, sendo as mais utilizadas aquelas mencionadas anteriormente neste trabalho: vizinho mais próximo, bilinear e bicúbica. O resultado da aplicação desta variação é mostrado na Figura 3.8.c. O método de interpolação utilizado foi o do vizinho mais próximo.

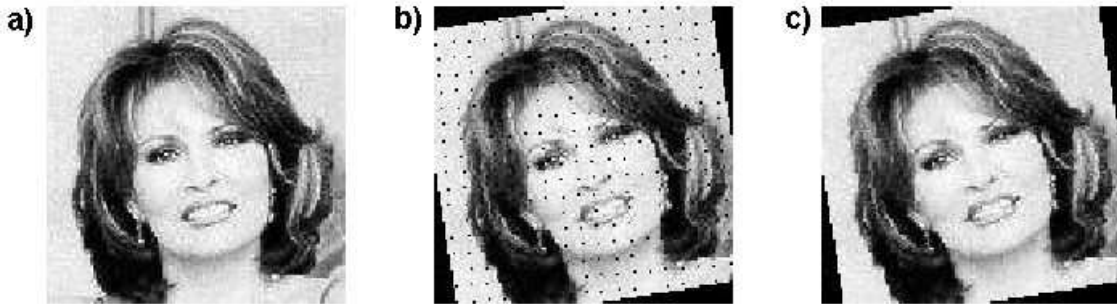


Figura 3.8: Exemplos de operações de rotação.

3.5 Transformações Radiométricas

As transformações radiométricas atuam sobre os *pixels* de uma imagem independentemente da posição dos mesmos, e alteram, de forma direta, o valor da luminância

destes *pixels*. Suas duas grandes aplicações em processamento de imagens correspondem ao realce de contraste ou detalhes, onde tenta-se salientar a distinção visual entre dois ou mais objetos em uma imagem, e a restauração, na qual tenta-se corrigir alguma distorção sofrida pela imagem (WINKLER & THOMÉ I, 2003). Neste trabalho, do conjunto de transformações radiométricas existentes, são abordadas três cujo objetivo direto é modificar o histograma de uma imagem, sendo elas a expansão e a equalização do mesmo, além da limiarização simples. Antes de mostrá-las, porém, apresenta-se brevemente o conceito de histograma.

3.5.1 Histograma de uma Imagem

O histograma corresponde a uma representação dos níveis de cinza contidos em uma imagem, podendo ser visto como uma função de densidade dos mesmos na imagem (WINKLER & THOMÉ I, 2003). Para uma imagem codificada em n bits, teria-se 2^n níveis de cinza possíveis, variando de 0 (zero) a $2^n - 1$. A construção do histograma é constituída do estabelecimento de uma relação entre cada nível de cinza possível e o número de vezes que ele aparece na imagem. A Figura 3.10 mostra o histograma da Figura 3.9, codificada em 8 bits.



Figura 3.9: Imagem base para construção do histograma.

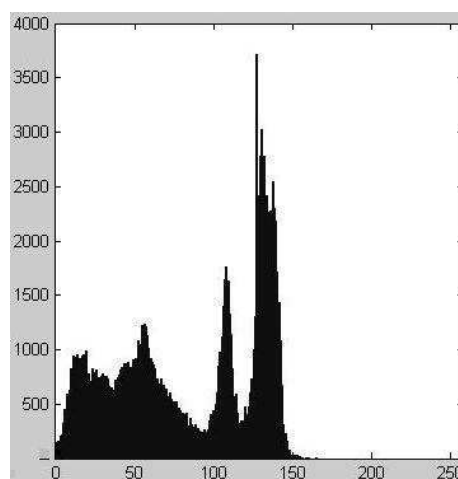


Figura 3.10: Histograma da imagem na Figura 3.9.

3.5.2 Limiarização Simples

A primeira técnica de modificação do histograma de uma imagem a ser apresentada neste trabalho é a limiarização simples, também conhecida como binarização, devido ao fato de que todos os *pixels* da imagem (representada em n bits) sobre a qual tal técnica é aplicada tem os valores de suas luminâncias modificados para 0 (zero) ou $2^n - 1$.

Inicialmente, estabelece-se um *threshold* θ , de forma que $0 \leq \theta \leq 2^n - 1$. Em seguida, são percorridos todos os *pixels* da imagem, denotada por f , aplicando a seguinte regra aos mesmos:

$$f(x, y) = \begin{cases} 0 & , \text{ se } f(x, y) \leq \theta \\ 2^n - 1 & , \text{ se } f(x, y) > \theta \end{cases}$$

Conforme pode-se ver na Figura 3.10, os níveis de cinza da Figura 3.9 estão todos concentrados entre 0 (zero) e 150 (cento e cinquenta). Na Figura 3.11, mostra-se o resultado da limiarização aplicada à Figura 3.9 com *threshold* igual a 50 (cinquenta).



Figura 3.11: Limiarização com $\theta = 50$ aplicada à Figura 3.9.

3.5.3 Expansão do Histograma

A técnica de expansão do histograma é baseada no fato de que quanto mais espalhados estiverem os níveis de cinza, maior é a discriminação visual entre os objetos de uma imagem (WINKLER & THOMÉ I, 2003).

Considere l a luminância do *pixel* (x, y) de uma imagem codificada em n bits. Sejam l_{min} , l_{max} e L o menor e maior valor de luminância contido na imagem, e a

máxima luminância possível ($2^n - 1$), respectivamente. Aplica-se a seguinte regra ao *pixel* (x, y) :

$$f(x, y) = \frac{l - l_{min}}{l_{max} - l_{min}} (L) \quad (3.8)$$

A técnica consiste da aplicação da regra mostrada na Equação 3.8 a todos os *pixels* da imagem. Na Figuras 3.12 e 3.13, vê-se, respectivamente, o resultado da expansão do histograma na Figura 3.9 e o histograma computado posteriormente ao processo de expansão. Nota-se que os níveis de cinza se encontram mais espalhados ao longo do domínio do histograma, e que o contraste entre as diferentes regiões da imagem foi ressaltado.



Figura 3.12: Imagem com histograma expandido.

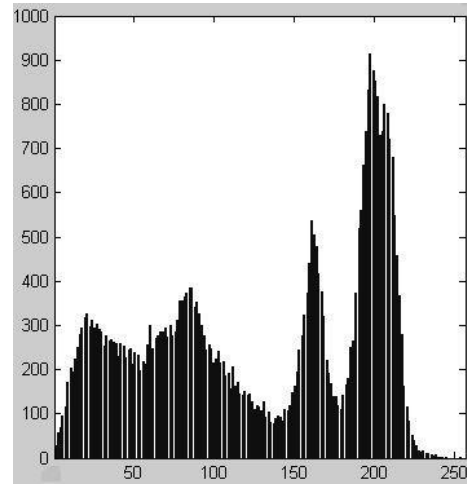


Figura 3.13: Histograma da imagem na Figura 3.12.

3.5.4 Equalização do Histograma

Outra técnica de modificação do histograma bastante utilizada é a equalização, que corresponde à tentativa de modificar a luminância dos *pixels* de uma determinada imagem para alterar seu histograma de forma que o mesmo se aproxima de uma distribuição uniforme.

Seja L o maior nível de cinza possível. Considerando r_k um determinado nível de cinza ($0 \leq r_k \leq L$), n_k o número de ocorrências do mesmo e n o número total de *pixels* da imagem, estabelece-se a seguinte relação:

$$p(r_k) = \frac{n_k}{n}$$

Tem-se em $p(r_k)$ uma função de densidade de probabilidade dos níveis de cinza da imagem. Constrói-se então a função distribuição $F(r_k)$, a qual é usada para mapear as luminâncias da imagem original naquelas da imagem equalizada. Sendo s_k um nível de cinza da imagem equalizada, tem-se:

$$s_k = F(r_k) = \sum_{i=0}^k \frac{n_i}{n} = \sum_{i=0}^k p(r_i) \quad (3.9)$$

Na Figura 3.14, temos o resultado da aplicação desta técnica na Figura 3.9, enquanto na Figura 3.15 é mostrado o histograma após a equalização. Pode-se ver que embora não seja exatamente uniforme, o histograma resultante da aplicação da equalização tem regiões com comportamento aproximadamente constante.



Figura 3.14: Imagem com histograma equalizada.

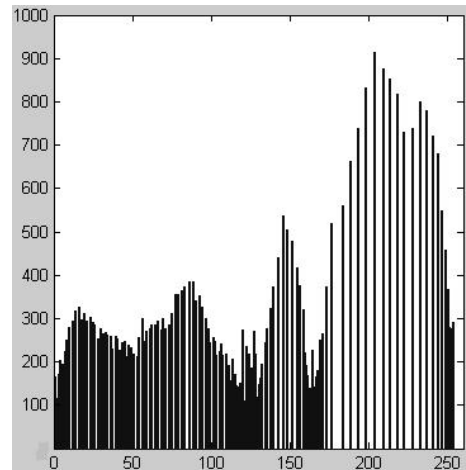


Figura 3.15: Histograma da imagem na Figura 3.14.

3.6 A Transformada de Fourier

Em processamento de imagens, lança-se mão freqüentemente de ferramentas matemáticas para realização de determinadas tarefas. Nesta seção, será apresentada uma das personagens mais importantes deste ferramental matemático: a transformada de Fourier. Embora o interesse prático deste trabalho resida apenas no caso

discreto e bidimensional, são apresentados os casos contínuo e discreto, uni e bidimensional com o objetivo de enriquecer e melhorar o entendimento do assunto.

3.6.1 Caso Contínuo

Considere $f(x)$ uma função contínua e integrável de uma variável real x . Denota-se a transformada de Fourier de $f(x)$ por $\mathfrak{F}(f(x))$. Sendo $j = \sqrt{-1}$, tem-se:

$$\mathfrak{F}(f(x)) = F(u) = \int_{-\infty}^{\infty} f(x) e^{-j2\pi ux} dx \quad (3.10)$$

A inversa da transformada de Fourier pode ser expressa da seguinte forma:

$$\mathfrak{F}^{-1}(F(u)) = f(x) = \int_{-\infty}^{\infty} F(u) e^{j2\pi ux} dx \quad (3.11)$$

O chamado “par de transformadas de Fourier”, formado pelas Equações 3.10 e 3.11, existe para $F(u)$ integrável e, conforme dito anteriormente, $f(x)$ contínua e integrável (WOLFRAM I, 2004; GONZALEZ & WOODS, 2000).

Também é comum se encontrar a transformada de Fourier expressa da seguinte forma (IORIO & IORIO, 2001):

$$\mathfrak{F}(f(x)) = F(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-j\omega x} dx$$

A transformada de Fourier é uma função complexa, conforme mostra a Equação 3.12.

$$F(u) = R(u) + jI(u) \quad (3.12)$$

Sendo $R(u)$ a componente real da transformada e $I(u)$ a componente imaginária, o *espectro* de Fourier $|F(u)|$, que é, na verdade, uma função de magnitude da transformada, é calculado da seguinte forma:

$$|F(u)| = \sqrt{R^2(u) + I^2(u)} \quad (3.13)$$

Para exemplificar a transformada de Fourier, considere a função $f(x) = c$ ($0 \leq x \leq T$), sendo c e T números reais. A Figura 3.16 mostra o gráfico de $f(x)$.

O cálculo da transformada de Fourier para $f(x)$ é realizado da seguinte forma:

$$F(u) = \int_0^T c e^{-j2\pi ux} dx$$

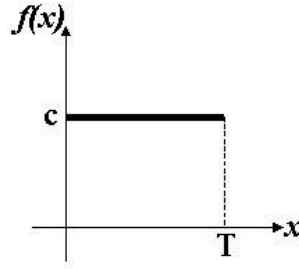


Figura 3.16: Função $f(x) = c$, definida no intervalo $[0, T]$.

$$\begin{aligned}
 &= \frac{-c}{j2\pi u} [e^{-j2\pi ux}]_0^T \\
 &= \frac{-c}{j2\pi u} [e^{-j2\pi uT} - 1] \\
 &= \frac{c}{j2\pi u} [e^{j\pi uT} - e^{-j\pi uT}] e^{-j\pi uT} \\
 &= \frac{c}{j2\pi u} [\cos(\pi uT) + j \sin(\pi uT) - \cos(\pi uT) - j \sin(\pi uT)] e^{-j\pi uT} \\
 &= \frac{c}{j2\pi u} [j2 \sin(\pi uT)] e^{-j\pi uT} \\
 &= \frac{c}{\pi u} [\sin(\pi uT)] e^{-j\pi uT}
 \end{aligned}$$

O *espectro* de Fourier $|F(u)|$ para a função $f(x)$ é então calculado conforme demonstrado abaixo:

$$|F(u)| = \left| \frac{c}{\pi u} \right| |\sin(\pi uT)| |e^{-j\pi uT}|$$

Sabendo-se que $|e^{-j\pi uT}| = |\cos(\pi uT) - j \sin(\pi uT)|$

$$= \sqrt{\cos^2(\pi uT) + \sin^2(\pi uT)} = 1, \text{ tem-se que}$$

$$|F(u)| = c \left| \frac{\sin(\pi uT)}{\pi u} \right|$$

Multiplicando-se $|F(u)|$ por $\frac{T}{T}$, obtém-se

$$|F(u)| = cT \left| \frac{\sin(\pi uT)}{\pi uT} \right|$$

A Figura 3.17 mostra o gráfico do *espectro* de Fourier da função $f(x)$.

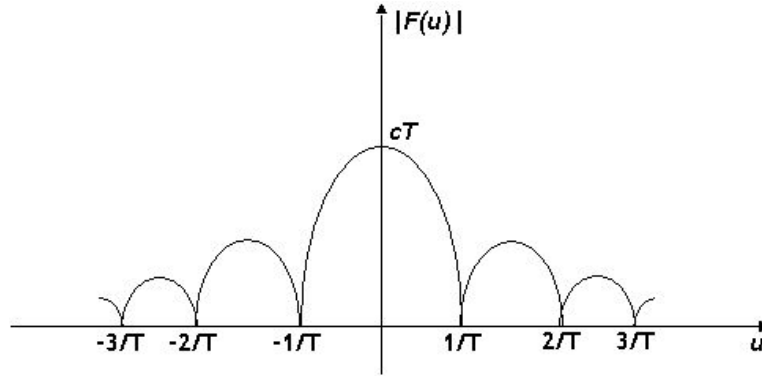


Figura 3.17: Espectro de Fourier da função $f(x) = c$, definida no intervalo $[0, T]$.

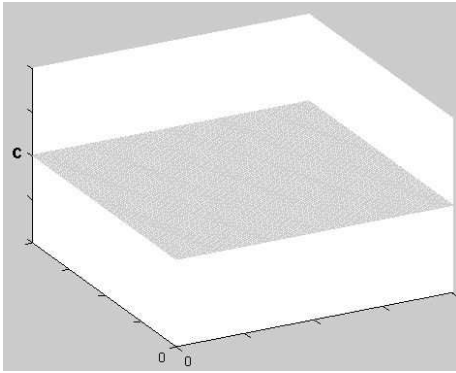


Figura 3.18: Gráfico da função $f(x, y) = c$.

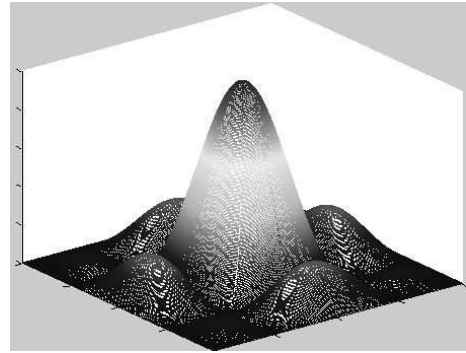


Figura 3.19: Espectro de Fourier da função $f(x, y) = c$.

É bastante simples a extensão da transformada de Fourier para o caso contínuo e bidimensional. As Equações 3.14 e 3.15 mostram, respectivamente, as transformadas direta e inversas para este caso (GONZALEZ & WOODS, 2000).

$$\mathfrak{F}(f(x, y)) = F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (3.14)$$

$$\mathfrak{F}^{-1}(F(u, v)) = f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} dx dy \quad (3.15)$$

O cálculo da transformada de Fourier para a função $f(x, y) = c$, ($0 \leq x \leq T$ e $0 \leq y \leq R$), cujo gráfico é mostrado na Figura 3.18, assim como o de seu *espectro*, é análogo ao demonstrado anteriormente, e o resultado é demonstrado a seguir:

$$\mathfrak{S}(f(x, y)) = F(u, v) = cTR \left[\frac{\text{sen}(\pi u T) e^{-j\pi u T}}{(\pi u T)} \right] \left[\frac{\text{sen}(\pi v R) e^{-j\pi v R}}{(\pi v R)} \right]$$

$$|F(u, v)| = cTR \left| \frac{\text{sen}(\pi u T)}{\pi u T} \right| \left| \frac{\text{sen}(\pi v R)}{\pi v R} \right|$$

3.6.2 Caso Discreto

Seja $f(x)$ uma função discreta e uniformemente espaçada (Δx) em N pontos.

$$f(x) = \{f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \dots, f(x_0 + [N - 1]\Delta x)\}$$

Sabendo-se que $x = 0, 1, 2, \dots, n - 1$, define-se $f(x) = f(x_0 + x \Delta x)$.

O par de transformadas de Fourier para $f(x)$ (direta e inversa, respectivamente) é mostrado nas Equações 3.16 e 3.17 (GONZALEZ & WOODS, 2000).

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N}, \text{ para } u = 0, 1, 2, \dots, n - 1 \quad (3.16)$$

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) e^{j2\pi ux/N}, \text{ para } x = 0, 1, 2, \dots, n - 1 \quad (3.17)$$

Em duas dimensões, teria-se $f(x, y)$ como uma função discreta e uniformemente espaçada (em Δx no eixo x e Δy no eixo y), além de $x = 0, 1, 2, \dots, M - 1$ e $y = 0, 1, 2, \dots, N - 1$.

O par de transformadas para $f(x, y)$ pode ser expresso da seguinte forma:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (3.18)$$

para $u = 0, 1, 2, \dots, m - 1$ e $v = 0, 1, 2, \dots, n - 1$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u, v) e^{-j2\pi(ux/M + vy/N)} \quad (3.19)$$

para $x = 0, 1, 2, \dots, M - 1$ e $y = 0, 1, 2, \dots, N - 1$

O cálculo do *espectro* de Fourier para o caso discreto é análogo ao do caso contínuo, seja para funções uni ou bidimensionais.

3.6.3 Compressão do Espectro

As aplicações da transformada de Fourier em processamento de imagens são inúmeras, sendo uma das mais importantes a análise visual do seu *espectro*. Porém, o intervalo de valores da transformada pode ser muito extenso, comparado com o intervalo de valores que pode ser representado em 8 *bits*. Sendo assim, aplica-se técnicas de compressão nos valores do *espectro* para uma melhor visualização do mesmo. A Equação 3.20 mostra uma dessas técnicas.

$$S(u, v) = c \log[1 + |F(u, v)|], \text{ sendo } c \text{ uma constante multiplicativa} \quad (3.20)$$

A aplicação do *log* nos valores do *espectro* reduz consideravelmente o intervalo em que os mesmos se encontram. Sendo assim, torna-se possível mapear o *espectro* para uma imagem de 8 *bits* de forma a se obter um nível de detalhamento visual mais satisfatório (GONZALEZ & WOODS, 2000). As Figuras 3.21 e 3.22 mostram, respectivamente, a visualização do *espectro* de Fourier da imagem na Figura 3.20 sem e com compressão, respectivamente. A diferença entre os *espectros* quanto a seus níveis de detalhamento é claramente perceptível.



Figura 3.20: Imagem base para visualização do *espectro*.

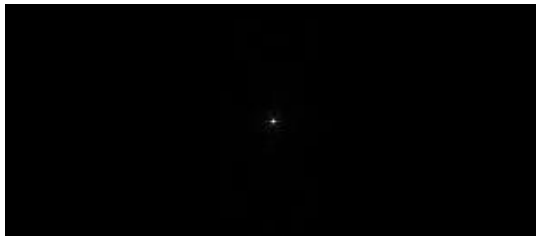


Figura 3.21: *Espectro* da imagem na Figura 3.20 sem compressão de intervalo.

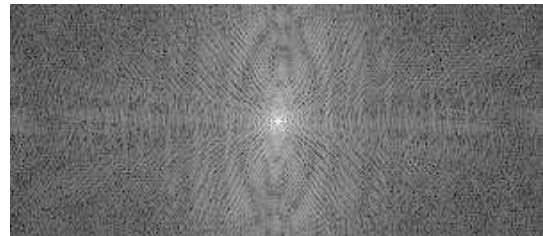


Figura 3.22: *Espectro* da imagem na Figura 3.20 com compressão de intervalo.

3.6.4 A Transformada Rápida de Fourier

O cálculo direto da transformada de Fourier discreta unidimensional (Equação 3.16) é uma operação bastante custosa ($O(N^2)$) em termos computacionais (GONZALEZ & WOODS, 2000). A transformada rápida de Fourier (*FFT - Fast Fourier Transform*) é uma forma mais eficiente de se realizar este cálculo, que reduz a complexidade do mesmo para $O(N \log N)$ (HOFFMAN, 2004) e é empregada em praticamente qualquer aplicação da transformada de Fourier em processamento de imagens. O desenvolvimento do método é apresentado a seguir (GONZALEZ & WOODS, 2000).

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-j2\pi ux/N}$$

Assume-se que $N = 2^n$. Sendo assim, $N = 2M$, onde M é um inteiro positivo.

$$\begin{aligned} F(u) &= \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) e^{-j2\pi ux/2M} \\ &= \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) e^{-j2\pi u(2x)/2M} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) e^{-j2\pi u(2x+1)/2M} \right] \\ &= \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) e^{-j2\pi u(x)/M} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) e^{-(j2\pi u(x)/M) - (j2\pi u/2M)} \right] \\ &= \frac{1}{2} \left[\frac{1}{M} \sum_{x=0}^{M-1} f(2x) e^{-j2\pi u(x)/M} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) e^{-j2\pi u(x)/M} e^{-j2\pi u/2M} \right] \end{aligned}$$

Define-se então, para $u = 0, 1, 2, \dots, M-1$, as seguintes relações:

$$\begin{aligned} F_{par}(u) &= \frac{1}{M} \sum_{x=0}^{M-1} f(2x) e^{-j2\pi u(x)/M} \\ F_{impar}(u) &= \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) e^{-j2\pi u(x)/M} \end{aligned}$$

Sendo assim, estabelece-se $F(u)$ conforme mostra a Equação 3.21. A prova da complexidade reduzida do algoritmo pode ser encontrada em (GONZALEZ & WOODS, 2000). Conforme foi mencionado anteriormente, a transformada de Fourier

bidimensional pode ser obtida a partir de uma composição de transformadas unidimensionais. Sendo assim, conclui-se que o algoritmo unidimensional apresentado pode ser utilizado em ambos os casos.

$$F(u) = \frac{1}{2} \left[F_{par}(u) + F_{impar}(u) e^{-j2\pi u/2M} \right] \quad (3.21)$$

O desenvolvimento apresentado acima considera o caso em $N = 2^n$. Para o caso genérico, veja (DINIZ, SILVA & NETTO, 2002). Para mais detalhes sobre a implementação do método, veja (GONZALEZ & WOODS, 2000).

3.7 Outras Ferramentas Matemáticas

Além da transformada de Fourier, outras duas técnicas matemáticas são bastante empregadas em processamento de imagens: a convolução e a correlação.

3.7.1 Convolução

Sejam $f(x)$ e $g(x)$ funções unidimensionais contínuas. A convolução entre as duas, denotada por $f(x) * g(x)$ é definida da seguinte forma (GONZALEZ & WOODS, 2000):

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha) g(x - \alpha) d\alpha \quad (3.22)$$

Para o caso discreto, realiza-se uma extensão das funções $f(x)$ e $g(x)$ para estabelecer a dimensão N do domínio da convolução entre as funções. O tamanho mínimo para que não haja sobreposição de períodos é $N \geq A + B - 1$ (GONZALEZ & WOODS, 2000), sendo A e B as dimensões de $f(x)$ e $g(x)$, respectivamente. A extensão de $f(x)$ e $g(x)$ pode ser realizada da seguinte forma:

$$f_e(x) = \begin{cases} f(x), & 0 \leq x \leq A - 1 \\ 0, & A \leq x \leq N - 1 \end{cases}$$

$$g_e(x) = \begin{cases} g(x), & 0 \leq x \leq B - 1 \\ 0, & B \leq x \leq N - 1 \end{cases}$$

A expressão para o caso discreto é apresentada na Equação 3.23 (GONZALEZ & WOODS, 2000).

$$f(x) * g(x) = \frac{1}{N} \sum_{n=0}^{N-1} f_e(n) g_e(x - n) \quad (3.23)$$

A Figura 3.24 mostra o resultado da convolução entre as funções $f_e(x)$ e $g_e(x)$ discretas, apresentadas na Figura 3.23.

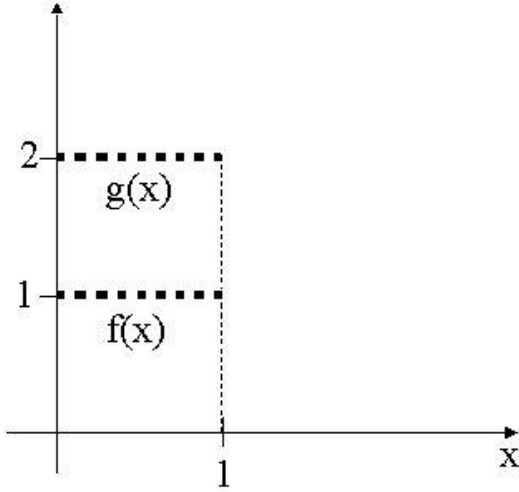


Figura 3.23: Funções discretas unidimensionais $f(x)$ e $g(x)$.

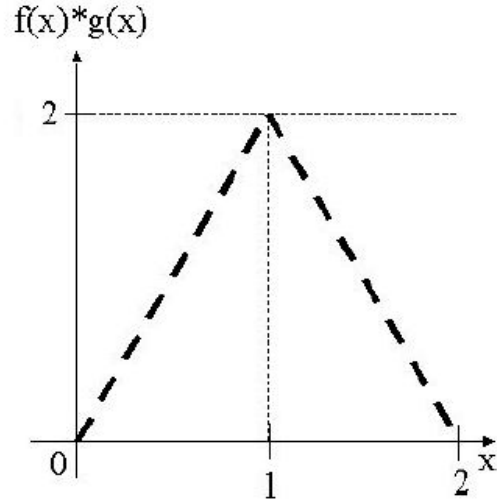


Figura 3.24: Convolução entre as funções discretas $f(x)$ e $g(x)$.

Para o caso discreto bidimensional, após a realização da extensão das funções $f(x, y)$ e $g(x, y)$ de forma análoga à mencionada no caso unidimensional, tem-se a seguinte expressão (GONZALEZ & WOODS, 2000):

$$f_e(x, y) * g_e(x, y) = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_e(n, m) g_e(x - n, y - m) \quad (3.24)$$

O Teorema da Convolução

Além de ser amplamente utilizada em operações de filtragem de imagens no domínio do espaço, a convolução tem considerável aplicabilidade na análise de imagens no domínio da frequência. Através do Teorema da Convolução, estabelece-se uma relação entre a convolução de duas funções e suas transformadas de Fourier, conforme mostra a Equação 3.25 (GONZALEZ & WOODS, 2000).

$$f(x) * g(x) \Leftrightarrow F(u)G(u) \quad (3.25)$$

A Equação 3.25 mostra que é possível obter a convolução entre duas funções através do produto de suas transformadas de Fourier. Em determinados casos específicos, utilizando o algoritmo da transformada rápida de Fourier (para casos discretos), tal procedimento é mais rápido do que o cálculo direto da convolução (GONZALEZ & WOODS, 2000).

