

UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

Pós-Graduação em Ciência da Computação

Gustavo Botelho de Souza

Análise de Formas Planas em Imagens Digitais:
Uma Nova Abordagem Baseada na
Transformada de Hough

UNESP

2013

Gustavo Botelho de Souza

Análise de Formas Planas em Imagens Digitais:
Uma Nova Abordagem Baseada na
Transformada de Hough

Orientador: Prof. Dr. Aparecido Nilceu Marana

Dissertação de Mestrado elaborada junto ao Programa de Pós-Graduação em Ciência da Computação – Área de Concentração em Computação Aplicada, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

UNESP

2013

Souza, Gustavo Botelho de.

Análise de formas planas em imagens digitais : uma nova abordagem baseada na Transformada de Hough / Gustavo Botelho de Souza. -- São José do Rio Preto, 2013

102 f. : il.

Orientador: Aparecido Nilceu Marana

Dissertação (mestrado) ã Universidade Estadual Paulista óJúlio de Mesquita Filho, Instituto de Biociências, Letras e Ciências Exatas

1. Computação gráfica. 2. Processamento de imagens - Técnicas digitais. 3. Reconhecimento de padrões. 4. Formas (Matemática). I. Marana, Aparecido Nilceu. II. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU ã 578.72:76

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Campus de São José do Rio Preto

Gustavo Botelho de Souza

**ANÁLISE DE FORMAS PLANAS EM IMAGENS DIGITAIS: UMA NOVA
ABORDAGEM BASEADA NA TRANSFORMADA DE HOUGH**

Área de Concentração: Computação Aplicada

Linha de Pesquisa: Processamento de Imagens e Visão Computacional

Banca Examinadora:

Prof. Adjunto Aparecido Nilceu Marana
Faculdade de Ciências
Universidade Estadual Paulista - Bauru
(Presidente)

Prof. Dr. João Paulo Papa
Faculdade de Ciências
Universidade Estadual Paulista - Bauru

Prof. Associado Odemir Martinez Bruno
Instituto de Física de São Carlos
Universidade de São Paulo - São Carlos

Resultado:

Bauru, 09 de agosto de 2013.

Agradecimentos

Agradeço primeiramente a Deus pelo dom da vida e pela graça de concluir mais esta etapa em minha história acadêmica.

Aos meus pais, irmã e familiares por todo o apoio e incentivo que deram e pela presença nos momentos fáceis e difíceis.

Também agradeço ao professor Nilceu (Aparecido Nilceu Marana), um exemplo de profissional e amigo, por sua orientação neste trabalho me ajudando a tomar as decisões mais acertadas durante seu desenvolvimento.

Agradeço ainda ao professor João Paulo Papa (Universidade Estadual Paulista “Júlio de Mesquita Filho” - UNESP) pela disponibilização dos códigos dos descritores de formas da literatura testados. À Dra. Sharon Brooks, ao Dr. James McNamara e à Universidade de Michigan, pela disponibilização da base de imagens radiográficas (pertencente ao projeto *Denver Growth Study*) utilizada em um dos experimentos deste trabalho, e à Dra. Flávia Casedei, do Hospital Estadual de Bauru, pela supervisão na segmentação destas imagens radiográficas.

Resumo

Atualmente, dada a difusão dos computadores pela sociedade, a tarefa de se reconhecer padrões visuais está sendo cada vez mais automatizada, em especial para tratar a vasta quantidade de imagens digitais existentes. Aplicações de diversas áreas como biometria, recuperação de imagens baseada em conteúdo e diagnóstico médico, se valem do processamento de imagens bem como de técnicas de extração e análise de características das mesmas a fim de identificar pessoas, objetos, ações, textos, etc. As características básicas que são utilizadas na análise de uma imagem digital são: cor, textura e forma. Neste trabalho são apresentados conceitos fundamentais sobre a análise de formas em imagens digitais, assim como métodos propostos na literatura para tal procedimento. Apresenta-se também um novo método para análise de formas planas, denominado HTS (*Hough Transform Statistics*), com três versões de descritores. Esses três descritores se valem de valores estatísticos extraídos do espaço de Hough para a caracterização das formas de objetos em imagens digitais. Experimentos realizados mostram que estes novos descritores são muito eficazes e eficientes quando comparados com outros propostos na literatura, apresentando desempenhos superiores em muitos casos, tanto no quesito acurácia quanto no tempo de processamento. Os descritores propostos se configuram como boas alternativas, particularmente quando se deseja reconhecer objetos pela forma em grandes bases de dados, contendo imagens com grandes dimensões ou com alta resolução, tendência verificada nos dias atuais dada a popularização e o barateamento dos sensores de captura e das mídias de armazenamento.

Palavras-chave: Reconhecimento de Padrões, Análise de Imagens, Análise de Formas, Transformada de Hough.

Abstract

Nowadays, given the widespread use of computers by society, the task of recognizing visual patterns are increasingly being automated, in particular to address the vast amount of digital images available. Applications of many areas such as biometrics, content-based image retrieval and medical diagnosis, make use of image analysis techniques in order to identify people, objects, actions, texts, etc. in the images. The basic features that can be extracted and analyzed from digital images are color, texture and shape. This paper presents some fundamental concepts about the shape analysis on digital images as well as some methods proposed in the literature. It also proposes a new shape analysis method called HTS (Hough Transform Statistics), with three versions of shape descriptors, which make use of statistical values extracted from the Hough space in order to characterize the shape of objects. Experiments conducted show that these new descriptors are very effective, presenting excellent performance regarding accuracy and processing time. These methods are good alternatives for the shape analysis task especially when working with large or high-resolution images, and with large databases, a tendency in our days, given the popularization and low costs of sensors and storage media.

Keywords: Pattern Recognition, Image Analysis, Shape Analysis, Hough Transform.

Sumário

1	Introdução.....	15
1.1	Objetivo do Trabalho.....	18
1.2	Organização do Trabalho.....	19
2	Análise de Formas	20
2.1	Contexto da Forma	22
2.1.1	Conceitos Básicos.....	22
2.1.2	Correspondência entre Contextos da Forma e Reconhecimento de Objetos.....	28
2.2	BAS (<i>Beam Angle Statistics</i>).....	29
2.2.1	O que são Momentos?	29
2.2.2	O Método BAS	30
2.2.3	Medindo a Similaridade entre Formas com OCS	33
2.3	Dimensão Fractal Multiescala	35
2.4	Saliências do Contorno	40
2.5	Descritores de Fourier	42
2.5.1	Assinaturas das Formas	43
2.5.2	Representação de Formas por Descritores de Fourier.....	45
2.6	O Descritor <i>Tensor Scale</i>	47
3	A Transformada de Hough.....	52
3.1	A Transformada de Hough com o Espaço θ - ρ	54
4	HTS: <i>Hough Transform Statistics</i>	56
4.1	Etapa de Pré-Processamento.....	57
4.2	Etapa de Extração de Características.....	57
4.3	Etapa de Reconhecimento	61
4.4	O Algoritmo do HTS	62
4.5	Invariância a Transformações Geométricas	65
4.6	HTSn: Histogramas Baseados na Vizinhança das Senóides	68
4.7	HTSh: Vetores de Histogramas	70
5	Experimentos, Resultados e Discussão	74
5.1	Primeiro Experimento - Base de Imagens Kimia-216.....	75
5.2	Segundo Experimento - Base de Imagens MPEG-7 Part B	84
5.3	Terceiro Experimento - Tempo de Processamento	89
5.4	Quarto Experimento - Base de Imagens de Sinus Frontais	94
5.5	Análise Geral dos Resultados	100
6	Conclusão	102
7	Trabalhos Futuros.....	104
8	Publicações Realizadas.....	106
	Referências	107

Lista de Figuras

Figura 1. Exemplos de espaços de Hough (matrizes acumuladoras) gerados a partir de silhuetas de maçãs (esquerda) e ossos (direita) da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999).	18
Figura 2. Etapas da análise de imagens: do processamento de imagens ao reconhecimento de padrões. Adaptado de (FALCÃO, 2011).	20
Figura 3. Cálculo dos descritores <i>shape contexts</i> e da correspondência entre os pontos das bordas das imagens originais (que se encontram em (a) e (b)). Extraído de (BELONGIE; MALIK, 2000).	25
Figura 4. Para a construção do histograma (chamado de <i>shape context</i>) de cada ponto selecionado da borda do objeto utiliza-se o diagrama polar-logarítmico em vermelho centrado no ponto sendo analisado.	26
Figura 5. Diagrama polar-logarítmico posicionado sobre um dos pontos da borda de um triângulo para calcular seu histograma (<i>shape context</i>) contando-se a quantidade de pontos presentes em cada compartimento do diagrama.	26
Figura 6. Histograma construído para o ponto sendo analisado na Figura 5. As posições em cinza correspondem aos compartimentos indicados pelas setas na Figura 5.	27
Figura 7. Comparação entre histogramas de três pontos distintos de duas bordas obtidas dos sinus frontal do mesmo indivíduo em momentos diferentes. Extraído de (FALGUERA; MARANA, 2008).	27
Figura 8. Os quatro primeiros pares de raios de um ponto p_i (cada par está representado em uma cor diferente).	31
Figura 9. Ângulo $C_3(i)$ entre os raios da vizinhança de ordem 3 do ponto p_i	31
Figura 10. Imagem binária de uma maçã da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). O ponto de borda indicado pela seta corresponde ao ponto inicial considerado na construção das funções unidimensionais de momentos.	33
Figura 11. Funções unidimensionais geradas a partir dos momentos associados aos pontos de borda da maçã da Figura 10.	33
Figura 12. Exemplo de um fractal: estrela de Koch.	36
Figura 13. Exemplo da Transformada de Distância Euclidiana (SAITO; TORIWAKI, 1994) para a silhueta de uma folha de árvore (em verde).	37
Figura 14. Em (a) exibe-se a silhueta de um morcego da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). Em (b) mostra-se a curva logarítmica (pontos azuis) de $\log(A_r)$ em função de $\log(r)$ da silhueta do morcego, e a reta (em laranja) que melhor representa os pontos desta curva (interpolação linear da curva logarítmica).	38

Figura 15. Dimensão fractal multiescala da silhueta do morcego da Figura 14(a), cuja curva <i>log-log</i> é apresentada na Figura 14(b).	39
Figura 16. Áreas de influência internas (em vermelho) e externas (em verde) de pontos de saliência.	41
Figura 17. Esquema da localização dos pontos de saliência de um polígono em (a) e da criação do vetor de características do mesmo em (b).	42
Figura 18. Normalização de uma forma: pontos são amostrados da mesma e o contorno é reconstruído com base nestes pontos.	46
Figura 19. Cálculo da elipse associada ao ponto vermelho da forma.	49
Figura 20. Os segmentos de reta opostos (em relação ao ponto vermelho) são percorridos simultaneamente no algoritmo proposto por Miranda, Torres e Falcão (2005).	50
Figura 21. Equação da reta na imagem. O parâmetro <i>b</i> (<i>intercept</i>) pode ser encontrado ao se fixar $x=0$ (corresponde ao valor de <i>y</i> quando $x=0$) e o <i>a</i> (<i>slope</i>) corresponde ao coeficiente angular da reta.	52
Figura 22. Nova representação de retas: por meio dos parâmetros θ e ρ	54
Figura 23. (a) Silhueta de uma maçã da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999); (b) Borda da maçã segmentada; (c) Espaço de Hough (matriz acumuladora) gerado a partir dos pontos da borda.	58
Figura 24. Histogramas associados a quatro pontos da borda de dois objetos de mesma classe (sinos) da base MPEG-7 Part B (JEANNIN; BOBER, 1999).	58
Figura 25. Funções obtidas para silhuetas da base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e da base MPEG-7 Part B (JEANNIN; BOBER, 1999).	60
Figura 26. Em (a) é exibida uma imagem da base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e em (b) são exibidas suas funções unidimensionais. As retas verticais tracejadas indicam as posições onde as funções foram amostradas para se gerar o vetor de características para o pássaro.	61
Figura 27. Em (a) exibe-se a imagem (silhueta) de uma maçã da base MPEG-7 Part B (JEANNIN; BOBER, 1999). O ponto de borda localizado na ponta do cabo da maçã (indicado pela seta vermelha) corresponde ao ponto cujo histograma foi verificado antes (b) e depois (c) de o objeto sofrer rotação em torno de seu centro de massa.	66
Figura 28. O histograma do ponto de borda indicado pela seta vermelha em (a) foi verificado antes (b) e após (c) a forma original ter sua escala alterada (dobrou-se a largura e altura da forma). O histograma do ponto praticamente não se alterou.	67
Figura 29. O histograma do ponto indicado pela seta vermelha em (a) foi verificado antes (b) e após (c) a forma (com tamanho de 250 pixels por 250 pixels) ser transladada	

aproximadamente dois mil pixels em relação ao eixo x e também ao y . Novamente pode-se constatar que o histograma não se alterou significativamente. 67

Figura 30. As matrizes com bordas verdes representam uma hipotética parte ampliada do espaço de Hough final de um dado objeto em quatro momentos distintos, de (a) a (d), no processo de construção (pelo descritor HTS) do histograma do ponto de borda associado à senóide cujas posições na matriz aparecem destacadas em vermelho..... 69

Figura 31. De forma análoga à Figura 30, as matrizes de (a) a (d) ilustram o processo de construção (agora pelo descritor HTSn) do histograma do ponto de borda associado à senóide cujas posições na matriz aparecem destacadas em vermelho..... 70

Figura 32. Construção do vetor de características (do método HTSh) para a imagem de um sino da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). 71

Figura 33. Curvas *Precision-Recall* do método HTS obtidas para a base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999) ao se variar o tamanho (número de posições) dos vetores de características..... 75

Figura 34. Exemplos de imagens (de 4 classes) presentes na base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). 76

Figura 35. Curvas *Precision-Recall* para os métodos testados na base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). 77

Figura 36. Curvas *Multiscale Separability* para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) dos quatro melhores métodos em termos de resultados *Precision-Recall*: HTSh, HTS, HTSn e *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003). 77

Figura 37. Exemplos das imagens da base MPEG-7 Part B (JEANNIN; BOBER, 1999). 84

Figura 38. Curvas *Precision-Recall* para os métodos testados na base MPEG-7 Part B (JEANNIN; BOBER, 1999). 85

Figura 39. Curvas *Multiscale Separability* para a base MPEG-7 Part B (JEANNIN; BOBER, 1999) dos quatro melhores métodos em termos de resultados *Precision-Recall*: HTSh, BAS (ARICA; VURAL, 2003), HTSn e HTS. 86

Figura 40. Matrizes de confusão e gráficos com as taxas de acerto e erro por classe (obtidas a partir destas matrizes) para a base MPEG-7 Part B (JEANNIN; BOBER, 1999). 87

Figura 41. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTS (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens. ... 90

Figura 42. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSn (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens. ... 90

Figura 43. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSh (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens. ... 91

Figura 44. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTS (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.....	92
Figura 45. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSn (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.....	92
Figura 46. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSh (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.....	93
Figura 47. (a) Imagem radiográfica original do crânio de um indivíduo e (b) a mesma imagem com a borda dos sinus frontais demarcada manualmente com o auxílio de uma médica radiologista.	95
Figura 48. Curvas ROC dos descritores testados no reconhecimento das pessoas da base de sinus frontais.....	96

Lista de Tabelas

Tabela 1. Exemplos de aplicações nas quais a análise de formas é fundamental (COSTA; JÚNIOR, 2000).....	17
Tabela 2. Matriz de confusão do BAS (ARICA; VURAL, 2003) para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).	78
Tabela 3. Taxas de acerto e erro por classe da matriz de confusão do BAS (ARICA; VURAL, 2003) sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).	79
Tabela 4. Matriz de confusão do descritor HTS para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).....	79
Tabela 5. Taxas de acerto e erro por classe da matriz de confusão do HTS sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).	80
Tabela 6. Matriz de confusão do HTSn para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).....	80
Tabela 7. Taxas de acerto e erro por classe da matriz de confusão do HTSn sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).	81
Tabela 8. Matriz de confusão do HTSh para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).....	81
Tabela 9. Taxas de acerto e erro por classe da matriz de confusão do HTSh sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).....	82
Tabela 10. Tempos aproximados em segundos tomados pelos quatro métodos em questão sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) nas etapas de Extração de Características e Reconhecimento.	83
Tabela 11. Tempos aproximados em segundos tomados pelos quatro métodos em questão sobre a base MPEG-7 Part B (JEANNIN; BOBER, 1999) nas etapas de Extração de Características e Reconhecimento.	88
Tabela 12. Taxa EER e valor “Top 1” (aproximados) para os descritores testados.....	97
Tabela 13. Matriz de confusão do BAS (ARICA; VURAL, 2003) para a base do projeto <i>Denver Growth Study</i>	98
Tabela 14. Matriz de confusão da abordagem HTSh para a base do projeto <i>Denver Growth Study</i>	99

Lista de Siglas

BAS - *Beam Angle Statistics*
CBIR - *Content-Based Image Retrieval*
CMC - *Cumulative Matching Curve*
CS - *Contour Saliency*
MFD - *Multiscale Fractal Dimension*
DNA - *Deoxyribonucleic Acid*
EER - *Equal Error Rate*
FD - *Fourier Descriptors*
FFT - *Fast Fourier Transform*
HTS - *Hough Transform Statistics*
HTSh - *Hough Transform Statistics Histograms*
HTSn - *Hough Transform Statistics Neighborhood*
NN - *Nearest Neighbor*
OCR - *Optical Character Recognition*
OCS - *Optimal Correspondent Subsequence*
PCA - *Principal Component Analysis*
RAM - *Random Access Memory*
ROC - *Receiver Operating Characteristic*
SC - *Shape Context*
TS - *Tensor Scale*
UNESP - *Universidade Estadual Paulista*

1 Introdução

A visão sempre foi um recurso poderoso e importante ao ser humano, uma vez que é principalmente através dela que o homem consegue perceber o mundo que o cerca. Os homens primitivos já reconheciam seus aliados e inimigos valendo-se principalmente da visão e, antes mesmo do surgimento da escrita, também registravam fatos ocorridos por meio de representações visuais (desenhos em cavernas).

A habilidade que o ser humano tem no trato com imagens é extremamente grande. Desde seu surgimento, o homem desenvolveu sistemas neurais e cognitivos altamente sofisticados voltados à tarefa de “processamento de imagens” e “reconhecimento de padrões visuais” (DUDA; HART; STORK, 2000).

Com o advento dos computadores, pesquisadores e empresas de todo o mundo direcionaram esforços para a automatização do processo de captura, processamento, extração de características e reconhecimento de padrões visuais presentes nas imagens digitais, dada a grande variedade de aplicações.

Inúmeras atividades realizadas por seres humanos podem ser executadas em menor tempo, e em alguns casos com maior precisão, se valendo do reconhecimento automático de padrões visuais, tais como a identificação de pessoas a partir de características biométricas passíveis de extração através de imagens digitais, o reconhecimento óptico de caracteres, a análise de sequências de DNA, a busca de imagens baseada em seus conteúdos, conhecida como CBIR (*Content-Based Image Retrieval*), dentre outras.

Atualmente, diversas técnicas de análise de imagens (que engloba o processamento de imagens, a extração de características e o reconhecimento de padrões) e *hardwares* mais eficientes estão disponíveis para uso na construção de sistemas de reconhecimento de padrões visuais. Vale mencionar que muitas das técnicas e *hardwares* criados apresentam princípios e formas de funcionamento baseados na natureza, em especial no sistema responsável pela visão e “processamento” de imagens do homem: as câmeras fotográficas digitais, por exemplo, apresentam sensores sensíveis à luz de forma similar ao que ocorre com os cones e bastonetes presentes no olho humano (DUDA; HART; STORK, 2000).

Ciências de diversas áreas, como a inteligência artificial, também têm contribuído muito com a análise de imagens disponibilizando técnicas e métodos mais eficazes e eficientes para se trabalhar com problemas complexos típicos desta atividade.

Apesar de tudo isto, os sistemas computacionais para reconhecimento de padrões visuais, em geral, ainda estão muito aquém do sistema visual humano em relação ao tratamento de imagens. Dentro deste contexto, novas técnicas de análise de imagens continuam surgindo, aprimorando as já existentes ou tratando os problemas a partir de novas abordagens na tentativa de automatizar o processo de análise de imagens.

As principais características que podem ser extraídas e analisadas a partir de uma imagem digital são a cor, a textura e a forma de objetos ou regiões. Em relação à cor, os métodos de análise de cor se voltam à distribuição das cores nos objetos sob análise (FALCÃO, 2008a). Em alguns casos leva-se em consideração a informação espacial, isto é, relaciona-se as cores às respectivas regiões dos objetos em que elas aparecem. Em outros métodos observa-se de forma geral as cores que aparecem nos objetos. O histograma de cor de uma imagem é o método de análise de cor mais conhecido e um dos menos eficazes (FALCÃO, 2008a).

Em relação à textura, muitos dos métodos de análise de imagens que visam extrair e reconhecer a textura de objetos nas imagens, em geral se baseiam em estatísticas acerca dos padrões de valores (tons de cinza) dos pixels da imagem, como por exemplo no caso das matrizes de co-ocorrência (FALCÃO, 2008a).

A forma, por sua vez, corresponde a uma das características mais significativas que podem ser utilizadas na análise de imagens, sendo em alguns casos a única característica disponível, por exemplo, no reconhecimento de objetos baseado em suas silhuetas. A análise de formas pode ser muito útil aos sistemas de análise de imagens, uma vez que fornece informações de fácil extração e representação, bem como de grande poder de distinção entre objetos. A Tabela 1 apresenta algumas aplicações onde a análise de formas pode ser imprescindível (COSTA; JÚNIOR, 2000).