Para provar o Teorema da Convolução, considere $h(x) = f(x) * g(x)$. Considere também $\mathfrak{F}(h(x)) = H(u)$, $\mathfrak{F}(f(x)) = F(u)$ e $\mathfrak{F}(g(x)) = G(u)$ (HOFFMAN, 2004).

$$\begin{aligned}
 H(u) &= \mathfrak{F}\{f(x) * g(x)\} \\
 &= \mathfrak{F}\left\{\int_{-\infty}^{\infty} f(\alpha) g(x - \alpha) d\alpha\right\} \\
 &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(\alpha) g(x - \alpha) d\alpha\right] e^{-j2\pi ux} dx \\
 &= \int_{-\infty}^{\infty} f(\alpha) \left[\int_{-\infty}^{\infty} g(x - \alpha) e^{-j2\pi ux} dx\right] d\alpha
 \end{aligned}$$

Fazendo-se uma substituição de variáveis, tem-se

$$y = x - \alpha \rightarrow dy = dx$$

$$\begin{aligned}
 H(u) &= \int_{-\infty}^{\infty} f(\alpha) \left[\int_{-\infty}^{\infty} g(y) e^{-j2\pi u(y+\alpha)} dy\right] d\alpha \\
 &= \int_{-\infty}^{\infty} f(\alpha) \left[\int_{-\infty}^{\infty} g(y) e^{-j2\pi uy} e^{-j2\pi u\alpha} dy\right] d\alpha \\
 &= \int_{-\infty}^{\infty} f(\alpha) e^{-j2\pi u\alpha} \left[\int_{-\infty}^{\infty} g(y) e^{-j2\pi uy} dy\right] d\alpha \\
 &= G(u) \int_{-\infty}^{\infty} f(\alpha) e^{-j2\pi u\alpha} d\alpha
 \end{aligned}$$

$$H(u) = F(u)G(u)$$

Uma parte complementar do Teorema da Convolução, que pode ser provada de forma similar à apresentada anteriormente, diz que

$$f(x)g(x) \Leftrightarrow F(u) * G(u) \quad (3.26)$$

3.7.2 Correlação

Denotada por $f(x) \circ g(x)$, a correlação entre duas funções contínuas $f(x)$ (sendo $f^*(x)$ o conjugado complexo de $f(x)$) e $g(x)$ é definida da seguinte forma (GONZALEZ & WOODS, 2000):

$$f(x) \circ g(x) = \int_{-\infty}^{\infty} f^*(x) g(x + \alpha) d\alpha \quad (3.27)$$

Na correlação, desloca-se $g(x)$ através de $f(x)$ e integra-se o produto de $-\infty$ a ∞ para cada um dos possíveis valores do deslocamento x . As Equações 3.28 e 3.29 mostram, respectivamente, os casos discreto unidimensional e discreto bidimensional da correlação.

$$f(x) \circ g(x) = \frac{1}{N} \sum_{n=0}^{N-1} f_e^*(n) g_e(x + n) \quad (3.28)$$

$$f_e(x, y) \circ g_e(x, y) = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f_e^*(n, m) g_e(x + n, y + m) \quad (3.29)$$

Umas das principais aplicações da correlação em Processamento de Imagens, e Processamento de Sinais em geral, é conhecida por *template matching*, e corresponde à consulta sobre a ocorrência de um determinado padrão g dentro do sinal f , de forma que se g ocorre em f , a região onde isto se dá oferecerá uma resposta alta, sendo possível então identificar onde começa e termina o padrão dentro do sinal (WOLFRAM II, 2005).

3.8 Realce de Imagens

As técnicas de realce de imagens tem objetivos variados, que dependem diretamente da aplicação específica na qual são empregadas. Porém, em geral, estes objetivos circulam em torno do destaque de uma determinada característica ou aspecto da imagem (GONZALEZ & WOODS, 2000). As técnicas aqui abordadas residem em um dos dois seguintes: espaço e frequência.

O domínio do espaço se refere ao próprio conjunto de *pixels* da imagem, de forma que as técnicas residentes deste domínio atuam diretamente sobre os valores das luminâncias dos *pixels*. Segundo (GONZALEZ & WOODS, 2000), considerando $f(x, y)$ a imagem original, $g(x, y)$ a imagem realçada e T um operador aplicável a f , tem-se

$$g(x, y) = T[f(x, y)]$$

O operador T , quando atua somente na posição (x, y) de f para obter $g(x, y)$, representa um simples mapeamento de níveis de cinza. Quando se baseia na vizinhança do *pixel* (x, y) de f para determinar a luminância do seu correspondente em g , sendo a vizinhança de um *pixel* geralmente definida como uma subimagem de forma quadrada ou retangular centrada no mesmo, T é dito operação de máscara ou filtragem (GONZALEZ & WOODS, 2000).

No domínio da frequência, as técnicas de processamento de imagens estão fundamentadas no Teorema da Convolução (Equação 3.25). Para entender melhor esta afirmativa, considere $f(x, y)$ uma imagem qualquer, e $t(x, y)$ um operador linear invariante com relação a posição, ou seja, o resultado de sua aplicação depende apenas do valor contido em uma determinada posição, e não da posição em si. Seja $g(x, y)$ a imagem obtida através da convolução entre $t(x, y)$ e $f(x, y)$. Considerando $G(u, v)$, $T(u, v)$ e $F(u, v)$ as transformadas de Fourier de g , t e f , respectivamente, e aplicando o Teorema da Convolução, tem-se as seguintes relações:

$$G(u, v) = T(u, v)F(u, v) \quad (3.30)$$

$$g(x, y) = \mathfrak{F}^{-1}[T(u, v)F(u, v)] \quad (3.31)$$

Diversas técnicas de realce no domínio da frequência, como os filtros de passa-alta e baixa, podem ser definidas conforme mostra a Equação 3.31.

Nas seções seguintes serão apresentadas algumas operações de filtragem, tanto no domínio do espaço, quanto no domínio da frequência. É importante notar que algumas técnicas de realce de imagens já foram apresentadas anteriormente neste texto, como expansão e equalização de histograma e limiarização simples.

3.8.1 Tipos de Filtragem

São dois os tipos de filtragem abordados neste trabalho: passa-alta e passa-baixa.

Embora os termos “alta” e “baixa” se refiram aos componentes de frequência da imagem, os mesmos são aplicados a filtros tanto do domínio da frequência, quanto do espaço. Os filtros de passa-alta atuam de forma a atenuar os componentes de baixa frequência da imagem, que correspondem a regiões de pouca alteração no valor da luminância dos *pixels*. O resultado é um reforço nas regiões onde há uma variação alta na luminância, o que qualifica este tipo de filtro para aplicações de detecção de bordas (GONZALEZ & WOODS, 2000). Já os filtros de passa-baixa atuam de forma inversa, atenuando as regiões de alta frequência. O resultado é uma certa diminuição do nível de detalhamento da imagem. Este tipo de filtro é utilizado em aplicações de suavização de imagens e remoção de ruídos (WINKLER & THOMÉ I, 2003).

3.8.2 Filtragem no Domínio do Espaço

Máscaras espaciais são geralmente matrizes simétricas, quadradas e de dimensões ímpares. Sua aplicação a uma determinada imagem ocorre da seguinte forma:

- Seja $f(x, y)$ a imagem original, $t(x, y)$ a máscara e $g(x, y)$ a imagem destino.
- Para se calcular $g(i, j)$, sobrepõe-se $t(x, y)$ sobre a imagem $f(x, y)$ de forma que o *pixel* central da máscara fique exatamente sobre a posição (i, j) de f .
- Calcula-se então a soma dos produtos entre *pixels* de f e da máscara que estão alinhados, sendo este o valor $g(i, j)$.
- Uma operação de máscara sobre uma determinada imagem corresponde a repetição deste procedimento para cada *pixel* da mesma, realizando-se um deslocamento da máscara por todo seu espaço de *pixels*.

Caso a máscara seja de fato uma matriz simétrica, o procedimento descrito acima corresponde à realização de uma convolução entre $t(x, y)$ e $f(x, y)$ (GONZALEZ & WOODS, 2000), conforme mostra a Equação 3.32.

$$g(x, y) = t(x, y) * f(x, y) \quad (3.32)$$

Passa-Baixa

Os filtros de passa-baixa no domínio do espaço têm como característica principal o fato de que todos os seus valores são positivos. Conforme mencionado anteriormente, os filtros de passa-baixa têm como objetivo a suavização da imagem. Uma das máscaras deste tipo mais utilizadas é mostrada abaixo.

$$t(x, y) = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3.33)$$

A Figura 3.25.b mostra o resultado da aplicação do operador apresentado na Equação 3.33 na Figura 3.25.a. As Figuras 3.25.c e 3.25.d mostram o resultado da aplicação de máscaras análogas àquela da Equação 3.33, porém, com dimensões 5X5 e 7X7, respectivamente.

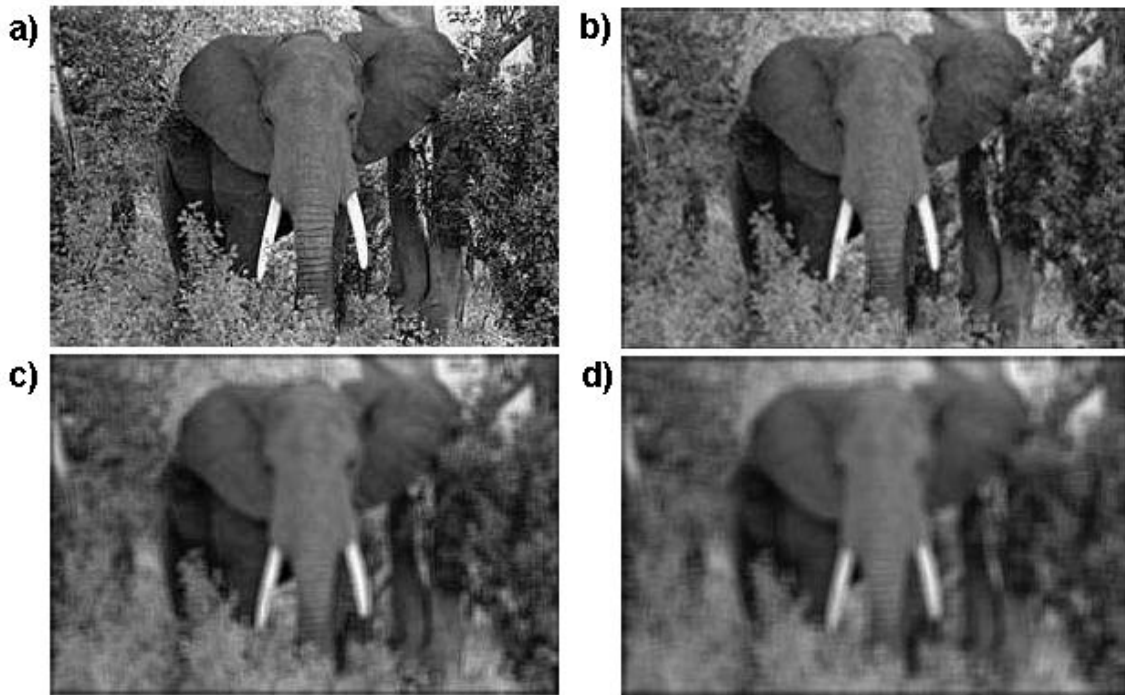


Figura 3.25: Exemplos de aplicações de filtros passa-baixa no domínio do espaço.

Outro tipo de operação de suavização no domínio do espaço é a filtragem por mediana. Esta é aplicada especialmente em casos de remoção de ruídos, diferindo dos filtros de média (como o apresentado na Equação 3.33) pelo fato de que não borra

bordas ou detalhes da imagem (GONZALEZ & WOODS, 2000). Na filtragem por mediana, cada *pixel* tem sua luminância substituída pela mediana das luminâncias dos *pixels* de sua vizinhança. A Figura 3.26.b mostra a Figura 3.26.a corrompida por ruído aleatório. As Figuras 3.26.c e 3.26.d mostram os resultados da remoção de ruídos com a aplicação da filtragem por mediana com vizinhanças de 3×3 e 5×5 , respectivamente.

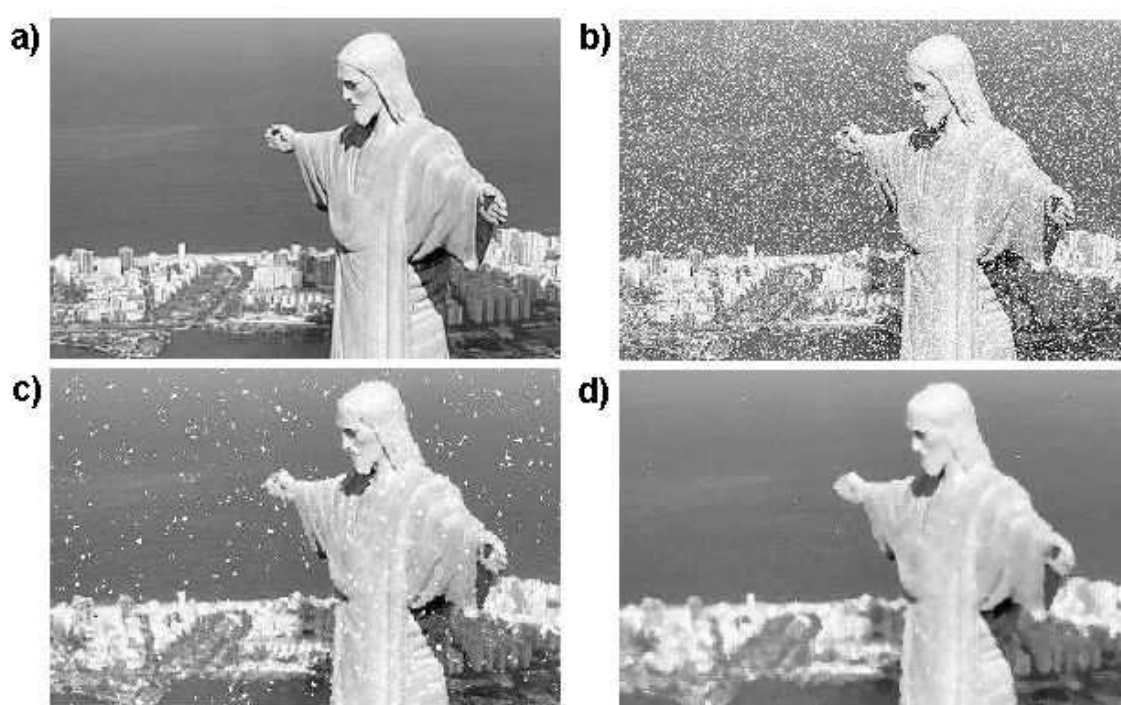


Figura 3.26: Exemplos de remoção de ruídos através da aplicação de filtragem passa-baixa por mediana.

Nota-se que a qualidade da imagem corrigida pela filtragem é inferior à da imagem original. Porém, localizar as bordas do Cristo Redentor na imagem corrigida tem complexidade muito menor do que na imagem corrompida.

Passa-Alta

As máscaras de filtragem passa-alta tem, em geral, valores positivos próximos ao centro e negativos na periferia (GONZALEZ & WOODS, 2000). A Equação 3.34 mostra uma máscara de dimensões 3×3 freqüentemente empregada neste tipo de operação.

$$t(x, y) = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (3.34)$$

Outra forma de se realizar uma filtragem que privilegie as regiões de alta frequência é conhecida por filtragem alto-reforço, e corresponde a computação da diferença entre a imagem original e o resultado da aplicação de uma filtragem passa-baixa na mesma. As Figuras 3.27.b e 3.27.c apresentam os resultados das aplicações da filtragem por máscara, utilizando a matriz na Equação 3.34, e da filtragem auto-reforço da Figura 3.27.a.

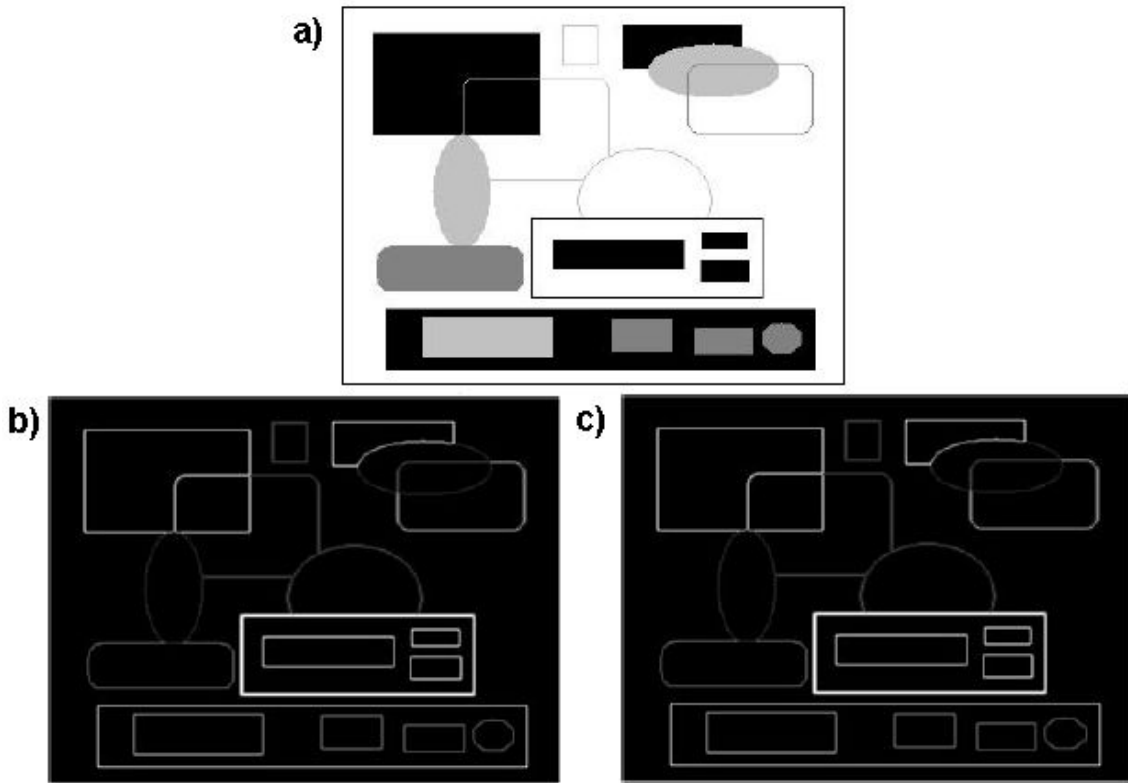


Figura 3.27: Exemplos de filtragem passa-alta no domínio do espaço.

Outras máscaras de filtragem passa-alta são os chamados operadores de Prewitt e Sobel. Na Equação 3.35, apresenta-se dois exemplos de operadores de Prewitt, enquanto na Equação 3.36 são mostrados dois operadores de Sobel. Ambos são baseados na magnitude do gradiente da imagem (GONZALEZ & WOODS, 2000).

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad (3.35)$$

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (3.36)$$

Existem diversos outros filtros passa-alta que atuam no domínio do espaço, como os filtros direcionais, o gradiente e o Laplaciano. Mais informações a respeito de alguns desses filtros podem ser encontradas em (WINKLER & THOMÉ, 2004).

3.8.3 Filtragem no Domínio da Frequência

Devido a sua simplicidade de implementação e menor custo computacional, filtros espaciais são aplicados com maior frequência do que aqueles do domínio da frequência. Todavia, a compreensão do domínio da frequência no âmbito de processamento de imagens é imprescindível para um melhor entendimento dos processos de realce de detalhes e suavização de imagens, além de ser absolutamente necessário para o desenvolvimento de idéias mais avançadas, como o emprego de filtros homomórficos e restauração de imagens (GONZALEZ & WOODS, 2000).

Conforme dito anteriormente no texto, as operações de filtragem no domínio da frequência correspondem ao cálculo da transformada de Fourier da imagem em questão e a posterior multiplicação da transformada por uma função filtro de dimensões iguais às da mesma. Então, calcula-se a transformada inversa de Fourier deste produto para se obter a imagem devidamente filtrada. Sendo $f(x, y)$ a imagem original, $g(x, y)$ a imagem resultante após a filtragem, $F(u, v)$ a transformada de Fourier de $f(x, y)$ e $H(u, v)$ a função filtro, estabelece-se a seguinte relação:

$$g(x, y) = \mathfrak{F}^{-1}[H(u, v)F(u, v)]$$

Antes de estudar o processo de filtragem no domínio da frequência, cabe destacar a forma de cálculo do *espectro* de potência de uma imagem, que permite que seja observada a distribuição da informação da mesma por suas diferentes frequências, sendo o ponto central do *espectro* correspondente a frequência 0 (zero). Desta forma, as regiões próximas ao centro representam as baixas frequências, enquanto nas quinas se encontram as altas frequências. A Equação 3.37 mostra de que forma o mesmo é

calculado. A Equação 3.38 mostra como se calcula a potência total de uma imagem (GONZALEZ & WOODS, 2000). A Figura 3.28.b mostra o *espectro* de potência da Figura 3.28.a, que apresenta alto contraste, enquanto a Figura 3.28.d apresenta o *espectro* de potência da Figura 3.28.c, na qual há pouco contraste.

$$P(u, v) = R^2(u, v) + I^2(u, v) \quad (3.37)$$

$$P_T = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} P(u, v) \quad (3.38)$$

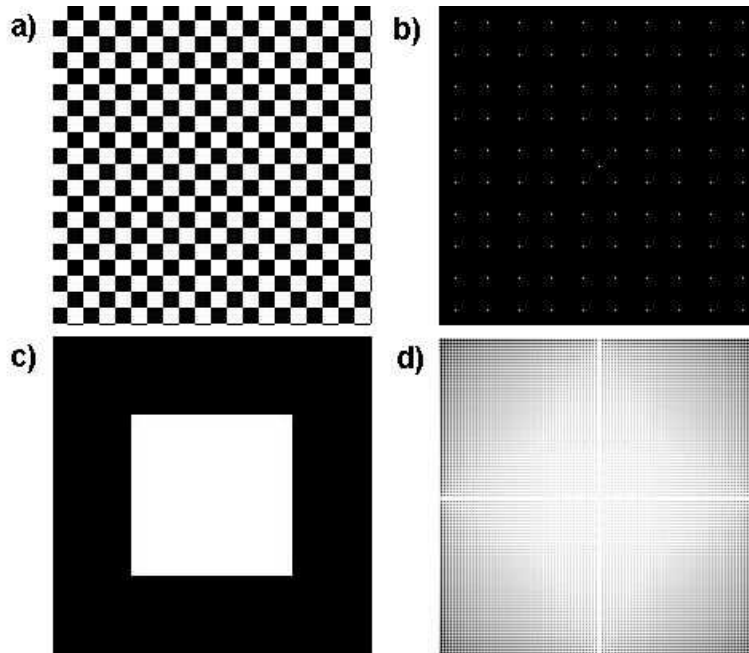


Figura 3.28: Exemplos de *espectros* de potência.

É possível notar na Figura 3.28.b que a informação está distribuída por diversas frequências da imagem, enquanto na Figura 3.28.d, nota-se que as frequências mais baixas concentram a maior parte da informação da imagem correspondente.

Passa-Baixa

Assim como no domínio do espaço, a filtragem passa-baixa no domínio da frequência tem como objetivo a suavização da imagem, desprivilegiando regiões de borda ou mudança abrupta de luminância. Nesta seção são mostrados dois tipos de função filtro com esta finalidade: ideal e de Butterworth.

Filtro Ideal Na Equação 3.39 é mostrado o filtro ideal, sendo R um valor não negativo qualquer e $D(u, v)$ a distância entre a posição (u, v) e a origem do espectro, ou seja, o ponto de frequência 0 (zero). A Figura 3.29 apresenta um gráfico tridimensional do filtro.

$$H(u, v) = \begin{cases} 1, & \text{se } D(u, v) \leq R \\ 0, & \text{se } D(u, v) > R \end{cases} \quad (3.39)$$

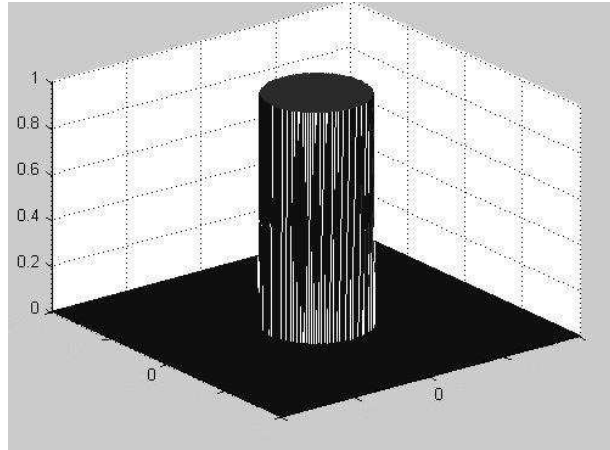


Figura 3.29: Gráfico de um filtro ideal passa-baixa.

Como a aplicação do filtro no domínio da frequência se resume à sua multiplicação ponto-a-ponto pela transformada de Fourier da imagem, conclui-se, considerando o centro da transformada como o ponto de frequência 0 (zero), que as frequências mais altas, representadas em pontos cuja distância à origem é maior que R , serão completamente atenuadas, restando apenas os pontos referentes às frequências mais baixas. O valor R é conhecido como frequência de corte (GONZALEZ & WOODS, 2000). Na Figura 3.30.b é mostrado o resultado da aplicação da filtragem ideal na Figura 3.30.a com $R = 60$.

Através do *espectro* de potência da imagem, é possível calcular qual a porcentagem de informação I que foi aproveitada na operação de filtragem, tendo, o restante, sido descartado. Para tanto, basta calcular a porcentagem da potência total da imagem contida no interior do círculo, conforme demonstra a Equação 3.40 (GONZALEZ & WOODS, 2000).

$$I = 100 \sum_u \sum_v \frac{P(u, v)}{P_T} \text{ para } u, v \text{ tal que } \sqrt{u^2 + v^2} \leq R. \quad (3.40)$$

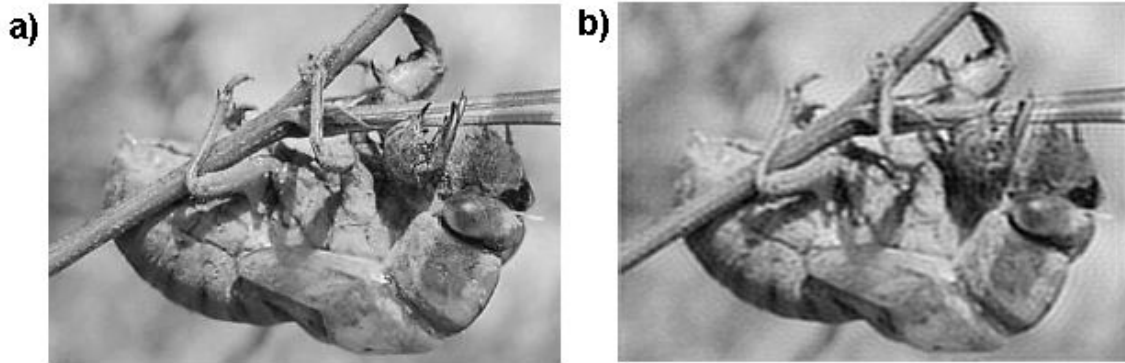


Figura 3.30: Exemplo de filtragem ideal passa-baixa no domínio da frequência.

Conforme aumenta-se o valor de R , maior se torna a porcentagem da potência da imagem englobada pelo círculo, aproximando o resultado cada vez mais da imagem original. Por outro lado, quanto menor R , maior será o borramento dos detalhes finos da imagem. No exemplo da Figura 3.30.b, obteve-se $I = 99,14\%$.

Filtro de Butterworth O filtro de Butterworth emprega uma atenuação de frequência gradual, de forma que, quanto maior for a distância entre o ponto que representa uma determinada frequência e a origem, maior será a atenuação desta frequência (GONZALEZ & WOODS, 2000). A Equação 3.41 mostra a função filtro passa-baixa de Butterworth de ordem n , enquanto a Figura 3.31 apresenta seu gráfico.

$$H(u, v) = \frac{1}{1 + [D(u, v)/R]^{2n}} \quad (3.41)$$

Na Figura 3.32.b mostra-se o resultado da aplicação do filtro passa-baixa de Butterworth na Figura 3.32.a, com $R = 60$ e $n = 1$.

Passa-Alta

Conforme foi dito antes, na filtragem passa-alta deseja-se salientar as regiões de borda e de alta variação de luminância. Analogamente à filtragem passa-baixa, os dois tipos de função filtro apresentados nesta seção são o ideal e o de Butterworth.

Filtro Ideal A Equação 3.42 mostra a função filtro utilizada neste caso. A Figura 3.33 apresenta o gráfico da função. A Figura 3.34.b mostra o resultado da aplicação

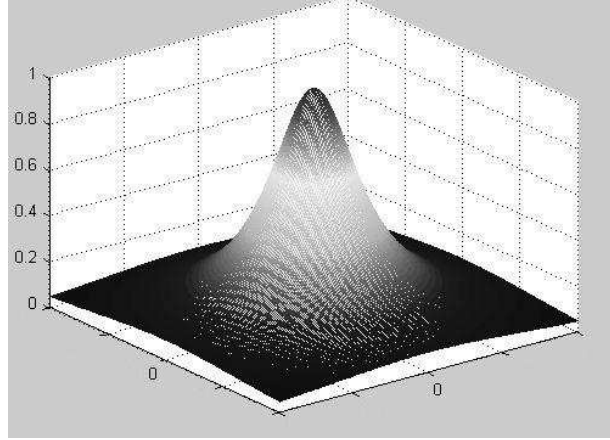


Figura 3.31: Gráfico de um filtro de Butterworth passa-baixa.

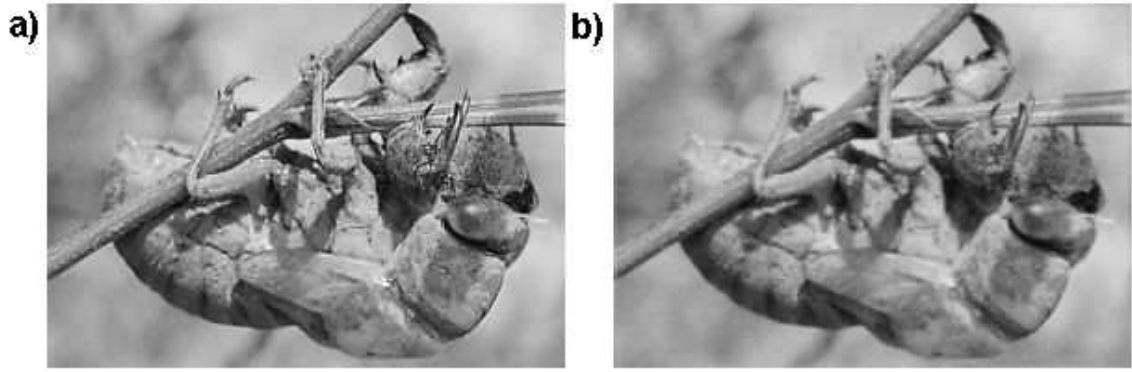


Figura 3.32: Exemplo de filtragem Butterworth passa-baixa no domínio da frequência.

do filtro ideal passa-alta na Figura 3.34.a com $R = 30$.

$$H(u, v) = \begin{cases} 0, & \text{se } D(u, v) \leq R \\ 1, & \text{se } D(u, v) > R \end{cases} \quad (3.42)$$

Filtro de Butterworth Analogamente ao seu correspondente na filtragem passa-baixa, o filtro de Butterworth passa-alta atenua as frequências baixas da imagem de forma gradual. A Equação 3.43 mostra a função filtro para este caso. A Figura 3.35 mostra o gráfico da mesma. A Figura 3.36.b mostra o resultado da aplicação do filtro de Butterworth passa-alta na Figura 3.36.a com $R = 30$.

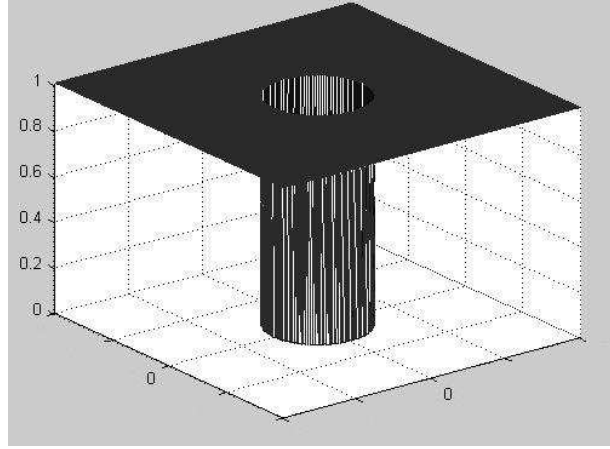


Figura 3.33: Gráfico de um filtro ideal passa-alta.

$$H(u, v) = \frac{1}{1 + [R/D(u, v)]^{2n}} \quad (3.43)$$

3.9 Rotulação de Componentes Conexas

Por muitos anos acreditou-se que o modelo de visão do ser humano fosse totalmente baseado na divisão da imagem em regiões de características comuns e posterior análise individual de cada uma destas regiões. Foi o movimento Gestalt (LEHAR, 2004) que combateu esta idéia do sistema visual, conhecida como reducionista. O movimento alegava que existia um elemento de forma no sistema visual, que unia as diferentes regiões e interpretava, de forma geral, o campo de visão.

Porém, em processamento de imagens, na grande maioria dos casos, o que se deseja é exatamente quebrar a imagem em partes. Geralmente, existem um ou mais objetos específicos de interesse na imagem. Sendo assim, para que seja possível classificá-los como tais ou analisá-los posteriormente, torna-se imprescindível separá-los do todo. A este processo dá-se o nome de segmentação.

A segmentação é, sem sombra de dúvida, um dos maiores desafios em processamento de imagens. Seu objetivo é extrair da imagem os objetos de interesse que nela constam, baseados em critérios de similaridade, proximidade, continuidade, entre outros (FELZENSZWALB & HUTTENLOCHER, 2004).

Uma das principais técnicas de segmentação é a rotulação de componentes conexas. O algoritmo apresentado a seguir foi extraído de (GONZALEZ & WOODS,

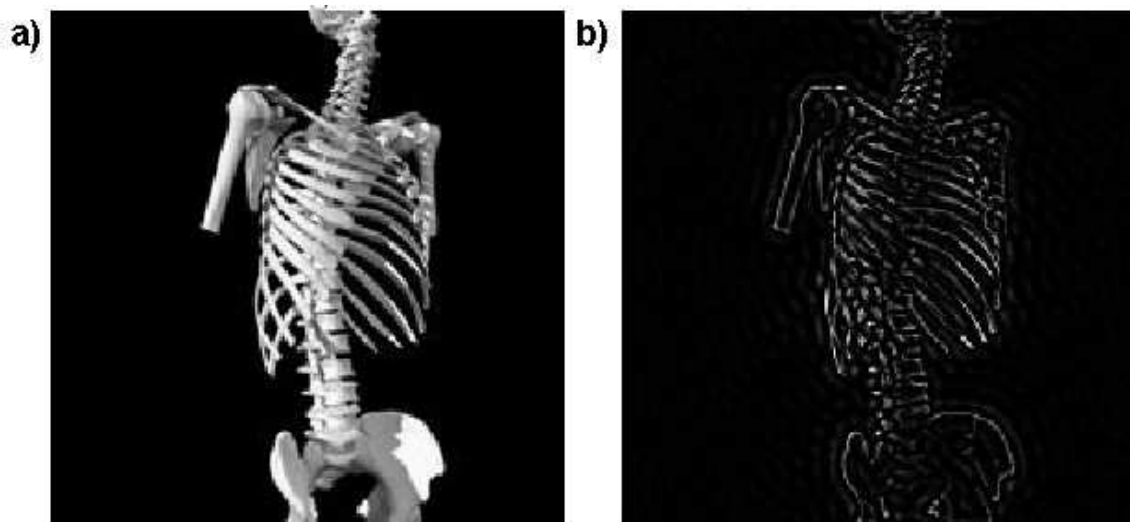


Figura 3.34: Exemplo de filtragem ideal passa-alta no domínio da frequência.

2000).

Embora o algoritmo apresentado aplique-se apenas a imagens binárias e utilize vizinhança-de-4, sua generalização, em ambos os casos, é intuitiva. As Figuras 3.37.b-f mostram as componentes conexas da Figura 3.37.a determinadas pelo algoritmo.

3.10 Detecção de Movimento

A movimentação de um determinado objeto em um ambiente é, por muitas vezes, essencial para sua identificação e extração, visando à análise e processamento individuais posteriores. Quando um dos objetivos de uma certa aplicação é segmentar um componente dos demais presentes, e é viável e possível obter-se uma seqüência de quadros do ambiente, a detecção de movimento se torna uma arma poderosa para a realização de tal tarefa. Tendo isto em vista, nesta seção serão apresentados dois métodos de detecção de movimento em imagens, ambos extraídos de (GONZALEZ & WOODS, 2000).

3.10.1 Comparação Simples

Este método baseia-se na computação da diferença absoluta entre duas imagens *pixel* a *pixel*. Para tanto, cria-se uma nova imagem *d* que represente um identificador de pontos distintos entre as duas imagens em questão. Considere $f(x, y, t_i)$ e $f(x, y, t_j)$

Algoritmo 3.1 Algoritmo de Rotulação de Componentes Conexas

Seja f uma imagem binária com altura n e largura m .

Seja r uma matriz $(n \times m)$ de zeros.

Seja l um inteiro igual a 1.

para y de 0 a $m - 1$ **faça**

para x de 0 a $n - 1$ **faça**

 Seja $p = f(x, y)$.

 Seja sup o vizinho superior de p .

 Seja esq o vizinho esquerdo de p .

se p igual a 1 **então**

se sup e esq iguais a 1 **então**

se $r(sup) == r(esq)$ **então**

$r(p) = r(sup)$

senão

 Registra relação entre $r(sup)$ e $r(esq)$

$r(p) = r(sup)$

fim se

senão se sup igual a 1 **então**

$r(p) = r(sup)$

senão se esq igual a 1 **então**

$r(p) = r(esq)$

senão

$r(p) = l$

$l = l + 1$

fim se

fim se

fim para

fim para

Calcula-se os conjuntos de rótulos relacionados.

Para cada conjunto, cria-se um novo rótulo.

Varre-se a matriz r substituindo-se os rótulos antigos pelos novos.

retorna r .

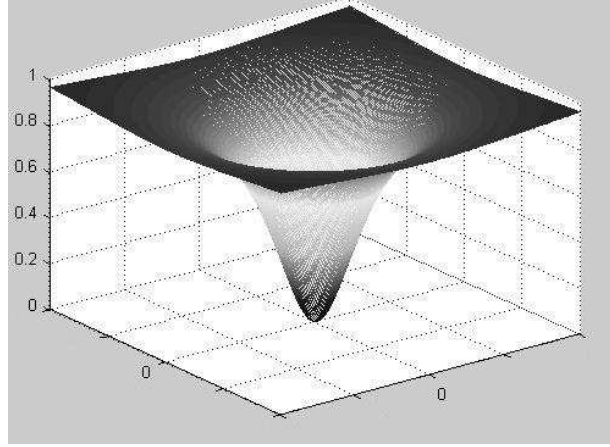


Figura 3.35: Gráfico de um filtro de Butterworth passa-alta.

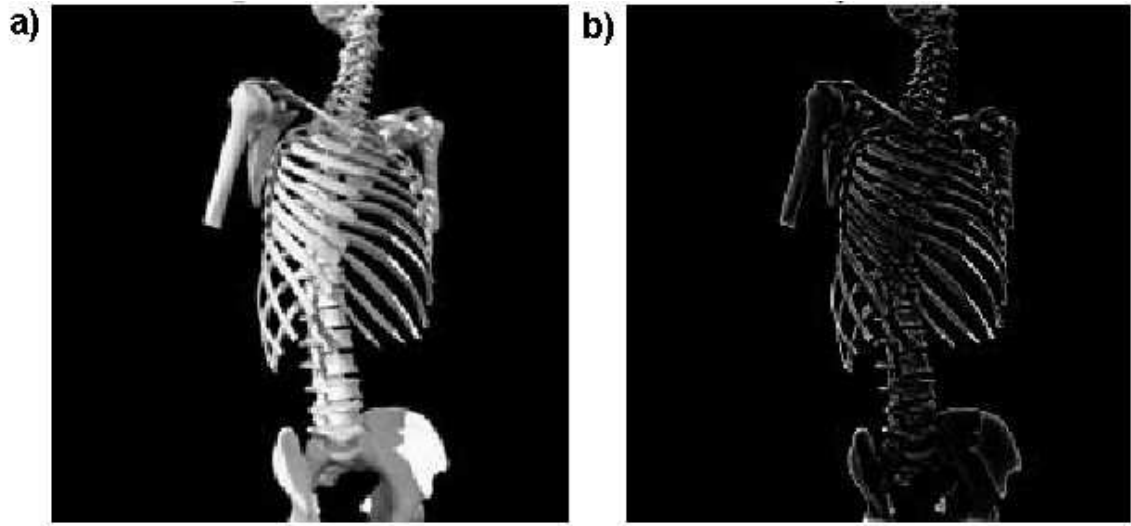


Figura 3.36: Exemplo de filtragem Butterworth passa-alta no domínio da frequência.

quadros capturados nos instantes t_i e t_j , respectivamente. Seja θ um valor positivo contido no intervalo de luminâncias representáveis nas duas imagens. Define-se $d(x, y)$ como

$$d(x, y) = \begin{cases} 1 & \text{se } |f(x, y, t_i) - f(x, y, t_j)| > \theta \\ 0 & \text{caso contrário} \end{cases} \quad (3.44)$$

As Figuras 3.38 e 3.39 mostram dois exemplos ilustrativos de imagens, enquanto a Figura 3.40 apresenta o resultado da aplicação da Equação 3.44, computando a diferença entre eles, com $\theta = 10$.

O cenário ideal para a aplicação desta técnica é aquele no qual é possível se obter

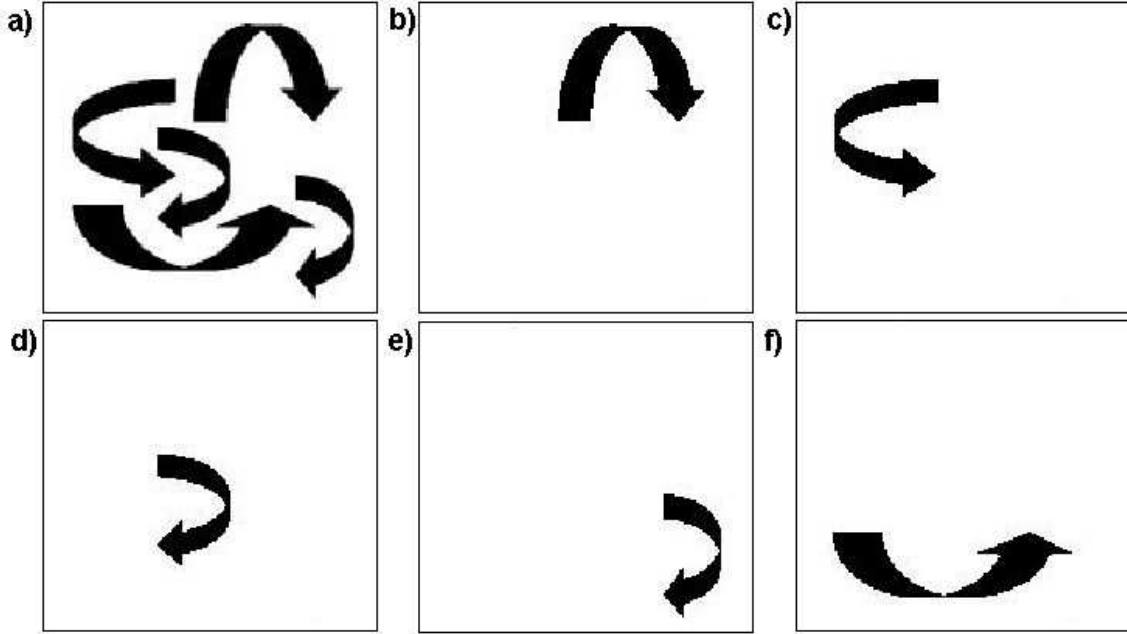


Figura 3.37: Exemplo de segmentação de imagens através de rotulação de componentes conexas.

$$\begin{bmatrix} 20 & 25 & 22 & 21 & 20 \\ 23 & 27 & 21 & 22 & 22 \\ 25 & 21 & 24 & 23 & 22 \\ 20 & 20 & 20 & 20 & 24 \\ 21 & 23 & 25 & 25 & 25 \end{bmatrix}$$

Figura 3.38: $f(x, y, t_i)$

$$\begin{bmatrix} 80 & 25 & 22 & 21 & 20 \\ 23 & 27 & 21 & 22 & 22 \\ 25 & 62 & 63 & 64 & 22 \\ 20 & 60 & 70 & 74 & 24 \\ 21 & 59 & 61 & 25 & 200 \end{bmatrix}$$

Figura 3.39: $f(x, y, t_j)$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Figura 3.40: $d(x, y)$

uma imagem do ambiente na qual constam apenas os elementos estacionários. Desta forma, utiliza-se esta imagem para comparação e identificação de qualquer objeto que se mova, saindo, portanto, de sua condição estacionária.

Esta técnica, embora amplamente utilizada, é bastante vulnerável à ocorrência de ruídos, sejam eles oriundos de variação de iluminação ou do próprio dispositivo de captação de imagens. O tratamento para este tipo de problema consiste da aplicação de filtros de suavização (passa-baixa) ou da abordagem por vizinhança, a qual corresponde a computar a diferença entre médias de luminâncias de regiões da imagem (TCSMITAIL, 2004). Desta forma, para uma imagem de dimensões $A \times A$, comparando regiões de *pixels* $b \times b$, teria-se $d(x, y)$ com dimensões $A \div b \times A \div b$.

No Apêndice I são mostrados resultados da aplicação desta técnica, comparando

quadros consecutivos. No Apêndice II, aplica-se esta técnica realizando a comparação dos quadros de movimento a um quadro estacionário. Para todas as comparações, foi utilizado $\theta = 40$.

3.10.2 Diferenças Acumulativas

Esta técnica é uma variação da Comparação Simples, na qual tenta-se, através da introdução de memória no processo, eliminar a ocorrência de ruídos. A comparação é realizada entre quadros consecutivos, porém, as diferenças computadas são armazenadas de forma acumulativa.

Considere $f(x, y, t_0)$ o primeiro de uma seqüência de quadros e $f(x, y, t_n)$ o último. A matriz de diferenças ac é computada conforme mostra a Equação 3.45.

$$ac(x, y) = \sum_{i=0}^{n-1} d_i(x, y) \quad (3.45)$$

$$d_i = \begin{cases} 1 & \text{se } |f(x, y, t_i) - f(x, y, t_{i+1})| > \theta \\ 0 & \text{caso contrário} \end{cases}$$

Desta forma, torna-se possível remover pontos de diferença esporádicos, levando em conta que o valor dos mesmos na matriz ac será relativamente baixo.

3.10.3 Atualização do Fundo

Um aspecto essencial da detecção de movimento diz respeito à integridade do fundo. Em ambientes controlados, ou seja, ambientes nos quais se conhece, *a priori*, a situação de repouso, e a mesma é mantida durante a ocorrência do evento, tal aspecto não tem relevância, já que o parâmetro de comparação para identificação de movimento é estável.

Porém, em ambientes onde a situação de repouso é instável, torna-se necessário estabelecer uma estratégia de atualização do fundo durante a detecção de movimento. A técnica apresentada a seguir foi retirada de (BHAT, SHAPTHARISH & KHOSLA, 2003). É importante destacar que a mesma só é aplicável a sistemas de detecção de movimento que empreguem o uso de uma imagem base para comparação, excluindo, portanto, métodos que utilizam comparações entre imagens consecutivas.

Seja $f_b(t)$ a imagem base no instante t , e $f_a(t)$ a imagem capturada também no instante t . Inicialmente, considera-se $f_b(t)$ como a primeira imagem capturada pelo sistema. A atualização de f_b , é realizada conforme mostra a Equação 3.46.

$$f_b(t+1) = f_b(t) + s \cdot (f_a(t) - f_b(t)), \quad 0 < s \leq 1 \quad (3.46)$$

O fator s determina a velocidade de incorporação de novos objetos ao quadro base, de forma que quanto mais próximo de 1 (um) estiver seu valor, mais drásticas serão as alterações do fundo. Considerando as Figuras 3.41.a o fundo e 3.41.b a imagem atual, as Figuras 3.41.c e 3.41.d mostram o resultado da aplicação da técnica para $s = 0.1$ e $s = 0.3$, respectivamente.

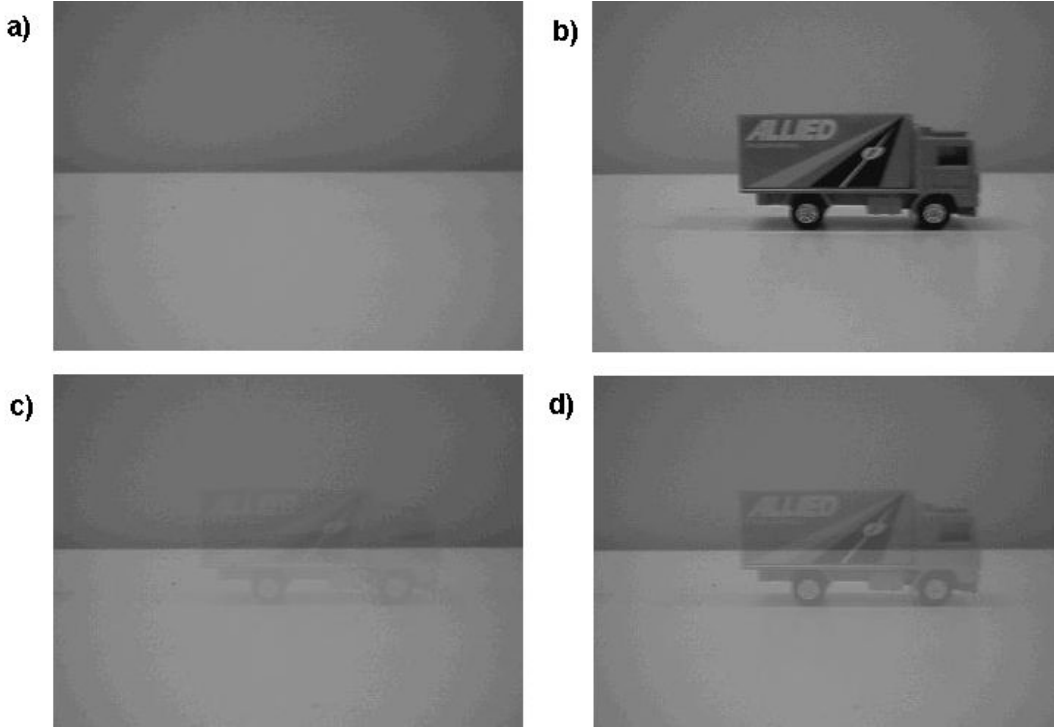


Figura 3.41: Exemplo da aplicação da técnica de atualização do fundo em situações de detecção de movimento.

3.10.4 Cálculo de Velocidade

Um outro aspecto inerente a aplicações de detecção de movimento é o que diz respeito ao cálculo da velocidade dos objetos em deslocamento na imagem, principalmente em

sistemas de controle de tráfego de veículos ou de determinados processos industriais. É importante ressaltar, porém, que a abordagem apresentada a seguir para o cálculo da velocidade é aplicável apenas a situações nas quais a distância focal dos objetos em movimento é conhecida e fixa.

Para realizar tal tarefa, deve-se, antes de qualquer coisa, estabelecer uma relação entre o número de *pixels* da imagem e a distância real no local alvo da filmagem. Desta forma, a velocidade de movimento pode ser calculada em *pixels/tempo*, sendo então realizada uma operação de conversão para a unidade desejada.

Para determinar o deslocamento de um objeto, o primeiro passo consiste da segmentação do mesmo da imagem, sendo possível então o cálculo do seu centro de massa. Considerando (x_0, y_0) o centro de massa do objeto segmentado na imagem $f(t_0)$, obtida no instante t_0 , e (x_1, y_1) o centro de massa do objeto segmentado na imagem $f(t_1)$, obtida no instante t_1 , o cálculo da velocidade seria realizado conforme indica a Equação 3.47.

$$v = \frac{|x_1 - x_0|}{t_1 - t_0} \text{ pixels/tempo} \quad (3.47)$$

Outra maneira de realizar o cálculo da velocidade reside no domínio da frequência, e é baseada na Transformada de Fourier. Maiores informações sobre a mesma podem ser encontradas em (GONZALEZ & WOODS, 2000).

Capítulo 4

Estudo de Casos

4.1 1º Caso: Detecção da Queda da Primeira Gota em um Processo de Destilação de Petróleo em Batelada

4.1.1 Introdução

Industrialmente, o petróleo é destilado em grandes quantidades, produzindo frações conhecidas como diesel, gasolina, nafta, etc. Cada uma destas frações têm uma faixa de temperatura característica e propriedades físicas e químicas padronizadas. No entanto, um dado petróleo, dependendo de sua origem geoquímica e do seu processo de formação, não necessariamente produz todas as frações ou não as produz nos valores padronizados das propriedades de cada fração. Sendo assim, uma mistura de petróleos é composta, acrescida de outras correntes produzidas na refinaria, de forma que as diferentes frações sejam produzidas de acordo com as especificações e em volume adequado à demanda do mercado (ASTM, 2002).

Para determinar, em laboratório, o volume que cada uma das faixas de corte (frações) do petróleo produz, utiliza-se a destilação em batelada, sendo um dos métodos para tal procedimento o padrão ASTM D5236 (ASTM, 2002). Com a destilação em batelada em laboratório pode-se caracterizar o petróleo pela determinação do volume que cada faixa de corte produz e as propriedades destas frações. Como o valor do petróleo é função do volume que o mesmo produz das frações mais nobres e da especificação natural das suas propriedades, tal procedimento tem extrema importância.