Neste trabalho é apresentada uma revisão da literatura sobre os principais métodos de análise de formas e é proposto um novo método, denominado HTS (*Hough Transform Statistics*), o qual se baseia em medidas estatísticas extraídas dos espaços de Hough (DUDA; HART, 1972) gerados a partir dos pontos das bordas dos objetos presentes nas imagens sendo analisadas. Estes espaços, conforme pode-se observar na Figura 1, são muito semelhantes quando gerados a partir de formas similares e bastante distintos, caso contrário. Com isto, eles podem ser usados como fontes de informação de fácil construção e simples consulta, uma vez que utiliza-se matrizes acumuladoras para representá-los, a fim de reconhecer formas, justificando seu uso. Duas outras versões do HTS denominadas HTSn (*HTS Neighborhood*) e HTSh (*HTS Histograms*) também são apresentadas.

Tabela 1. Exemplos de aplicações nas quais a análise de formas é fundamental (COSTA; JÚNIOR, 2000).

Área do Conhecimento	Exemplos de Aplicações
Neurociência	Taxonomia morfológica de neurônios, investigações sobre a relação função e forma das células, comparação entre células de diferentes áreas corticais e de diferentes espécies, modelagem de células biologicamente realistas.
Análise de Documentos	Análise de documentos eletrônicos (textuais ou visuais) como em aplicações CBIR, sistemas OCR (<i>Optical Character Recognition</i>), análise de dados em banco de dados multimídia.
“Artes Visuais”	Restauração de vídeo, efeitos especiais, monitoramento de vídeo, jogos, computação gráfica, síntese de imagens.
Medicina	Identificação de tumor, quantificação de mudança e deformação de estruturas anatômicas, análise numérica de cromossomos, identificação de doenças genéticas.
Biologia	Identificação de espécies, taxonomia, relação entre forma e função, comparativo de anatomias, citologia, identificação e contagem de células (como glóbulos brancos no sangue), caracterização de células e formas nucleares, crescimento e modificação de formas de estruturas, análise da forma de caminhar dos seres vivos.
Física	Aplicações envolvendo microscopia como a análise da trajetória, comportamento e distribuição de partículas em um meio ou material, análise de estruturas como polímeros e cristais, caracterização de grupos de estrelas ou análise de propriedades de corpos celestes em astronomia, análise do movimento de objetos macroscópicos (velocidade, aceleração, etc.).
Engenharia	Controle visual de qualidade de produtos em linha de produção, detecção de perigo, interpretação (pelas máquinas) de desenhos feitos à mão, automação, robótica, sensoriamento remoto.
Segurança	Detecção de impressão digital, face e íris, verificação de assinatura e modo de andar de uma pessoa (biometria).
Agricultura	Controle de colheita, contagem de sementes, análise da maturação de frutos.

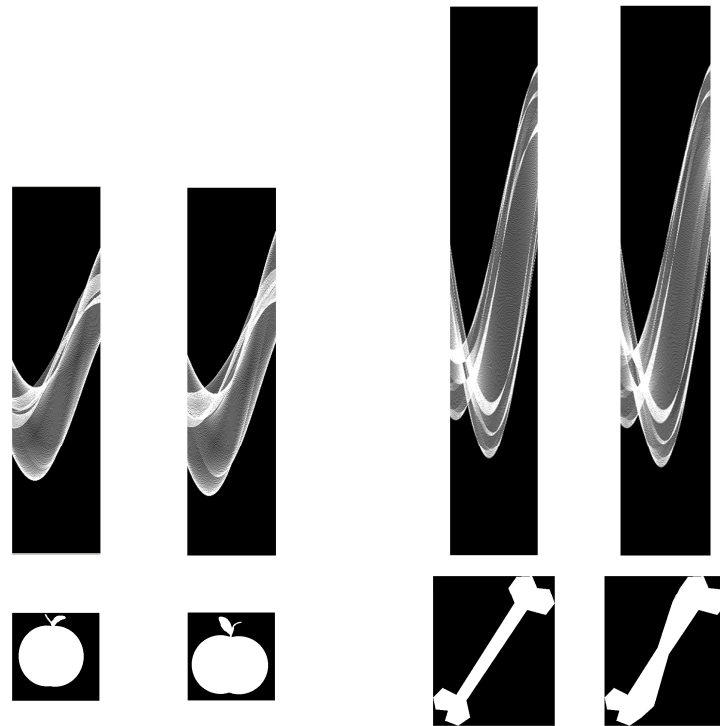


Figura 1. Exemplos de espaços de Hough (matrizes acumuladoras) gerados a partir de silhuetas de maçãs (esquerda) e ossos (direita) da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). Cada espaço de Hough foi gerado a partir da borda da respectiva forma. As regiões mais claras do espaço de Hough indicam grande concentração de senóides as quais representam cada um dos pontos da borda da forma da imagem.

1.1 Objetivo do Trabalho

Esta dissertação de mestrado tem os seguintes objetivos:

1. **Objetivo Geral:** fazer uma revisão da literatura sobre a análise de formas em imagens digitais, apresentando alguns dos principais descritores de forma encontrados, visando compreender seus princípios, suas propriedades, suas peculiaridades e seus desempenhos;
2. **Objetivos Específicos:** i) propor novos descritores de formas planas baseado na Transformada de Hough, visando obter maior precisão e menor tempo de processamento; ii) desenvolver algoritmos para os novos descritores propostos; iii) avaliar os desempenhos dos novos descritores propostos, comparando-os com os desempenhos dos descritores encontrados na literatura, utilizando-se bases de imagens disponíveis publicamente.

1.2 Organização do Trabalho

Esta dissertação de mestrado está organizada em oito capítulos. No Capítulo 1 apresenta-se uma introdução ao tema da dissertação, bem como seus objetivos gerais e específicos.

No Capítulo 2 apresenta-se uma revisão da literatura descrevendo-se alguns dos principais métodos de análise de formas.

No Capítulo 3 descreve-se a Transformada de Hough em sua versão original e em sua versão “melhorada”.

No Capítulo 4 apresenta-se o novo método para análise de formas planas baseado na Transformada de Hough, denominado HTS (*Hough Transform Statistics*), proposto nesta dissertação de mestrado, bem como dois outros diferentes descritores de formas derivados deste novo método.

No Capítulo 5 apresenta-se resultados experimentais obtidos com os novos descritores de formas propostos nesta dissertação de mestrado, bem como algumas considerações.

No Capítulo 6 apresenta-se a conclusão desta dissertação de mestrado e no Capítulo 7 apresenta-se alguns trabalhos futuros.

Por fim, no Capítulo 8, cita-se os artigos elaborados e publicados durante o desenvolvimento desta dissertação, relacionados ao seu tema.

2 Análise de Formas

A análise de imagens digitais tem como objetivo extrair e analisar informações do conteúdo de uma dada imagem de modo a poder subsidiar uma tomada de decisão como, por exemplo, a identificação de pessoas para fins de segurança por meio de imagens faciais ou imagens de impressões digitais, a identificação de tumores ou outras estruturas biológicas em imagens médicas facilitando e auxiliando o diagnóstico, o reconhecimento de objetos em imagens a fim de descobrir informações sobre seus conteúdos, etc. (FALCÃO, 2008b).

Resumidamente, a análise de imagens engloba não só técnicas de processamento de imagens e extração de conteúdo, mas também técnicas de representação, descrição e classificação de padrões. A Figura 2 ilustra as etapas da análise de imagens dentre um conjunto de procedimentos.

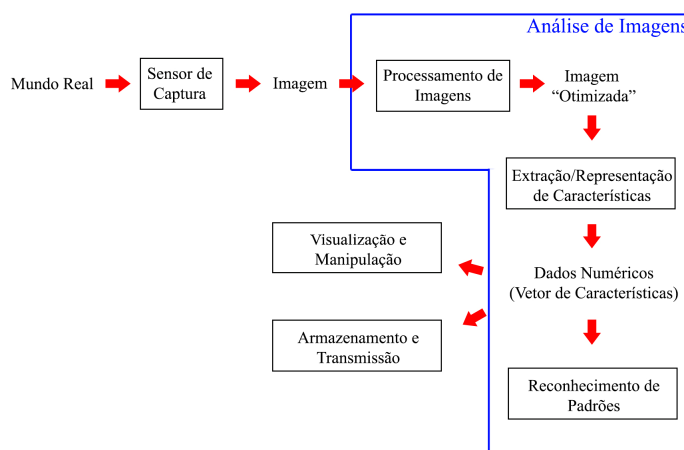


Figura 2. Etapas da análise de imagens: do processamento de imagens ao reconhecimento de padrões. Adaptado de (FALCÃO, 2011).

Como mostrado na Figura 2, a partir da captura da imagem digital por algum sensor, muitas vezes é necessário que se melhore sua qualidade ressaltando aspectos da mesma, eliminando ruídos, etc. a fim de facilitar e tornar mais eficaz e eficiente a atividade que será realizada sobre a imagem posteriormente. Esta tarefa de adequação da imagem original obtida compete ao processamento de imagens digitais.

Após esta otimização da imagem, algoritmos de extração e representação de características são aplicados sobre ela a fim de se identificar características nela existentes e compor, a partir destas, um vetor de características que representará a mesma. A partir dos vetores de características da imagem sob consulta e das imagens da base de dados, e também

de alguma função de comparação entre estes, realiza-se o processo de reconhecimento de padrões identificando o conteúdo da imagem sendo analisada.

Como dito, dentre as diversas características que podem ser usadas no reconhecimento de objetos presentes em imagens digitais encontra-se a forma. A análise de formas nada mais é do que a própria análise de imagens, mas voltada ao estudo da característica “forma” de objetos e regiões das imagens a fim de classificá-los.

Apesar de o ser humano lidar com o reconhecimento de formas a todo momento, a definição do conceito “forma” não é fácil de ser elaborada. Uma importante definição é apresentada por Costa e Júnior (2000) para os quais uma forma pode ser entendida como um conjunto de pontos conectados (quer no espaço discreto ou no contínuo) os quais apresentam entre si um padrão de distribuição.

Como dito, na análise de imagens, após a imagem ter sido capturada por um sensor, otimizada pelo processamento de imagens (melhorando o contraste, segmentando os objetos a ser identificados, detectando as bordas dos objetos, etc.), tem-se as fases de representação e descrição da mesma. No caso da análise de formas, no início destas etapas os objetos presentes nas imagens podem, por exemplo, estar segmentados (representados por todos seus pixels) ou ainda estar representados apenas por suas bordas, dependendo das operações realizadas na etapa de processamento de imagens. Dada a representação dos objetos na imagem, para identificá-los em relação às suas formas é necessário quantificar algumas de suas propriedades, processo chamado de caracterização da forma (COSTA; JÚNIOR, 2000). Esta caracterização pode ser feita em termos de suas medidas e características como área, perímetro, número de concavidades, número de extremidades, curvatura da borda, pertinência a alguma classe (como polígonos), posição dos seus pontos de borda, dentre outras.

Ao escolher as características que serão analisadas das formas na aplicação sob estudo, se está, na verdade, definindo o vetor de características das mesmas o qual será usado posteriormente pelo algoritmo de reconhecimento de padrões (etapa posterior à fase de representação e descrição na análise de imagens). Se os vetores têm n características se operará no espaço $\mathcal{R}^n$.

Uma cuidadosa seleção das características que irão compor os vetores de características pode beneficiar muito o processo de reconhecimento de padrões uma vez que, ao se escolher as características adequadas, os agrupamentos de vetores de características de classes diferentes estarão bem concentrados em algumas posições do espaço de características e bem espaçados entre si, facilitando e tornando mais precisa a classificação dos padrões (COSTA; JÚNIOR, 2000). Entretanto, a escolha das características adequadas é um processo delicado e

difícil e deve observar toda a informação disponível sobre as peculiaridades e variabilidades das formas nas imagens sendo tratadas.

A partir dos vetores de características pode-se definir uma medida de distância entre eles criando-se assim um descritor de formas. Com isto, inicia-se a etapa de reconhecimento de padrões que se utiliza de classificadores para o reconhecimento das formas dos objetos.

É importante observar que existem duas abordagens na representação de formas: as que trabalham com regiões do objeto e as que trabalham somente com o contorno (borda) do mesmo. Em ambas as abordagens, é necessário que o algoritmo de processamento de imagens segmente corretamente os objetos na imagem.

Vale ressaltar ainda que, segundo Latecki e Lakämper (2000), independente do método usado na análise de formas, ele deve funcionar como funciona o sistema de percepção visual humano. Isto significa que ele deve apresentar as seguintes características: (i) a medida de similaridade entre formas deve permitir o reconhecimento de objetos visualmente similares, mesmo que não sejam matematicamente idênticos; (ii) deve abstrair distorções, como ruídos da digitalização e erros de segmentação; (iii) deve dar atenção às partes visuais significantes dos objetos e (iv) não deve depender da escala, orientação e posição dos mesmos.

Em relação ao último item, vale observar que em alguns casos a invariância a certas transformações como rotação, translação e mudança de escala pode atrapalhar o desempenho do sistema, por exemplo: um sistema de reconhecimento de caracteres invariante à rotação seria incapaz de distinguir o dígito “6” do “9”. Desta forma, ao se criar um método de análise de formas deve-se analisar a aplicação e as características das imagens que serão tratadas a fim de definir se ele deve ser invariante ou não a cada uma destas transformações mencionadas.

2.1 Contexto da Forma

Um dos tradicionais descritores de formas, proposto por Belongie e Malik (2000), é o *Shape Context* (Contexto da Forma), que caracteriza cada ponto da borda da forma com base nas posições dos demais pontos de borda em relação ao mesmo.

2.1.1 Conceitos Básicos

Belongie e Malik (2000) apresentam um descritor de formas (planas) chamado de *Shape Context* (Contexto da Forma). Este descritor, após a seleção de alguns pontos da borda

segmentada de um objeto, atua vinculando a cada um deles um histograma que representa a distribuição de orientação e distância dos demais pontos selecionados da borda em relação ao ponto sendo tratado no momento. Assim, pontos correspondentes em duas bordas (segmentadas) semelhantes apresentam histogramas parecidos.

Segundo Belongie e Malik (2000), a análise de formas baseada no *Shape Context* tem início tomando-se N pontos da borda, previamente detectada e segmentada através do algoritmo de Canny (1986), por exemplo, do objeto sob análise. Estes pontos podem ser de contornos internos ou externos do objeto e não precisam ser pontos-chave da borda (como pontos de inflexão ou de curvatura máxima). Os autores aconselham que os pontos selecionados estejam de certa forma espaçados uniformemente pela borda. Como exemplo, dada a imagem da letra “A” na Figura 3(a), a segmentação das bordas da mesma bem como a seleção de alguns pontos destas bordas (que aparecem no início de cada vetor tangente à borda) são mostradas na Figura 3(c).

Considerando os vetores originados em um dos pontos selecionados da borda da letra “A” em sentido aos demais pontos selecionados da mesma borda (Figura 3(d)), pode-se dizer que estes vetores expressam a aparência da forma toda em relação ao ponto de referência. Obviamente, este conjunto de vetores (trabalhados um a um) é uma descrição muito rica, e caso se tome N pontos da borda sendo analisada, com N suficientemente grande (igual ao número total de pontos da borda detectada e segmentada), tem-se uma representação exata da forma em relação ao ponto de referência através deste conjunto. Entretanto, como os autores desejam detectar formas semelhantes, mas que podem apresentar certas variações de uma instância para outra, o uso deste conjunto de vetores como um descritor de formas é inadequado.

Para tratar este problema, Belongie e Malik (2000) propõem um descritor compacto para cada ponto (dos N selecionados da borda do objeto) por meio da geração de um histograma das coordenadas relativas dos outros pontos selecionados da borda em relação ao ponto de referência. Para diferenciar mais precisamente as variações de posição dos pontos selecionados da borda (“términos” dos vetores) mais próximos ao ponto de referência no momento e de forma mais grosseira as variações dos pontos mais distantes uma vez que podem haver distorções, ruídos, etc. é apresentado para a construção do histograma associado a cada ponto um sistema de coordenadas polar-logarítmicas cuja origem é o ponto sob análise no momento (ponto de referência) como mostrado na Figura 4 (no caso há 60 compartimentos no diagrama polar-logarítmico, isto é, 5 intervalos de distâncias por 12 intervalos de ângulos).

Um exemplo de um histograma computado com base em um ponto de referência e no conjunto de vetores associados ao mesmo (na realidade nos pontos de borda amostrados, os quais se encontram nas terminações dos vetores) da Figura 3(c) é mostrado na Figura 3(e). Pixels mais escuros representam valores maiores no respectivo compartimento (*bin*) do histograma. Neste exemplo, emprega-se 30 compartimentos no diagrama polar-logarítmico (6 intervalos de ângulo e 5 de distância). Este histograma é denominado de *shape context* (contexto da forma) do ponto de referência, ele incorpora informação global da forma em relação ao ponto ao armazenar a distribuição dos demais pontos selecionados nos compartimentos do diagrama polar-logarítmico.

Como observação, os autores propõem o uso da orientação de referência absoluta (eixo positivo x) na construção do diagrama polar-logarítmico (em relação aos compartimentos de ângulos).

Em síntese o histograma de um ponto p_i pode ser definido por:

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{compartimento}(k)\} \quad (1)$$

onde q e p_i são pontos selecionados da borda do objeto, k varia de 1 ao total de compartimentos existentes no diagrama polar-logarítmico, $\text{compartimento}(k)$ é um dos compartimentos do diagrama polar-logarítmico e $\#$ indica o número de pontos q da borda que atendem à restrição dado p_i .

A Figura 3 ilustra o cálculo dos descritores *shape context* e da correspondência entre os pontos das bordas de duas formas (que se encontram em (a) e (b)). Na Figura 3(c), como dito, é exibida a borda da forma de (a) e os vetores tangentes à mesma nos 50 pontos selecionados desta. Na Figura 3(d) tem-se o conjunto de vetores originados em um dos pontos selecionados da borda da forma de (a) (ponto na região central e à esquerda da letra “A”) e terminados nos outros pontos selecionados da borda (descrição rica porém inapropriada, como explicado). A distância média λ de todos os pares de pontos selecionados é mostrada na parte inferior da imagem. Na Figura 3(e) mostra-se o histograma $\log r, \theta$ associado ao mesmo ponto de referência de (d), com 5 e 6 compartimentos para $\log r$ e θ , respectivamente. As cores mais escuras significam valores mais altos. Na Figura 3(f) mostra-se as correspondências encontradas entre os pontos das bordas de (a) e (b) usando o Método Húngaro (PAPADIMITRIOU; STIEGLITZ, 1982), com pesos dados pela soma de dois termos: diferença entre os histogramas do par de pontos e diferença entre os ângulos tangentes dos mesmos. Ainda na Figura 3, em (g) e (h) são apresentados os *shape contexts* “completos” das

duas letras “A” de (a) e (b), formados pela junção dos histogramas (um por linha) de todos os pontos selecionados nas bordas de (a) e (b), desta forma cada *shape context* completo tem 50 linhas (pois existem 50 pontos selecionados em cada borda) e 30 colunas (por cada histograma apresentar $5 \cdot 6 = 30$ compartimentos).

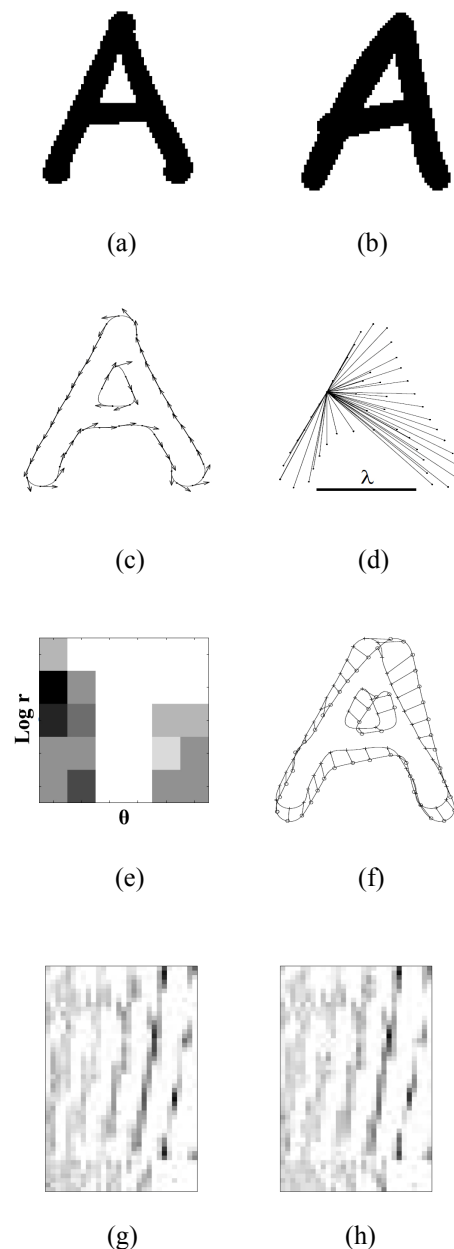


Figura 3. Cálculo dos descritores *shape contexts* e da correspondência entre os pontos das bordas das imagens originais (que se encontram em (a) e (b)). Em (c) são mostrados vetores tangentes aos pontos amostrados da borda de (a). Em (d) mostra-se vetores originados em um ponto da borda da forma de (a) em direção aos demais pontos selecionados da borda, em (e) mostra-se o histograma (*shape context*) deste ponto. Em (f) são apresentadas as correspondências encontradas entre pontos das duas bordas de (a) e (b) e em (g) e (h) os *shape contexts* completos das duas formas (letras “A”) de (a) e (b). Extraído de (BELONGIE; MALIK, 2000).

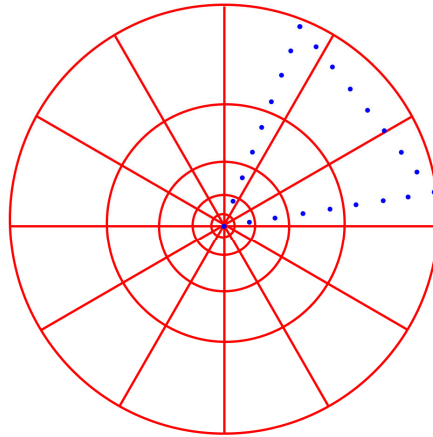


Figura 4. Para a construção do histograma (chamado de *shape context*) de cada ponto selecionado da borda do objeto utiliza-se o diagrama polar-logarítmico em vermelho centrado no ponto sendo analisado. Basicamente o que se faz é verificar quantos pontos existem em cada compartimento do diagrama atribuindo esta soma ao respectivo *bin* do histograma (ao *bin* referente aos respectivos valores de distância, $\log r$, e ângulo, θ).

Resumindo, após escolher pontos da borda segmentada de um objeto usa-se os *shape contexts* (histogramas baseados no diagrama polar-logarítmico) para caracterizar a distribuição dos demais pontos selecionados em relação ao ponto sendo tratado.

A Figura 5 apresenta uma imagem onde aparecem os pontos amostrados da borda de um triângulo e na qual se ilustra o cálculo do histograma de um dos pontos da borda (vértice do triângulo) com base em seu diagrama polar-logarítmico. Para tanto, conta-se o número de pontos da borda em cada compartimento do diagrama. A Figura 6 ilustra o histograma (*shape context*) do ponto da borda analisado na Figura 5 com destaque para os *bins* associados aos 4 compartimentos do diagrama polar-logarítmico apontados na Figura 5.

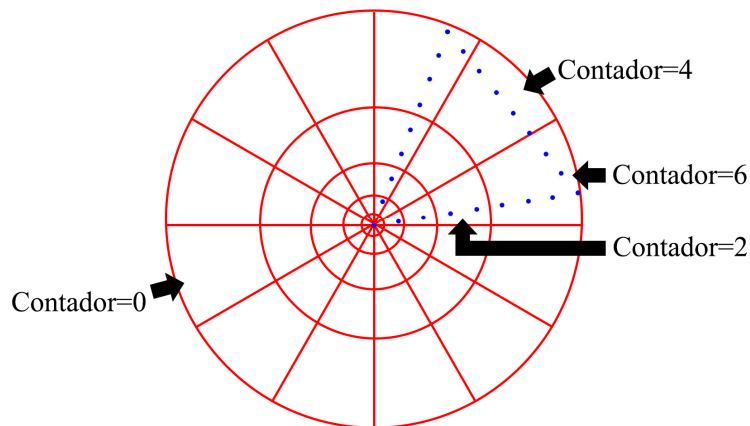


Figura 5. Diagrama polar-logarítmico posicionado sobre um dos pontos da borda de um triângulo para calcular seu histograma (*shape context*) contando-se a quantidade de pontos presentes em cada compartimento do diagrama.

	6	4	5	0	0	0	0	0	0	0	0	0
	2	0	2	0	0	0	0	0	0	0	0	0
Log r	1	0	1	0	0	0	0	0	0	0	0	0
	1	0	1	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0
	θ											

Figura 6. Histograma construído para o ponto sendo analisado na Figura 5. As posições em cinza correspondem aos compartimentos indicados pelas setas na Figura 5.

A Figura 7 ilustra pontos de borda (e seus histogramas) de duas imagens do sinus frontal, cavidade existente no osso frontal do crânio humano e que pode ser vista em imagens radiográficas, de um mesmo indivíduo obtidas momentos diferentes. Nota-se que os pontos das bordas localizados em regiões semelhantes apresentam histogramas semelhantes. Nota-se também que o terceiro ponto de borda (o localizado mais à direita), posicionado em uma região da borda bastante diferente das dos anteriores, apresenta um histograma completamente diferente. Para facilitar a comparação entre os histogramas, seus valores foram substituídos por níveis de cinza: quanto mais escura a cor da posição no histograma, maior é a ocorrência de pontos da borda naquele intervalo de ângulo e distância no diagrama polar-logarítmico (FALGUERA; MARANA, 2008).

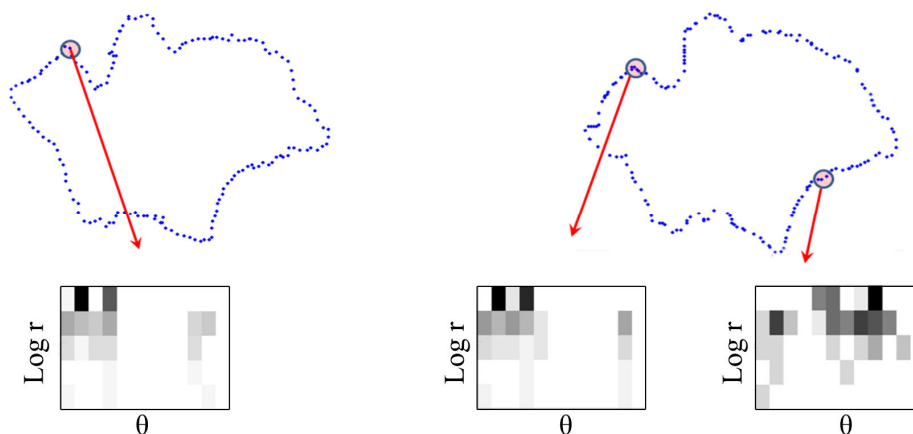


Figura 7. Comparação entre histogramas de três pontos distintos de duas bordas obtidas dos sinus frontais do mesmo indivíduo em momentos diferentes. Percebe-se que os dois pontos localizados em regiões semelhantes das bordas (os dois mais acima na imagem) apresentam histogramas parecidos enquanto que o terceiro ponto, por estar em outra região da borda, apresenta histograma bastante diferente dos outros dois. Extraído de (FALGUERA; MARANA, 2008).

2.1.2 Correspondência entre Contextos da Forma e Reconhecimento de Objetos

Para determinar as correspondências entre duas formas (entre seus pontos) baseado no uso do descritor apresentado (*Shape Context*), Belongie e Malik (2000) propõem dois critérios: (i) pontos correspondentes nas formas devem ter descritores (histogramas polar-logarítmicos) semelhantes, e (ii) as correspondências entre pontos devem ser únicas.

Em relação ao primeiro critério, o custo do “casamento” de dois pontos (um de cada forma) é composto de dois termos: um relacionado à similaridade dos *shape contexts* dos pontos (C_S) e outro à aparência local dos mesmos (C_A). O termo C_S , relacionado aos *shape contexts* (histogramas polar-logarítmicos) dos pontos, é dado pela distância χ^2 entre os histogramas dos pontos. Denotando os dois *K-bins* histogramas (já normalizados) dos pontos sendo comparados por $g(k)$ e $h(k)$, a distância χ^2 é dada por:

$$C_S = \frac{1}{2} \sum_{k=1}^K \frac{[g(k) - h(k)]^2}{g(k) + h(k)} \quad (2)$$

Já o termo C_A relacionado à comparação da aparência local entre os dois pontos, que no caso de (BELONGIE; MALIK, 2000) corresponde a uma medida de diferença dos ângulos tangentes aos pontos, é definido por:

$$C_A = \frac{1}{2} \cdot \left\| \begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} - \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix} \right\| \quad (3)$$

A partir de C_S e C_A pode-se obter o custo total de “casamento” entre os dois pontos sendo tratados no momento (um de cada borda) através da soma ponderada (escolhido um valor para β no intervalo $[0; 1]$) de C_S e C_A da seguinte forma:

$$C = (1-\beta) \cdot C_S + \beta \cdot C_A \quad (4)$$

Vale observar que $C_S \in [0; 1]$ e $C_A \in [0; 1]$ e assume-se que as duas bordas sendo comparadas têm o mesmo número N de pontos selecionados.

Dado o conjunto de custos C_{ij} de todos os pares de pontos possíveis (onde sempre um pertencente a uma das bordas sendo analisada e outro à outra) e satisfazendo o segundo critério (que afirma que as correspondências entre os pontos das duas formas devem ser únicas), para encontrar a correspondências dos pontos das duas formas sendo analisadas,

minimiza-se o custo total do “casamento” entre os pontos das formas sujeito à restrição de que cada ponto de uma forma só pode ser “casado” com um ponto da outra (o “casamento” é sempre “um para um”). Este problema pode ser resolvido pelo Método Húngaro (PAPADIMITRIOU; STIEGLITZ, 1982) com complexidade $O(n^3)$ e a entrada fornecida ao Método Húngaro é uma matriz de custos quadrada ($N \times N$) cujos valores C_{ij} representam os custos de se “casar” o ponto i da primeira borda sendo analisada com o ponto j da segunda borda.

O resultado da aplicação deste método de correspondências entre formas a partir das formas das Figuras 3(a) e 3(b) pode ser visto, como dito, na Figura 3(f) tendo sido usado um valor de $\beta=0,3$.

O reconhecimento de formas baseado no uso dos *shape contexts* é apresentado por Belongie, Malik e Puzicha (2002). O processo de reconhecimento de objetos com base em suas formas (bordas) pode ser definido em três etapas: (i) resolver o problema de correspondência entre os pontos selecionados das duas bordas sendo comparadas (a borda de consulta e certa borda já conhecida da base de dados) usando a análise dos *shape contexts* dos pontos das bordas, (ii) usar as correspondências para estimar uma transformação de alinhamento dos objetos (de suas bordas), (iii) calcular a distância (que indica o grau de similaridade) entre as duas formas como uma soma dos erros entre os pontos “casados” das bordas das mesmas juntamente com um termo que mede a magnitude da transformação de alinhamento utilizada. Baseado nas distâncias entre as formas da base de dados e a forma de teste pode-se reconhecer a última.

2.2 BAS (*Beam Angle Statistics*)

Outro método que pode ser usado na análise de formas foi proposto por Arica e Vural (2003) e é denominado BAS (*Beam Angle Statistics*). Este método se baseia em medidas estatísticas (momentos) de ângulos gerados a partir de cada ponto da borda para caracterizar e identificar o objeto sob análise pela sua forma.

2.2.1 O que são Momentos?

Dentro da análise estatística define-se o que são os momentos de uma variável. Segundo o Instituto Nacional de Pesquisas Espaciais (1996), considerando que $x_1, x_2, \dots, x_N$ são os N valores assumidos pela variável X , o momento de ordem r de X é definido por:

$$m'_r = \frac{1}{N} \sum_{i=1}^N x_i^r \quad (5)$$

O momento de ordem 1 (isto é, com $r=1$) é a própria média dos valores de X . Por sua vez, o momento de ordem r centrado na média é definido por:

$$m_r = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^r \quad (6)$$

onde $\bar{x} = m'_1$, isto é, a média de X . Como observação o momento centrado na média de ordem 2 é a variância a qual mede a dispersão dos dados em relação à média.

Em geral, quando se trata de momentos de ordem r , com $r \geq 2$, apesar de dizer-se que se está calculando “os momentos de ordem r ”, na realidade emprega-se a equação do “momento centrado na média de ordem r ”.

2.2.2 O Método BAS

O método BAS (ARICA; VURAL, 2003) baseia-se na ideia da representação do contorno de um objeto (que é uma função bidimensional) por meio de uma função unidimensional a qual deve registrar ao máximo as convexidades e concavidades da borda do objeto com o mínimo de distorções. O BAS se vale dos raios (*beams*), que são segmentos de reta (vetores) que conectam cada ponto de borda de referência aos demais pontos da borda, a fim de caracterizar as formas (ARICA; VURAL, 2003).

Uma borda $B = \{p_1, \dots, p_N\}$ é representada por uma sequência de pontos conectados $p_i = (x_i, y_i)$ com $i=1, 2, \dots, N$, sendo N o número de pontos da borda e $p_i = p_{i+N}$, isto é, considerando-se a sequência de pontos de borda como sendo circular. Para cada ponto p_i , os raios de p_i são representados pelo conjunto de vetores $L(p_i) = \{V_{i+j}, V_{i-j}\}$, onde V_{i+j} e V_{i-j} são os vetores posterior e anterior que conectam p_i aos pontos p_{i+j} e p_{i-j} , respectivamente, na borda B , com $j=1, 2, \dots, N/2$.

O sistema de vizinhança de ordem K de um ponto p_i é definido por $p_{i \pm K} \in \eta_K(p_i)$ para todo ponto de borda p_i , com $i=1, 2, \dots, N$ e $K=1, 2, \dots, N/2$. Os vetores (raios) V_{i+K} e V_{i-K} conectam p_i a $p_{i \pm K}$. A Figura 8 ilustra os pares de raios que conectam p_i aos outros pontos da borda.

A inclinação de cada raio que se origina em p_i no sistema de vizinhança K pode ser definida por:

$$\theta_{V_{i+l}} = \tan^{-1} \frac{\Delta y_{i+l}}{\Delta x_{i+l}} \quad (7)$$

onde $l = \pm K$.

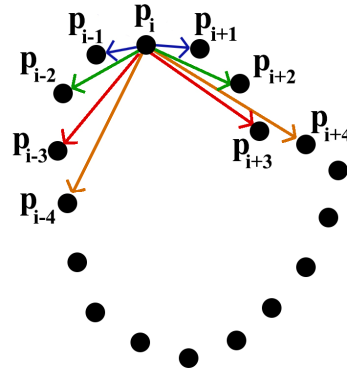


Figura 8. Os quatro primeiros pares de raios de um ponto p_i (cada par está representado em uma cor diferente). Como exemplo, os pontos p_{i+1} e p_{i-1} pertencem à vizinhança $\eta_i(p_i)$ e estão conectados a p_i por meio dos vetores em azul denominados V_{i+1} e V_{i-1} , respectivamente. Como observação, apesar de os pontos da borda da forma (pontos pretos) estarem espaçados entre si, isto ocorre apenas para facilitar a visualização no exemplo, o método BAS (ARICA; VURAL, 2003) não realiza a amostragem de pontos da borda para encontrar os vetores (raios) associados a cada ponto de borda.

O ângulo entre os raios posterior e anterior de p_i no sistema de vizinhança de ordem K , é dado por:

$$C_K(i) = (\theta_{V_{i-K}} - \theta_{V_{i+K}}) \quad (8)$$

A Figura 9 exibe o ângulo entre os raios da vizinhança de ordem 3 ($C_3(i)$) do ponto p_i .

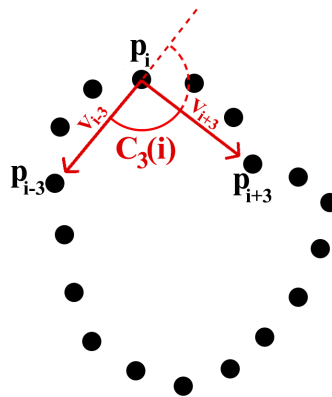


Figura 9. Ângulo $C_3(i)$ entre os raios da vizinhança de ordem 3 do ponto p_i .

Considerando-se, para cada ponto p_i da borda B sendo analisada, o ângulo $C_K(i)$ entre os pares de raios do sistemas de vizinhança $\eta_K(p_i)$ como sendo uma variável aleatória com função de densidade de probabilidade $P_K(C_K(i))$, no BAS (ARICA; VURAL, 2003), associa-se os m primeiros momentos desta variável, isto é, do conjunto de ângulos $C_K(i)$ com $K=1, 2, \dots, N/2$, ao respectivo ponto de borda. O m -ésimo momento da variável aleatória $C_K(i)$ de um ponto p_i da curva (borda) B é definido por:

$$E[C_K^m(i)] = \sum_K C_K^m P_K(C_K(i)), \quad m=1,2,3,\dots \quad (9)$$

onde E indica o valor de operador esperado e $P_K(C_K(i))$, como dito, é a função de densidade de probabilidade de $C_K(i)$. Como já mencionado anteriormente o máximo valor de K é $N/2$, onde N corresponde o número total de pontos da borda. Durante as implementações, segundo os autores de (ARICA; VURAL, 2003), a função $P_K(C_K(i))$ pode ser aproximada pelo histograma de $C_K(i)$ em cada ponto p_i .

Cada ponto de borda então passa a ter associado a si um vetor cujos elementos são os primeiros momentos do seu conjunto de ângulos $C_K(i)$ (com $K=1, 2, \dots, N/2$):

$$V(p_i) = [E[C^1(i)], E[C^2(i)], \dots] \quad (10)$$

Percorrendo-se a borda a partir de um ponto inicial em um dado sentido e observando-se os valores de momentos associados a cada ponto da borda obtém-se funções unidimensionais para representar a borda bidimensional sob análise (cada momento gera uma “função”).

A Figura 11 ilustra as curvas unidimensionais obtidas a partir dos momentos (primeiro, segundo e terceiro momentos) associados aos pontos de borda da Figura 10. A seta vermelha na Figura 10 indica o ponto de borda inicial. A partir dele, a fim de se gerar as funções unidimensionais, percorre-se a borda em sentido horário observando os valores de momentos associados a cada ponto de borda (cada valor de momento gera uma das função). Como dito, os momentos descrevem o comportamento dos ângulos $C_K(i)$ associados a cada ponto p_i . Na prática, para representar uma forma, geralmente são necessários apenas os três ou dois primeiros momentos (ARICA; VURAL, 2003).

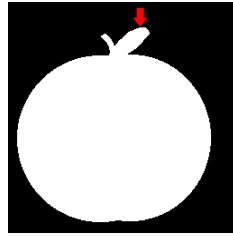


Figura 10. Imagem binária de uma maçã da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). O ponto de borda indicado pela seta corresponde ao ponto inicial considerado na construção das funções unidimensionais de momentos. Percorre-se a borda em sentido horário a partir do ponto inicial observando os valores de momentos associados aos pontos de borda.

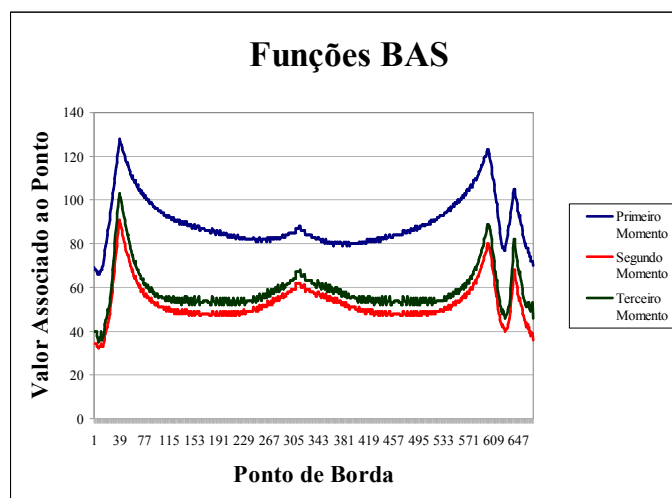


Figura 11. Funções unidimensionais geradas a partir dos momentos associados aos pontos de borda da maçã da Figura 10.

Tendo encontrado as funções unidimensionais que representam a borda bidimensional da forma sob consulta e as funções unidimensionais de certa borda conhecida pode-se comparar as duas bordas com base nas suas funções. Como observação, esta representação da forma proposta no BAS é invariante à escala, orientação e translação (ARICA; VURAL, 2003).

2.2.3 Medindo a Similaridade entre Formas com OCS

A similaridade entre as funções unidimensionais associadas aos momentos obtidas a partir do método BAS (ARICA; VURAL, 2003) para duas formas distintas pode ser estimada pelo algoritmo de subsequência correspondente ótima (OCS - *Optimal Correspondent Subsequence*) apresentado em (WANG; PAVLIDIS, 1990).

O OCS (WANG; PAVLIDIS, 1990) foi desenvolvido para resolver o problema de “casamento ótimo” de *strings* (encontrar correspondências entre vetores) estimando a

similaridade entre estas. Tem-se como objetivo na estimativa de similaridade minimizar a distância entre os dois vetores sob análise permitindo a deformação dos mesmos (inserção de elementos nulos) se necessário. Depois de resolver o problema da correspondência ótima entre os elementos dos vetores (*strings*), considera-se a distância (grau de similaridade) entre eles como sendo a soma dos erros de casamento de seus elementos.

No BAS (ARICA; VURAL, 2003), dadas as três (ou duas) funções unidimensionais de uma forma desconhecida, elas são amostradas em posições igualmente espaçadas gerando-se o vetor de características da forma. Em cada posição do vetor ficam armazenados os três (ou dois) valores de momentos amostrados das funções em uma dada posição. Para reconhecer a forma sob análise, seu vetor de características sofre o processo de “casamento ótimo” através do OCS (WANG; PAVLIDIS, 1990) com os vetores da base de dados a fim de verificar a similaridade entre eles.