No processo de destilação em batelada ao qual este trabalho se aplica, as di-

ferentes frações do petróleo são separadas em frascos localizados em um carrossel na parte inferior da unidade de destilação utilizada. As frações são coletadas em intervalos determinados a partir do instante em que se desprende a primeira gota, indicando que o processo de destilação teve início. A cada intervalo, o carrossel sofre um movimento de rotação, trocando o frasco recipiente das diferentes frações do petróleo destilado.

Além disto, é preciso determinar a temperatura do início da destilação, ou seja, a temperatura mínima da primeira faixa ou fração. Como a retirada é contínua e pela própria dinâmica do processo, o início da destilação corresponde ao instante da formação da primeira gota no condensador lateral (ASTM, 2002).

Com a instrumentação atual, a identificação da formação da primeira gota apresenta algumas dificuldades em termos de instrumentação. Os dispositivos existentes se baseiam na interrupção de um feixe de luz, e apresentam três grandes inconvenientes: a) são específicos para cada unidade (eletrônica dedicada), b) a regulação do feixe ao caminho da gota é difícil e instável, pois a queda da gota é sujeita a perturbações e c) os instrumentos não apresentam recursos que facilitem a regulação. É importante ressaltar que a perda da primeira gota equivale à perda de todo o processo de destilação até aquele instante e que um erro na determinação da gota pode comprometer todo o trabalho de medida.

Este estudo de caso corresponde ao desenvolvimento de um sistema que tem como objetivo realizar a tarefa acima mencionada utilizando técnicas de processamento de imagens. As vantagens de se empregar tal solução são as seguintes: a) uso de instrumentação padrão de mercado (câmera e computador), sendo as diferenças entre cada equipamento de destilação tratadas por programa, b) a regulação é muito mais fácil, não sendo necessário o ajuste do equipamento ao caminho da gota, podendo o operador do laboratório utilizar a própria imagem gerada pelo dispositivo para enquadrar/regular o posicionamento da câmera.

Nesta seção serão apresentadas as técnicas empregadas na realização das tarefas inerentes ao cumprimento das metas estabelecidas para o sistema. O mesmo foi totalmente desenvolvido em C/C++, tendo ainda alguns testes sido realizados anteriormente em Matlab. Na Figura 4.4 mostra-se a interface do sistema construído.

4.1.2 O Processo de Destilação

Cabe neste ponto uma breve descrição da unidade de destilação utilizada, assim como do processo de destilação em si. O equipamento de destilação em batelada em escala de laboratório (Figura 4.1.2) é construído totalmente em vidro e consiste de um balão refeedor (A), cujo volume se situa no intervalo de 5 a 50 litros, uma torre de destilação (B), que pode ter de um a dois metros de altura, um condensador de topo (refluxo) (C), um condensador de saída lateral (D) e um dispositivo de coleta das frações. Este equipamento fica instalado dentro de uma capela de cerca de quatro metros de altura e opera com exaustão forçada.

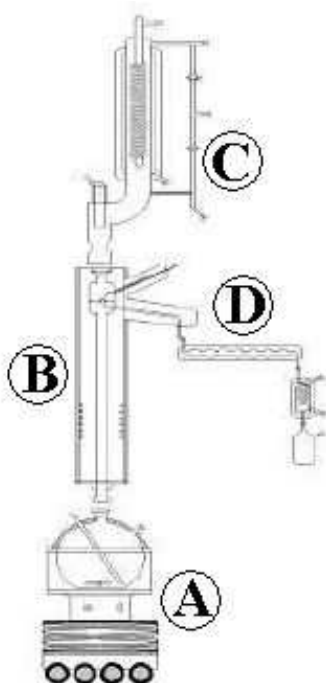


Figura 4.1: Unidade de destilação em batelada.

O início do processo se dá com o aquecimento controlado do volume de petróleo armazenado no balão refeedor. Com o aquecimento, os compostos voláteis à temperatura do balão passam para a fase gasosa e sobem pela coluna em direção ao condensador de topo. Ao atingirem o condensador de topo, estes compostos perdem energia, se condensam e retornam ao balão refeedor. Este fluxo de compostos voláteis entre o balão e o condensador de topo provoca um transporte de energia que estabiliza a temperatura, ou seja, como o balão perde energia, a temperatura

tem seu crescimento interrompido.

Os compostos voláteis são gradativamente retirados da coluna por uma saída lateral, se condensam no condensador lateral e são recolhidos no dispositivo de coleta. Com isso, o transporte de voláteis diminui, assim como a perda de energia. Desta forma, energia e temperatura voltam a subir, até atingirem uma nova faixa de volatilização de componentes mais pesados. Este processo se repete até que as frações de interesse tenham sido separadas.

4.1.3 O Cenário

O cenário é composto por seis elementos principais: a unidade de destilação em batelada, uma câmera de vídeo, uma fonte de luz, um anteparo de cor preta e material fosco, um computador com uma placa de captura de vídeo, um software específico e um dispositivo de atuação física na unidade de destilação. O trabalho aqui descrito foca-se apenas no desenvolvimento do software específico que detecta a formação e queda da primeira gota do processo de destilação e interage com os outros elementos do ambiente, estando o funcionamento da unidade de destilação, da câmera de vídeo e do atuador fora do escopo aqui abordado.

A estrutura funciona da seguinte forma: a câmera de vídeo, localizada na altura do bico de onde se desprendem as gotas provenientes do processo de condensação, envia um sinal de vídeo para uma placa de captura instalada no computador. A fonte de luz incide diretamente sobre o bico. A placa de cor preta é posicionada no lado oposto à câmera, servindo de fundo e impedindo que outros movimentos sejam confundidos com a queda da gota. O sistema interage com a placa de captura e acessa os quadros da imagem fornecida pela câmera.

Inicialmente, o sistema deve localizar com precisão o bico do aparelho de destilação, o que lhe permite destacar uma região de interesse mais específica para análise dos diferentes quadros.

Iniciado o processo de destilação, o sistema analisa os quadros (resolução de 320 X 240) obtidos através da placa de captura tentando identificar a formação e o desprendimento da primeira gota. Ao detectar a queda da gota, o sistema envia um sinal elétrico ao atuador, o qual calcula os intervalos adequados e interage com o aparelho de destilação realizando os movimentos de rotação do carrossel de frascos. Na Figura 4.2, tem-se uma representação gráfica do cenário apresentado.

Devido a restrição de tempo inerente ao sistema, este pode ser classificado como

um sistema de tempo real. Um atraso significativo na comunicação do evento do desprendimento da primeira gota por parte do sistema pode comprometer a qualidade das amostras coletadas.

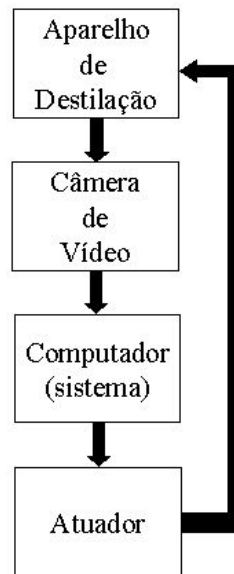


Figura 4.2: Cenário do sistema.

4.1.4 O Sistema

O sistema desenvolvido tem seu funcionamento baseado em 4 *threads*. São elas:

- *Thread* de Captura: responsável por obter os quadros a partir da placa de captura e colocá-los na fila de análise.
- *Thread* de Análise: responsável por analisar cada quadro da fila de análise e enviar o sinal elétrico ao atuador caso detecte o desprendimento da primeira gota.
- *Thread* de Log: responsável por registrar todos os eventos do sistema, além de imagens e vídeos para análises posteriores.
- *Thread* Principal: responsável por coordenar o processo de localização do bico e as três *threads* mencionadas, iniciando-as, suspendendo-as ou reiniciando-as conforme necessário e por realizar os procedimentos gerais da aplicação.

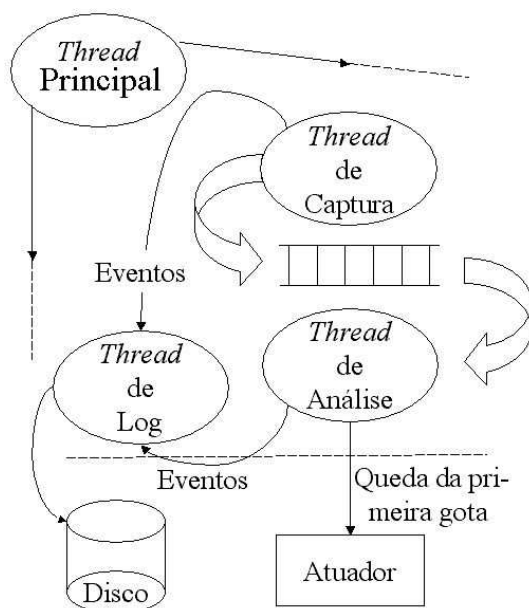


Figura 4.3: Esquema das *threads* do sistema.

A Figura 4.3 apresenta um esquema do funcionamento do sistema. Nas seções seguintes, será mostrado como foram concebidas as duas tarefas chaves deste trabalho: a localização do bico e a detecção da formação e queda da primeira gota. A Figura 4.4 apresenta uma foto da interface do sistema.

4.1.5 Localização do Bico

Conforme mencionado anteriormente, a localização do bico do qual se desprendem as gotas de óleo durante a destilação é de fundamental importância para o sucesso da aplicação, visto que os algoritmos de tratamento de imagens tendem a ser bastante custosos em termos computacionais. Com a localização precisa da região do bico, torna-se possível uma aplicação mais específica e localizada dos algoritmos de detecção da queda da gota. Devido a própria natureza do sistema, que deve atuar a uma taxa de captura bastante elevada, tal operação é considerada primordial.

A Figura 4.5 apresenta o bico do aparelho de destilação em mais detalhes. Repara-se que o bico é interno ao artefato, havendo ainda uma outra cobertura de vidro externa ao mesmo, o que dificulta ainda mais sua localização, devido a manchas que surgem ao longo do tempo nesta cobertura externa.

A estratégia adotada para localizar o bico lança mão de técnicas tradicionais

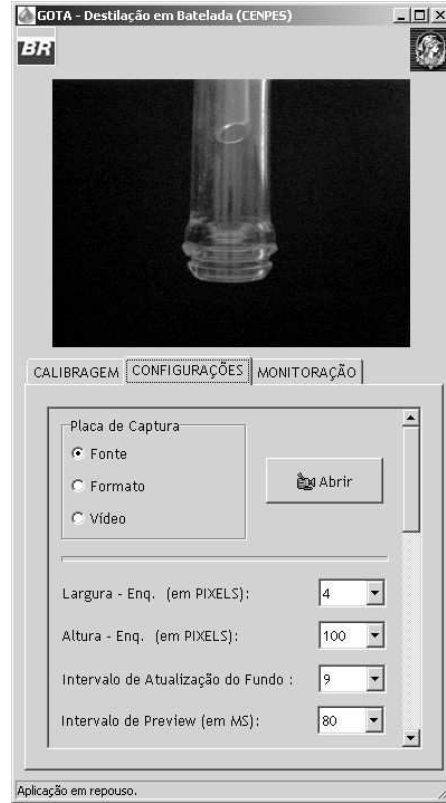


Figura 4.4: Foto da interface do sistema de detecção da queda da primeira gota no processo de destilação de petróleo.

de processamento de imagens: limiarização, rotulação de componentes conexas e projeções vertical e horizontal (Gonzalez & Woods, 2000; Jain, 1988). Com a iluminação gerada pela fonte de luz, instalada de forma a iluminar artificialmente a região do bico, as paredes do tubo interno refletem uma quantidade de luz suficiente para ressaltá-lo em meio ao campo de visão da câmera, conforme mostra a Figura 4.6.

Considerando que

$$0 \leq f(x, y) \leq 255, \quad (4.1)$$

sendo $f(x, y)$ o valor da luminância do pixel na posição (x, y) da imagem, utilizando-se um limiar relativamente alto, de aproximadamente 250, elimina-se boa parte dos outros elementos contidos na imagem. O bico corresponde à extremidade (borda) da maior região conexa de todas as remanescentes na imagem limiarizada, conforme

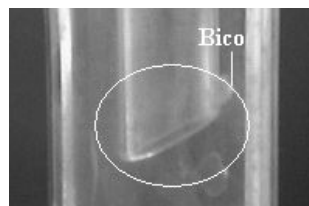


Figura 4.5: Imagem em detalhe do bico do aparelho de destilação.



Figura 4.6: Bico do aparelho de destilação iluminado.



Figura 4.7: Imagem limiarizada do bico do aparelho de destilação.

mostra a Figura 4.7.

Com uma pequena modificação no algoritmo de rotulação das componentes conexas de uma imagem sugerido por (Gonzalez & Woods, 2000) torna-se possível, além de rotular as diferentes regiões conexas, identificar aquela com maior área, que, neste caso, corresponde ao tubo interno do dispositivo.

De posse de um rótulo para cada região conexa e do rótulo da região com maior número de *pixels*, o que pode ser obtido com uma pequena modificação no algoritmo de rotulação das componentes conexas da imagem, é possível destacar a maior região de todas, correspondente ao bico. O resultado desta operação pode ser visto na Figura 4.8.



Figura 4.8: Região conexa correspondente ao bico encontrado.

Conforme pode-se ver, respectivamente, nas Figuras 4.9 e 4.10, realizando as projeções vertical e horizontal, torna-se possível identificar as bordas do tubo interno

e, em consequência, determinar a posição do bico (Figura 4.11). A Figura 4.12 mostra a região a ser monitorada pelo algoritmo de detecção da formação e queda da primeira gota, definida como um retângulo que abrange a região por onde passa a gota ao se desprender do bico.

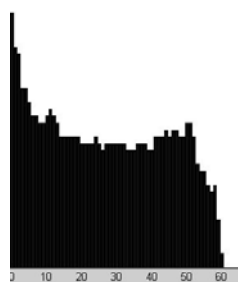


Figura 4.9: Projeção vertical do bico separado.

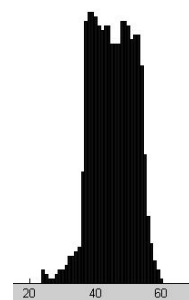


Figura 4.10: Projeção horizontal do bico separado.

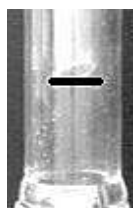


Figura 4.11: Localização do bico.

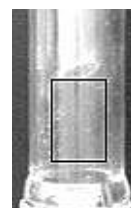


Figura 4.12: Marcação da região de interesse.

4.1.6 Detecção da Primeira Gota

Antes de apresentar a estratégia empregada na detecção da queda da primeira gota, convém mencionar que o ambiente monitorado pelo sistema é controlado, ou seja, sua situação de repouso é conhecida. Sendo assim, qualquer movimentação, exceto possíveis tremidas de câmera ou reflexos de objetos ou pessoas no vidro, pode ser tomada como formação de gota.

O sistema funciona em três estados possíveis. São eles: a) Pré-gota, correspondente ao período anterior à formação da gota; b) Formação-gota, referente ao período em que a primeira gota aparece nas imagens obtidas pelo sistema; e c) Pós-gota, período de repouso após o aparecimento da gota, indicando que a gota se desprendeu do bico. Sendo assim, cada quadro é analisado e ao mesmo tempo é atribuído um destes

três estados, baseado na classificação do mesmo como um quadro de movimento ou de repouso e na classificação do quadro anterior. Para classificar os quadros como situações de movimento ou repouso, utilizou-se a técnica de detecção de movimento Comparações Simples, apresentada na Seção 3.10.1.

A princípio, no momento em que se iniciava a monitoração, o primeiro quadro obtido pelo sistema era definido como a imagem base (fundo), sendo todo quadro subsequente classificado como situação de movimento ou não. A Tabela 4.1 apresenta a relação entre o estado do quadro imediatamente anterior e a classificação do atual para determinação do seu estado.

Est(Anterior)	Cl(Atual)	Est(Atual)
PRÉ	MOV	FORMAÇÃO
PRÉ	REP	PRÉ
FORMAÇÃO	MOV	FORMAÇÃO
FORMAÇÃO	REP	PÓS
PÓS	MOV	PÓS
PÓS	REP	PÓS

Tabela 4.1: Relação de classificação entre dois quadros subsequentes.

O desprendimento da gota é caracterizado pela passagem do estado Formação-gota para o estado Pós-Gota. No momento em que essa mudança é constatada, o sistema envia um sinal elétrico ao atuador para que o mesmo possa dar início a seus procedimentos de controle do processo.

A Figura 4.13 apresenta três quadros correspondentes à queda de massa líquida a partir do bico da unidade. A Figura 4.14 mostra os resultados da detecção de movimento nos mesmos.

Os pontos isolados no alto das imagens são ruídos gerados pela própria câmera, problema que é tratado pelo emprego do limite mínimo para o número de *pixels* diferentes para que o quadro seja considerado como situação de movimento.

4.1.7 Trepidações e Reflexos

Os dois principais problemas enfrentados durante o projeto eram provenientes de dois fenômenos inerentes ao ambiente: a) trepidação da câmera, causada por esbarrões de transeuntes e b) reflexão de objetos e pessoas no vidro que envolve o bico da

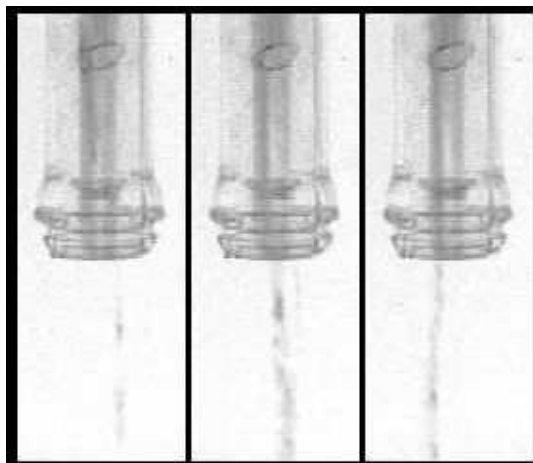


Figura 4.13: Quadros correspondentes à queda de massa líquida.

unidade de destilação.

O problema da trepidação da câmera foi tratado com uma pequena alteração na técnica Comparações Simples. Ao invés de comparar cada *pixel* da imagem atual com seu correspondente na imagem de fundo, compara-se o mesmo com seu correspondente e os vizinhos-de-8 (CASTLEMAN, 1995) do mesmo. Desta forma, reduz-se a sensibilidade do sistema para o deslocamento de uma linha de *pixels* em qualquer sentido ou direção. Para que seja considerado diferente, o *pixel* em questão deve ser suficientemente diferente dos 9 *pixels* mais próximos a ele na imagem de fundo.

Já o problema da reflexão de objetos e pessoas foi abordado com iluminação. A fonte de luz foi posicionada no lado oposto ao fundo, próxima ao vidro. Devido a esta proximidade, a luz incidente na unidade ofusca qualquer possível reflexo a partir do vidro, impedindo que o mesmo seja confundido com o desprendimento de uma gota a partir do bico.

4.1.8 Resultados

A primeira versão do sistema atuava a uma taxa de 4 quadros por segundo, a qual se mostrava absolutamente insuficiente para a realização da tarefa em questão. Um grande esforço de programação foi realizado com o intuito de aumentar a eficiência das rotinas e permitir que o sistema atingisse uma taxa de captura satisfatória. Conseguiu-se então atingir a taxa de 10 quadros por segundo, na qual ocorreu uma

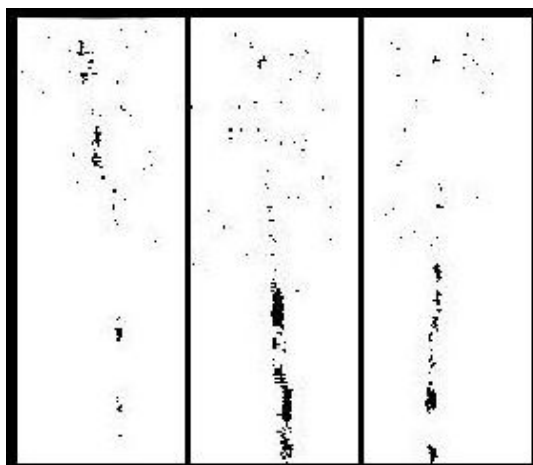


Figura 4.14: Resultados da aplicação da rotina de detecção de movimento nos quadros da Figura 4.13.

estagnação por um período de aproximadamente um mês, durante o qual foram tentadas diversas modificações nos procedimentos do sistema em busca de um ganho significativo na taxa de captura de quadros sem sucesso.

A partir dos resultados dos primeiros testes (Tabela 4.2 / Máquina I), verificou-se que a taxa de captura até então atingida não era satisfatória para a realização da referida tarefa. Foi tentada então a utilização de uma máquina de configuração superior àquela na qual os testes iniciais haviam sido realizados, de forma a descobrir se o baixo desempenho do sistema decorria da qualidade do software ou do hardware.

A segunda bateria de testes mostrou, através de seus resultados (Tabela 4.2 / Máquina II), que o hardware utilizado nos primeiros testes era responsável pela taxa de captura insuficiente que os mesmos apresentaram, o que nos levou a concluir que a qualidade do software era satisfatória.

Nº Máquina	I	II
Config.	DURON 1.0	PENTIUM IV 1.8
Quad./Seg.	10	15
nº Testes	100	100
nº Acertos	98	100

Tabela 4.2: Resultados dos testes realizados para o sistema.

No primeiro teste realizado (Máquina I), das 100 tentativas, apenas duas não foram bem sucedidas. Como provável explicação para as duas falhas, acreditou-

se na perda da gota pelo baixo número de quadros capturados por segundo pelo sistema. Dada a velocidade do desprendimento da gota, a taxa de 10 quadros por segundo não seria suficiente para garantir que a queda da mesma fosse registrada pelo sistema. Sendo assim, resolveu-se tentar melhorar a configuração da máquina de teste, passando de um DURON 1.0 GHz com 256 Mb de memória para um Pentium IV 1.8 GHz com memória de 512 Mb (Máquina II). A mudança surtiu o efeito desejado, de forma que os mesmos 100 testes foram realizados, obtendo-se, desta vez, desempenho de 100 %.

No Apêndice III são apresentadas imagens obtidas na unidade de destilação para a qual o sistema foi concebido, assim como resultados da aplicação de algumas técnicas mencionadas anteriormente na tentativa de se detectar a queda da primeira gota.

4.2 2º Caso: Controle do Nível de Petróleo em Frasco Destino de um Processo de Destilação em Batelada

4.2.1 Introdução

O segundo estudo de caso realizado neste trabalho corresponde a uma extensão do primeiro apresentado. O objetivo é a construção de um sistema capaz de monitorar a variação do volume de petróleo armazenado no frasco coletor da unidade de destilação apresentada no estudo de caso anterior. Com isso, ganha-se a capacidade de se manter um controle mais específico sobre o volume produzido de cada uma das frações da amostra de petróleo no decorrer do processo de destilação.

O frasco (Figura 4.15) tem formato cilíndrico, sendo sua altura e a distância entre o mesmo e a câmera valores conhecidos, além de possuir marcas brancas indicadoras de seus limites inferior e superior. Ao fundo, encontra-se um anteparo de cor negra e material fosco. O problema consiste, mais especificamente, da localização do frasco, detecção das marcas brancas, conseqüentemente determinando suas extremidades, e posteriormente, a monitoração do movimento da superfície do líquido acumulado, através da qual calcula-se, em tempo real, seu volume, conferindo ao sistema a capacidade de controlar a variação desta grandeza durante o processo de destilação.



Figura 4.15: Imagem do frasco utilizado nos testes em laboratório para o sistema de monitoração da variação de volume no processo de destilação de petróleo em batelada.

4.2.2 Categorias do Problema

Este problema se divide em duas categorias, determinadas pelas condições espaciais do cenário em questão, apresentado na Figura 4.16. A câmera encontra-se a uma distância conhecida d do frasco, localizada ao nível do fundo do mesmo.

A primeira categoria, na qual se enquadra este estudo de caso, corresponde à situação em que a distância d é grande o suficiente para que não seja necessário tratar a distorção do frasco provocada pela projeção do ambiente 3-D em um espaço bidimensional.

A segunda categoria é aquela na qual a relação *pixel/distância* é variável ao longo do frasco, devido à distorção mencionada no parágrafo anterior. Embora as distâncias $y_1 - y_0$ e $y_3 - y_2$ sejam iguais, elas diferem no número de *pixels* que as representam na imagem, por estarem localizadas em regiões distintas do frasco. Existem duas formas de se abordar este problema, as quais serão brevemente apresentadas ao final deste estudo de caso, embora não se enquadrem especificamente no mesmo. Tal apresentação será feita em virtude da importância deste aspecto dentro de problemas da mesma classe.

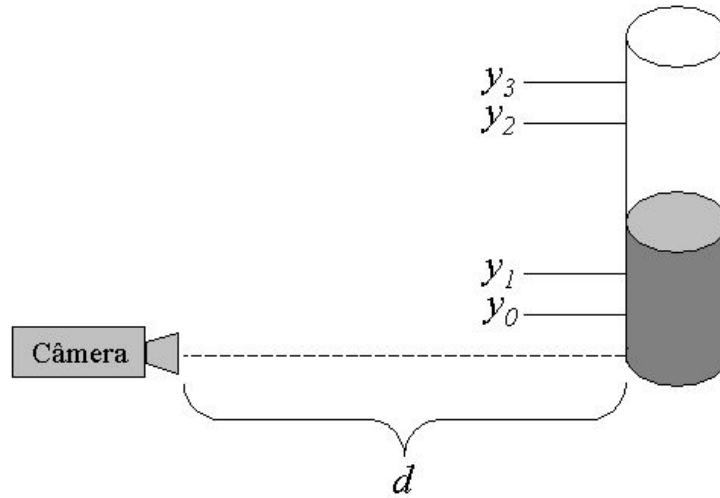


Figura 4.16: Cenário básico do sistema de controle do nível de líquido no frasco da unidade de destilação.

4.2.3 Posicionamento da Câmera

A câmera localiza-se ao fundo do frasco, para minimizar o número de casos diferentes com os quais o sistema terá que lidar.

Se estivesse localizada no meio do frasco, seria necessário abordar dois tipos de visão, sendo uma correspondente aos momentos em que a superfície do líquido se encontra abaixo do nível da câmera, e a outra aos que a superfície se encontra acima do mesmo. É importante ressaltar que as duas situações são extremamente distintas, já que, na primeira, enxerga-se toda a superfície do líquido (Figura 4.17.a), enquanto na segunda enxerga-se apenas a borda da superfície (Figura 4.17.b). Para que se pudesse evitar a necessidade de tratar duas situações diferentes, concluiu-se que o mais indicado seria a definição da posição da câmera ao nível de uma das extremidades do frasco.

A extremidade inferior foi escolhida por motivos de simplicidade espacial. No caso em que a distância entre câmera e frasco é grande o suficiente para que possíveis distorções sejam desprezadas, não haveria grande diferença entre posicionar a câmera na extremidade superior ou inferior. Porém, a montagem do experimento em laboratório é extremamente facilitada pela escolha da última. Desta forma, determinou-se que a câmera seria posicionada ao nível do fundo do frasco.

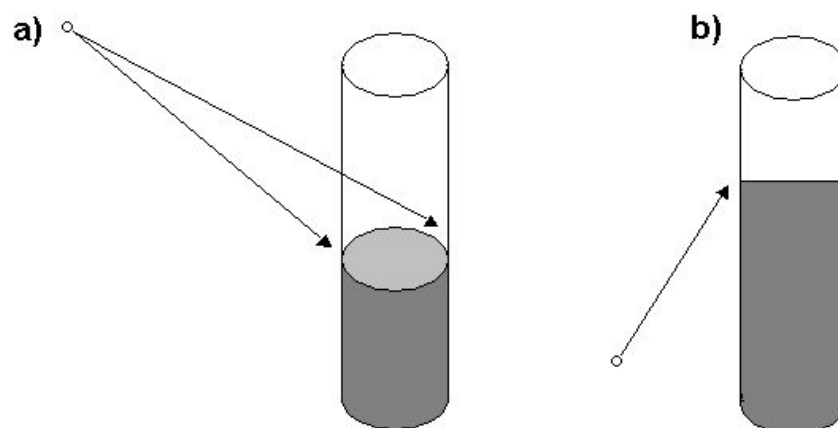


Figura 4.17: Diferentes tipos de visão do líquido baseados no posicionamento da câmera.

4.2.4 Sub-tarefas do Sistema

A tarefa chave do sistema é baseada na realização de três sub-tarefas, as quais são abordadas separadamente nas seções seguintes. São elas: a) localização das paredes do frasco, b) localização das marcas delimitadoras dos limites inferior e superior do frasco, e c) determinação da altura da superfície do líquido acumulado no frasco. Obtendo-se sucesso na realização destas tarefas, assunto das seções seguintes, a construção do sistema torna-se meramente a organização das mesmas de forma a atender à finalidade principal do estudo de caso.

4.2.5 Localização das Paredes do Frasco

A primeira tarefa do sistema consiste da determinação dos limites laterais do frasco no qual o petróleo é acumulado. Conforme mencionado anteriormente, ao fundo, encontra-se um anteparo de cor negra e material fosco, o qual deve ter altura e largura maiores do que o frasco. No cenário ideal, a imagem captada pelo sistema inclui apenas o frasco e o anteparo ao fundo, não havendo outros objetos enquadrados que possam dificultar o destacamento do recipiente. Para tanto, o sistema deve permitir ao usuário que indique, com o uso do *mouse*, a região onde se encontra o frasco (ver Figura 4.18).

O sistema recorta a região indicada pelo usuário e utiliza a mesma como imagem base para determinação dos limites laterais. Em seguida, a imagem é submetida

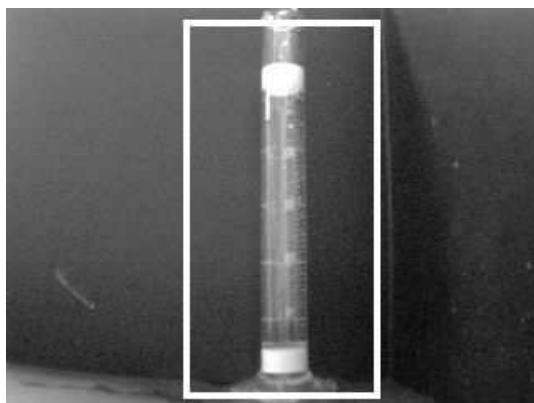


Figura 4.18: Região do frasco destacada com o uso do *mouse* pelo usuário.

à limiarização simples, sendo o limiar calculado como a média da luminância de todos os *pixels* não brancos (Figura 4.19). Ao resultado deste processo é aplicado um filtro passa-baixa, que tem por objetivo remover da imagem pequenos ruídos, que, por sua grande quantidade, atrapalhariam a localização do frasco, e tornar a imagem mais homogênea, conforme pode ser visto na Figura 4.20.

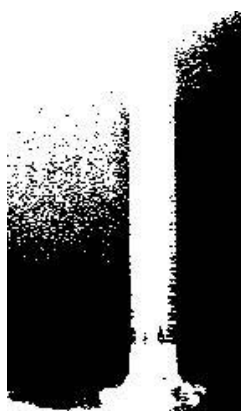


Figura 4.19: Resultado do processo de limiarização simples aplicado à região de interesse indicada na Figura 4.18.



Figura 4.20: Aplicação de um filtro passa-baixa para remoção de ruídos da Figura 4.19.

O próximo passo corresponde a aplicação do algoritmo de rotulação de componentes conexas (Seção 3.9), e a posterior segmentação da maior região identificada pelo mesmo, que corresponde, neste caso, ao frasco. Com isto, alguns dos ruídos que não haviam sido removidos pelo filtro-passa baixa são finalmente removidos

(ver Figura 4.21). Aplica-se então um filtro passa-alta para determinar os contornos do frasco. É aplicada ainda, sobre a imagem que contém os contornos do frasco, uma dilatação da imagem, de forma que todo pixel branco tem seus vizinhos-de-8 pintados de branco (ver Figura 4.22).



Figura 4.21: Imagem contendo o resultado da aplicação do algoritmo de rotulação de componentes conexas na Figura 4.20.

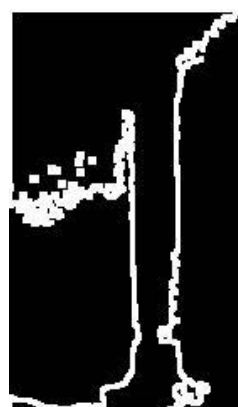


Figura 4.22: Figura 4.21 após processo de dilatação.

Foram concebidos dois métodos para determinar os limites laterais do frasco a partir da imagem obtida no passo anterior. O sistema executa ambos e toma como resultado a interseção entre os resultados apresentados pelos mesmos. Nas seções seguintes, os dois métodos são apresentados, assim como seus resultados para o caso-exemplo.

1º Método

Neste método a imagem é subdividida em partições de tamanhos iguais (Figura 4.23), com exceção da última, que pode ter largura menor que as demais. A largura das partições utilizada no trabalho foi de 10 *pixels*. Outros valores foram testados, como 15, 20 e até 30 *pixels*, porém, os resultados mais precisos foram obtidos com largura 10. Em cada uma destas partições são contabilizados os *pixels* de cor branca. Um gráfico ilustrando o resultado desta contagem pode ser visto na Figura 4.24, no qual cada valor do eixo *x* corresponde ao fim de cada partição, enquanto os valores no eixo *y* correspondem ao número de *pixels* brancos em cada partição.

O objetivo do próximo passo do método é determinar quais partições da imagem

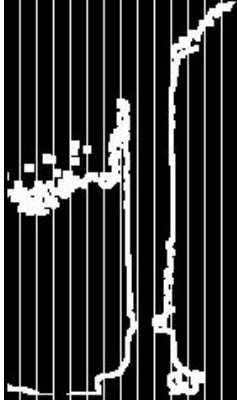


Figura 4.23: Imagem contendo os contornos do frasco dividida em partições de largura 10.

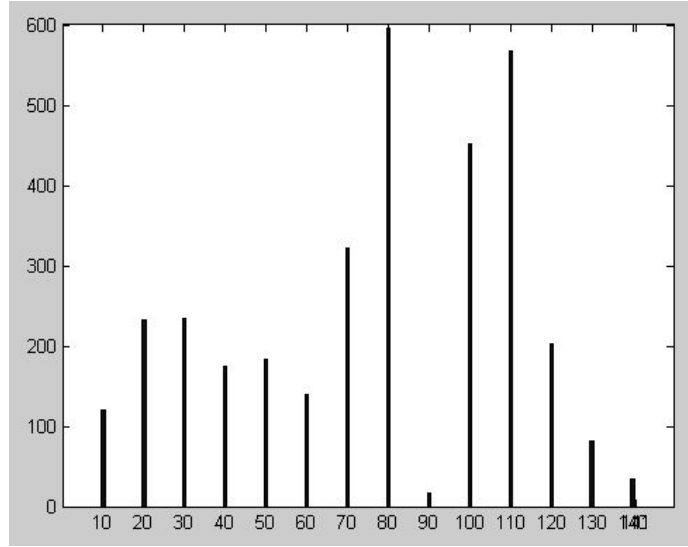


Figura 4.24: Gráfico que ilustra a relação entre cada uma das partições apresentadas na Figura 4.23 e o número de *pixels* brancos contidos em cada uma.

são regiões candidatas a conterem os limites laterais do vidro. Para tanto, realiza-se a seguinte operação: seja m a média das quantidades de *pixels* brancos em cada partição, d o desvio padrão para a mesma coleção de valores e q_i a quantidade de *pixels* brancos encontrados na partição i . São consideradas regiões candidatas aquelas que atenderem ao critério apresentado na Equação 4.2, o qual foi obtido de forma empírica.

$$q_i \geq m + \frac{d}{2} \quad (4.2)$$

O critério foi baseado no fato de que, pela orientação vertical dos limites laterais do vidro, as partições que contivessem *pixels* pertencentes a eles teriam, obrigatoriamente, um alto número de *pixels* de cor branca. Para a Figura 4.23, as partições 7, 8, 10 e 11, ressaltadas na Figura 4.25, foram apontadas como candidatas pelo método.

A seguir é apresentado o 2º método, desenvolvido na tentativa de resolver algumas deficiências do 1º.

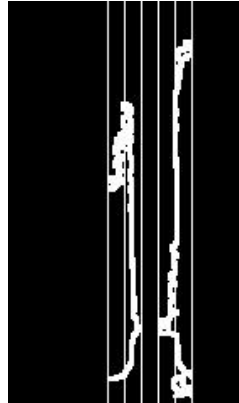


Figura 4.25: Partições candidatas eleitas pelo método 1.

2º Método

O segundo método foi concebido para corrigir possíveis erros do primeiro, que ocorriam devido a ruídos muito concentrados em determinadas regiões, que não só não eram removidos pelo filtro passa-baixa, como eram também ressaltados pelo processo de dilatação, conforme ocorreu na região 7 do exemplo mostrado na Figura 4.25.

Esta segunda abordagem baseia-se no fato de que as bordas do frasco devem apresentar um comprimento vertical considerável dentro da partição nas quais se encontram. Sendo assim, realiza-se uma projeção vertical das partições e verifica-se a proporção de valores diferentes de zero nas mesmas. Observe o exemplo mostrado nas Figuras 4.26 (projeção da partição 7) e 4.27 (projeção da partição 8). Repare que a proporção de valores diferentes de zero na projeção da partição 8 é bem maior do que na partição 7, cuja projeção apresentou-se mais esparsa com relação a valores não nulos.

Para o exemplo em questão, as proporções de valores não nulos para as partições 7 e 8 são, respectivamente, 26 % e 68 %. Segundo este método, os candidatos a conter uma das bordas, considerando p_i a proporção de valores não nulos na partição i , m a média das projeções e d o seu desvio padrão, são aqueles que repetitem a condição apresentada na Equação 4.3, a qual também foi obtida de forma empírica.

$$p_i \geq m + \frac{d}{2} \quad (4.3)$$

Os candidatos determinados pelo método 2 para o exemplo em questão foram as

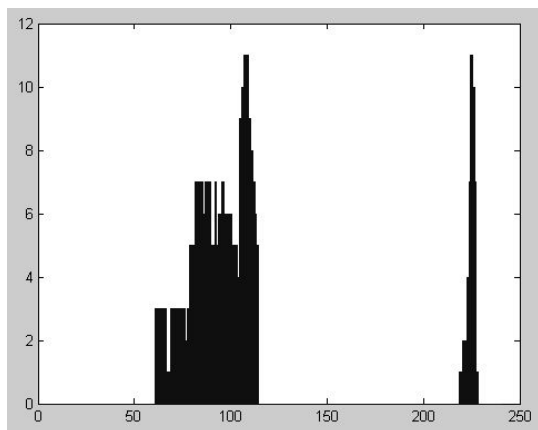


Figura 4.26: Projeção vertical da partição de número 7.

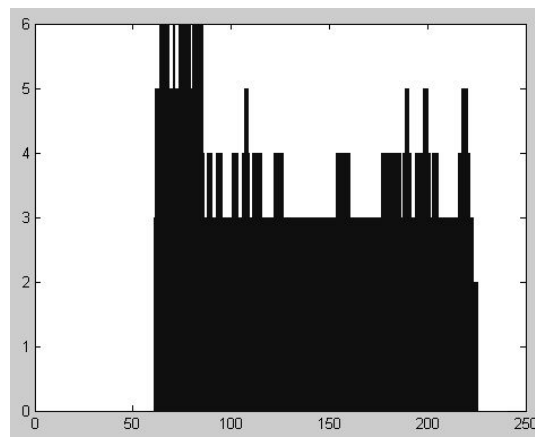


Figura 4.27: Projeção vertical da partição de número 8.

partições 8, 10 e 11. Conforme foi mencionado anteriormente, realiza-se a interseção entre as candidatas dos dois métodos, tomando-se o resultado desta interseção como as partições que contém as bordas do frasco, no caso 8, 10 e 11. Partições adjacentes, como 10 e 11 no exemplo, são unificadas. Caso haja apenas uma ou mais do que duas partições não adjacentes no resultado, considera-se que o método falhou em localizar os limites laterais do frasco, devendo ser executado novamente com uma leve variação de seus parâmetros. Em último caso, o sistema oferece controles manuais para que o próprio operador possa informar os limites do frasco.

De posse de duas regiões candidatas, calcula-se, para cada uma, o centro de massa dos *pixels* de cor branca e toma-se suas ordenadas como limites laterais do frasco, conforme mostrado na Figura 4.28.

4.2.6 Localização das Marcas

Conforme mencionado anteriormente, o frasco deve ter um par de marcas brancas, indicando os limites inferior e superior do frasco (ver Figura 4.29). Antes de mostrar como são localizadas estas marcas, é importante ressaltar que, a esta altura, o sistema já determinou os limites laterais do tubo, podendo então recortar a região de interesse (o frasco propriamente dito), reduzindo, desta forma, a complexidade da localização das marcas. A Figura 4.30 mostra a região de interesse já separada do restante da imagem original.

O primeiro passo corresponde à expansão do histograma da imagem, para me-

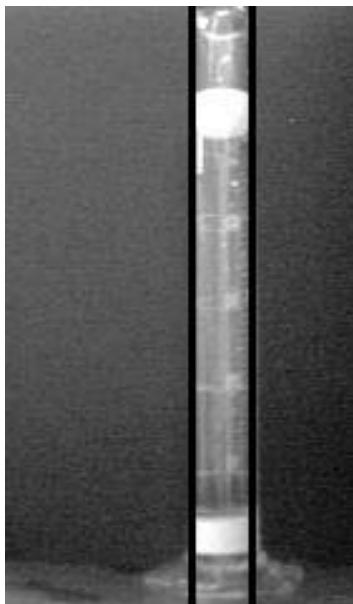


Figura 4.28: Limites laterais do frasco de coleta de petróleo.

lhorar a qualidade da mesma (Figura 4.31.a) e facilitar a localização das marcas. Em seguida, limiariza-se a imagem (Figura 4.31.b) e aplica-se um filtro de mediana para remover ruídos (Figura 4.31.c).

As marcas brancas correspondem às maiores regiões remanescentes na imagem. Para localizá-las, realiza-se uma projeção vertical da imagem (ver Figura 4.32.a), removendo-se em seguida as regiões de pequena representatividade na projeção (ver Figura 4.32.b).

Os limites internos das duas regiões remanescentes são os limites inferior e superior do frasco. Pode-se então recortar da imagem as regiões desnecessárias para a tarefa de monitorar o nível de líquido, localizadas acima da marca superior e abaixo da marca inferior.

4.2.7 Determinação do Nível do Líquido

A tarefa chave deste sistema consiste da determinação da altura em que se encontra a superfície do líquido armazenado no frasco. Sendo atingido este objetivo, o cálculo de valores como o volume, a velocidade de deslocamento da superfície do líquido e a própria vazão da destilação torna-se imediato, dado que, conforme mencionado anteriormente, diâmetro e altura do frascos são conhecidos *a priori*.



Figura 4.29: Marcas indicadoras dos limites inferior e superior do frasco.



Figura 4.30: Frasco segmentado do restante da imagem.

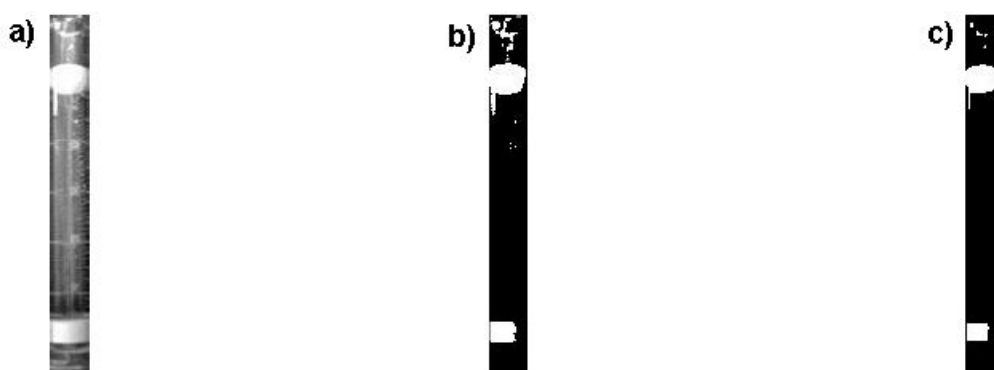


Figura 4.31: Imagens do processo de localização das marcas indicadoras dos limites inferior e superior do frasco.

A estratégia concebida neste trabalho para realizar tal tarefa é baseada na técnica de detecção de movimento abordada na seção 3.10.1 (Comparação Simples). No início do processo, toma-se uma imagem vazia do frasco como referência para a detecção de movimento. A cada nova imagem obtida pelo sistema através da placa de captura, realiza-se uma comparação entre esta e a imagem base, conforme pode ser visto na Figura 4.33.c, que corresponde ao resultado da aplicação da técnica mencionada nas Figuras 4.33.a e 4.33.b.

Em seguida, realiza-se a projeção vertical, mostrada na Figura 4.34, do resultado da comparação para que se possa identificar a região na qual se encontra a superfície do líquido.

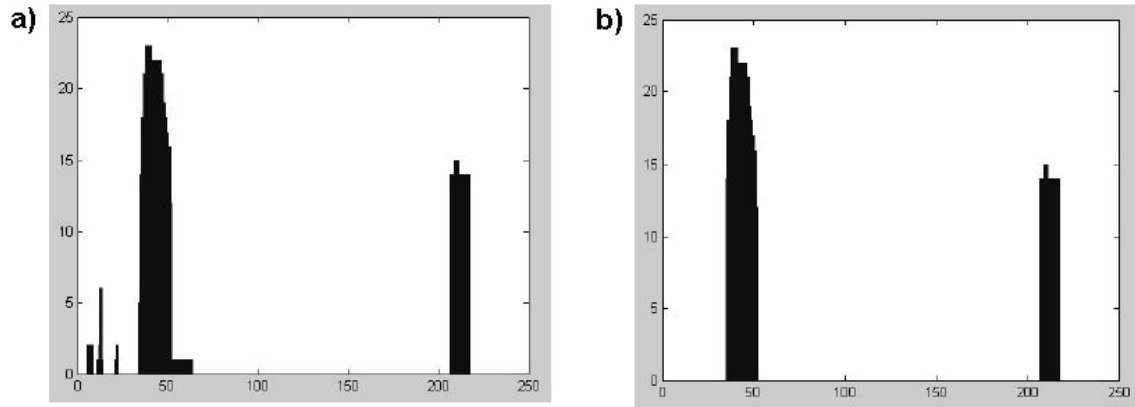


Figura 4.32: Projeção vertical da Figura 4.31.c e a posterior remoção de regiões de pouca relevância para localização das marcas.

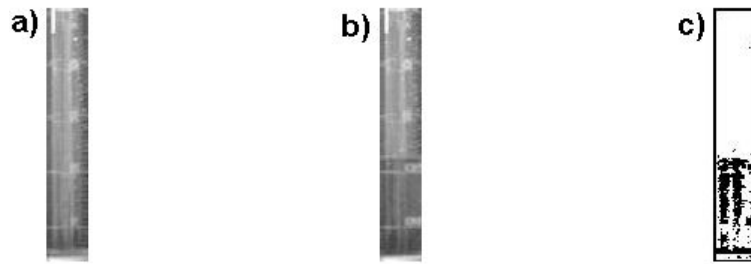


Figura 4.33: Imagens do processo de determinação do nível de líquido armazenado no frasco.

Para determinar a região na qual se encontra a superfície do líquido, calcula-se, primeiro, a média m e o desvio padrão d da projeção realizada. Varre-se então o resultado da projeção a partir de zero até que se encontre um ponto y , sendo $p(y)$ o valor da projeção em y , que atenda à condição expressa na Equação 4.4.

$$p(y) > m + d \quad (4.4)$$

Sendo lim_sup e lim_inf os limites superior e inferior, respectivamente, da região que contém a superfície do líquido, define-se os mesmos conforme indicado abaixo.

$$\begin{aligned} lim_sup &= y - \delta \\ lim_inf &= y + \delta \end{aligned}$$

onde δ é um valor que determina a abrangência da região a ser destacada, de forma que, quanto maior o valor de δ , maior será a área da região.

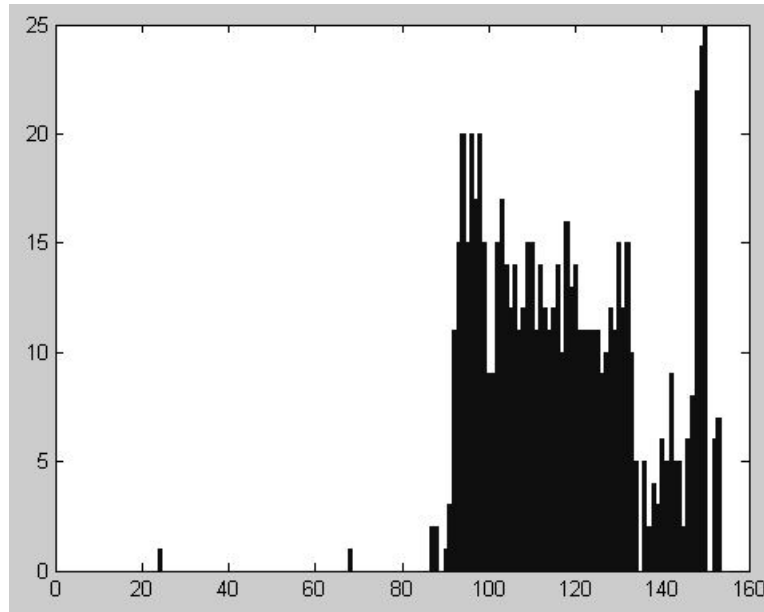


Figura 4.34: Projeção vertical do resultado da aplicação da técnica Comparação Simples na Figura 4.33.c.

Neste trabalho utilizou-se $\delta = 10$, valor este obtido de forma empírica.

De posse das coordenadas desta nova região de interesse, recorta-se a mesma a partir da imagem gerada pela técnica da Comparação Simples (Figura 4.33.c), que na verdade contém a diferença entre a imagem base e a imagem atual, e aplica-se novamente a projeção vertical, para então determinar o ponto correspondente à localidade da superfície do líquido com maior acurácia, já que, com a restrição do espaço de busca, é consideravelmente reduzida a possibilidade de ruídos isolados dificultarem a localização da região de interesse. A Figura 4.35.a mostra a nova região de interesse acima mencionada. Já a projeção da mesma é vista na Figura 4.35.b.

Novamente varre-se a projeção a partir de zero e define-se como posição da superfície o primeiro ponto y para o qual o valor da projeção $p(y)$ respeita a condição apresentada na Equação 4.4. A Figura 4.36 mostra a localização da superfície para o exemplo utilizado ao longo deste estudo de caso.

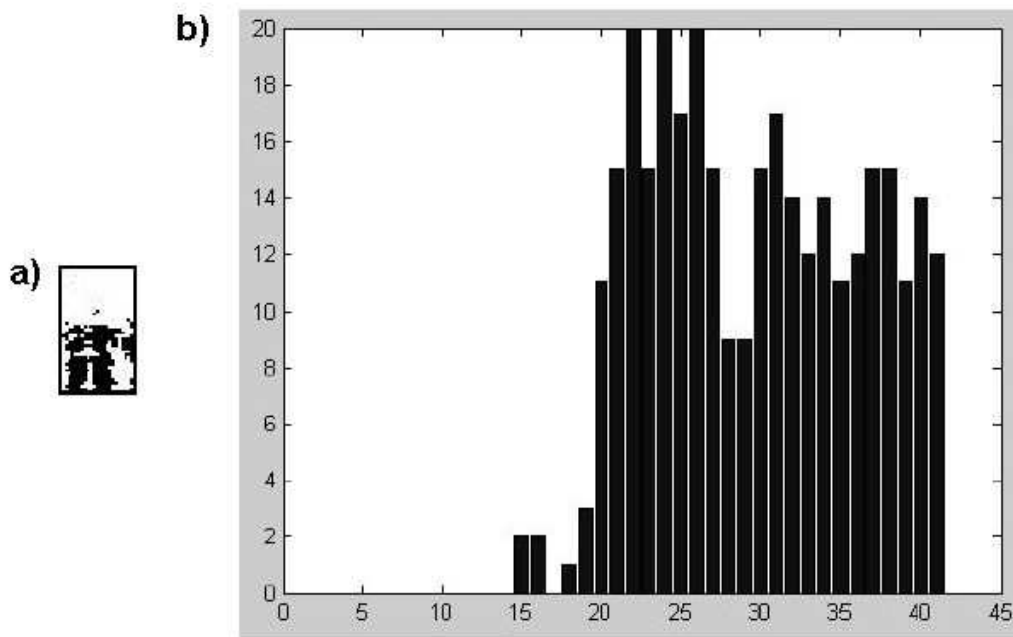


Figura 4.35: a) Região contendo a superfície do líquido segmentada do restante do frasco e b) projeção vertical da mesma.

4.2.8 Tratamento de Distorções Projetivas

Conforme mencionado anteriormente, dependendo da distância e do ângulo entre a câmera e o frasco, a relação número de *pixels*-distância real pode se tornar variável ao longo da imagem. É importante enfatizar que o estudo de caso da determinação do nível realizado neste trabalho não é afetado por esse tipo de problema, já que a distância entre a câmera e o frasco é suficientemente grande para que se descarte tal ocorrência. A abordagem do referido problema se dá devido a sua grande relevância para casos de naturezas similares ao apresentado neste texto.

São sugeridas duas abordagens para o problema em questão, sendo a primeira a geração de uma imagem corrigida, na qual as distorções de inclinação e distância são atenuadas em altíssimo grau, e a segunda o emprego de um modelo neural que aprenda a realizar o mapeamento número de *pixels*-distância de acordo com a região da imagem.

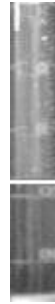


Figura 4.36: Localização da superfície do líquido determinada pela estratégia apresentada.

1ª Abordagem - Correção da Imagem

A abordagem apresentada nesta seção foi proposta em (HEMERLY & MICOSKY, 2004), e consiste da determinação de um conjunto de parâmetros do cenário real da filmagem via o emprego de um método de otimização e a posterior utilização desses parâmetros na construção de uma imagem na qual as distorções anteriormente presentes são atenuadas, sendo qualquer erro residual na relação número de *pixels*-distância decorrente do caráter discreto da imagem.

Os parâmetros a serem determinados pela técnica em questão são a altura da câmera em relação ao plano de filmagem (h), a distância focal (f) e os ângulos de *tilt* (τ) e *pan* (θ) da câmera. A Figura 4.37 apresenta um cenário ilustrativo, através do qual são apresentados outros elementos essenciais para a compreensão do método.