Em termos matemáticos, dados dois vetores de características Γ e Π extraídos de duas bordas sob comparação a partir de suas funções unidimensionais, geradas pelo BAS (ARICA; VURAL, 2003) e relacionadas aos momentos de seus pontos de borda, com N e M elementos respectivamente, o OCS (WANG; PAVLIDIS, 1990) define uma correspondência $t(i)$ de modo que cada elemento do primeiro vetor, $\Gamma(i)$, é mapeado para um elemento do segundo vetor, $\Pi(t(i))$. A minimização do custo total de casamento dos dois vetores é definida por:

$$C_{total}(\Gamma, \Pi) = \min_{t(i)} \sum_i dist(\Gamma(i), \Pi(t(i))) \quad (11)$$

onde $dist(\Gamma(i), \Pi(t(i)))$ representa a distância entre os elementos $\Gamma(i)$ e $\Pi(t(i))$ e é definida por:

$$dist(\Gamma(i), \Pi(j)) = \sum_l (\Gamma_l(i), \Pi_l(j))^2 \quad (12)$$

onde l indica a ordem dos valores de momentos sendo comparados. Em outras palavras, a distância $dist(\Gamma(i), \Pi(j))$ entre os elementos $\Gamma(i)$ e $\Pi(j)$ dos vetores Γ e Π é definida como a soma das distâncias euclidianas ao quadrado dos três (ou dois) valores de momentos presentes nestas posições.

O problema da minimização do custo total $C_{total}(\Gamma, \Pi)$ de casamento entre dois vetores de características resolvido pelo OCS (WANG; PAVLIDIS, 1990) pode ser descrito através das relações de recorrência:

$$D(i, j) = \min \{ [D(i-1, j-1) + \text{dist}(\Gamma(i), \Pi(j))], [D(i-1, j) + r], [D(i, j-1) + r] \} \quad (13)$$

Nesta relação, $D(i, j)$ é a distância cumulativa entre os elementos $\Gamma(i)$ e $\Pi(j)$ e r é uma constante. Considerando-se $D(0, 0) = 0$, a relação de recorrência da distância cumulativa é usada para obter a correspondência ótima entre os elementos dos vetores. Os três termos da relação podem ser explicados por:

- O primeiro termo, caso selecionado, indica que se “casou” $\Gamma(i)$ e $\Pi(j)$, assim o custo cumulativo é igual à distância entre $\Gamma(i)$ e $\Pi(j)$ mais o custo de se “casar” o próximo elemento de Γ , $\Gamma(i-1)$, com o próximo elemento de Π , $\Pi(j-1)$;
- O segundo termo indica que o elemento $\Gamma(i)$ fica sem se “casar” com elementos do vetor Π (casa-se este elemento com um elemento nulo), assim o custo relativo é igual a uma penalização r por não ter “casado” $\Gamma(i)$ com elementos de Π mais o custo de se casar o próximo elemento de Γ , $\Gamma(i-1)$ e o elemento $\Pi(j)$;
- O terceiro termo indica que se deixa $\Pi(j)$ sem se “casar” com elementos do vetor Γ (casa-se este elemento com um elemento nulo), passando-se já para a análise do próximo elemento de Π . Com isto o custo relativo é igual a uma penalidade r (por não ter “casado” $\Pi(j)$) mais o custo de se casar o próximo elemento de Π , $\Pi(j-1)$, com $\Gamma(i)$.

Iterando em i e j , $D(N, M)$, que é o custo total de casamento entre os vetores sendo comparados ($C_{total}(\Gamma, \Pi)$), é computado. Conforme implementado por Wang e Pavlidis (1990), o algoritmo que resolve o problema do custo total $C_{total}(\Gamma, \Pi)$ através da fórmula de recorrência citada tem complexidade quadrática.

Com base no custo total de casamento de um vetor extraído a partir de uma forma sendo analisada e dos vetores da base de dados pode-se classificar a forma sob análise, por exemplo, como sendo da mesma classe que a forma da base cujo vetor de características apresenta o melhor casamento (menor custo) em relação ao vetor extraído da imagem de teste.

2.3 Dimensão Fractal Multiescala

Diferentemente da geometria euclidiana, onde se trabalha com um conjunto de formas geométricas perfeitas (círculos, quadrados, retângulos, triângulos, etc.) na representação dos elementos do mundo, a geometria fractal se volta ao estudo de formas complexas conhecidas

como fractais, as quais não podem ser explicadas dentro da geometria euclidiana. Os fractais são objetos matemáticos que existem apenas na teoria e apresentam as seguintes propriedades: auto-semelhança em escala, complexidade infinita e dimensão fractal (BACKES; BRUNO, 2010). A Figura 12 ilustra a estrela de Koch, exemplo de um fractal.

Em relação a auto-semelhança em escala, ela se resume na propriedade de que, ao se ampliar a visualização do contorno dos fractais, nota-se que a parte ampliada é muito similar ao contorno maior (sem a ampliação) (BACKES; BRUNO, 2006, 2010; TORRES; FALCÃO; COSTA, 2004).

Em se tratando da complexidade infinita, são necessárias infinitas iterações na composição do contorno dos objetos (“réplica” do padrão ou regra principal) para se chegar ao objeto fiel ao modelo matemático. Diz-se que os fractais apresentam estruturas finas em qualquer escala, isto é, um detalhamento infinito: caso se amplie sucessivamente a imagem do fractal, mais detalhes poderão ser enxergados em cada ampliação (e este processo não termina) (FALCONER, 1997; MURR et al., 2005). Vale observar que na natureza existem muitas formas similares aos fractais porém com complexidade finita e auto-semelhança em escala apenas em um número finito de níveis (BACKES; BRUNO, 2010). Como observação, para poder desenhar a estrela de Koch da Figura 12, também foi necessário limitar a composição de seu contorno a apenas algumas iterações (não é possível desenhar a estrela perfeita segundo seu modelo matemático pois o processo de composição de sua borda não termina).

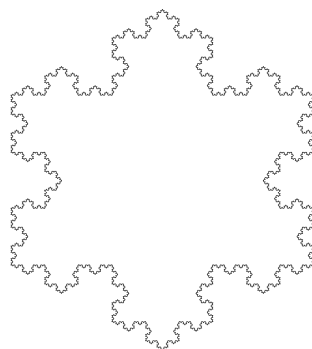


Figura 12. Exemplo de um fractal: estrela de Koch.

Por fim, a dimensão fractal, disseminada por Mandelbrot (1982), corresponde a uma medida da irregularidade da forma do objeto fractal. Ela indica como o fractal ocupa o espaço. Esta medida é uma importante fonte de informação e pode ser usada na caracterização das formas de objetos (fractais ou não) em imagens a fim de identificá-los (BACKES; BRUNO, 2006, 2010; MANDELBROT, 1982). Dentre as definições matemáticas encontradas na

literatura para o cálculo da dimensão fractal de objetos, em geral, o método de Minkowski-Bouligand é o que produz os resultados mais precisos (BACKES; BRUNO, 2006; TRICOT, 1995).

Dada uma forma representada pelo conjunto S de todos seus pontos (tanto de borda como interiores) e com área A (que corresponde ao total de pontos em S), a dilatação exata da mesma por um raio r (denominada S_r), consiste na união dos pontos de S com os pontos da imagem que estão a uma distância menor ou igual a r em relação a algum dos pontos de S . De maneira simplificada, a ideia para se obter a nova forma S_r a partir de uma forma S consiste em se percorrer todos os pontos da forma inicial com um círculo (centrando o mesmo em cada ponto do objeto). Os pontos da imagem que ficarem sob o círculo em algum momento durante seu deslocamento são incluídos na nova forma S_r . A área desta versão dilatada da forma é indicada por A_r .

Vale observar que pode-se empregar a Transformada de Distância Euclidiana (SAITO; TORIWAKI, 1994) para encontrar a nova forma S_r , isto é, os pontos da imagem pertencentes a este conjunto. Nesta transformada, dados os pontos do objeto na imagem (ou apenas do seu contorno), associa-se a cada pixel da imagem o valor de distância euclidiana entre ele e o ponto do objeto mais próximo de si. Para encontrar o conjunto de pontos S_r da versão dilatada da forma por um raio r , basta aplicar um limiar (igual ao valor de r) nos valores associados aos pixels da imagem pela Transformada de Distância Euclidiana selecionando os pixels com valores menores ou iguais ao limiar. A Figura 13 ilustra esta transformada para uma forma de uma folha (de árvore).



Figura 13. Exemplo da Transformada de Distância Euclidiana (SAITO; TORIWAKI, 1994) para a silhueta de uma folha de árvore (em verde). Cada pixel da imagem não pertencente à silhueta tem associado a si um valor de distância que corresponde à distância euclidiana entre ele e o pixel da silhueta mais próximo. Os tons de cinza associados aos pixels da imagem não pertencentes à silhueta são proporcionais às distâncias associadas aos mesmos pela transformada em questão. Quanto mais claro o pixel na imagem, menor a distância associada ao mesmo, ou seja, mais perto dos pixels da silhueta ele está.

Baseado nestas definições, a dimensão fractal de Minkowski-Bouligand (F), operando-se em um espaço bidimensional, é definida por:

$$F = 2 - \lim_{r \rightarrow 0} \frac{\log(A_r)}{\log(r)} \quad (14)$$

Deste modo a dimensão fractal é um número no intervalo $[0; 2]$. Na realidade ela também pode ser calculada de maneira aproximada com base no coeficiente angular da reta que melhor se aproxima da curva logarítmica da área A_r em função do raio de dilatação r (realiza-se uma interpolação linear da curva). Após encontrar esta a reta, seu coeficiente angular A'_r pode ser utilizado para encontrar o valor de F aproximado através da equação:

$$F = 2 - A'_r \quad (15)$$

A Figura 14(b) ilustra esta interpolação para a curva logarítmica da silhueta de um morcego da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999), apresentada na Figura 14(a).

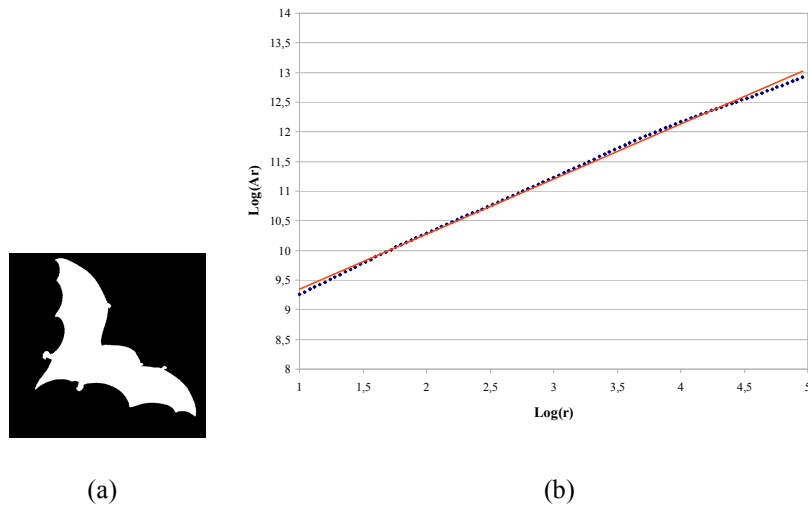


Figura 14. Em (a) exibe-se a silhueta de um morcego da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). Em (b) mostra-se a curva logarítmica (pontos azuis) de $\log(A_r)$ em função de $\log(r)$ da silhueta do morcego, e a reta (em laranja) que melhor representa os pontos desta curva (interpolação linear da curva logarítmica).

No caso da estimativa da dimensão fractal através do método de Minkowski-Bouligand, o valor da dimensão fractal (F) indica como a forma restringe sua dilatação (ao se aumentar o valor do raio r). Formas mais complexas tendem a limitar mais suas dilatações e por isto

apresentam valores de F mais elevados uma vez que, seguindo a Equação 15, apresentam coeficientes A'_r baixos. Formas mais simples como retas e pontos limitam menos suas dilatações e portanto apresentam coeficientes A'_r mais elevados e, conseqüentemente, dimensões fractais menores (TORRES; FALCÃO; COSTA, 2004). É bom ressaltar também que deve-se escolher valores de r relativamente pequenos na análise das curvas logarítmicas, pois, ao se tomar valores sucessivos grandes para r , a área A_r da forma tende a permanecer aproximadamente a mesma (BACKES; BRUNO, 2010).

A caracterização de formas em imagens por meio da dimensão fractal, apesar de ser relativamente simples, faz com que se perca muita informação sobre o objeto em análise visto que o mesmo passa a ser representado por um único valor. Em outras palavras, a interpolação linear da curva *log-log* é muito grosseira e faz com que se perca muitos dados relevantes (BACKES; BRUNO, 2006, 2010; TORRES; FALCÃO; COSTA, 2004)

Para tratar este problema desenvolveu-se a dimensão fractal multiescala. Diferentemente da caracterização de formas pela dimensão fractal tradicional, onde se encontra um único coeficiente angular para representar toda a curva *log-log* do objeto em análise, na caracterização pela dimensão fractal multiescala obtém-se, a partir da curva *log-log*, sua função de derivada a qual é usada para representar o objeto sendo analisado observando a Equação 15 (BACKES; BRUNO, 2006, 2010; TORRES; FALCÃO; COSTA, 2004). A Figura 15 ilustra o conceito de dimensão fractal multiescala para a silhueta do morcego da Figura 14(a), cuja curva *log-log* é exibida na Figura 14(b).

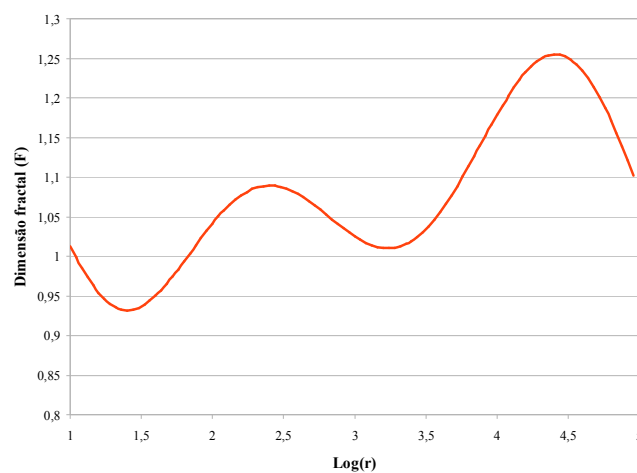


Figura 15. Dimensão fractal multiescala da silhueta do morcego da Figura 14(a), cuja curva *log-log* é apresentada na Figura 14(b). Para se obter a dimensão fractal multiescala do objeto utiliza-se a derivada da curva *log-log* em cada ponto desta e observa-se a Equação 15.

Como exemplo, a abordagem utilizada em (TORRES; FALCÃO; COSTA, 2004) aproxima a curva *log-log* por meio de uma curva polinomial de grau nove obtida através de regressão. A partir desta curva polinomial obtém-se sua função de derivada observando a Equação 15, denominada curva de dimensão fractal multiescala. Então amostra-se esta curva em 100 posições para gerar o vetor de características da forma em análise. Os autores alegam que o grau do polinômio e o número de amostragens foram obtidos através de testes de modo que os resultados de testes feitos com mais de 100 amostras não apresentaram melhora. Dados dois vetores de características de duas formas, a comparação entre eles para verificar a similaridade dos mesmos, gerando assim uma distância, usa a métrica L_2 .

Como observação, Florindo et al. (2012) apresentam diversos outros métodos que se valem da dimensão fractal multiescala na caracterização e reconhecimento de formas e comparam seus desempenhos sobre algumas bases de imagens. Dada a curva de dimensão fractal multiescala de uma forma, estes métodos utilizam diferentes princípios a fim de extrair as melhores características da mesma na geração do vetor de características da forma.

2.4 Saliências do Contorno

Costa, Campos e Manoel (2001) propuseram o uso das saliências das formas como meio de representá-las. Diz-se que as saliências de uma forma (contorno) são definidas como as áreas máximas de influência de seus pontos de borda de maiores curvaturas (TORRES; FALCÃO, 2007).

O algoritmo proposto por Costa, Campos e Manoel (2001) para determinar as saliências de uma forma atua da seguinte maneira: dado um conjunto S de sementes (todos os pixels da borda do objeto) ele associa a cada pixel p da imagem um valor $C(p)$ e um rótulo $R(p)$ os quais são, respectivamente, a distância euclidiana mínima entre p e uma das sementes e o rótulo (identificação) da semente mais próxima de p . Como observação, a Transformada de Distância Euclidiana (SAITO; TORIWAKI, 1994) pode ser usada para encontrar os dois parâmetros citados para cada pixel da imagem.

Para cada pixel q da borda do objeto (semente) verifica-se dois valores: o tamanho da área interna e da área externa ao contorno da forma que contêm os pixels p da imagem que apresentam $R(p)=q$. As áreas de influência de pontos de maiores curvaturas na borda, chamados de pontos de saliência, em geral, são maiores que as áreas de outros pixels da borda da forma (TORRES; FALCÃO, 2007), deste modo, Costa, Campos e Manoel (2001)

propuseram associar um valor de saliência a cada pixel de borda (semente) o qual corresponde à maior área de influência do mesmo entre a interior e exterior. Eles sugeriram localizar então os pontos de saliência no contorno através de uma limiarização nos valores de saliência (tamanho da maior área associada) dos pixels de borda (selecionando os pontos com valores acima de um limiar especificado). A Figura 16 ilustra as áreas de influência associadas a pontos de saliência (vértices) de um polígono.

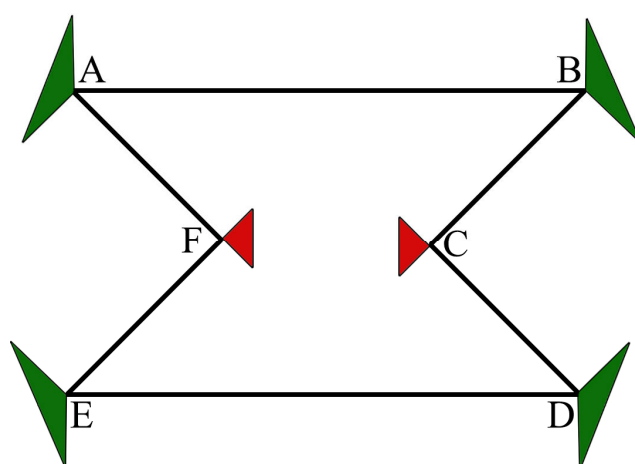


Figura 16. Áreas de influência internas (em vermelho) e externas (em verde) de pontos de saliência. Os pontos A, B, D e E são chamados de pontos convexos, isto é, cujas áreas de influência externas são maiores que as internas (no caso eles sequer apresentam áreas internas). Os pontos C e F são chamados de côncavos (caso contrário).

No descritor Saliência do Contorno (*Contour Saliency*) (TORRES; FALCÃO, 2007), após determinar os pontos de saliência do contorno, pontos de saliência côncavos, isto é, cuja área de influência interna é maior que a externa, têm seus valores de saliência tomados como negativos e pontos de saliência convexos (caso contrário) continuam com seus valores de saliência positivos. Um ponto de saliência aleatório é tomado como referência e o método calcula as posições relativas de cada ponto de saliência do contorno em relação ao ponto de referência (percorrendo-se o contorno em sentido anti-horário).

Então, os valores de saliências dos pontos de saliência do contorno e suas posições relativas no mesmo formam dois vetores de características de mesmo tamanho, os quais são usados pelo descritor *Contour Saliency*. A Figura 17(b) ilustra estes vetores para um contorno poligonal. O contorno do polígono, os pontos de saliência do mesmo e o ponto tido como referência (ponto A) são exibidos na Figura 17(a). O gráfico mostrado na Figura 17(b) indica

os valores de saliência dos pontos de saliência do contorno pelas posições relativas dos pontos no contorno (os dois vetores de características gerados).

Sempre que dois contornos de um mesmo objeto aparecerem em diferentes posições (devido a rotações, translações e mudança de escala) eles estarão representados pelos mesmos pontos de saliência (não há alteração nos pontos e portanto há invariância à estas transformações). Entretanto pode ser que os pontos tomados como referência não sejam os mesmos nos dois casos. Também, os vetores de características de diferentes objetos podem ter diferentes tamanhos. Para tratar isto o descritor *Contour Saliency* usa um algoritmo de casamento heurístico de vetores de características de dois contornos o qual registra seus vetores de características usando os pontos de referência e calcula suas similaridades considerando a diferença de tamanhos entre eles. Este algoritmo é baseado no algoritmo proposto por Mokhtarian e Abbasi (2002) e é descrito em (ABBASI; MOKHTARIAN; KITTLER, 2000; TORRES; FALCÃO; COSTA, 2004).

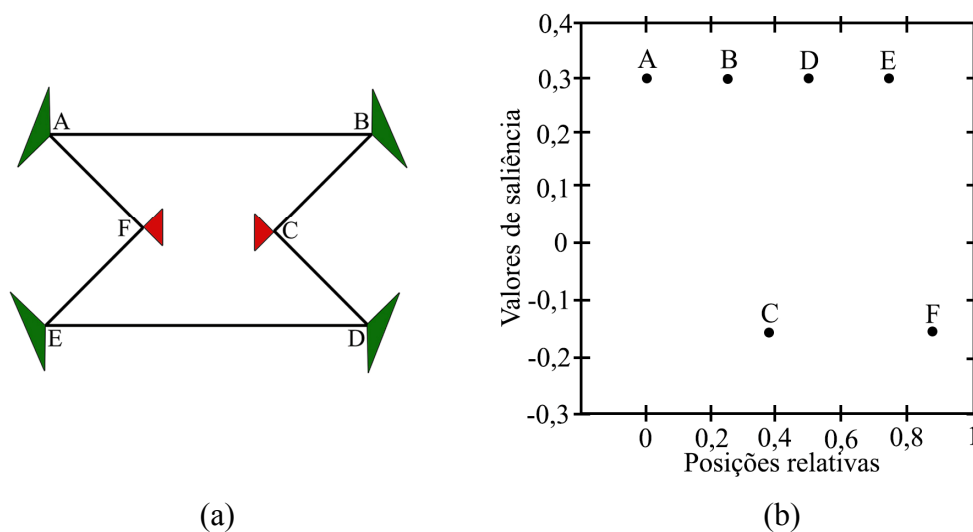


Figura 17. Esquema da localização dos pontos de saliência de um polígono em (a) e da criação do vetor de características do mesmo em (b).