Três sistemas de coordenadas são utilizados, sendo eles:

- C - Sistema 3D, formado pelos eixos x_c, y_c, z_c , cuja origem está localizada na câmera.
- W - Sistema 3D, formado pelos eixos x_w, y_w, z_w , cuja origem está localizada no plano, representando o sistema de coordenadas do mundo real.
- I - Sistema 2D, formado pelos eixos x_i, y_i , representando o sistema de coordenadas da imagem.

A relação entre as coordenadas dos sistemas C e I é obtida a partir do sistema *pin hole* da câmera. Já a relação entre as coordenadas dos sistemas C e W são obtidas a partir de uma série de transformações geométricas que tem por intuito alinhar os sistemas em questão. Passada esta fase, torna-se possível relacionar variáveis

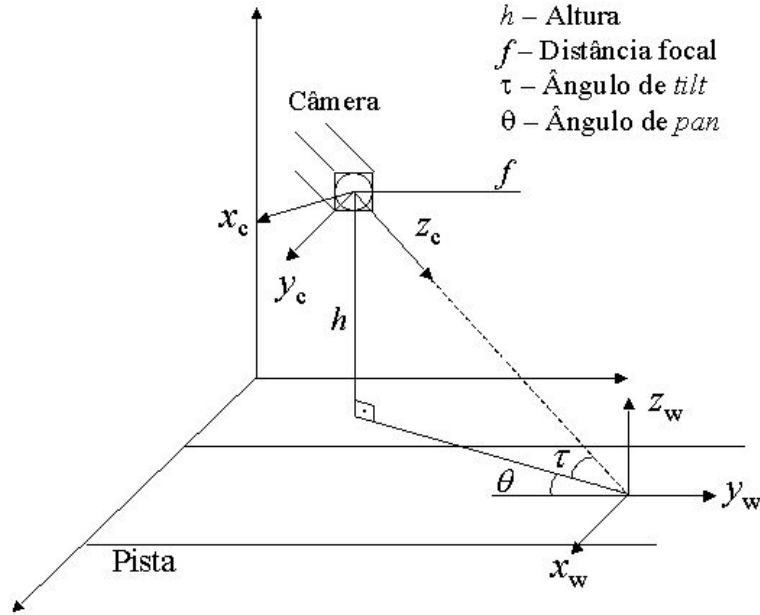


Figura 4.37: Cenário ilustrativo da primeira abordagem para tratamento das distorções projetivas da imagem.

dos espaços W e I . Observe que, neste método, assume-se que $z_w = 0$, ou seja, o interesse da filmagem reside apenas no plano representado pelo eixo (x_w, y_w) . Por fim, obtém-se as seguintes relações (HEMERLY & MICOSKY, 2004):

$$x_i = f \cdot \frac{x_w \cos \theta - y_w \sin \theta}{x_w \cos \tau \sin \theta + y_w \cos \tau \cos \theta + \frac{h}{\sin \tau}} \quad (4.5)$$

$$y_i = f \cdot \frac{-x_w \sin \tau \sin \theta - y_w \sin \tau \cos \theta}{x_w \cos \tau \sin \theta + y_w \cos \tau \cos \theta + \frac{h}{\sin \tau}} \quad (4.6)$$

$$x_w = -h \cdot \frac{y_i \sin \theta + x_i \sin \tau \cos \theta}{\sin \tau \cdot (y_i \cos \tau + f \sin \tau)} \quad (4.7)$$

$$y_w = -h \cdot \frac{y_i \cos \theta + x_i \sin \tau \cos \theta}{\sin \tau \cdot (y_i \cos \tau + f \sin \tau)} \quad (4.8)$$

A técnica depende da demarcação, por parte de um usuário, de quatro pontos P, Q, R e S, de forma que \overline{PQ} e \overline{RS} formem segmentos de reta paralelos entre si no mundo real. De posse destes pontos, fica fácil calcular o ponto de fuga destes segmentos na imagem, denotado por $PF = (x_{pf}, y_{pf})$. A Figura 4.38 ilustra esse procedimento, considerando que os eixos (x_i, y_i) estão posicionados no centro

da imagem, e não no canto superior esquerdo, como de costume (HEMERLY & MICOSKY, 2004).

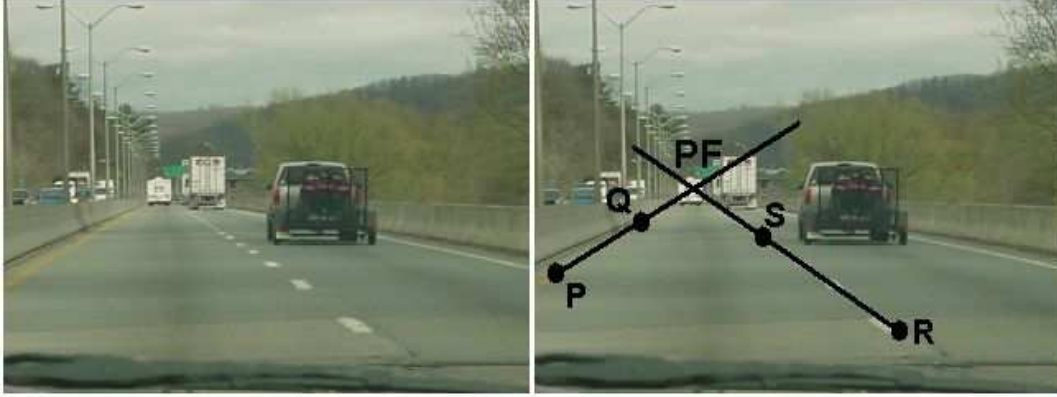


Figura 4.38: Determinação do ponto de fuga através das marcações dos segmentos de reta PQ e RS .

Tem-se então as seguintes relações (HEMERLY & MICOSKY, 2004):

$$x_{pf} = -f \cdot \frac{\tan \theta}{\cos \tau} \quad (4.9)$$

$$y_{pf} = -f \cdot \tan \tau \quad (4.10)$$

Com estas relações, torna-se simples o cálculo de valores para f e τ a partir de um valor inicial para θ .

A relação que envolve o parâmetro h é introduzida através da segunda, e última, intervenção do usuário no procedimento, e consiste da demarcação de dois segmentos de reta \overline{AB} e \overline{CD} cujos comprimentos L_{AB} e L_{CD} ele conhece e informa ao sistema (HEMERLY & MICOSKY, 2004).

Considerando-se que

$$L'_{AB} = \sqrt{(y_{wA} - y_{wB})^2 + (x_{wA} - x_{wB})^2}$$

$$L'_{CD} = \sqrt{(y_{wC} - y_{wD})^2 + (x_{wC} - x_{wD})^2}$$

cria-se uma função de erro definida da seguinte forma (HEMERLY & MICOSKY, 2004):

$$e = (L_{AB} - L'_{AB})^2 + (L_{CD} - L'_{CD})^2$$

Considerando que valores para y_{wA} , y_{wB} , x_{wA} , x_{wB} , y_{wC} , y_{wD} , x_{wC} e x_{wD} são calculado através das equações de retro-projeção (Equações 4.7 e 4.8), tem-se estabelecida uma relação entre h e os demais parâmetros mencionados anteriormente (HEMERLY & MICOSKY, 2004).

Estimando-se valores iniciais para h e τ e aplicando-se um algoritmo de otimização (no caso, o Nelder-Mead) tendo como base a função de erro e , encontra-se valores aproximados para f e θ , além de valores, também aproximados, para os parâmetros h e τ (HEMERLY & MICOSKY, 2004).

No passo seguinte, emprega-se os valores computados na geração de uma imagem retro-projetada a partir das Equações 4.7 e 4.8, sendo esta a imagem corrigida e praticamente livre de distorções projetivas. Considerando-se que a altura do frasco é conhecida, torna-se possível calcular o nível do óleo através da proporção da imagem do frasco ocupada pelo líquido.

2ª Abordagem - Modelo Neural

A 2ª abordagem consiste do emprego de uma rede neural do modelo MLP (Multi-Layer Perceptron) (RUMELHART, HINTON & WILLIAMS, 1986) de forma que a mesma aprenda a realizar o mapeamento direto entre altura em *pixels* e centímetros a partir da imagem original.

Para gerar a massa de dados a ser utilizada no treinamento da rede, tornam-se necessárias marcações ao longo do frasco, sendo a altura de cada marca conhecida *a priori*. O problema maior desta abordagem reside na determinação de quantas marcas devem ser feitas e o quão espaçadas as mesmas devem estar entre si para que se consiga representar com sucesso e acurácia o frasco. Não foi feito nenhum estudo a este respeito neste trabalho, e no exemplo, a ser demonstrado na seção seguinte, tanto a quantidade de marcas, quanto seu espaçamento, foram determinados de forma empírica.

Exemplos Práticos

Nesta seção são apresentadas aplicações das duas abordagens mencionadas na seção anterior. Os exemplos demonstrados são baseados na imagem de um frasco exibida

na Figura 4.39, sendo que as marcas horizontais, espaçadas de um centímetro, foram acrescentadas manualmente. As marcas são utilizadas em ambos os exemplos, seja para a determinação do comprimento dos segmentos \overline{AB} e \overline{CD} na correção da imagem, seja para a geração da massa de treinamento para a rede neural na segunda abordagem.

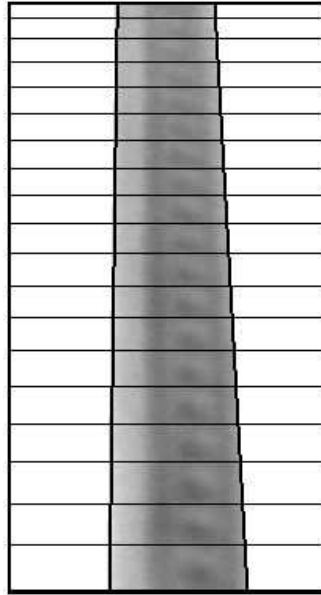


Figura 4.39: Imagem de um frasco na qual podem ser observadas as distorções projetivas.

Abordagem Corretiva Na abordagem corretiva, implementou-se a técnica de determinação dos parâmetros h , f , τ e θ com uma diferença em relação àquela apresentada em (HEMERLY & MICOSKY, 2004), que correspondeu à utilização, ao invés do método Nelder-Mead, de um algoritmo genético (HOLLAND, 1975).

Enquanto no método proposto em (HEMERLY & MICOSKY, 2004) havia a necessidade da estimativa de valores iniciais para os parâmetros h e τ , na implementação com o emprego de um algoritmo genético isso não ocorria, já que o próprio modelo de cromossomo utilizado representava os dois valores mencionados, os quais eram variados pelo próprio funcionamento do algoritmo.

Foi utilizado um modelo de cromossomo apenas com valores binários, sendo a primeira metade deles correspondente a um valor real representando a altura h em centímetros, e a segunda metade outro valor real para representar o ângulo τ em

Parâmetro	Valor
$h(cm)$	25
$f(pixels)$	300
$\tau(graus)$	80,34
$\theta(graus)$	0,83

Tabela 4.3: Resultados obtidos pela abordagem corretiva para o problema das distorções projetivas.

graus. Cada um destes valores era formado por 19 bits, sendo 9 para a parte inteira e 10 para a parte decimal, totalizando 38 bits como comprimento do cromossomo. O algoritmo foi executado 10 vezes, de forma que cada execução foi delimitada por 200 gerações, sendo de 100 cromossomos o tamanho das populações. Ao final, tomou-se a melhor solução dentre aquelas encontradas por cada execução. É importante ressaltar que os valores mencionados acima foram obtidos de forma empírica.

O emprego de um algoritmo genético na implementação do método proposto em (HEMERLY & MICOSKY, 2004) conferiu ao mesmo, como vantagem, a eliminação da necessidade de se estimar valores iniciais para os parâmetros h e τ . Como desvantagem, considerando-se a natureza aleatória do algoritmo genético, fez-se necessário que se executasse o algoritmo por diversas vezes, ocasionando uma razoável queda no desempenho temporal do método. Porém, sabendo-se que o mesmo é executado apenas uma vez, a perda em desempenho não apresenta gravidade considerável. O algoritmo de geração da imagem retro-projetada foi obtido através de (HEMERLY, 2004).

Na obtenção da Figura 4.39, os valores reais de h e θ foram de aproximadamente 20 cm e 2º, respectivamente. Os valores determinados pela abordagem corretiva são apresentados na Tabela 4.3.

Para averiguar a qualidade do resultado, computou-se a diferença em número de *pixels* entre cada par de linhas vizinhas da imagem original (Figura 4.39), calculando-se, em seguida, a média e o desvio padrão destes valores, obtendo-se 20,3 e 4,3, respectivamente. Posteriormente, construiu-se a imagem retro-projetada (Figura 4.40) e, portanto, corrigida, e repetiu-se o cálculo. A média obtida foi de 19,7 *pixels* e o desvio padrão de 1,6. Este resultado confirma que a imagem gerada a partir da aplicação do método apresenta fidelidade altíssima ao objeto real, sendo o erro de cerca de 1 (um) *pixel* atribuído à natureza discreta da imagem.

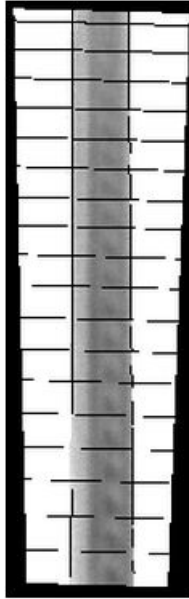


Figura 4.40: Imagem do frasco mostrado na Figura 4.39 alterada pela abordagem corretiva.

Abordagem Neural Na abordagem neural, utiliza-se a mesma imagem da primeira abordagem, porém, levando-se em conta as marcações ao longo do frasco. A Tabela 4.4 mostra a relação número de *pixels*-distância para cada uma das marcas, além da diferença entre elas. Observe que a diferença em número de *pixels* entre marcas consecutivas varia ao longo da imagem do frasco.

Alt. (cm)	Alt. (<i>Pixels</i>)	Dif.	Alt. (cm)	Alt. (<i>Pixels</i>)	Dif.
1	13		10	172	21
2	28	15	11	192	20
3	44	16	12	213	21
4	61	17	13	236	23
4	78	17	14	260	24
6	96	18	15	284	24
7	114	18	16	311	27
8	132	18	17	337	26
9	151	19	18	367	30

Tabela 4.4: Relação distância - número de *pixels* para as marcações da Figura 4.39.

A massa de treinamento da rede consistiu das 18 observações apresentadas na Tabela 4.4. A arquitetura da rede, obtida de forma empírica, tinha 2 camadas, com

30 neurônios na primeira (função de transferência não-linear) e 1 na camada de saída (função de transferência linear). O treinamento foi realizado em aproximadamente 2000 *epochs*, quando foi atingido o erro mínimo de 10^{-6} . A Tabela 4.5 apresenta os resultados obtidos pela rede, nos quais se nota que o erro cometido pela rede é mínimo para fins práticos.

Alt. (cm) Real	Alt. (cm) Rede	Alt. (cm) Real	Alt. (cm) Rede
1	0.99697	10	9.9994
2	2.0022	11	11.0006
3	2.9992	12	11.9995
4	4	13	13.0003
5	5.0002	14	13.9999
6	5.9997	15	15
7	7.0004	16	16
8	7.9994	17	17
9	9.0006	18	18

Tabela 4.5: Resultado da rede neural para o problema das distorções projetivas.

4.3 3º Caso: Controle Anti-Furto de Objetos

O terceiro estudo de caso apresentado neste trabalho corresponde a monitoração de um objeto específico localizado no interior de uma sala. O objetivo do estudo é conceber, de forma preliminar, uma estratégia de identificação da retirada do objeto em questão do recinto em observação. A aplicação de um sistema dotado de tal capacidade é óbvia e imediata, residindo nas áreas de vigilância e controle anti-furto.

É importante ressaltar que a complexidade deste problema é elevadíssima, não sendo portanto o objetivo desta abordagem a resolução por completo do mesmo, e sim, a concepção de uma estratégia cujos frutos sejam idéias preliminares constituintes de uma solução mais abrangente a ser perseguida em estudos futuros.

Sendo assim, diversas condições de contorno são conferidas ao problema de forma a simplificá-lo e focar os esforços aqui realizados em experimentos e idéias simples para a construção de uma base de conhecimento desta classe de problemas. Tal postura foi adotada neste estudo de caso devido à falta de tempo para o estabelecimento de uma meta mais ambiciosa.

As condições de contorno incorporadas ao problemas são listadas a seguir:

- O conjunto de objetos que não estão sendo vigiados presentes no recinto é invariável.
- Considera-se apenas a situação em que **uma** pessoa entra na sala.
- A pessoa que adentrar o recinto pode mover todos os objetos presentes.
- Considera-se que todos os objetos são menores em área na imagem do que qualquer pessoa que venha adentrar a sala.
- O objeto a ser monitorado deve ter **uma** cor consideravelmente predominante

4.3.1 Abordagem Proposta

A abordagem escolhida neste trabalho para garantir que seja identificada a retirada de um determinado objeto de uma sala, caso tal evento ocorra, pode ser dividida nas seguintes fases:

1. Amostragem do Objeto
2. Determinação dos Intervalos de Cor
3. Construção do Módulo de Reconhecimento
4. Detecção de Movimento no Recinto
5. Localização do Objeto

As seções seguintes apresentarão o procedimento correspondente a cada uma das fases mencionadas.

4.3.2 Amostragem do Objeto

Nesta etapa, o objeto a ser monitorado pelo sistema deve ser colocado em um ambiente com fundo neutro. Em seguida devem ser capturadas e armazenadas diversas imagens do objeto em diferentes posições e a variados ângulos.

Conforme será mostrado na próxima seção, o método de extração de características utilizado na fase de construção do módulo de reconhecimento é invariante

a rotação, escalamento e translação. Por este motivo, as imagens mencionadas no parágrafo anterior, embora não necessariamente geradas no ambiente a ser monitorado pelo sistema, podem ser utilizadas no reconhecimento do objeto, já que o que será considerado para realizar tal tarefa é a forma do mesmo.

4.3.3 Determinação dos Intervalos de Cor

Embora tenha sido estipulado como condição de contorno que o objeto a ser monitorado deva ter uma cor predominante, em nível de *pixel*, e mesmo a olho nu em certos casos, o tom da cor predominante do objeto pode variar, seja por efeitos de iluminação, seja por erros cometidos ou ajustes realizados pela câmera. Sendo assim, torna-se necessário que se informe ao sistema alguns valores de cores que fazem parte da imagem do objeto já no ambiente em observação.

Para tanto, o sistema exibe a imagem do ambiente e um operador, utilizando o *mouse*, clica em n_c pontos. A partir das cores destes pontos, serão determinados intervalos a serem utilizados na tentativa de localização e segmentação nas fases posteriores. Estes intervalos int são definidos conforme mostra a Equação 4.11, considerando-se $c_i = [r_i, g_i, b_i]$ a cor do i -ésimo ponto indicado pelo operador e δ_r , δ_g e δ_b valores de tolerância para cada componente do RGB, calculados a partir do desvio padrão dos valores de cor presentes na vizinhança do i -ésimo ponto clicado pelo operador.

$$int_i = c_i \pm \delta = [r_i \pm \delta_r, g_i \pm \delta_g, b_i \pm \delta_b], \quad 1 \leq i \leq n_c \quad (4.11)$$

4.3.4 Construção do Módulo de Reconhecimento

Nesta seção será mostrado o processo de preparação do módulo de reconhecimento do sistema. Inicialmente, foram utilizadas três abordagens para o problema, sendo duas delas baseadas em descritores de Fourier. Por este motivo será feita uma breve introdução a este ferramental matemático empregado na realização da tarefa em questão. Ao longo do estudo, deu-se a necessidade da introdução de uma quarta abordagem, a qual será apresentada mais adiante em respeito ao andamento cronológico do estudo.

Descritores de Fourier

Seja $B = [(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)]$ a sequência de pontos que formam a borda de um objeto em uma imagem qualquer. Considerando-se os índices do conjunto B como elementos do domínio t (tempo) e os pontos constituintes da borda como resultados de uma função complexa $Y(t)$, pode-se interpretar o contorno como uma função discreta e periódica em t , tendo em vista que tomando-se um ponto aleatório p de B e percorrendo-se o conjunto em qualquer direção, p será atingido novamente ao final do percurso. Desta forma, torna-se adequada a aplicação da transformada de Fourier na tentativa de se representar o contorno (JAMUNDÁ, 2002). Para calcular os descritores de Fourier desta borda aplica-se ao conjunto B a transformação Y , mostrada na Equação 4.12.

$$c_t = Y(t) = x_t + i \cdot y_t \quad (4.12)$$

Em seguida, calcula-se a transformada de Fourier dos elementos do vetor complexo c , sendo os coeficientes desta transformada os descritores de Fourier. Os descritores referentes a frequências mais baixas são responsáveis por descrever a forma bruta do objeto, enquanto os componentes de frequências mais elevadas descrevem detalhes finos do contorno.

Porém, como todo método de descrição de objetos deve ser invariante a operações geométricas, como rotação, translação e escalamento, alguns ajustes são feitos a estes descritores para que fiquem imunes a efeitos das mesmas.

Em primeiro lugar, despreza-se o primeiro elemento de c , já que o mesmo corresponde à componente DC da transformada e é dependente da posição do objeto na imagem (GONZALEZ & WOODS, 2000).

Sabe-se que o resultado da transformada de Fourier é um conjunto de números complexos. Sendo cada um dos elementos resultantes da transformada do tipo $a + bi$, considera-se apenas seus valores absolutos, conforme mostra a Equação 4.13.

$$a + b \cdot i \rightarrow \sqrt{a^2 + b^2} \quad (4.13)$$

Desta forma, descarta-se a influência da inclinação do objeto na imagem, deixando a representação imune à rotação (GONZALEZ & WOODS, 2000). Por fim, divide-se todos os valores pelo maior valor de todos, normalizando a representação e eliminando possíveis efeitos de escalamento (GONZALEZ & WOODS, 2000).

Determinação da Seqüência de Pontos da Borda

Para empregar os descritores de Fourier na representação de objetos neste estudo de caso foi necessário o desenvolvimento de um algoritmo para obtenção da seqüência de pontos da borda de um objeto, o qual é apresentado como Algoritmo 4.1.

Algoritmo 4.1 Obtenção do Contorno de um Objeto

```
// Observações:
// obterVizinhos(p) retorna uma lista com os vizinhos de p.
// eliminaPontoVisitado remove da lista de vizinhos aquele que já foi
// visitado.

Var
    contorno, vizinhos: lista
    p0, p: ponto
    visitados: matriz

1) Seja f uma imagem binária contendo apenas o objeto.
2) Obtenha a imagem b com a borda do objeto,
   garantindo que a borda tem largura 1.
3) Seja visitados uma matriz de zeros com as mesmas dimensões de f.
4) Selecione um ponto p0 aleatório.
5) visitados[p0.linha, p0.coluna] = 1
6) insere(p0, contorno)
7) p = p0
8) enquanto não visitar novamente p0 faça
    8.1) vizinhos = concatena(obterVizinhos4(p), obterVizinhosD(p));
    8.2) vizinhos = eliminaPontoVisitado(vizinhos, visitados);
    8.3) p = vizinhos[0]; // ‘vizinhos’ só tem um elemento.
    8.4) visitados[p0.linha, p0.coluna] = 1
    8.5) insere(p, contorno)
9) fim enquanto
10) retorna contorno
```

A Figura 4.41.b mostra o contorno detectado pelo algoritmo quando aplicado sobre a Figura 4.41.a.

O Reconhecimento

A princípio, a tarefa de diferenciar os diversos objetos contidos na imagem do objeto de interesse foi implementada de três formas, sendo duas delas com o emprego

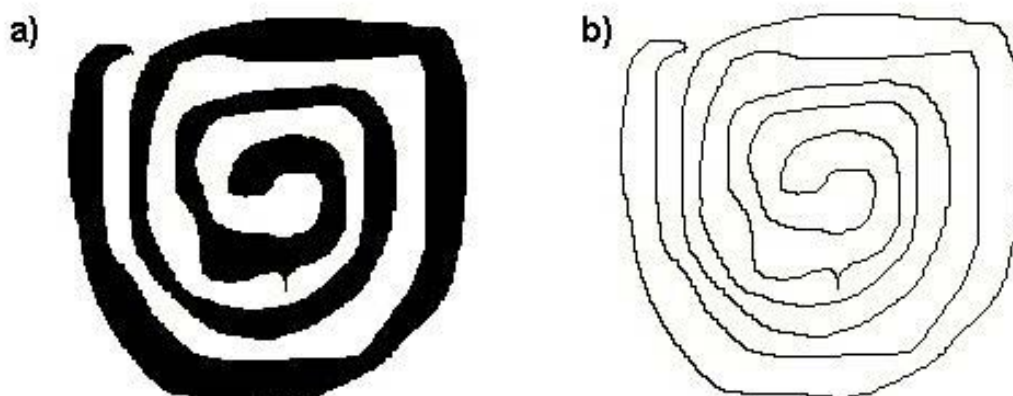


Figura 4.41: Exemplo da aplicação do algoritmo de obtenção do contorno de um objeto.

de uma rede neural do modelo MLP (HAYKIN, 1998), e a restante baseada em uma medida de erro calculada a partir das cores presentes no objeto. A seguir são apresentadas essas três abordagens, utilizadas para representar os objetos em cena, e permitir, posteriormente, sua classificação no módulo de reconhecimento.

Pontos do Contorno Esta abordagem faz a utilização direta do que foi explicado anteriormente sobre descritores de Fourier. As coordenadas da seqüência de pontos das bordas dos objetos são transformadas em valores complexos a partir dos quais os descritores são obtidos. Porém, na obtenção da borda, não são levados em consideração todos os pontos determinados pelo algoritmo, tendo em vista que a representação dos objetos deve ter o mesmo tamanho e objetos diferentes teriam bordas de comprimentos diferentes. Neste trabalho, foram aproveitados 100 pontos igualmente espaçados ao longo do contorno, de forma que a seqüência estabelecida para eles confere implicitamente à representação a característica de forma do objeto.

Para construir a massa de dados a ser utilizada no treinamento da rede neural, utiliza-se as imagens do objeto em observação obtidas na fase de amostragem como exemplos positivos. Os exemplos negativos são gerados a partir da aplicação, na imagem em repouso da sala, de uma rotulação de componentes conexas para os intervalos de cor calculados na fase de determinação dos intervalos de cor. Segmenta-se então todas as regiões conexas apontadas pelo algoritmo e acrescenta-se as mesmas à massa de treinamento como contra-exemplos, com exceção do objeto de inte-

resse, o qual é adicionado como exemplo positivo. É importante ressaltar que é possível identificá-lo neste cenário inicial uma vez que o operador clicou em pontos do mesmo. Sendo assim, o objeto de interesse corresponde à região segmentada mais próxima dos pontos indicados. Para compensar possíveis diferenças numéricas entre as massas de exemplos positivos e negativos, são replicados exemplos já presentes no conjunto de treinamento, porém, com a incorporação de um ruído aleatório de 10 % da amplitude dos valores de suas características.

Distâncias do Centróide A segunda abordagem também emprega descritores de Fourier, porém, eles não são calculados a partir dos pontos que compõe o contorno do objeto, mas sim de suas respectivas distâncias ao centróide do mesmo. A vantagem desta representação é que seu vetor de características tem a metade do tamanho daquele gerado pela abordagem anterior, já que a transformada de Fourier é calculada apenas sobre valores reais, de forma que seu *espectro* é simétrico em torno da origem.

Quanto à construção da massa de treinamento da rede neural, segue-se o mesmo padrão descrito na abordagem anterior.

Medida de Erro Baseada nas Cores do Objeto A terceira abordagem é a mais simples de todas e emprega apenas o conhecimento que se tem sobre as cores do objeto, assumindo-se que o operador, no momento de determinar os intervalos de cor, clicou em pontos suficientes do objeto para representar de forma relevante sua distribuição de cores.

Conforme mostrado na Equação 4.11, cada elemento do conjunto *int* corresponde a um intervalo de cores encontrado no objeto. Considerando-se ϕ o conjunto de cores de um determinado objeto segmentado da imagem, o cálculo da medida de erro é demonstrado no Algoritmo 4.2.

Espera-se que o erro calculado para o objeto de interesse seja o menor de todos.

4.3.5 Detecção de Movimento no Recinto

Tendo sido executadas todas as fases anteriormente apresentadas, o sistema entra em estado de monitoração, e passa a tentar identificar a entrada de uma pessoa no recinto em observação. Enquanto tal evento não ocorre, não há necessidade de se monitorar o objeto, pois assume-se que apenas uma pessoa possa retirá-lo da sala.

Algoritmo 4.2 Cálculo da Medida de Erro Baseada nas Cores do Objeto

```

Var
  n: inteiro
  p: ponto
  soma, v, w: tuplaRGB // [r: 0 → 255, g: 0 → 255, b: 0 → 255]

1) soma = [0, 0, 0]
2) para v ∈ φ faça
  2.1) Seja p a posição de v = [vr, vg, vb].
  2.2) Seja inti o intervalo mais próximo de p.
  2.3) Seja pi a posição do ponto correspondente ao centro do
intervalo inti.
  2.4) Seja w = [wr, wg, wb] a cor do ponto pi.
  2.5) soma = soma + |v - w|
  2.6) n = n + 1
9) fim para
10) retorna soma / n

```

Quando o sistema identifica a entrada de uma pessoa, ele passa a tentar detectar a ocorrência de outro evento: a saída da pessoa. Assume-se que enquanto a pessoa não sai, não é possível que o objeto seja retirado do recinto. No momento em que se detecta a saída da pessoa, o sistema tenta então encontrar o objeto, tarefa a ser discutida na próxima seção.

Sendo assim, esta fase se resume a identificação dos eventos entrada e saída de uma pessoa. A abordagem proposta para resolver este problema baseia-se na técnica de detecção de movimento Comparação Simples, apresentada na seção 3.10.1. Para lidar com efeitos de mudança de iluminação ao longo do dia, a imagem referência é submetida a uma atualização bem lenta e gradual, de forma a reduzir os efeitos de possíveis reflexos e sombras lentamente gerados pela luz do sol na detecção de movimento.

Conforme será mostrado a seguir, a abordagem proposta neste trabalho para detecção da entrada e saída de uma pessoa no recinto em observação é baseada na determinação de limites entre as situações de presença e ausência de uma pessoa na sala, o que confere ao procedimento o caráter particular da sala na qual foram realizados os testes (ver Figura 4.42). Estabelece-se então que para aplicação em diferentes ambientes, seria necessária uma nova determinação de valores limites para

cada um deles.



Figura 4.42: Sala de testes para a monitoração anti-furto.

No teste utilizado para ilustrar a abordagem, uma pessoa entrou na sala em observação e mudou alguns objetos de lugar, se retirando em seguida. A Figura 4.43.a mostra uma imagem na qual a pessoa está presente na sala, enquanto a Figura 4.43.b apresenta o resultado limiarizado da detecção de movimento.

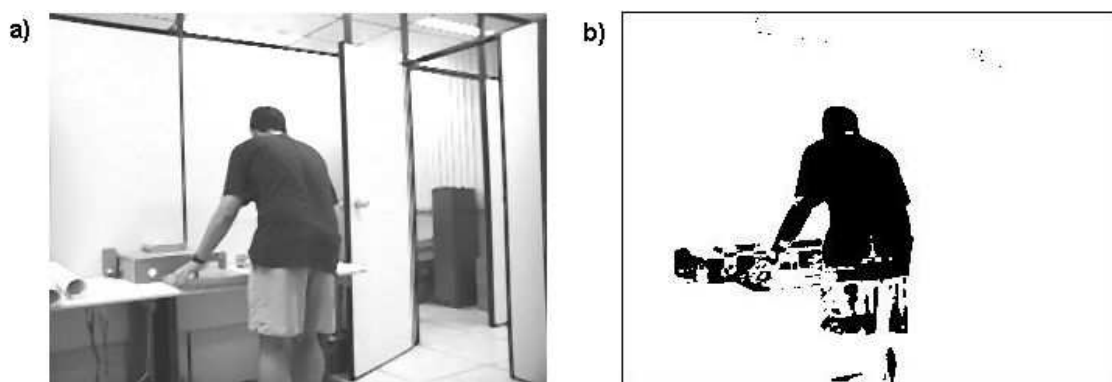


Figura 4.43: Exemplo de detecção de movimento com uma pessoa na sala.

Através do cálculo da projeção horizontal da imagem resultante da detecção de movimento, calcula-se a área de cada entidade apontada pelo algoritmo. Estatisticamente, levantou-se que a área (em *pixels*) média ocupada por uma pessoa quando enquadrada pela câmera correspondia a 6841 *pixels* (cerca de 10 % do tamanho da imagem), e o desvio-padrão a 2394.

Estabeleceu-se inicialmente que, no momento em que fosse detectada um objeto de área maior que $6841 - 2394 = 4447$, seria acusada a presença da pessoa no recinto. Porém, esta medida não se mostrou suficiente para discriminar as duas

situações em questão. Um exemplo de situação em que esta medida não funcionaria pode ser facilmente imaginada. Suponha que uma pessoa entra na sala e agrupe diversos objetos em um mesmo local. Ao se retirar, a detecção de movimento, que no caso funciona como um detector de diferenças entre a situação inicial de repouso e a situação atual, identificará uma entidade conexa de área significativa que pode extrapolar o limite estabelecido mesmo sem a presença da pessoa. A Figura 4.44.a mostra uma imagem que não contém a pessoa, mas que, porém, apresenta uma entidade conexa de área maior do que o limite determinado. O resultado da detecção de diferenças é mostrado na Figura 4.44.b.

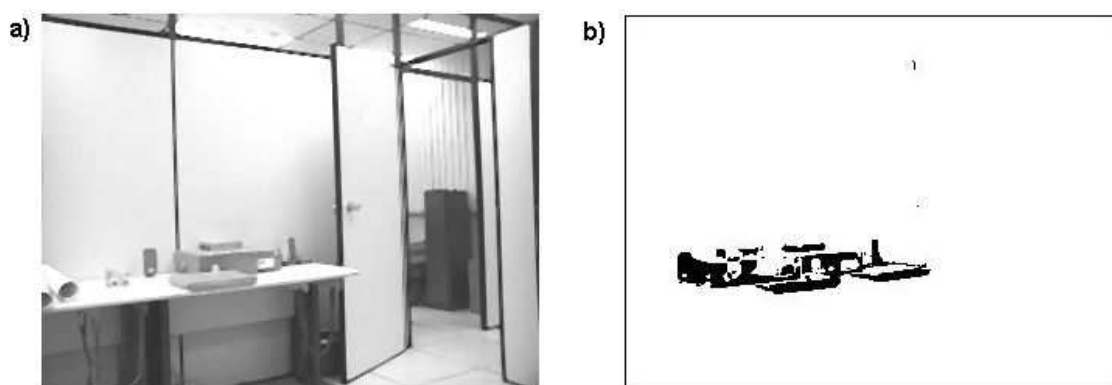


Figura 4.44: Exemplo de detecção de movimento sem uma pessoa na sala.

Por este motivo, decidiu-se que seria necessária a combinação da área das entidades detectadas com uma segunda medida de controle. Observando-se as projeções verticais das imagens, notou-se que nas situações em que a pessoa estava presente no recinto a porção do domínio para qual a projeção apresentava valores diferentes de zero era mais extensa e concisa, conforme pode ser visto nas Figuras 4.45.a, 4.45.b, 4.45.c, 4.45.d.

Assim como no caso da área das entidades detectadas pelo algoritmo, foi realizado um levantamento estatístico da extensão dos intervalos de valores diferentes de zero na projeção vertical para os casos em que a pessoa estava presente no recinto. Nota-se que a extensão desses intervalos correspondia à altura do objeto em questão. Para eliminar os ruídos, as projeções foram submetidas a um filtro de média. O valor obtido como média da altura dos objetos foi de 187 (cerca de 80 % da altura da imagem). O desvio-padrão obtido foi de 21. Sendo assim, estabeleceu-se que imagens com um objeto cuja área fosse superior ao limite citado anteriormente e

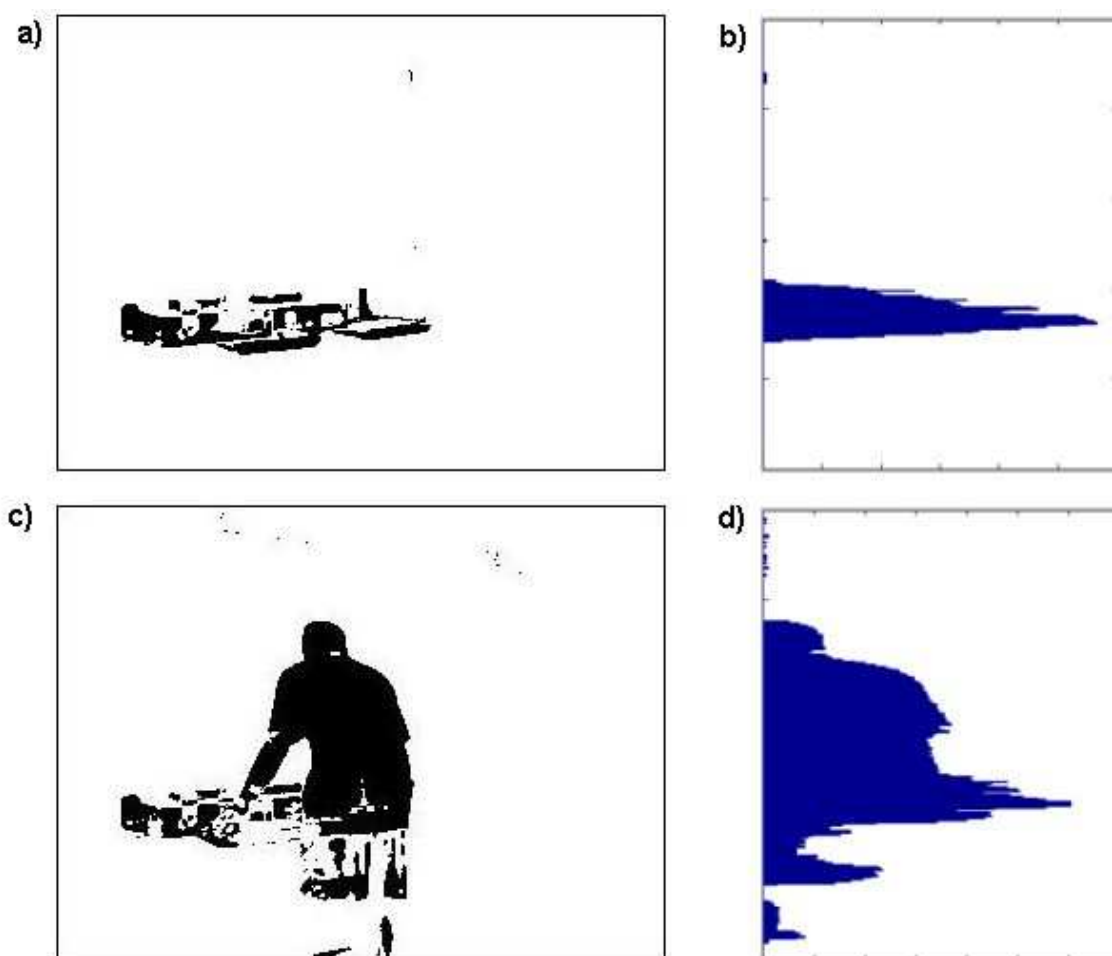


Figura 4.45: Projeções verticais de situações com a ausência e presença de uma pessoa.

cuja projeção vertical apontasse um objeto com altura maior que $187 - 21 = 166$ indicariam a presença de uma pessoa na sala. Já a saída da pessoa é determinada quando, após extrapolados os limites que acusam sua presença na sala, atinge-se uma situação em que algum dos critérios de presença é violado.

Supondo que uma pessoa entre no recinto em observação e empilhe diversos objetos até que a pilha atinja a altura média de uma pessoa, o método falhará. Uma maneira de endereçar este problema é o acompanhamento da movimentação da entidade de maior área, controlando, desta forma, a trajetória do indivíduo, e determinando, através de sua posição, sua saída da sala. Este controle, porém, não foi implementado neste trabalho, e será citado com maiores detalhes na seção de

propostas futuras no capítulo de conclusões.

4.3.6 Localização do Objeto

No momento em que se determina a saída da pessoa do recinto, a fase que se inicia é a de localização e reconhecimento do objeto em observação. Para realizar tal tarefa realiza-se uma rotulação de componentes conexas empregando-se os intervalos de cor estebelecidos, obtendo-se como resultados diversas entidades conexas. Submete-se então cada uma destas entidades ao módulo de reconhecimento.

A proposta inicial correspondia a construir a lógica do módulo de reconhecimento com base em uma combinação das três estratégias de reconhecimento apresentadas, porém, em testes específicos com cada uma constatou-se que o desempenho das mesmas era muito irregular, e até mesmo incoerente (a ser discutido na seção 4.3.7). Por este motivo, foi concebida uma quarta estratégia de classificação, a qual é apresentada a seguir.

Descritores de Baixas Freqüências

Conforme será mostrado com maiores detalhes na seção 4.3.7, concluiu-se que o mau desempenho das estratégias de classificação se deveu à escolha de um subconjunto de tamanho fixo dos pontos do contorno de um objeto para o cálculo dos descritores de Fourier, tanto na abordagem das distâncias ao centróide, quanto na de utilização direta das coordenadas dos *pixels* que compunham o contorno do objeto.

Segundo foi mencionado na seção 4.3.4, os descritores de Fourier referentes a freqüências mais baixas são responsáveis por representar a forma geral do objeto, enquanto os descritores de freqüências mais altas guardam informações sobre detalhes finos do mesmo.

Decidiu-se então que, ao invés de se usar um subconjunto de pontos do contorno para calcular os descritores, seriam utilizados todos os pontos determinados pela algoritmo de obtenção da borda do objeto, porém, somente um conjunto de tamanho fixo de descritores seria utilizado para todo e qualquer objeto ou entidade conexa segmentada, sendo estes descritores aqueles de freqüências mais baixas. Vale ressaltar que, considerando que os pontos do contorno são transformados em valores conexos, tem-se que *espectro* de Fourier não é simétrico em torno da origem. Sendo assim, torna-se necessário considerar valores das bandas positiva e negativa.

Esperava-se com isso conferir ao módulo de reconhecimento uma maior coerência do que a apresentada pelo mesmo quando construído com base nas estratégias anteriores.

É importante ressaltar que a construção da massa de treinamento seguiu o mesmo padrão apresentado para as abordagens anteriores, empregando-se a replicação de amostras e introdução de ruído aleatório quando necessário.

4.3.7 Exemplo Prático

Nesta seção será apresentado um exemplo completo da aplicação da abordagem proposta em uma situação simulada, com exceção da fase de detecção de entrada e saída de uma pessoa no recinto, considerando-se que as imagens utilizadas para ilustrar o procedimento realizado pelo sistema já foram empregadas na seção 4.3.5. É importante lembrar o leitor de que nunca foi objetivo deste trabalho resolver o problema por completo, mas sim, entender suas dificuldades e acumular conhecimento sobre o mesmo, permitindo desta forma a identificação de possíveis caminhos para trabalhos a serem realizados no futuro. A Figura 4.46 mostra a sala na qual foi realizado o experimento, estando o objeto a ser monitorado destacado pelo retângulo preto que o envolve. Embora a imagem seja exibida em tons de cinza, ela é considerada pelo sistema em cores.



Figura 4.46: Sala na qual foi realizado o experimento da monitoração anti-furto de um objeto, estando o mesmo destacado por um retângulo que o envolve.

Amostragem do Objeto

A Figura 4.47 mostra imagens do objeto em diferentes posições e utilizadas como exemplos positivos no treinamento do módulo de classificação.



Figura 4.47: Imagens do objeto de interesse a serem utilizadas na construção do módulo de reconhecimento.

Determinação dos Intervalos de Cor

A Figura 4.48 mostra a imagem do objeto em *zoom* com os pontos clicados pelo usuário destacados em branco. A partir destes pontos e suas vizinhanças (5 X 5) foram determinados os intervalos de cor que representam o objeto a ser monitorado.



Figura 4.48: Pontos clicados pelo usuário para a determinação dos intervalos de cor do objeto.

Construção do Módulo de Reconhecimento

A construção do módulo de reconhecimento do sistema se deu conforme especificado na seção 4.3.4, de forma que, para este caso, a massa de treinamento foi composta por 40 amostras, sendo 20 positivas e 20 negativas. Inicialmente, o resultado oferecido pelo módulo para cada objeto a ele submetido era formado pelas saídas de cada uma das três estratégias inicialmente apresentadas. O objetivo era conceber, através dos resultados individuais de cada uma, uma lógica que as combinasse e oferecesse maior robustez aos resultados do sistema.

Porém, devido ao desempenho irregular e inconsistente das três técnicas, foi testada uma quarta (descritores de baixa frequência), cujos resultados foram mais satisfatórios.

A arquitetura das rede neurais, obtida de forma empírica, através de alguns experimentos, utilizadas nas três estratégias foi a mesma, sendo o número de camadas igual a 2, de forma que a primeira camada continha 32 neurônios e a segunda 2. O algoritmo utilizado para o treinamento foi o do gradiente descendente com taxa de aprendizado adaptativa, tendo taxa de aprendizado inicial igual a 0,4, 0.5 de momentum e 10^{-5} como objetivo, valores os quais foram obtidos de forma empírica.

Detecção de Movimento no Recinto

A execução da fase de detecção da entrada e saída de uma pessoa na sala sob monitoração para este experimento já foi apresentada como exemplo ilustrativo na seção 4.3.5.

Localização do Objeto

Conforme explicado na seção 4.3.6, o primeiro passo desta fase correspondia à segmentação de entidades conexas candidatas, a qual se realizava a partir da aplicação de um algoritmo de rotulação de componentes conexas com base nos intervalos de cor estabelecidos. A Figura 4.49 mostra os objetos segmentados de uma imagem seguinte a detecção da saída da pessoa e suas respectivas posições. São definidos também rótulos para cada uma das entidades conexas segmentadas, sendo **b** o rótulo do objeto de interesse.

A Tabela 4.6 mostra os resultados das abordagens de classificação I (Pontos do Contorno), II (Distâncias do Centróide), III (Erro de Cor) e IV (Descritores de Baixa

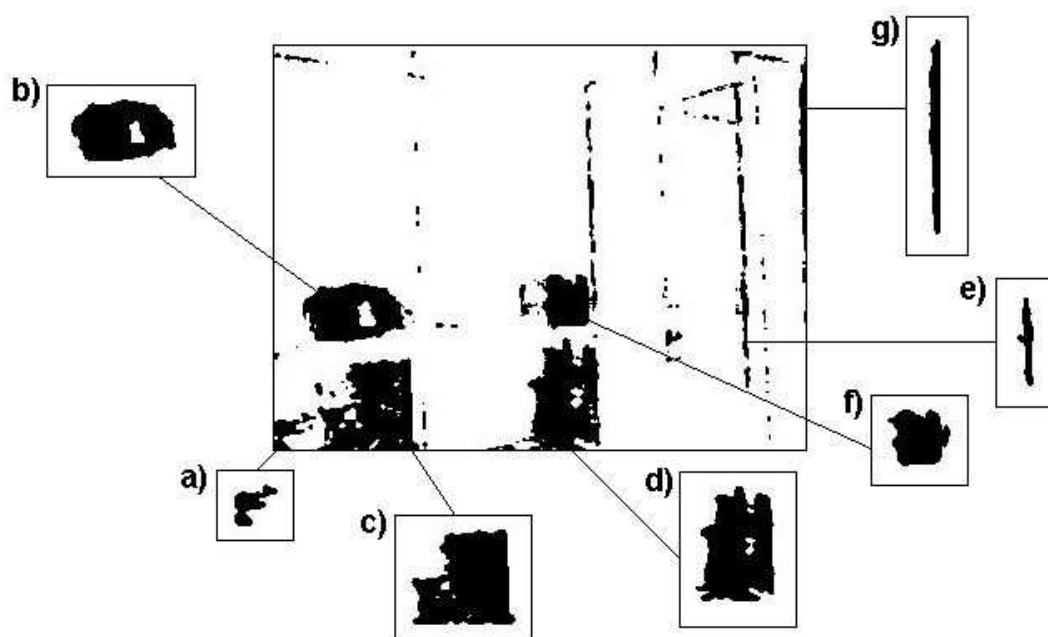


Figura 4.49: Entidades conexas segmentadas de uma imagem posterior à saída da pessoa da sala e seus respectivos rótulos.

Frequência) para cada um dos objetos exibidos na Figura 4.6, sendo \mathbf{P} correspondente a uma identificação positiva do objeto de interesse e \mathbf{N} a uma identificação negativa.

Analisando-se os resultados apresentados na Tabela 4.6, percebe-se que nenhuma das estratégias iniciais foram capazes de reconhecer corretamente o objeto de interesse. Os erros do tipo falso-positivo cometidos por essas abordagens mostraram não só ineficiência, mas também incoerência, já que as formas das entidades conexas, com exceção daquela de rótulo **d**, são completamente diferentes, em termos de contorno, do objeto sob monitoração. Por este motivo, concluiu-se que a amostragem em 100 pontos igualmente espaçados do contorno da imagem realizada na extração de características prejudicou a representação da forma dos objetos.

Quanto à medida de cor, suspeitava-se que efeitos de iluminação diferenciados dependendo da posição do objeto na sala pudessem comprometer a estratégia. Confirmando-se essa suspeita, concluiu-se que em ambientes nos quais a iluminação não pode ser controlada estratégias de discriminação baseadas somente em cor têm pouca confiabilidade.

Sendo assim, definiu-se a quarta estratégia, na qual, conforme já foi explicado,

Objeto	Abord. I	Abord. II	Abord. III	Abord. IV
a)	P	N	13.58	N
b)	N	N	8.86	P
c)	P	P	8.33	N
d)	N	P	7.60	P
e)	P	N	8.25	N
f)	N	N	7.32	N
g)	N	N	9.00	N

Tabela 4.6: Resultados atingidos pelo módulo de reconhecimento.

Legenda:Abord. I - *Pontos do Contorno.*Abord. II - *Distâncias do Contróide.*Abord. III - *Medida de Erro Baseada nas Cores do Objeto.*Abord. IV - *Descritores de Baixas Frequências.*

calculava-se os descritores de Fourier com base em todos os pontos do contorno do objeto e tomava-se apenas os descritores de baixa frequência (no caso, os 40 primeiros). Pode ser visto na Tabela 4.6 que essa estratégia foi a única capaz de reconhecer o objeto corretamente, embora tenha cometido um erro do tipo falso-positivo com a entidade conexa de rótulo **d**.

Embora este resultado demonstre que a estratégia IV também não é eficiente, ele mostra coerência, já que a forma bruta dos objetos **b** e **d** é razoavelmente parecida, se forem desconsideradas suas inclinações. Considerando que o critério de busca e localização do objeto é baseado em sua cor (segmentação) e no seu contorno (reconhecimento), conclui-se que é necessária a introdução de uma ou mais novas medidas, a serem combinadas com as anteriores. Como sugestões, podem ser testadas características de textura ou detalhes finos da região interna do objeto.

Conclui-se desta forma o 3º estudo de caso desta dissertação. No capítulo seguinte são apresentadas as conclusões do trabalho, além de propostas para desenvolvimento de trabalhos futuros.

Capítulo 5

Conclusões

5.1 Cumprimento de Objetivos

Abrindo o capítulo de conclusões, segue a reafirmação dos objetivos estabelecidos no início do trabalho.

“O objetivo deste trabalho consiste da realização de 3 tarefas: a) o estudo de técnicas de Processamento de Imagens e suas aplicações a sistemas de monitoração de processos baseados em imagens; b) a realização de três estudos de caso nos quais problemas práticos de monitoração são abordados e soluções propostas com o emprego das técnicas mencionadas e c) a análise da adequação de sistemas baseados em imagens à atividade de monitorar um processo qualquer, sendo sub-objetivos o levantamento de dificuldades inerentes ao uso de imagens para este fim e a identificação da possibilidade de melhorias e avanços naquilo que for implementado neste trabalho.”

As seções seguintes correspondem a comentários e conclusões específicas relacionadas a cada um dos objetivos mencionados.

5.1.1 Estudo de Processamento de Imagens

O aprofundamento nas técnicas de processamento de imagens foi realizado ao longo do curso de Mestrado e sintetizado, de forma prática e sucinta, no Capítulo 3. Embora diversas outras técnicas tenham sido estudadas, considerou-se que aquelas inseridas no trabalho estão entre as mais importante e representativas do campo.

5.1.2 Aplicação à Monitoração de Processos

Esta parte do primeiro objetivo enunciado foi abordada em duas frentes, sendo o resultado da primeira delas o Capítulo 2, no qual apresenta-se uma revisão bibliográfica de vários papéis desempenhados por sistemas de imagens em diferentes campos de aplicação. A segunda foi a própria realização dos estudos de caso, nos quais foi possível a aquisição de conhecimento significativo para trabalhos futuros a serem conduzidos que envolvam o emprego de técnicas de processamento de imagens.

5.1.3 Estudos de Caso

As observações e conclusões atingidas a partir de cada estudo de caso realizado são apresentadas a seguir, estando separadas em seções correspondentes a cada um dos problemas abordados.

1º Caso: Detecção da Queda da Primeira Gota em um Processo de Destilação de Petróleo em Batelada

Sendo este um experimento de difícil simulação, o primeiro problema significativo enfrentado neste estudo de caso foi fruto de um erro de julgamento com relação à velocidade de queda da gota. Acreditou-se, inicialmente, que a gota se formava lentamente no bico do condensador lateral, aumentando de volume de forma gradual, até se desprender completamente. Na verdade, a gota escorria pelo corpo do condensador e se desprendia a uma velocidade bastante alta.

A primeira versão do sistema atuava a uma taxa de 5 quadros por segundo, de forma que, em algumas simulações, a imagem da gota não era capturada pela câmera e, conseqüentemente, não era analisada pelo sistema. Segundo o manual da placa de captura utilizada no experimento, a capacidade máxima da mesma era de 15 quadros por segundo. Por este motivo, trabalhou-se intensamente na melhoria da qualidade do software, procurando-se otimizar todos os procedimentos realizados pelo mesmo que não estivessem atrelados ao sistema operacional. Lentamente ganhou-se terreno, até que os resultados ficaram estagnados em 10 quadros por segundo.