2.5 Descritores de Fourier

Os descritores de Fourier propostos por Zang e Lu (2001) se valem de um método de extração de características responsável por extrair alguma característica da forma (como por exemplo, a distância dos pontos do contorno ao centro de massa do objeto, a curvatura da forma em cada posição do contorno, etc.) gerando funções. A partir destas funções alguns

coeficientes de Fourier são originados e com base neles os descritores de Fourier são estimados.

Dada uma forma (silhueta) obtida a partir de objetos reais em imagens ou a partir de desenhos de usuários, uma etapa de pré-processamento (processamento de imagens) é realizada na imagem da forma (em tons de cinza) com o intuito de extrair a borda e as coordenadas de cada pixel da mesma. O primeiro passo na etapa de pré-processamento é a binarização, isto é, um limiar é utilizado para converter a imagem em tons de cinza em uma imagem binária (forma em branco, isto é, com seus pixels valendo “1” e fundo em preto, ou seja, com seus pixels valendo “0”, ou ao contrário).

Geralmente, após a binarização, ruídos acabam se formando no contorno da forma (ou próximo a ele) especialmente se o objeto sendo analisado está em uma imagem real (fotografia, etc.), então um processo de eliminação de ruído é aplicado à imagem. Este processo elimina os pixels isolados ou pequenos grupos de pixels isolados na imagem. Um algoritmo responsável por detectar a borda da forma e preencher vazios entre pixels da mesma é executado para deixar a borda da forma totalmente conectada, ou seja, sem ser composta por várias partes descontínuas (ZHANG; LU, 2001).

2.5.1 Assinaturas das Formas

Considera-se como a assinatura de uma dada forma, qualquer função unidimensional que represente as características internas da forma ou de sua borda (que por natureza são bidimensionais). Zhang e Lu (2001) apresentam quatro métodos capazes de gerar assinaturas para uma dada forma (extrair de características da mesma): distância central, coordenadas complexas, curvatura e função cumulativa de ângulos. Estes métodos são descritos resumidamente a seguir. Como observação, considera-se que já se tem a borda do objeto segmentada na imagem (tarefa realizada no pré-processamento) além de uma sequência estabelecida entre os pixels de borda (iniciando em determinado pixel e percorrendo a borda em sentido horário, por exemplo). Indica-se as coordenadas de cada pixel de borda por $(x(t), y(t))$, com $t=0, 1, \dots, L-1$, onde L corresponde ao número de pontos de borda.

Coordenadas Complexas

Uma função de coordenadas complexas é obtida para a forma a partir das coordenadas de cada pixel do contorno da mesma:

$$z(t)=x(t)+y(t).i \quad (16)$$

A fim de tornar a função invariante à translação do objeto na imagem desloca-se o centro de massa (centróide) do objeto para a origem (juntamente com todo objeto) e então a função de coordenadas complexas é obtida através de:

$$z(t)=[x(t)-x_c]+[y(t)-y_c].i \quad (17)$$

onde (x_c, y_c) são as coordenadas do centro de massa da forma, isto é, x_c equivale à média de todas as coordenadas $x(t)$ dos pontos da borda do objeto (com t variando de 0 até $L-1$) e y_c é a média de todas as coordenadas $y(t)$ dos pontos da borda do objeto (com t variando de 0 até $L-1$).

Distância do Centro de Massa

A função de distância do centro de massa é expressa pelas distâncias euclidianas dos pontos do contorno da forma ao centro de massa da mesma com coordenadas (x_c, y_c) :

$$r(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} \quad (18)$$

Esta função é invariante à translação e rotação.

Curvatura

Basicamente o que se analisa neste método de extração de características da forma (e geração de uma função a partir dela) é a curvatura da borda em cada ponto da mesma. Basicamente trabalha-se com a segunda derivada do contorno da forma (a primeira derivada das tangentes da borda). Em (KAUPPINEN; SEPPANEN; PIETIKAINEN, 1995) é descrito um exemplo de função de curvatura. Esta função é invariante à translação e rotação.

Ângulos Cumulativos

Baseado no trabalho de Zahn e Roskies (1972), uma função de curvatura cumulativa corresponde ao montante (soma) “líquido” dos ângulos de curvatura do contorno da forma iniciando em um ponto $z(0)$ (ponto inicial da borda) e terminando em um ponto $z(t)$ do contorno. Também há invariância a rotação e translação.

2.5.2 Representação de Formas por Descritores de Fourier

As quatro assinaturas de formas (funções unidimensionais) apresentadas são derivadas das coordenadas dos pontos do contorno da forma. Uma Transformada de Fourier pode ser aplicada nas assinaturas encontradas, dada uma forma, a fim de reconhecê-la. Os coeficientes da Transformada de Fourier originam os descritores de Fourier para a forma sendo analisada. Estes descritores representam a forma do objeto em um domínio de frequência. Os descritores relacionados às menores frequências apresentam informações sobre as características gerais da forma e os descritores de alta frequência apresentam informações sobre detalhes e peculiaridades da forma. Apesar de o número de coeficientes gerados através da transformada ser em geral grande, um subconjunto destes coeficientes é suficiente para se registrar características da forma capazes de identificá-la (ZHANG; LU, 2001).

As informações obtidas a partir de frequências muito altas descrevem detalhes muito pequenos do contorno da forma, os quais, muitas vezes, não auxiliam no processo de identificação da mesma e assim podem ser descartados. Com isto, as dimensões dos descritores de Fourier usados para representar formas são bem reduzidas.

A fim de facilitar o casamento de assinaturas de duas formas (evitando que elas tenham tamanhos diferentes), suas bordas ou funções (assinaturas) são amostradas em um número de pontos. Este número é em geral inteiro e uma potência de 2 a fim de facilitar a aplicação da Transformada Rápida de Fourier (*Fast Fourier Transform* - FFT). A amostragem de pontos do contorno da forma (associados aos seus respectivos valores na assinatura da mesma) normaliza o tamanho das formas bem como de certa forma “suaviza” seus contornos. Esta suavização pode eliminar ruídos ou detalhes muito pequenos (sem utilidade) do contorno da forma.

Ao variar o número de pontos amostrados do contorno a precisão da representação do mesmo é escolhida. Quanto maior o número de pontos amostrados, mais detalhes da forma são representados e consequentemente o resultado dos casamentos dela com outras formas se torna mais preciso. Deve-se escolher um número de pontos capaz de tornar a classificação da forma precisa, mas sem deixar que este número fique tão grande a ponto de que pequenos detalhes sem importância na forma atrapalhem o reconhecimento da mesma (ou façam este processo ficar lento demais).

Entre os métodos de amostragem apresentados em por Zhang e Lu (2001) o que os autores consideram o melhor e usam na criação de descritores de Fourier a partir de uma forma é o método de amostragem baseado em comprimento de arco. Neste método, pontos espaçados através de comprimentos de arcos de mesmo tamanho no contorno da forma são escolhidos. O

espaço entre dois pontos amostrados consecutivos é dado por P/K onde P é o perímetro do contorno e K equivale ao número de pontos amostrados. Nos experimentos de Zhang e Lu (2001), são amostrados 64 pontos das bordas das formas para então compará-las. A Figura 18 mostra uma normalização de uma forma (amostragem de pontos e reconstrução da mesma apenas com base nestes pontos). Como pode-se observar a normalização elimina ruídos e pequenos detalhes (sem muita importância) da forma ao passo que extrai de forma eficaz os aspectos gerais da mesma, inclusive preservando pontos de saliência (pontos de alta curvatura) importantes à representação da forma.

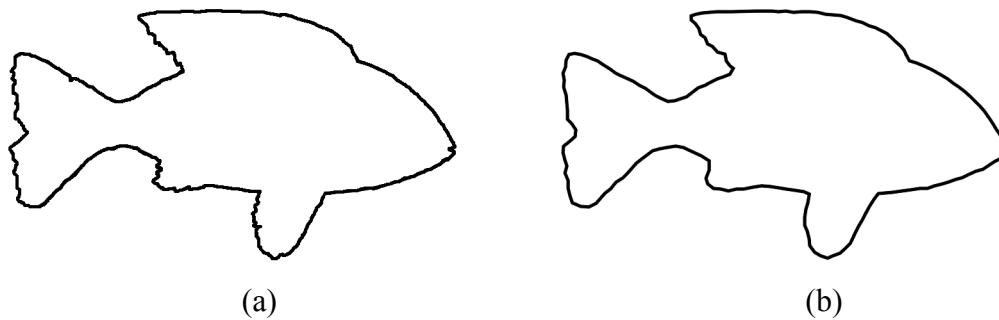


Figura 18. Normalização de uma forma: pontos são amostrados da mesma e o contorno é reconstruído com base nestes pontos. Em (a) a forma original e em (b) a forma normalizada.

Após a amostragem dos pontos do contorno da forma, e consequentemente da assinatura desta, aplica-se uma Transformada de Fourier sobre a mesma. Dada uma assinatura $s(t)$ de uma forma (obtida a partir de um dos quatro métodos citados anteriormente), onde $t=0, 1, \dots, L-1$, após amostrá-la em N pontos conforme citado, sua transformada discreta de Fourier é dada por:

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) \exp\left(\frac{-j2\pi nt}{N}\right) \quad (19)$$

com $n=0, 1, \dots, N-1$ conforme (ZHANG; LU, 2001). Os coeficientes u_n , $n=0, 1, \dots, N-1$, são chamados pelos autores de descritores de Fourier (FD – *Fourier Descriptors*) da forma (ZHANG; LU, 2001) e são denotados por FD_n , $n=0, 1, \dots, N-1$.

Usando a assinatura obtida pelo método das coordenadas complexas o vetor de características (invariante à translação, rotação e escala) gerado a partir dos coeficientes mencionados para representar a forma é dado por:

$$f = \left[\frac{|FD_2|}{|FD_1|}, \frac{|FD_3|}{|FD_1|}, \dots, \frac{|FD_{N-1}|}{|FD_1|} \right] \quad (20)$$

Para a assinatura obtida pelo método da distância ao centro de massa e pelo método da curvatura o vetor de características é dado por:

$$f = \left[\frac{|FD_1|}{|FD_0|}, \frac{|FD_1|}{|FD_0|}, \dots, \frac{|FD_{N/2}|}{|FD_0|} \right] \quad (21)$$

E para a assinatura obtida pelo método de ângulos cumulativos o vetor é definido como:

$$f = [|FD_0|, |FD_1|, \dots, |FD_{N/2}|] \quad (22)$$

A partir do vetor de características gerado para uma forma sendo analisada $f_d = [f_d^1, f_d^2, \dots, f_d^{N_c}]$ e do vetor de características de uma forma da base de dados $f_m = [f_m^1, f_m^2, \dots, f_m^{N_c}]$, os mesmos podem ser comparadas através da distância euclidiana entre seus elementos:

$$d = \sqrt{\sum_{i=1}^{N_c} |f_m^i - f_d^i|^2} \quad (23)$$

onde N_c equivale ao número de elementos no vetor de característica usado (escolhido com base no método de assinatura empregado). Em alguns testes realizados por Zhang e Lu (2001) ao se usar a assinatura baseada na distância dos pontos do contorno da forma em relação ao centro de massa os resultados foram os melhores que quando usando os três outros métodos para gerar a assinaturas das formas.

2.6 O Descritor *Tensor Scale*

Saha (2005) apresenta um método local para análise de imagens, chamado *Tensor Scale*, o qual se resume a encontrar, para cada ponto p da forma (tanto no seu interior como na sua borda), a maior elipse centrada em p e dentro da região homogênea deste ponto na forma sendo analisada.

Este método define a elipse centrada em p através de três fatores:

- Orientação: ângulo entre $t_l(p)$ e o eixo x ;

- Anisotropia: $\sqrt{1 - \frac{|t_2(p)|^2}{|t_1(p)|^2}}$;
- Espessura: $|t_2(p)|$;

onde $|t_1(p)|$ e $|t_2(p)|$, com $|t_1(p)| \geq |t_2(p)|$, correspondem aos comprimentos dos dois semi-eixos $t_1(p)$ e $t_2(p)$ da elipse centrada em p (ANDALÓ et al., 2007, 2010).

O primeiro passo para se encontrar a elipse do ponto p corresponde a traçar pares de segmentos de reta opostos entre si em relação ao ponto e iniciando nele, variando suas inclinações em relação ao eixo x no intervalo $[0^\circ; 180^\circ[$ conforme mostra a Figura 19(a). Um vetor unitário é usado para representar a direção de cada segmento, existindo assim m pares destes vetores (onde m não equivale a 180 necessariamente, isto é, pode-se “espaçar” mais ou menos os segmentos de reta em relação aos seus ângulos) os quais são representados por $(\vec{\tau}_1, \vec{\tau}'_1), (\vec{\tau}_2, \vec{\tau}'_2), \dots, (\vec{\tau}_m, \vec{\tau}'_m)$.

O segmento com direção $\vec{\tau}_i$ e originado em p é denotado por $\pi_i^{(p)}: [0; L] \rightarrow \mathbb{R}^2$ e o segmento oposto com direção $\vec{\tau}'_i$ por $\pi_i'^{(p)}$. A Equação 24 representa o segmento $\pi_i^{(p)}$ com base no ponto p , no seu respectivo vetor $\vec{\tau}_i$ e em um valor λ que varia de 0 a L onde L representa o tamanho máximo (o último ponto) do segmento, o qual é definido pelo programador ou usuário do método. Em experimentos feitos constatou-se que L deve ser um número entre 8 e 15 pontos a fim de evitar que o custo computacional fique muito alto (ANDALÓ et al., 2007; MIRANDA; TORRES; FALCÃO, 2005).

Equação do segmento $\pi_i^{(p)}$:

$$\pi_i^{(p)}(\lambda) = p + \lambda \vec{\tau}_i \quad (24)$$

onde $\lambda \in [0; L]$.

Após traçar os segmentos a partir do ponto p , verifica-se para cada um deles, o ponto da borda da forma mais próximo de p por onde o segmento passa como mostrado na Figura 19(b). Para encontrar estes pontos a intensidade (nível de cinza) presente ao longo de cada segmento de reta é verificada. Como uma observação, caso a distância entre o ponto de borda a ser encontrado e o ponto p ultrapasse o valor de L escolhido (valor máximo de distância), o ponto escolhido para o segmento é o que está à distância L de p no mesmo (último ponto no segmento). Isto ocorre com seis segmentos na parte superior direita da Figura 19(b), os pontos associados a estes segmentos estão em verde para facilitar a identificação dos mesmos.

Depois de encontrar o ponto da borda associado a cada segmento, em cada par de segmentos de reta opostos, o ponto de borda com maior distância em relação ao ponto p é “deslocado” (sobre o segmento) de forma a ficar com a mesma distância em relação a p que o ponto de borda oposto (Figura 19(c)). Isto é realizado para se garantir a simetria da elipse a ser gerada para o ponto p . Com os pontos todos rearranjados, uma elipse (que melhor se aproxima da posição dos mesmos) é obtida através da Análise do Componente Principal (*Principal Component Analysis* – PCA) como mostrado na Figura 19(d).

Na verdade a técnica PCA é usada para encontrar a orientação da elipse (ângulo entre seu eixo principal e o eixo x) e após isto uma função de erro é usada para encontrar os tamanhos dos eixos da elipse (esta função de erro atua basicamente somando os erros entre as posições dos pontos selecionados em cada segmento de reta e os pontos da elipse) (ANDALÓ et al., 2007; MIRANDA; TORRES; FALCÃO, 2005).

Este processo de encontrar a elipse é realizado para todo pixel (ponto) da forma (ou seja, todos os pixels são tomados, um por vez, como sendo o ponto p , isto é, o centro de massa elipse sendo buscada).

Um problema grave desta abordagem é o custo computacional da mesma que faz com que ela se torne inapropriada para a análise de imagens (em especial para sistemas de recuperação de imagens por conteúdo). Miranda, Torres e Falcão (2005), visando solucionar este problema, propuseram uma implementação eficiente do método original que difere do mesmo nos aspectos mencionados a seguir.

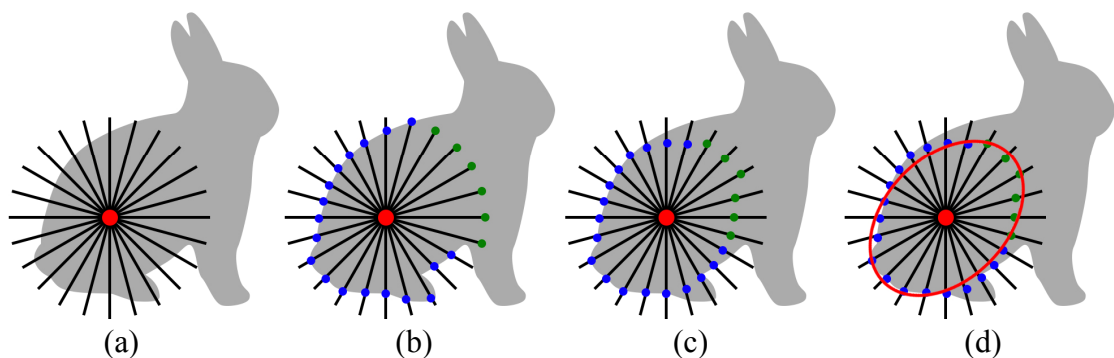


Figura 19. Cálculo da elipse associada ao ponto vermelho da forma. Em (a) são traçados os m pares de segmentos de reta opostos entre si. Em (b) os pontos de borda do objeto são localizados sobre os segmentos (pontos em azul). Quando não há ponto de borda sobre o segmento em análise associa-se ao mesmo seu último ponto (pontos em verde). Em (c) os pontos são rearranjados e em (d) a elipse que mais se aproxima da disposição dos pontos é traçada.

A primeira mudança se encontra na fase de localização dos pixels de borda sobre os segmentos de reta gerados a partir do ponto sendo analisado. Ao invés de se percorrer um segmento de reta procurando um ponto de borda e depois o segmento oposto, realiza-se a busca dos pontos de borda simultaneamente nos segmentos opostos, ou seja, vai-se percorrendo, a partir do ponto sob análise, simultaneamente os dois segmentos de reta opostos. Ao se encontrar o primeiro ponto de borda em um dos segmentos de reta, se registra a posição de busca no momento nos dois segmentos de reta já deixando os dois pontos de borda associados aos mesmos com a menor distância entre eles e o ponto sob análise (que corresponde à centróide da elipse sendo buscada). A Figura 20 ilustra este processo. Com isto, a fase de deslocamento dos pontos de borda sobre os segmentos de reta a fim de ficarem com mesma distância em relação ao ponto sendo analisado que os pontos de borda dos segmentos de reta opostos não é necessária.

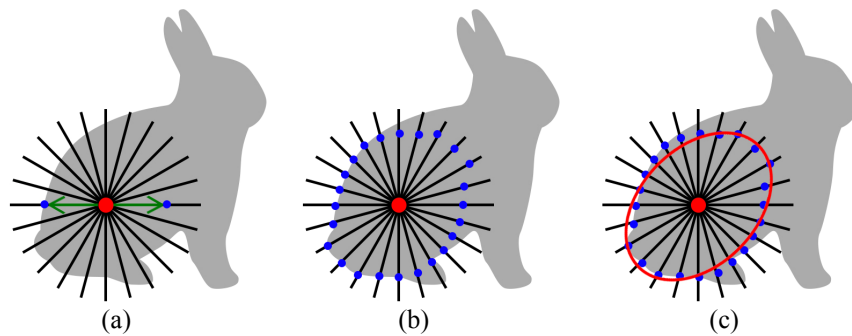


Figura 20. Os segmentos de reta opostos (em relação ao ponto vermelho) são percorridos simultaneamente no algoritmo proposto de Miranda, Torres e Falcão (2005) e ao se encontrar o primeiro pixel de borda em um deles já se posiciona os dois pontos azuis (um em cada segmento), como mostrado em (a), para então obter a configuração final dos pontos sem ter de ficar deslocando os mesmos (b). Em (c) mostra-se a elipse encontrada que melhor se aproxima dos pontos azuis.

A segunda mudança diz respeito em uma localização de pontos de borda através de limiares de forma mais simples e a terceira mudança, por sua vez, corresponde a uma melhoria na fase de encontrar a elipse que melhor se aproxima dos pontos obtidos. Miranda, Torres e Falcão (2005) propõem uma função g que permite encontrar o ângulo de inclinação da elipse diretamente ao invés de se usar o PCA. A orientação da elipse é obtida a partir do valor γ que minimiza a função g :

$$g(\gamma) = \sum_{i=1,2,\dots,2m} [x_{i_\gamma}^2 - y_{i_\gamma}^2] \quad (25)$$

onde $x_{i_\gamma} = x_i \cdot \cos(\gamma) - y_i \cdot \sin(\gamma)$, $y_{i_\gamma} = x_i \cdot \sin(\gamma) + y_i \cdot \cos(\gamma)$, (x_i, y_i) são as coordenadas relativas dos pontos de borda encontrados em cada um dos $2 \cdot m$ segmentos de reta em relação ao pixel central da elipse (sob análise) p de coordenadas (x_p, y_p) , e $(x_{i_\gamma}, y_{i_\gamma})$ são as novas coordenadas dos pontos após aplicar uma rotação pelo ângulo γ .

Considerando estas otimizações, Miranda, Torres e Falcão (2005) propuseram o descritor *Tensor Scale* para imagens em tons de cinza. Diferentes objetos em geral apresentam diferentes orientações locais (elipses com diferentes orientações centradas em cada ponto da forma).

Com isto o *Tensor Scale* calcula o histograma das orientações locais dos pontos da forma sendo investigada (orientações das elipses centradas nos pontos da forma e encontradas como descrito anteriormente) e usa o mesmo como vetor de característica da forma. No cálculo dos histogramas só se considera pixels com elipses que apresentam alta anisotropia e espessura em um certo intervalo de valores relativamente altos.

O casamento de duas imagens pelo *Tensor Scale* é realizado através da diferença absoluta das áreas de seus histogramas após corrigir o deslocamento (rotações no objeto causam “deslocamentos” nos valores do histograma).

3 A Transformada de Hough

Um problema recorrente na análise de imagens automatizada é a detecção de retas. Ele se resume em detectar a presença de conjuntos de pontos colineares ou aproximadamente colineares em uma dada imagem.

Este problema pode ser resolvido testando-se todas as retas formadas a partir de todos os possíveis pares de pontos presentes na imagem. Entretanto, testar todos os pares de pontos a fim de encontrar retas é uma estratégia pouco inteligente e demorada, principalmente se o número de pontos for muito grande (HOUGH, 1962).

Utilizando-se a Transformada de Hough, proposta originalmente por Paul Hough em 1962 (HOUGH, 1962), é possível detectar retas nas imagens de uma maneira mais eficiente tanto em termos de recursos quanto em tempo de processamento. A Transformada de Hough define um mapeamento entre o espaço da imagem e o espaço de parâmetros. Hough propôs, originalmente, o espaço de parâmetros *slope-intercept* para a detecção de retas, no qual cada reta na imagem regida pela equação:

$$y=a \cdot x+b \quad (26)$$

é representada no espaço de parâmetros por um ponto de coordenadas (a,b) , em que o parâmetro a corresponde à inclinação da reta e o parâmetro b corresponde ao valor de y quando $x=0$. A Figura 21 ilustra uma reta no espaço da imagem com sua respectiva equação.

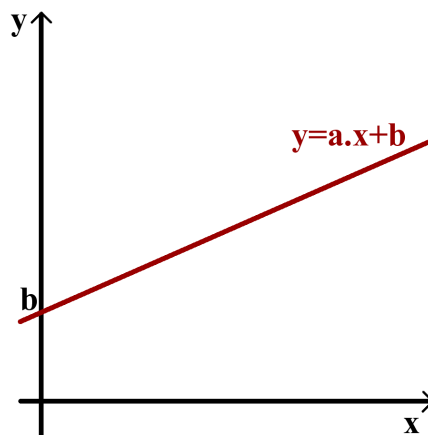


Figura 21. Equação da reta na imagem. O parâmetro b (*intercept*) pode ser encontrado ao se fixar $x=0$ (corresponde ao valor de y quando $x=0$) e o a (*slope*) corresponde ao coeficiente angular da reta.

Ainda usando a Equação 26, cada ponto no espaço da imagem, com suas coordenadas x (coluna) e y (linha) fixas, é representado como uma reta no espaço de parâmetros *slope-intercept* (ao se fixar o valor de x e y na Equação 26 obtém-se a equação de uma reta cujas variáveis são a e b).

Uma importante propriedade da Transformada de Hough (HOUGH, 1962) é que pontos colineares na imagem correspondem a retas concorrentes no espaço de parâmetros que se interceptam na posição do espaço que representa a reta da imagem à qual os pontos pertencem. Com isto, a fim de detectar retas em uma imagem, após representar todos os pontos presentes na imagem no espaço de parâmetros, basta encontrar nesse espaço as posições onde há um grande número de retas que se interceptam. Estas posições indicarão retas na imagem que satisfazem a Equação 26 e cujos valores de a e b correspondem às coordenadas dos pontos de intersecção no espaço de parâmetros (HOUGH, 1962).

Nos computadores, deve-se usar um espaço de parâmetros discreto o qual é representado por uma matriz acumuladora. Para detectar retas na imagem é necessário encontrar posições na matriz com altos valores acumulados (posições onde várias retas se interceptam).

Como observação, Hough (1962) propôs inicialmente a transformada em questão para resolver o problema de se encontrar retas (pontos colineares) em “fotografias” rudimentares a fim de determinar o trajeto de partículas subatômicas carregadas em um campo de observação onde as imagens das mesmas são capturadas.

A observação de partículas subatômicas carregadas não é uma tarefa fácil por estas serem extremamente pequenas. Uma ferramenta desenvolvida capaz de permitir a realização de tal atividade é a câmara de bolhas. Este equipamento contém um material líquido devidamente superaquecido (em geral, hidrogênio líquido) no qual as partículas carregadas adentram e criam rastros de bolhas ao longo de suas trajetórias (fazem o líquido passar para o estado gasoso ao redor de si). Pode-se empregar um pistão para variar a pressão na câmara a fim de viabilizar a formação e visualização das bolhas. Estes rastros de bolhas formam padrões complexos e são fotografados com o uso de um plano de fundo preto.

Dada uma partícula em movimento na câmara de bolhas, diversas “fotografias” são capturadas da mesma (de seu rastro) e a análise das imagens por um especialista humano é realizada a fim de descobrir padrões no trajeto da partícula sendo observada. Este processo de análise pode levar muitas horas e se tornar até inviável caso o número de imagens aumente muito.

Com o intuito de agilizar tal atividade, Hough formulou a transformada em questão a fim de detectar de maneira mais eficiente pequenos segmentos de reta formados pelas bolhas

geradas pelas partículas ao longo de suas trajetórias. Esta detecção é realizada como explicado anteriormente: encontrando no espaço *slope-intercept* pontos de intersecção de muitas retas (as quais representam os pontos, isto é, as bolhas das imagens capturadas).

3.1 A Transformada de Hough com o Espaço θ - ρ

A Transformada de Hough (HOUGH, 1962) com o espaço de parâmetros *slope-intercept* apresenta um problema: estes parâmetros são ilimitados (DUDA; HART, 1972), o que inviabiliza seu uso em computadores. Por exemplo, tem-se um problema quando se tenta representar retas paralelas ao eixo y da imagem no espaço *slope-intercept*.

Para tratar esse problema, Duda e Hart (1972) propuseram a utilização da equação da reta definida por coordenadas polares. Como ilustrado na Figura 22, esta nova parametrização representa cada reta da imagem por meio do ângulo θ de seu vetor normal ao eixo das abscissas e de sua distância ρ em relação à origem.

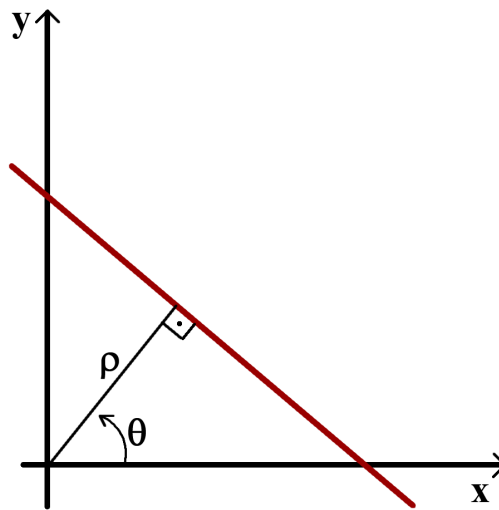


Figura 22. Nova representação de retas: por meio dos parâmetros θ e ρ .

A equação da reta de acordo com esta parametrização, é:

$$\rho = x \cdot \cos(\theta) + y \cdot \sin(\theta) \quad (27)$$

Se restringirmos θ ao intervalo $[0; \pi[$, então os parâmetros ρ e θ de uma dada reta são únicos. Com esta restrição, toda reta no espaço x - y corresponde a um único ponto no espaço

de parâmetros θ - ρ . Cada ponto da imagem com suas coordenadas (x_i, y_i) é representado por uma curva senoidal obtida a partir das coordenadas do ponto e da Equação 27, no espaço θ - ρ .

Pode-se verificar que as senóides que representam pontos colineares da imagem no espaço de parâmetros θ - ρ têm um ponto de intersecção comum (ρ_0, θ_0) o qual representa a reta da imagem à qual os pontos pertencem (DUDA; HART, 1972).

Desta forma, para detectar retas na imagem, como no espaço de Hough original, também basta encontrar pontos de intersecção das senóides no espaço de parâmetros.

4 HTS: *Hough Transform Statistics*

Baseado na Transformada de Hough proposta por Duda e Hart (1972), um novo método para análise de formas planas de objetos ou regiões presentes em imagens digitais é apresentado neste trabalho, o qual denominamos HTS (*Hough Transform Statistics*). Este método se volta à propriedade da Transformada de Hough (DUDA; HART, 1972) relacionada à detecção de retas em imagens, que faz com que pontos colineares da imagem sejam representados por senóides no espaço de Hough que se interceptam exatamente no ponto do espaço que representa o segmento de reta da imagem ao qual os pontos pertencem. Desta forma, pontos pertencentes a grandes segmentos de reta na imagem tendem a apresentar alto número de intersecções (picos) em algumas posições ao longo de suas senóides no espaço de Hough (nas posições que representam, no espaço de Hough, os segmentos de reta da imagem aos quais os pontos pertencem). Pontos que não pertencem a segmentos de reta na imagem tendem a apresentar quantidade de intersecções bem distribuída ao longo de suas senóides e não possuem picos. Com base nisto, é possível diferenciar pontos pertencentes e não pertencentes a segmentos de reta nas imagens.

Esta diferenciação pode ser utilizada na análise de formas: dados os pontos de borda de um objeto sendo analisado e o espaço de Hough gerado com base nos mesmos, pontos pertencentes a segmentos de reta da borda do objeto apresentarão picos ao longo de suas senóides no espaço de Hough, enquanto que pontos de regiões arredondadas da borda apresentarão quantidade de intersecções bem distribuída ao longo de suas senóides. Com isto pode-se caracterizar os pontos de borda e assim diferenciar regiões da borda da forma.

O HTS se vale desta diferenciação de pontos de borda: dada a caracterização dos pontos de borda de um objeto desconhecido e a caracterização dos pontos de borda dos objetos da base de dados pode-se comparar as bordas dos mesmos (características de seus pontos) a fim de identificar o objeto de teste.

Basicamente o método proposto apresenta, assim como outros descritores de formas, três etapas: Pré-Processamento, Extração de Características e Reconhecimento. Estas etapas são descritas nas próximas seções.

4.1 Etapa de Pré-Processamento

Esta etapa corresponde, no contexto da análise de imagens, ao processamento de imagens. A partir de uma dada imagem contendo um objeto, realiza-se a segmentação da borda do mesmo a fim de identificá-lo pela forma (através das senóides de seus pontos de borda).

A segmentação de pixels de borda em imagens pode ser realizada através de muitos métodos. Uma forma simples de se realizar a segmentação da borda de objetos quando se trabalha com imagens binárias (caso dos experimentos deste trabalho) corresponde a verificar para cada pixel da imagem sendo analisada a cor do objeto (branco ou preto, pois a imagem é binária) se ele possui vizinhos de cor diferente ou se ele se encontra nos limites da imagem. Caso o pixel atenda a uma destas duas restrições ele é considerado como pixel de borda, caso contrário não.

Ao fim desta etapa os pixels de borda ficam, por exemplo, destacados em branco (com valor “1”) na imagem e os demais pixels se tornam pretos (valor “0”).

4.2 Etapa de Extração de Características

Como descrito na seção 3.1, cada ponto da borda do objeto gera uma senóide no espaço de parâmetros θ - ρ , o qual é representado por uma matriz acumuladora, contendo 180 colunas e um número fixo de linhas (que depende do tamanho da imagem), inicialmente zerada em todas suas posições. Cada senóide é registrada no espaço de parâmetros incrementando-se uma unidade nas posições da matriz acumuladora por onde ela passa: a partir da Equação 27 e das coordenadas x e y do ponto de borda na imagem, varia-se o ângulo θ de 0° a 179° (com $\Delta\theta=1^\circ$), encontra-se um valor arredondado de ρ (com $\Delta\rho=1$) para cada valor de θ , e incrementa-se em uma unidade a posição (ρ, θ) da matriz acumuladora. A Figura 23 ilustra o espaço de Hough obtido para a borda de uma maçã após representar todos seus pontos de borda por meio de senóides no mesmo. Como mostrado na Figura 1, objetos semelhantes geram espaços de Hough similares enquanto que objetos diferentes produzem espaços de Hough muito diferentes.

Após registrar todas as senóides, as quais representam todos os pontos da borda do objeto no espaço de Hough, a cada ponto p da borda é associado um histograma contendo a quantidade de intersecções encontradas ao longo de sua senóide no espaço de parâmetros

(percorre-se novamente a senóide associada ao ponto p verificando-se os valores presentes nas respectivas posições da matriz acumuladora).

Como dito, pontos pertencentes a segmentos de reta na borda do objeto apresentarão picos em algumas posições de seus histogramas, enquanto que pontos pertencentes a regiões arredondadas da borda apresentarão histogramas com valores homogeneamente distribuídos por todas suas posições. A Figura 24 ilustra os histogramas associados a quatro pontos de borda de dois objetos de mesma classe (sinos) da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999).

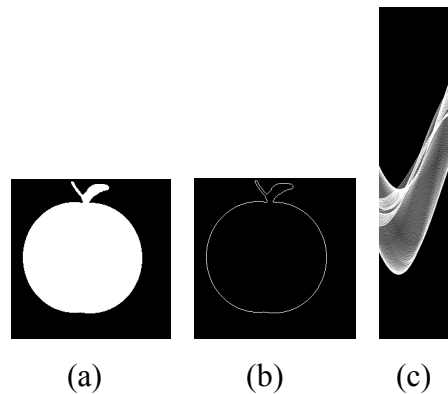


Figura 23. (a) Silhueta de uma maçã da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999); (b) Borda da maçã segmentada; (c) Espaço de Hough (matriz acumuladora) gerado a partir dos pontos da borda, quanto mais clara é a posição desta última imagem, maior é o valor presente na matriz acumuladora na respectiva posição (mais senóides passam pela posição).

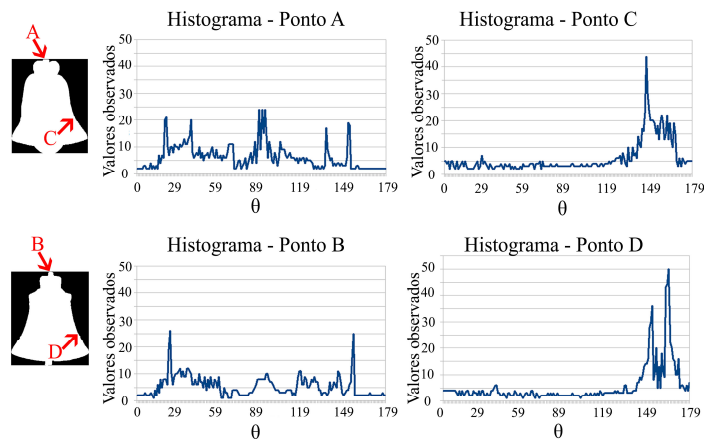


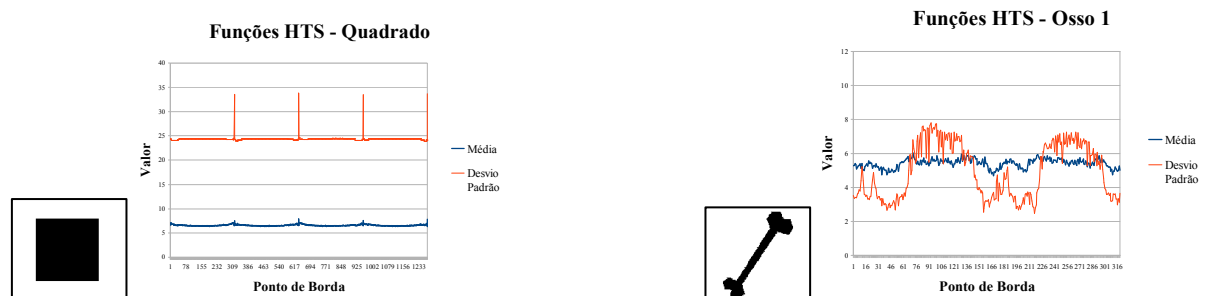
Figura 24. Histogramas associados a quatro pontos da borda de dois objetos de mesma classe (sinos) da base MPEG-7 Part B (JEANNIN; BOBER, 1999). Pode-se perceber que histogramas de pontos localizados na borda de um mesmo objeto, mas em regiões diferentes (A-C e B-D), são diferentes, e que pontos que pertencem a regiões respectivas em imagens de dois objetos de mesma classe (A-B e C-D) apresentam histogramas similares. Também é possível observar que pontos de borda que pertencem a grandes segmentos relativamente retos das bordas (C e D) apresentam picos bastante sobressalentes em seus histogramas, enquanto que outros pontos das bordas (como A e B) apresentam histogramas com valores mais homogeneamente distribuídos.

Associa-se então a cada ponto de borda, dado seu respectivo histograma, a média e o desvio padrão desta sua distribuição. Com isto, cada ponto de borda passa a ter apenas dois valores associados a si (os histogramas são descartados). Pontos de borda que pertencem a segmentos de reta da borda na imagem, por apresentarem picos em seus histogramas, tendem a apresentar desvios padrão maiores que os demais pontos de borda.

Dado um ponto inicial da borda (no nosso caso o ponto mais ao topo da borda na imagem) e percorrendo esta em sentido horário, pode-se construir duas funções unidimensionais para o objeto sendo analisado a partir dos dois valores (média e desvio padrão) associados aos pontos de borda (cada valor gera uma função). A Figura 25 ilustra as funções obtidas para alguns objetos. Pode-se perceber que objetos de mesma classe apresentam funções similares enquanto que objetos de diferentes classes apresentam funções bem diferentes.

Ainda na Figura 25 pode-se observar que a função de desvio padrão da silhueta do quadrado apresenta quatro picos. Estes picos correspondem aos quatro vértices do quadrado: por pertencerem a dois grandes segmentos de reta da borda do quadrado eles possuem histogramas com dois grandes picos, gerando alto desvio padrão para os mesmos. Os demais pontos de borda pertencem a apenas um grande segmento de reta da borda e, portanto, apresentam desvios elevados, porém menores que os dos vértices do quadrado. Além disto, como os quatro segmentos de reta (lados do quadrado) são do mesmo tamanho, os desvios padrão dos quatro vértices são praticamente iguais e os desvios padrão dos outros pontos de borda também são muito próximos entre si.

Após encontrar as funções unidimensionais para um objeto, elas são amostradas em posições igualmente espaçadas entre si para se gerar o vetor de características do mesmo. Cada posição do vetor de características contém dois valores: a média e o desvio padrão amostrados das funções unidimensionais em uma dada posição. A Figura 26 ilustra o processo de amostragem das funções unidimensionais de uma silhueta da base de imagens Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).



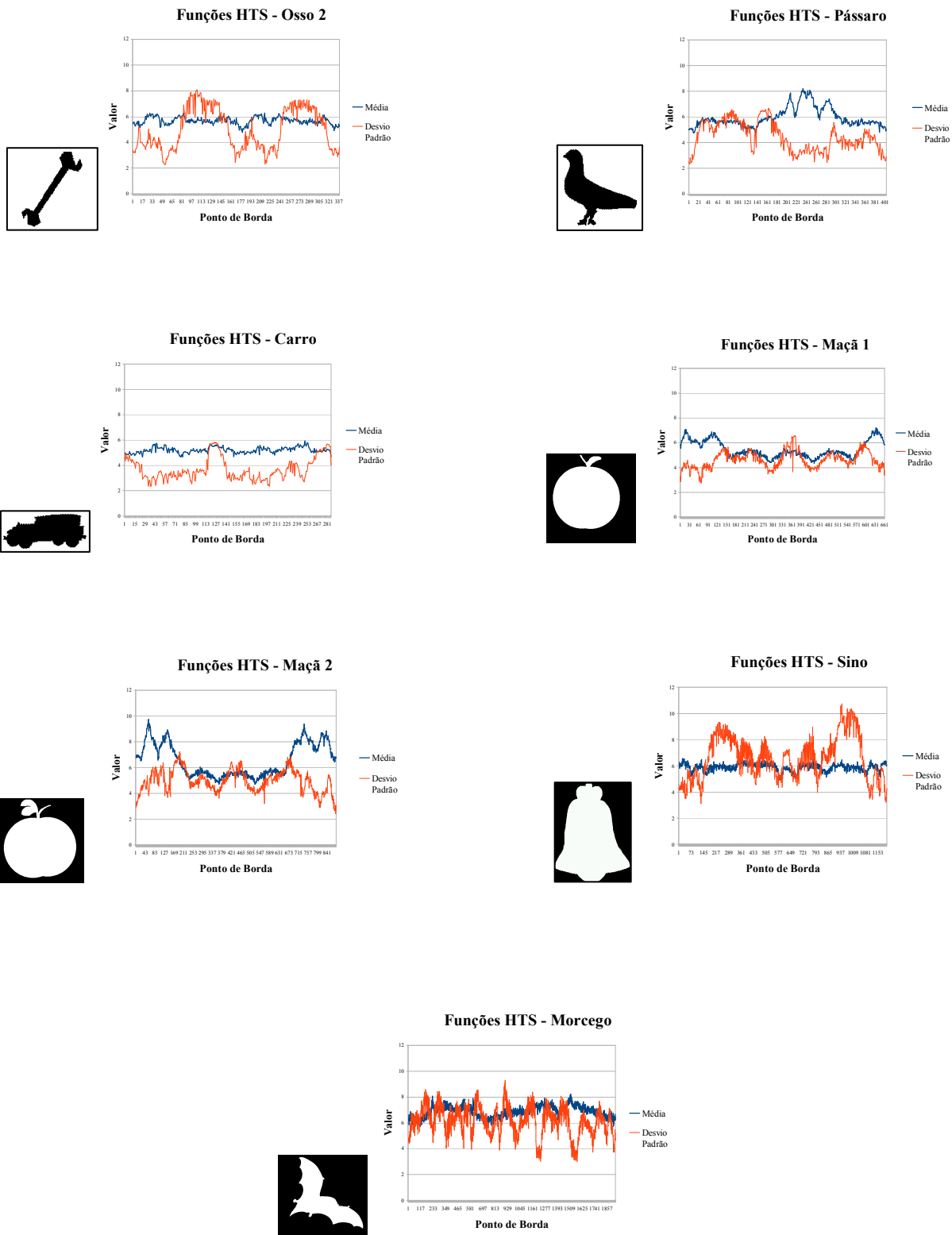


Figura 25. Funções obtidas para silhuetas da base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) (as 4 primeiras sem considerar o quadrado) e da base MPEG-7 Part B (JEANNIN; BOBER, 1999) (as 4 últimas). Pode-se notar que as funções de objetos diferentes são bastante diferentes enquanto que funções de objetos de mesma classe são similares.