Até então trabalhava-se com uma máquina com processador DURON 1.0 GHz e 256 Mb de memória. O software foi então testado em uma máquina com processador Pentium III 1.8 GHz e 512 Mb de memória, na qual atingiu-se a taxa de 15 quadros por segundo estabelecida pelo manual. Nos testes realizados com esta

taxa de amostragem o sistema foi capaz de detectar a queda da gota com 100 % de eficácia.

A situação descrita ilustra bem a importância da qualidade de ambos software e hardware em um projeto deste tipo, sendo recomendado o emprego de linguagens de programação que não sobrecarreguem o sistema operacional com camadas extras de leitura, além da utilização de hardware de alta capacidade de desempenho para garantir taxas de amostragem satisfatórias e suficientes.

Outra observação interessante feita a partir deste estudo de caso diz respeito ao quesito iluminação. Conforme será visto na seção correspondente ao estudo de caso da monitoração anti-furto de um objeto em uma sala, a iluminação pode ser uma grande vilã nos sistemas que empregam imagens. Neste caso, porém, a iluminação, aplicada de forma orientada e específica, foi de suma importância para o sucesso do projeto.

Pode ser visto nas imagens exibidas no Apêndice III que a qualidade visual das mesmas, devido à iluminação ostensiva no local, é muito baixa, dificultando inclusive a visualização da gota. Porém, o esquema de iluminação concebido tem função essencial no sistema, evitando que reflexos de transeuntes na camisa de vidro que circunda o bico do condensador lateral sejam confundidos com o movimento da gota.

Conclui-se então que, embora seja, na maioria dos sistemas de imagens, um aspecto que confere complicações à tarefa objetivo, a iluminação, em casos específicos, pode ser aplicada de forma a facilitar e até mesmo viabilizar o cumprimento da meta.

Por fim, pode-se dizer que a realização do estudo de caso foi bem sucedida, tendo o sistema, até onde foi testado, funcionado de forma satisfatória.

2º Caso: Controle do Nível de Petróleo em Frasco Destino de um Processo de Destilação em Batelada

Embora, por problemas operacionais das entidades envolvidas, este projeto não possa ter sido desenvolvido com as bases ideais, ainda assim foi possível a obtenção de alguns resultados interessantes.

A possibilidade de se corrigir as distorções causadas por angulação e distância da câmera através do método de calibração da mesma apresentado em (HEMERLY & MICOSKY, 2004) permite que se resolva um problema de extrema complexidade

de forma direta e eficiente, estando sujeita, inclusive, a comprovação visual, uma vez que é possível a geração de uma imagem na qual as distorções mencionadas se apresentam consideravelmente atenuadas.

Outra constatação importante diz respeito ao emprego bem sucedido de técnicas de detecção de movimento como computadoradas de diferenças, forma através da qual se determinou o nível do líquido acumulado no frasco.

Dentro das limitações impostas pela impossibilidade de se desenvolver este estudo em um cenário mais próximo do real, pode-se considerar que ele foi bem sucedido, dado que, para a situação proposta, o mesmo funcionou sem grandes problemas.

3º Caso: Controle Anti-Furto de Objetos

A primeira observação a ser feita referente ao estudo de monitoração anti-furto de um objeto específico em uma sala diz respeito aos quesitos iluminação e segmentação. Para realizar a segmentação de objetos ou entidades conexas no estudo de caso em questão levou-se em consideração apenas a cor dos mesmos. Tal fator conferiu à tarefa um alto grau de dificuldade em virtude dos efeitos consideráveis da iluminação no ambiente.

A iluminação age não só no controle automático da íris da câmera, gerando variações bruscas na tonalidade geral do ambiente, o que causa sérios problemas a qualquer iniciativa de controle de movimento, mas também influi diretamente na cor dos objetos, de forma que, localizado em posições diferentes, um mesmo objeto pode apresentar constituições de cor visivelmente diferentes, devido a reflectância diferenciada e sombreamento.

Outra conclusão tirada neste estudo foi a de que tal tarefa não pode ser realizada por apenas uma câmera, considerando-se, obviamente, um ambiente no qual não haja controle sobre o quão próximo da câmera alguém possa chegar. Abordagens que empreguem apenas uma câmera, tendo portanto somente uma perspectiva, ficarão sujeitas à completa e total impotência diante de situações como por exemplo o encobrimento do objeto de interesse por um outro objeto qualquer no recinto.

Outra dificuldade detectada para estudos nesta classe de problemas se refere à ocorrência de dois objetos de mesma cor, sendo um deles o objeto de interesse, estarem encostados um no outro. Quando isto ocorre, torna-se visualmente impossível a separação e segmentação individual dos mesmos, o que dificulta, e muito, o reconhecimento do objeto sob monitoração.

Em relação à detecção da entrada e saída de uma pessoa no recinto, torna-se necessário dizer que o presente estudo não tratou casos particulares, como a possibilidade da mesma se sentar, o que, com certeza, afetaria o desempenho do critério referente à projeção horizontal. A abordagem também não é satisfatória quanto à possibilidade de se tentar adaptá-la para o controle da circulação de mais de uma pessoa no ambiente. Tais problemas poderiam ser atacados com a introdução do acompanhamento das entidades em movimento dentro da sala, controlando suas trajetórias e determinando o momento em que cada uma saiu da sala.

Vale ressaltar que a lógica estabelecida pela abordagem facilita o emprego deste tipo de controle, já que enquanto não se detecta que uma pessoa saiu do recinto não é necessária a localização do objeto de interesse. Desta forma não se faz necessário o controle individual da movimentação de cada pessoa que adentra o recinto, mas sim da ocorrência do evento saída, seja de uma ou mais pessoas.

O reconhecimento de objetos deve ser baseado na combinação de diversos critérios, tendo dois deles sido considerados neste estudo: as cores do objeto e sua forma. Porém, conclui-se que apenas estes critérios não são suficientes para que seja feita uma distinção eficiente entre objetos em um cenário complexo. A incorporação de informações de textura e detalhes da região interna ao contorno do objeto são candidatos imediatos.

Conforme havia sido especificado anteriormente, este estudo de caso em particular não almejava solucionar o problema por completo, mas sim estudá-lo de forma a conhecê-lo com mais propriedade, conceber e testar técnicas preliminares e destacar dificuldades que indicassem caminhos a serem seguidos em estudos futuros que envolvam esta classe de problemas. Sendo assim, conclui-se que a abordagem foi bem sucedida em relação a seus objetivos iniciais, tendo oferecido a oportunidade de se entender e estudar aspectos chave para a construção de uma solução mais genérica e aplicável para o problema no futuro.

5.1.4 Análise de Adequação

A análise da adequação de sistemas baseados em imagens à atividade de monitorar um processo qualquer deve ser considerada com base em alguns aspectos chave, os quais conferem os maiores obstáculos e restrições para a realização da tarefa em questão. São eles:

- **Velocidade do fenômeno:** questão de suma importância a ser considerada no desenvolvimento de um sistema de monitoração baseado em imagens, já que o funcionamento do mesmo será fundamentado em uma amostragem do fenômeno em si, cujos elementos correspondem aos quadros obtidos a partir da placa de captura. Sendo assim, deve ser questionada a viabilidade de se registrar o evento de interesse dentro dos limites temporais dessa amostragem. Embora existam placas de captura de capacidade muito superior àquela utilizada neste trabalho, tal equipamento confere ao orçamento do projeto um custo consideravelmente mais elevado. Assim como seu alto desempenho exige do software e hardware empregados uma resposta bem mais eficiente do que exigiria uma placa de captura de qualidade inferior.
- **Iluminação:** em ambientes nos quais a iluminação é controlada e uniforme, sistemas baseados em imagens não devem encontrar problemas para atuar, podendo-se inclusive orientar a iluminação de forma a facilitar a tarefa a ser realizada, como foi feito no estudo de caso da gota. Já em ambientes em que a iluminação tem efeitos diferenciados ao longo do recinto, caso o sistema em questão tenha algum procedimento baseado na cor dos objetos ou entidades, haverá um risco considerável de falha, especialmente em processos que envolvam segmentação seguida de reconhecimento, considerando-se que, pela irregularidade no quesito cor, a primeira tarefa é bastante prejudicada, comprometendo, conseqüentemente, o desempenho da segunda.
- **Perspectiva:** conforme foi observado no estudo de caso da monitoração anti-furto de um objeto, o sistema pode ser induzido ao erro pelo posicionamento do objeto fora do seu campo de visão, mesmo que dentro do recinto. Sendo assim, torna-se necessário para a aplicação segura e eficiente de um sistema desse tipo que se garanta que o mesmo tenha diferentes perspectivas, conferindo-lhe acesso a todas as regiões relevantes do fenômeno. Cabe ao projetista analisar o conjunto ambiente-fenômeno e determinar o número de câmeras a serem utilizadas, além de sua distribuição e posicionamento.
- **Isolamento e Proteção:** sistemas baseados em imagens realizam procedimentos de altíssimo custo computacional. Aliando-se a isso os requerimentos temporais e o fato de que esses sistemas podem realizar tarefas complementares em paralelo com suas funções de análise e processamento de imagens,

tem-se, em geral, a necessidade de um cenário de hardware de dedicação exclusiva à tarefa de monitoração, sendo importante impedir que pessoas que tenham acesso ao ambiente no qual se encontra a máquina *host* do sistema interfiram de forma indevida na mesma. Da mesma forma, deve ser considerado o isolamento da câmera ou câmeras do sistema, especialmente em casos de detecção de movimento, nos quais diferenciar uma entidade em movimento do ruído provocado por uma perturbação da câmera é tarefa extremamente difícil.

Como maior vantagem dos sistemas de monitoração de processos baseados em imagens está o fato de que, dificilmente, eles requerem alterações de grande porte nos sistemas aos quais se aplicam, além de empregarem equipamentos padronizados de fácil aquisição. Porém, para todo projeto no qual for cogitada a possibilidade de se empregar tal metodologia, devem ser considerados os fatores acima mencionados, de forma a garantir que a tarefa proposta seja viável.

5.1.5 Trabalhos Futuros

Os avanços e melhorias potenciais naquilo que foi implementado e estudado são apresentados a seguir.

- Desenvolvimento e estudo de técnicas para cálculo automático do valor a ser usado em procedimentos de limiarização.
- Desenvolvimento e estudo de algoritmos de correção ou eliminação de efeitos de iluminação (como a filtragem homomórfica), de forma a oferecer à imagem um aspecto mais uniforme em relação a objetos de mesma cor em diferentes pontos de um recinto qualquer, possibilitando um melhor desempenho do processo de segmentação baseado em cor.
- Implementação do acompanhamento (*tracking*) de entidades conexas em um cenário de detecção de entrada e saída de pessoas em uma sala, de forma a tratar casos particulares, referentes a possíveis movimentos diferenciados que possam ser realizados pela pessoa, ou até mesmo como abordagem preliminar à situação em que se considera a possibilidade do controle se estender a casos nos quais mais de uma pessoa pode adentrar o recinto.

- Desenvolvimento e estudo de técnicas para descrição de textura e detalhes internos de objetos.
- Estudo de modelos de monitoração com mais de uma câmera.

Referências Bibliográficas

- [1] ASTM, *D 5236-99a: Standard Test Method for Distillation of Heavy Hydrocarbon Mixtures (Vacuum Potstill Method)*, American Society for Testing and Materials, 2002.
- [2] AW, *How a Video Camera Works*, Alternative Worship, acessado através do endereço seaspray.trinity-bris.ac.uk/~altwfaq/graphics/video/hvw1.html em novembro de 2003.
- [3] BHAT, SHAPTHARISH & KHOSLA, *Motion Detection Segmentation Using Image Mosaics*, BHAT, Kiran S.; SHAPTHARISH, Mahesh; KHOSLA, Pradeep K., obtido através do endereço <http://www-2.cs.cmu.edu/~kiranb/motionDetection/p263.pdf> em abril de 2003.
- [4] BOURKE, *Bicubic Interpolation for Image Scaling*, BOURKE, Paul, acessado através do endereço <http://astronomy.swin.edu.au/~pbourke/colour/bicubic/> em fevereiro de 2004.
- [5] CASTLEMAN, *Digital Image Processing*, CASTLEMAN, Kenneth, Prentice Hall, 1995.
- [6] DIGICAMHISTORY, *DigiCamHistory.com*, acessado através do endereço <http://www.digicamhistory.com/> em julho de 2004.
- [7] DINIZ, SILVA & NETTO, *Digital Signal Processing - System Analysis and Design*, DINIZ, Paulo S. R.; SILVA, Eduardo A. B.; NETTO, Sergio L., Cambridge University Press, 2002.
- [8] EL-SHEIMY & WRIGHT, *Real Time Forest Fire Fighting through the Integration of Infrared Video, GPS and INS Systems*, EL-SHEIMY, Naser;

- WRIGHT, D. Bruce, FIG Working Week 2003, obtido através do endereço www.fig.net/figtree/pub/fig_2003/TS_19/TS19_1_Elsheimy_Wright.pdf em novembro de 2003.
- [9] FARINES & FRAGA & OLIVEIRA, *Sistemas de Tempo-Real*, FARINES, Jean-Marie; FRAGA, Joni da Silva; OLIVEIRA, Rômulo Silva de, acessado através do endereço <http://www.lcmi.ufsc.br/~romulo/discipli/cad-cbtisa/escola-1.pdf> em novembro de 2004.
- [10] FELZENSZWALB & HUTTENLOCHER, *Efficient Graph-Based Image Segmentation*, FELZENSZWALB, Pedro F.; HUTTENLOCHER, Daniel P., acessado através do endereço <http://www.ai.mit.edu/people/pff/papers/seg-ijcv.ps.gz> em abril de 2004.
- [11] FILHO, *Evolução dos Sistemas de Controle*, FILHO, Constantino Seixas, obtido via web através do endereço www.cpdee.ufmg.br/~seixas/PaginaII/Download/DownloadFiles/HistóriaControladores.PDF em outubro de 2003.
- [12] GOMES & VELHO, *Computação Gráfica*, Volume 1, GOMES, Jonas; VELHO, Luiz, Instituto de Matemática Pura e Aplicada, 1998.
- [13] GONZALEZ & WOODS, *Processamento de Imagens Digitais*, GONZALEZ, Rafael; WOODS, Richard, Ed. Edgard Blucher, 2000.
- [14] GONZALEZ, WOODS & EDDINS, *Digital Image Processing Using Matlab*, GONZALEZ, Rafael; WOODS, Richard, EDDINS, Stephen, Prentice Hall, 2004.
- [15] HAYKIN, *Neural Networks: A Comprehensive Foundation*, HAYKIN, Simon, Pearson Education, 1998.
- [16] HEMERLY, Re:Calibração da Câmera, Correspondência pessoal, e-mail de (hemerly@ita.br) para (homerwells1980@yahoo.com.br) [mensagem capturada em 03 de novembro de 2004.
- [17] HEMERLY & MICOSKY, *Calibração de Câmera Via Otimização para Aplicações em Tráfego*, HEMERLY, Elder; MICOSKY, Maurício, XV Congresso Brasileiro de Automática, 2004.

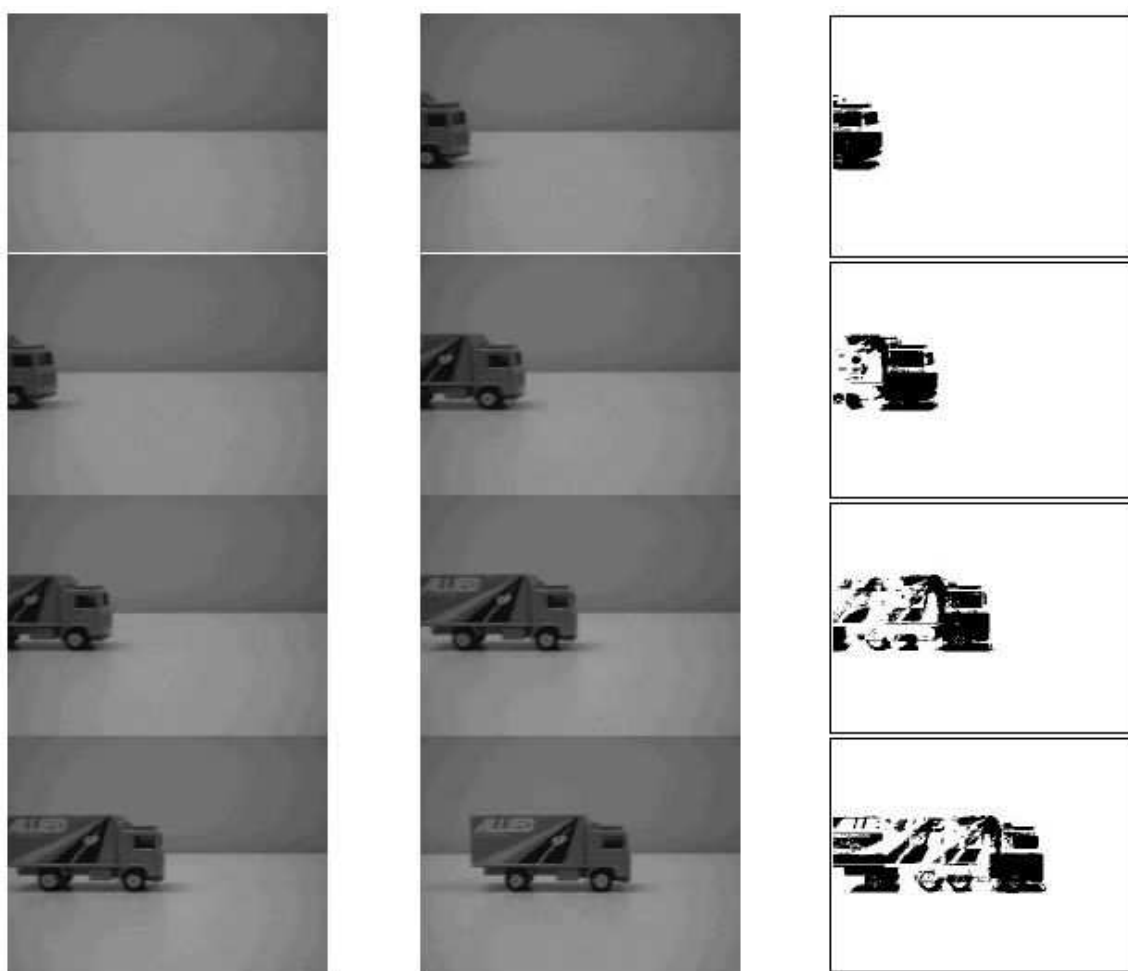
- [18] HOFFMAN, *An Introduction to Fourier Theory*, HOFFMAN, Forest, acessado através do endereço <http://aurora.phys.utk.edu/~forrest/papers/fourier/#introduction> em março de 2004.
- [19] HOLLAND, *Adaptation in Natural and Artificial Systems*, HOLLAND, John, The University of Michigan Press, 1975.
- [20] HORNAK, *The Basics of MRI*, HORNAK, Joseph P., acessado através do endereço www.cis.rit.edu/htbooks/mri/ em novembro de 2003.
- [21] HOWSTUFFWORKS I, *How MRI Works*, acessado via web através do endereço www.howstuffworks.com/mri.htm em outubro de 2003.
- [22] HOWSTUFFWORKS II, *How Ultrasound Works*, acessado via web através do endereço www.howstuffworks.com/ultrasound.htm em outubro de 2003.
- [23] IPAC, Infrared Processing and Analysis Center - California Institute of Technology, acessado via web através do endereço www.ipac.caltech.edu em outubro de 2003.
- [24] IORIO & IORIO, *Fourier Analysis and Partial Differential Equations*, IORIO, Rafael; IORIO, Valéria de Magalhães, Cambridge University Press, 2001.
- [25] JAIN, *Fundamentals of Digital Image Processing*, JAIN, Anil K., Prentice Hall, 1989.
- [26] JAMUNDÁ, *Um Sistema de Vigilância com Detecção de Intrusão Utilizando Inteligência Artificial*, JAMUNDÁ, Teobaldo, Dissertação de Mestrado, Universidade Federal de Santa Catarina, 2002.
- [27] JIANG, *Bilinear Interpolation (Digital Image Processing)*, JIANG, Ming, acessado via web através do endereço <http://ct.radiology.uiowa.edu/~jiangm/courses/dip/html/node67.html> em abril de 2004.
- [28] LEHAR, *Computational Implications of Gestalt Theory I: A Multi-Level Reciprocal Feedback (MLRF) to Model Emergence and Reification in Vi-*

- sual Processing*, LEHAR, Steven, acessado através do endereço <http://cns-alumni.bu.edu/pub/slehar/webstuff/orivar/orivar1.pdf> em abril de 2004.
- [29] NSSDC, *Lunar Exploration*, National Space Science Data Center, acessado através do endereço <http://nssdc.gsfc.nasa.gov/planetary/lunar/> em julho de 2004.
- [30] PANASYUK, *Deblurring*, PANASYUK, Svetlana, acessado através do endereço <http://cfauvcs5.harvard.edu/lana/rs.ip/deblur.htm> em julho de 2004.
- [31] RUMELHART, HINTON & WILLIAMS, *Learning Internal Representations by Error Propagation*, RUMELHART, D.E; HINTON, G. E & WILLIAMS, R. J., Mit Press Computational Models Of Cognition And Perception Series, MIT Press, 1986.
- [32] SATCT, Sidney Airport - Air Traffic Control Tower, acessado via web através do endereço www.acay.com.au/~willt/yssy/enthus/tower.html em novembro de 2003.
- [33] SWRI I, *Real Time X-Ray Imaging Used to Detect Foreign Matter in Food Products*, SouthWest Reasearch Institute, obtido através do endereço www.swri.edu/4org/d14/NDEDept/Monitor/promonit/images/X-Ray.pdf em outubro de 2003.
- [34] SWRI II, *Magnetic Resonance: Effective Process Monitoring Technology*, SouthWest Reasearch Institute, obtido através do endereço www.swri.edu/4org/d14/nddept/monitor/promonit/images/MR.pdf em outubro de 2003.
- [35] TCSMITAIL, *Motion Detection*, The Cog Shop - MIT Artificial Intelligence Laboratory, acessado através do endereço http://www.ai.mit.edu/projects/cog/VisionSystem/motion_detection.html em abril de 2004.
- [36] UCB, *X-Rays*, University of Colorado, acessado via web através do endereço www.colorado.edu/physics/2000/xray/ em outubro de 2003.

- [37] WINKLER & THOME I, *Tratamento de Imagens - Transformações (cont.)*, WINKLER, Letícia; THOMÉ, Antonio Carlos Gay, acessado via web através do endereço http://equipe.nce.ufrj.br/thome/p_grad/nn/transp/c3_tratam_v3b.pdf em março de 2003.
- [38] WINKLER & THOME II, *Tratamento de Imagens - Filtros*, WINKLER, Letícia; THOMÉ, Antonio Carlos Gay, acessado via web através do endereço http://equipe.nce.ufrj.br/thome/p_grad/nn/imgtranspc4_filtros_v4.pdf em maio de 2004.
- [39] WOLFRAM I, *Fourier Transform*, Mathworld, Wolfram Research, acessado via web através do endereço <http://mathworld.wolfram.com/FourierTransform.html> em março de 2004.
- [40] WOLFRAM II, *Basics of LSI Systems: Convolution and Correlation*, Mathworld, Wolfram Research, acessado via web através do endereço <http://documents.wolfram.com/applications/digitalimage/UsersGuide/5.2.html> em janeiro de 2005.

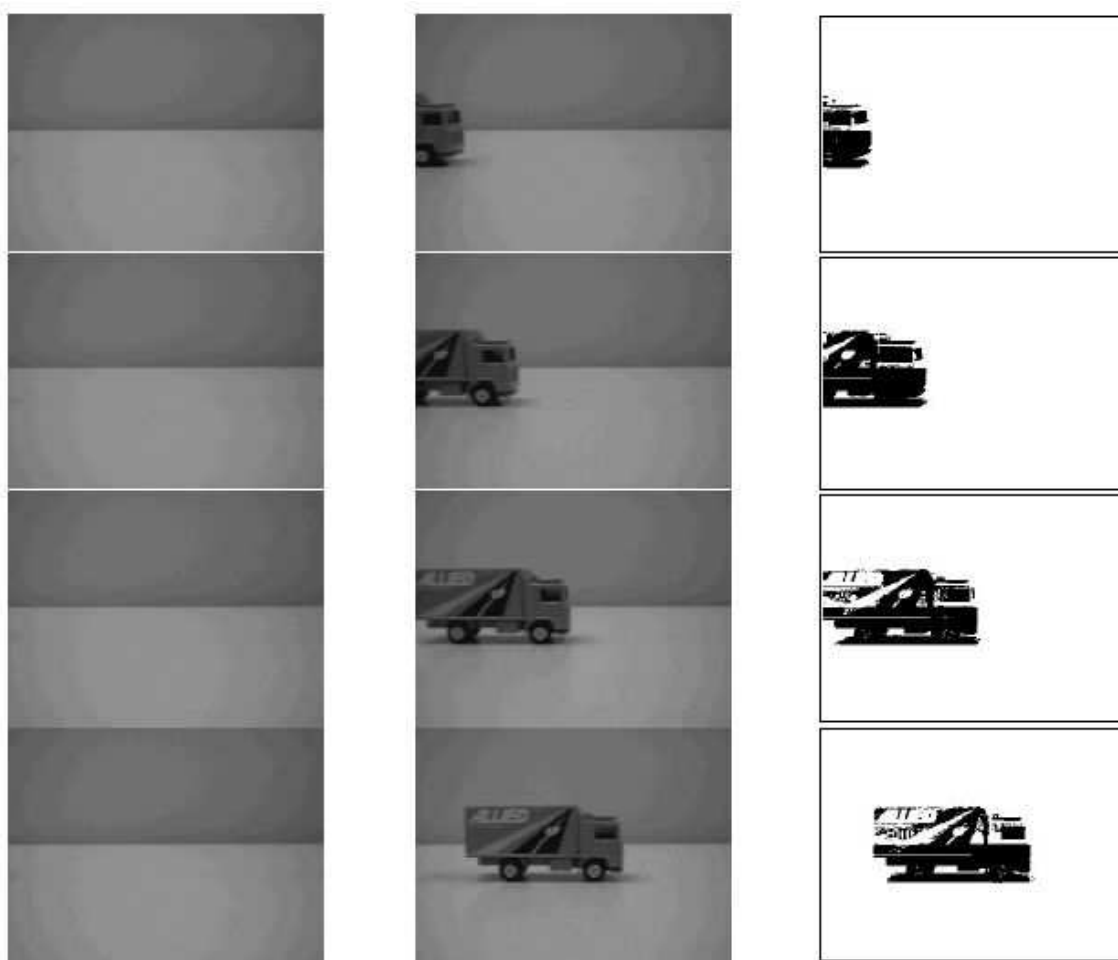
Apêndice I

Neste apêndice, cada quadro da extrema direita apresenta a diferença entre o central e o da extrema esquerda determinada pelo algoritmo de Detecção de Movimento apresentado na Seção 3.10.1. As comparações são realizadas entre quadros consecutivos.



Apêndice II

Este apêndice é análogo ao anterior. Porém, as comparações são realizadas em relação a um quadro estacionário, localizado sempre à extrema esquerda.



Apêndice III

Imagens obtidas na unidade de destilação para a qual o sistema apresentado no 1º estudo de caso deste trabalho, a detecção da queda da primeira gota em um processo de destilação em batelada, foi desenvolvido. À esquerda são mostradas imagens originais, enquanto à direita mostra-se imagens resultantes da técnica de detecção de movimento empregada. A elipse em cada imagem explicita a posição da gota.