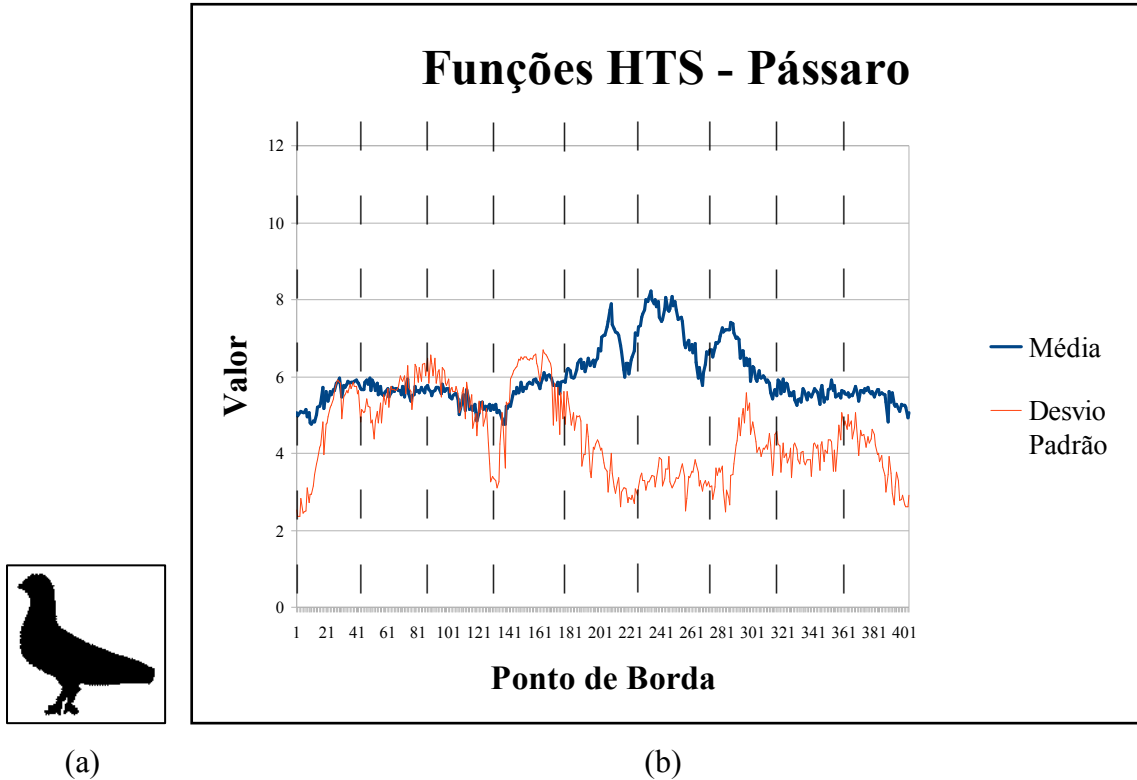


Figura 26. Em (a) é exibida uma imagem da base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e em (b) são exibidas suas funções unidimensionais. As retas verticais tracejadas indicam as posições onde as funções foram amostradas para se gerar o vetor de características para o pássaro. Como exemplo, a primeira posição do vetor de características conterá os valores 4,9 (média) e 2,4 (desvio padrão), a segunda posição conterá os valores 5,7 (média) e 5,0 (desvio padrão), e assim por diante. O vetor de características neste exemplo terá 9 posições pois as funções foram amostradas em 9 posições.

4.3 Etapa de Reconhecimento

Dados dois vetores de características, a comparação dos mesmos é realizada por meio da distância L_1 de forma que valores de média nas posições dos vetores de características são comparados com valores de média e valores de desvio padrão são comparados com valores de desvios padrão:

$$L_1(s, r) = \sum_{i=1}^N [|s_{i_m} - r_{i_m}| + |s_{i_d} - r_{i_d}|] \quad (28)$$

onde s e r são os vetores de características, s_{i_m} e r_{i_m} são os valores de média presentes nas posições s_i e r_i dos vetores, e s_{i_d} e r_{i_d} são os valores de desvio padrão presentes nas posições s_i e r_i dos vetores de características. O valor de N corresponde ao tamanho dos vetores.

Para determinar a classe de um objeto desconhecido, compara-se seu vetor de características com os vetores da base de dados. O objeto da base de imagens mais parecido (com menor distância) com o de teste pode ser utilizado então para determinar a classe deste (outros classificadores podem ser utilizados).

Na verdade, a comparação de dois vetores de características é realizada k vezes, onde k corresponde ao tamanho dos mesmos, isto é, ao número de posições amostradas das funções unidimensionais, rotacionando-se um dos vetores a cada comparação (trocando de posição de forma circular os dados presentes nas posições do vetor). A menor distância obtida nos k casamentos é tomada como a distância final entre os dois vetores. Isto é realizado para evitar que, caso o objeto esteja rotacionado em uma dada imagem, seu vetor de características seja penalizado em seus casamentos com outros vetores por também estar rotacionado (“deslocado” em algumas posições) visto que sempre se começa a construção das funções unidimensionais pelo ponto de borda que está mais no topo da forma na imagem (e continuando em sentido horário).

4.4 O Algoritmo do HTS

O Algoritmo 1 apresenta os passos do descritor HTS. A etapa de Pré-Processamento corresponde à linha 1. A etapa de Extração de Características encontra-se da linha 2 à linha 5 e a etapa de Reconhecimento está representada nas linhas 6 e 7.

O HTS foi criado inspirado no funcionamento do método BAS (ARICA; VURAL, 2003) (representação da borda da forma por meio de funções unidimensionais), descritor de formas tido como parâmetro de bom desempenho em muitos trabalhos de análise de formas. Entretanto, uma vantagem do HTS consiste no fato de que ele é mais rápido que o método BAS (ARICA; VURAL, 2003): o HTS apresenta complexidade linear em seu algoritmo enquanto que o BAS (ARICA; VURAL, 2003) apresenta complexidade quadrática.

Em relação à etapa de Pré-Processamento não há diferença entre os dois métodos: HTS e BAS (ARICA; VURAL, 2003). O mesmo algoritmo de processamento de imagens pode ser usado para se segmentar os pontos de borda do objeto na imagem.

Já em relação à etapa de Extração de Características, no método BAS (ARICA; VURAL, 2003), descritor apresentado anteriormente na seção 2.2, para se associar a cada um dos n pontos de borda a média, o desvio padrão e o momento de terceira ordem de seu conjunto de ângulos a fim de se obter as funções unidimensionais e de se gerar o vetor de características

do objeto sob análise, são realizados $n/2$ cálculos (medições de ângulos entre vetores). Com isto, para se associar estes valores estatísticos aos n pontos de borda são necessárias $n \cdot n/2$ operações obtendo-se uma complexidade $\theta(n^2)$.

Algoritmo 1:

Entrada: Uma imagem I de um objeto (sua silhueta).

Resultado: Classificação do objeto da imagem.

Estrutura de dados auxiliar: Uma matriz acumuladora (espaço de Hough) inicialmente zerada.

1. Segmente a borda do objeto a ser analisado na imagem;
2. Para cada pixel da borda faça:
 - 2.1. “Registre” sua senóide na matriz acumuladora incrementando as posições desta por onde a senóide passa;
3. Para cada pixel da borda faça:
 - 3.1. Percorra a respectiva senóide verificando os valores encontrados na matriz ao longo da mesma, gerando um histograma;
 - 3.2. Associe a média e o desvio padrão do histograma encontrado ao pixel (ponto) de borda;
4. Defina uma sequência entre os pixels da borda, iniciando em um deles e continuando em sentido horário, a fim de gerar duas funções unidimensionais com base nos valores de média e desvio padrão associados aos pontos;
5. Amostre as funções unidimensionais em posições igualmente espaçadas gerando o vetor de característica da forma de modo que cada posição do vetor armazene a média e o desvio padrão amostrados em certa posição das funções;
6. Realize a comparação deste vetor com todos os vetores da base de dados (através da distância L_1) gerando valores de distâncias aos mesmos;
7. Classifique a forma sendo analisada como sendo da mesma classe que a forma da base de dados mais similar (vetor com menor distância).

Na geração das funções unidimensionais, e a partir delas do vetor de características, pelo HTS, é preciso apenas que cada ponto da borda seja representado por uma senóide no espaço de Hough. Para isto são necessárias 180 operações por ponto de borda: varia-se θ de 0° a 179° e calcula-se, de acordo com a Equação 27, para cada valor de θ , o respectivo valor de ρ , incrementando-se a respectiva posição da matriz acumuladora (lembrando que $\Delta\rho=1$ e

$\Delta\theta=1^\circ$). Após registrar todas as senóides (de todos os pontos da borda) no espaço de Hough, basta percorrer novamente, para cada ponto do contorno, sua senóide, observando os valores presentes em cada posição da matriz por onde esta passa. Com isto para cada ponto da borda são necessárias mais 180 operações. Desta forma, para associar a média e desvio padrão a cada ponto de borda são necessárias 360 (180+180) operações. Como a borda tem n pontos, para se gerar as funções unidimensionais são necessárias $360 \cdot n$ operações. Assim, o HTS nesta etapa de seu algoritmo apresenta complexidade $\theta(n)$, muito menor que a complexidade da mesma etapa no BAS (ARICA; VURAL, 2003).

Por sua vez, na etapa de Reconhecimento, o BAS (ARICA; VURAL, 2003) emprega o algoritmo OCS (WANG; PAVLIDIS, 1990) na comparação de vetores de características, o qual apresenta complexidade quadrática. Já o HTS utiliza a distância L_1 para comparar os vetores, procedimento com complexidade linear.

Vale observar ainda que para se chegar na versão “final” do método HTS apresentada nesta dissertação, foram realizados diversos testes modificando-se partes do algoritmo, alterando-se estruturas de dados, etc. a fim de obter um melhor desempenho, mas tomando-se sempre o cuidado para não deixar o método lento. Dentre alguns dos testes realizados destacam-se:

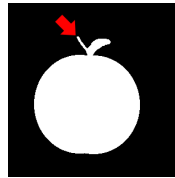
- **Geração e utilização de uma terceira função unidimensional para os objetos a partir do momento de terceira ordem dos histogramas dos pontos de borda:** os resultados foram piores além de tornar o método mais lento, também foram realizados testes variando-se as funções consideradas na construção dos vetores de características (ora considerando apenas a função de média, ora apenas a de desvio padrão, ora apenas a do terceiro momento, ora a de média e desvio padrão, ora a de média e do terceiro momento e ora a do desvio padrão e do terceiro momento), os melhores resultados foram os obtidos quando se trabalhou com as funções de média e desvio padrão na caracterização dos pontos de borda;
- **Comparação das próprias funções unidimensionais sem amostrá-las:** ao invés de se gerar o vetor de características para o objeto amostrando-se suas funções unidimensionais, considerou-se as próprias funções como vetor de características e comparou-se as mesmas utilizando distâncias como L_1 , χ^2 , Canberra, etc.; não houve melhora significativa e o método ficou muito mais lento;

- **Amostragem não homogênea das funções unidimensionais para se gerar os vetores de características:** foram testadas diversas formas de amostragem, tais como selecionando-se pontos de máximo globais, pontos de máximos locais, pontos de máximo em intervalos regulares, etc.; em todos os casos os resultados foram piores e o método ficou mais lento;
- **Casamento dos vetores de características usando outras distâncias e algoritmos (ao invés da distância L_1):** foram testados algoritmos e distâncias como o OCS (WANG; PAVLIDIS, 1990), distância de Bhattacharyya (BHATTACHARYYA, 1943), etc., porém os resultados não foram significativamente melhores e o descritor ficou bem mais lento;
- **Aumento no tamanho das matrizes acumuladoras variando com isto a “resolução” dos espaços de Hough (fazendo $0 < \Delta\rho < 1$ e também $0^\circ < \Delta\theta < 1^\circ$):** houve ganho de desempenho não muito expressivo, no entanto o método ficou mais lento e consumiu mais memória;
- **Diminuição no tamanho das matrizes acumuladoras (fazendo $\Delta\rho > 1$ e $\Delta\theta > 1^\circ$):** espaços de Hough menores também foram utilizados na tentativa de se realizar uma primeira etapa de comparações menos detalhadas e mais rápidas entre a forma de teste e as da base de dados a fim de só comparar com mais precisão (usando os espaços de Hough padrão) a forma de teste com as imagens mais parecidas da base de dados, evitando assim casamentos detalhados desnecessários; os resultados foram um pouco piores e o ganho de velocidade não foi muito significativo.

4.5 Invariância a Transformações Geométricas

O descritor HTS é invariante à rotação. Como mostrado na Figura 27, ao se rotacionar o objeto, os valores dos histogramas dos pontos de borda permanecem os mesmos, eles são apenas deslocados. Isto não muda a média e o desvio padrão dos histogramas e, portanto, os valores associados aos pontos de borda continuam os mesmos. Como dito, apenas é necessário rotacionar k vezes um dos dois vetores de características sendo comparados porque caso o objeto esteja rotacionado de uma imagem para outra, suas funções unidimensionais começarão a ser construídas em posições diferentes da borda (elas ficarão deslocadas entre si). Caso se deseje um descritor que não seja invariante à rotação basta não realizar a rotação

de um dos vetores de características k vezes na comparação entre vetores. Neste caso realiza-se a comparação entre dados dois vetores uma única vez nas suas configurações iniciais, isto é, sem rotacionar os mesmos.



(a)

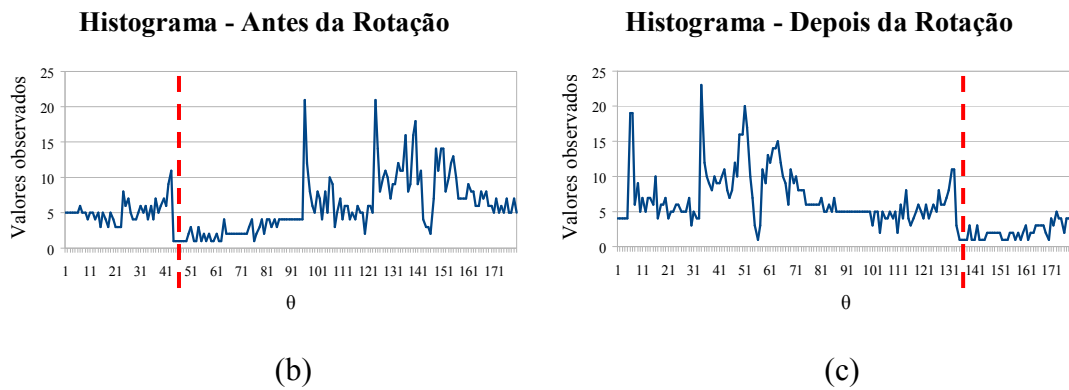
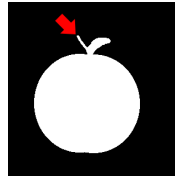


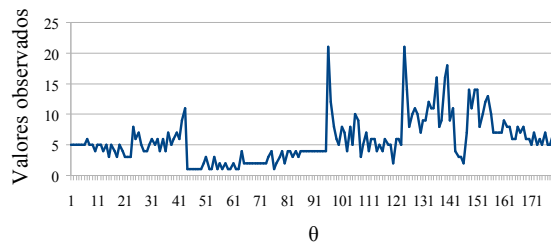
Figura 27. Em (a) exibe-se a imagem (silhueta) de uma maçã da base MPEG-7 Part B (JEANNIN; BOBER, 1999). O ponto de borda localizado na ponta do cabo da maçã (indicado pela seta vermelha) corresponde ao ponto cujo histograma foi verificado antes (b) e depois (c) de o objeto sofrer rotação em torno de seu centro de massa de aproximadamente 90° em sentido horário (o objeto foi deslocado até a origem por meio de seu centro de massa para a construção dos histogramas). Percebe-se que os valores dos histogramas são muito similares, há apenas uma mudança na disposição (“deslocamento”) dos mesmos ao longo dos histogramas. As retas vermelhas tracejadas indicam posições correspondentes dos histogramas. Rotacionando-se um dos histogramas a fim de deixar as duas retas vermelhas na mesma posição (mesmo valor de θ) obtém-se dois histogramas praticamente iguais.

Em relação à translação e mudança de escala, o HTS é invariante a estas transformações conforme mostram as Figuras 28 e 29. Ao aplicar estas transformações nos objetos, os histogramas de seus pontos de borda continuam praticamente os mesmos e com isto as funções unidimensionais e o vetores de características dos objetos não se alteram.



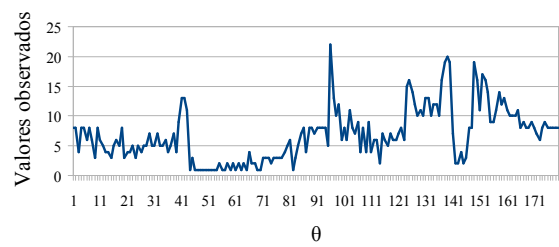
(a)

Histograma - Antes da Mudança de Escala



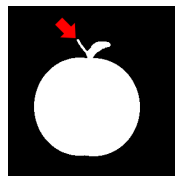
(b)

Histograma - Depois da Mudança de Escala



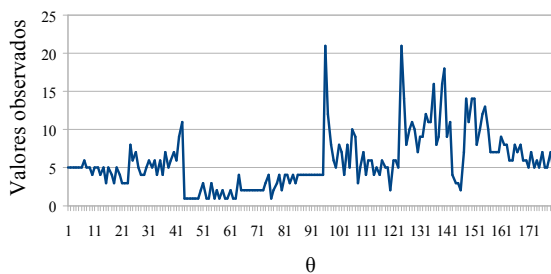
(c)

Figura 28. O histograma do ponto de borda indicado pela seta vermelha em (a) foi verificado antes (b) e após (c) a forma original ter sua escala alterada (dobrou-se a largura e altura da forma). O histograma do ponto praticamente não se alterou.



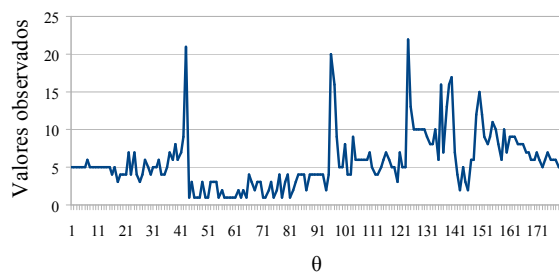
(a)

Histograma - Antes da Translação



(b)

Histograma - Depois da Translação



(c)

Figura 29. O histograma do ponto indicado pela seta vermelha em (a) foi verificado antes (b) e após (c) a forma (com tamanho de 250 pixels por 250 pixels) ser transladada aproximadamente dois mil pixels em relação ao eixo x e também ao y . Novamente pode-se constatar que o histograma não se alterou significativamente.

4.6 HTSn: Histogramas Baseados na Vizinhança das Senóides

Uma modificação na etapa de Extração de Características do método HTS fez com que suas taxas de acerto aumentassem consideravelmente. Em resumo, esta adaptação consiste em, ao invés de se percorrer a senóide de um dado ponto de borda verificando-se apenas os valores presentes nas posições da matriz acumuladora por onde a mesma passa a fim de construir o histograma do ponto, observa-se também, para cada posição desta curva, os valores nas posições adjacentes do espaço de Hough. Desta forma, o valor final observado para uma dada posição da senóide corresponde à soma do valor presente na respectiva posição do espaço de Hough com todos os valores presentes nas oito posições adjacentes do espaço. Como observação, demos o nome de HTSn (*HTS Neighborhood*) à esta nova versão do HTS.

As Figuras 30 e 31 ilustram o processo de construção do histograma de um dado ponto de borda pelo método HTS e pelo HTSn, respectivamente, evidenciando a diferença entre eles neste procedimento. No caso do HTS (Figura 30), o valor na célula com borda azul em cada matriz verde, de (a) a (d), corresponde ao valor armazenado no respectivo *bin* do histograma do ponto de borda associado à senóide cujas posições aparecem em vermelho na matriz, isto é: $\text{Histograma}[i] = 12$; $\text{Histograma}[i+1] = 13$; $\text{Histograma}[i+2] = 10$; $\text{Histograma}[i+3] = 12$. Vale ressaltar que a construção do histograma do ponto de borda pelo HTS continua de forma análoga percorrendo-se todas posições da senóide no espaço de Hough (matriz acumuladora) até o seu fim.

Já no HTSn (Figura 31), para se encontrar o valor de um *bin* do histograma do ponto de borda associado à senóide em vermelho, soma-se os valores presentes na respectiva posição da senóide e nas oito posições adjacentes a esta no espaço de Hough (conforme indicado pelas bordas azuis). Desta forma: $\text{Histograma}[i] = 10 + 12 + 12 + 14 + 12 + 13 + 10 + 10 + 16 = 109$; $\text{Histograma}[i+1] = 12 + 12 + 16 + 12 + 13 + 15 + 10 + 16 + 10 = 116$; $\text{Histograma}[i+2] = 13 + 15 + 15 + 16 + 10 + 12 + 10 + 15 + 13 = 119$; $\text{Histograma}[i+3] = 15 + 15 + 16 + 10 + 12 + 13 + 15 + 13 + 15 = 124$. A construção do histograma do ponto de borda continua de forma análoga percorrendo-se todas posições da senóide até o seu fim.

É importante mencionar que a observação dos valores presentes nas posições adjacentes às senóides no espaço de Hough na construção dos histogramas dos pontos de borda se justifica pelo fato de que, devido ao emprego de um espaço discreto (matriz acumuladora) na tentativa de representação de um espaço contínuo, trabalha-se com muitas aproximações e algumas vezes verificou-se que pontos pertencentes a segmentos de reta na imagem

apresentam senóides que não passam exatamente nas posições do espaço correspondentes aos segmentos (às vezes acabam passando muito próximos, em posições vizinhas). Com a observação das posições adjacentes da matriz, este problema é amenizado.

Vale ressaltar também que, assim como o HTS, o HTSn também apresenta complexidade linear no seu algoritmo (que é quase idêntico ao do HTS) e é invariante à rotação, translação e escala.

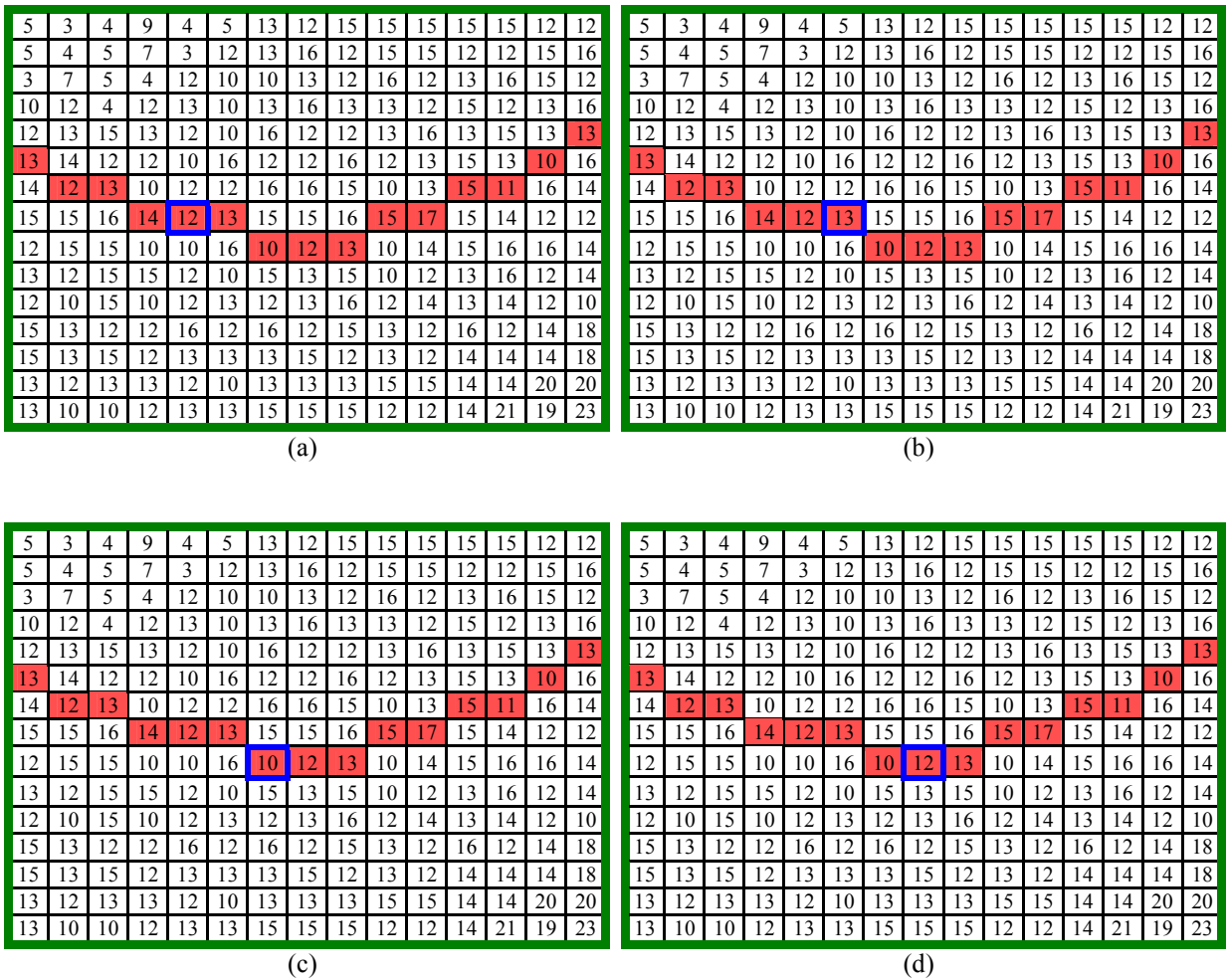


Figura 30. As matrizes com bordas verdes representam uma hipotética parte ampliada do espaço de Hough final de um dado objeto em quatro momentos distintos, de (a) a (d), no processo de construção (pelo descritor HTS) do histograma do ponto de borda associado à senóide cujas posições na matriz aparecem destacadas em vermelho. Os valores nas células com borda azuis são armazenados, obedecendo a ordem entre as matrizes verdes, isto é, de (a) a (d), nos *bins* do histograma do ponto.

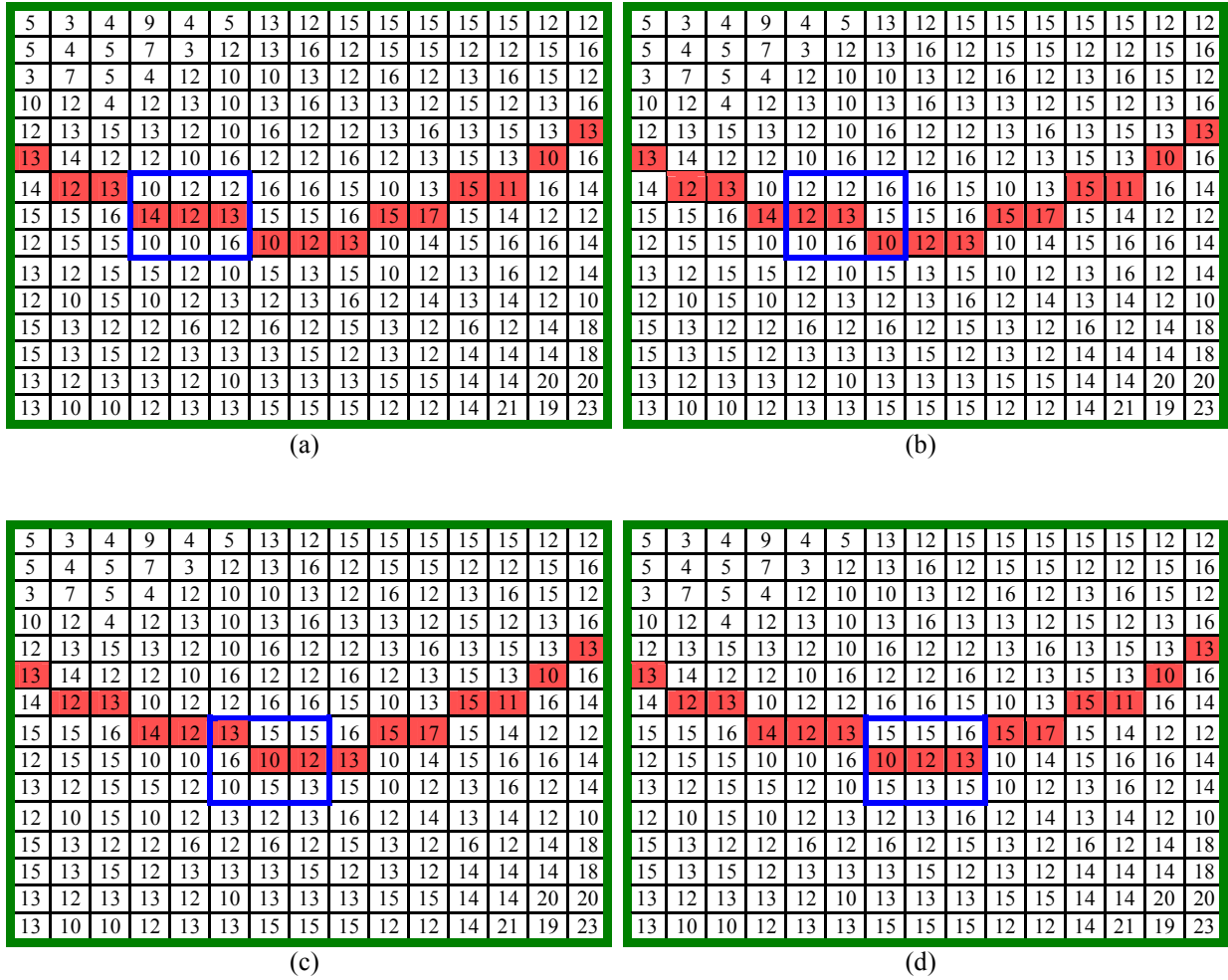


Figura 31. De forma análoga à Figura 30, as matrizes de (a) a (d) ilustram o processo de construção (agora pelo descritor HTSn) do histograma do ponto de borda associado à senóide cujas posições na matriz aparecem destacadas em vermelho. O valor armazenado em cada *bin* do histograma do ponto de borda corresponde à soma do valor na respectiva posição da senóide com os valores nas oito posições adjacentes a esta no espaço de Hough (valores dentro da borda azul).

4.7 HTSh: Vetores de Histogramas

Uma terceira versão do HTS que também apresentou excelentes resultados é proposta nesta dissertação de mestrado, à qual denominamos HTSh (HTS *Histograms*). Nesta versão, realiza-se a caracterização da forma através dos próprios histogramas dos pontos de borda, isto é, os próprios histogramas permanecem associados aos pontos de borda e não mais os valores da média e do desvio padrão dos mesmos, para a construção do vetor de características do objeto sob análise.

Depois de associar a cada ponto de borda um histograma (como no HTS original), são amostrados k pontos regularmente espaçados da borda da forma sob análise, percorrendo-se a borda em sentido horário a partir do topo da mesma na imagem. O vetor de características da forma então é composto pelos histogramas dos pontos amostrados (cada posição do vetor armazena um histograma). A Figura 32 ilustra a construção do vetor de características de um objeto com base nos histogramas dos pontos amostrados de sua borda.

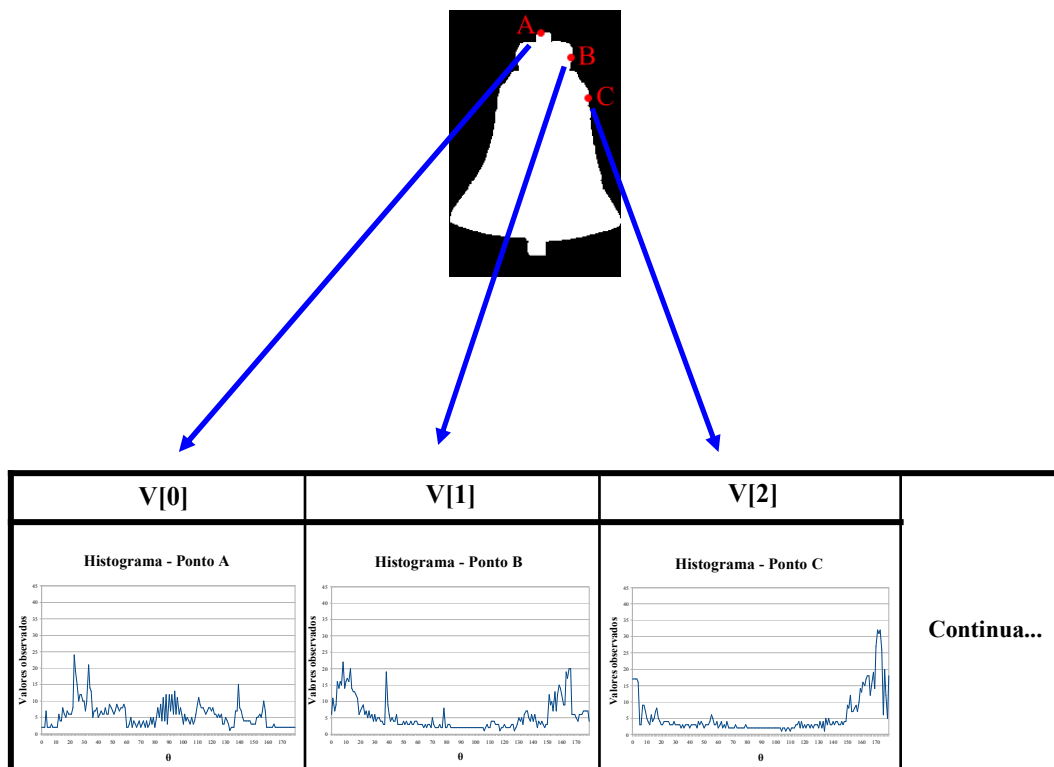


Figura 32. Construção do vetor de características (do método HTSh) para a imagem de um sino da base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999). Após associar a cada ponto da borda do sino um histograma (como no método HTS), k pontos regularmente espaçados da borda, começando pelo ponto A e continuando em sentido horário, são amostrados e seus histogramas são armazenados no vetor de características respeitando esta ordem (sequência) entre os pontos.

Dados dois vetores de características que representam as formas de dois objetos, eles são comparados a fim de verificar a similaridade dos mesmos. Tal atividade é realizada através da comparação dos pares de histogramas respectivos (um de cada vetor) nos vetores de características usando-se a distância de Canberra, definida na Equação 29:

$$C(s, r) = \sum_{i=1}^N \frac{|s_i - r_i|}{|s_i| + |r_i|} \quad (29)$$

onde s e r são os histogramas sendo comparados e N corresponde ao número de *bins* dos histogramas (no nosso caso $N=180$). A distância final entre os dois vetores de características corresponde à soma dos custos C da comparação de todos os pares de histogramas em posições respectivas nos vetores. Vale mencionar que outras distâncias como a χ^2 e L_1 também foram testadas na comparação entre os histogramas porém produziram piores resultados.

Para reconhecer a classe de um objeto, como no HTS, pode-se utilizar o classificador 1-*Nearest Neighbor* (1-NN). Compara-se o vetor de características do objeto de teste com os vetores de todos os objetos da base de dados, sendo que o vetor da base que apresenta a menor distância (primeiro vizinho mais próximo) é utilizado para determinar a classe do objeto desconhecido.

Na verdade, assim como no HTS, a comparação entre dois vetores de características é realizada k vezes, onde k corresponde ao tamanho dos mesmos (ao número de pontos amostrados das bordas), rotacionando-se um dos vetores a cada comparação, isto é, trocando de posição de forma circular os dados, no caso histogramas, presentes nas posições do vetor. A menor distância obtida nas k comparações é tomada como sendo a distância final entre os dois vetores. Como dito, isto é realizado para evitar que caso o objeto esteja rotacionado em uma dada imagem, seu vetor de características seja penalizado em suas comparações com outros vetores por estar rotacionado (“deslocado” em algumas posições), visto que sempre se começa a amostragem da borda do objeto pelo ponto de borda que está mais ao topo na imagem.

Apesar de utilizar o recurso de se rotacionar um dos vetores de características, o descritor HTSh não é invariante à rotação. Isto ocorre porque, como mostrado na Figura 27, ao se rotacionar um objeto na imagem, os valores encontrados nos *bins* dos histogramas de seus pontos de borda são deslocados (permanecem aproximadamente os mesmos, mas em disposições diferentes). Desta forma os custos de casamento C (distância de Canberra) dos respectivos histogramas de dois vetores de características obtidos de um mesmo objeto, mas antes e depois de este sofrer uma rotação, serão altos. Caso se deseje invariância à rotação uma técnica adicional precisa ser empregada.

Em relação à translação e mudança de escala, o HTSh é invariante da mesma forma que o HTS uma vez que estas transformações não alteram os histogramas dos pontos de borda como mostrado nas Figuras 28 e 29.

O Algoritmo 2 apresenta os passos do descritor HTSh. A etapa de Pré-Processamento corresponde à linha 1. A etapa de Extração de Características encontra-se da linha 2 à linha 5 e a etapa de Reconhecimento está representada nas linhas 6 e 7.

O HTSh também apresenta complexidade linear em seu algoritmo, no entanto a quantidade de operações realizadas na fase de Reconhecimento para a comparação de elementos de dois vetores de características é muito maior que no HTS e HTSn e por isto ele tende a ser mais lento nesta etapa. Vale dizer que o descritor HTSh foi criado na última etapa do desenvolvimento desta dissertação e ainda precisa ser otimizado para melhorar seus tempos de processamento.

Algoritmo 2:

Entrada: Uma imagem I de um objeto (sua silhueta).

Resultado: Classificação do objeto da imagem.

Estrutura de dados auxiliar: Uma matriz acumuladora (espaço de Hough) inicialmente zerada.

1. Segmente a borda do objeto a ser analisado na imagem;

2. Para cada pixel da borda faça:

2.1. “Registre” sua senóide na matriz acumuladora incrementando as posições desta por onde a senóide passa;

3. Para cada pixel da borda faça:

3.1. Percorra a respectiva senóide verificando os valores encontrados na matriz ao longo da mesma, gerando um histograma;

3.2. Associe o histograma encontrado ao pixel (ponto) de borda;

4. Defina uma sequência entre os pixels da borda, iniciando em um deles e continuando em sentido horário, e amostrando k pixels regularmente espaçados sobre a borda;

5. Gere o vetor de característica de modo que cada posição do vetor armazene o histograma de um dos pontos amostrados (obedecendo a sequência entre os pontos de borda);

6. Realize a comparação deste vetor com todos os vetores da base de dados utilizando a distância de Canberra entre os histogramas e gerando valores de distâncias aos mesmos;

7. Classifique a forma sendo analisada como sendo da mesma classe que a forma da base de dados mais similar (vetor com menor distância).

5 Experimentos, Resultados e Discussão

Os três novos descritores de formas propostos e apresentados no Capítulo 4 foram testados sobre algumas bases de imagens para avaliar seus desempenhos e compará-los com os resultados obtidos por importantes descritores de formas encontrados na literatura. Além dos testes realizados sobre duas importantes bases públicas de imagens denominadas Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER, 1999), onde são analisados tanto as taxas de acerto quanto os tempos de processamento dos descritores, no terceiro experimento realiza-se uma comparação mais detalhada sobre os tempos de processamento dos métodos propostos e do descritor BAS (ARICA; VURAL, 2003), os quatro melhores métodos segundo os resultados obtidos.

Em relação ao HTSh, apesar de não ser invariante a rotação, ele foi aplicado sobre as bases de imagens (que apresentam objetos rotacionados) sem nenhum procedimento adicional de registro dos objetos apenas para se ter uma ideia aproximada de seu desempenho. Futuramente, métodos de registro de formas deverão ser estudados e incorporados ao descritor caso se deseje deixá-lo invariante à rotação.

Como já mencionado anteriormente neste trabalho, vale ressaltar, entretanto, que em alguns casos a invariância a transformações geométricas pode não ser desejada (ou não ser necessária). Para exemplificar tal situação, na qual o registro dos objetos não é necessário (e portanto na qual o descritor HTSh não sofre nenhum tipo de prejuízo em seu desempenho), realizamos um quarto experimento sobre uma base de imagens radiográficas dos sinus frontais (cavidades do osso frontal do crânio) de pessoas, a fim de identificá-las por meio das formas dos mesmos. Nesta base, não há variação em termos de rotação, translação e escala de uma imagem para outra, assim, os resultados observados do HTSh são resultados “reais” e não apenas aproximados como nos outros experimentos. Comprovou-se assim o bom desempenho do método apresentado anteriormente nos primeiros experimentos.

Vale observar que na realização dos experimentos utilizou-se um computador Macintosh Power Mac G5 com processador PPC970 de 2.0 GHz (dois núcleos), 4 GB de memória RAM e sistema operacional Debian 6.0.4. Os descritores de formas foram implementados em linguagem C.

É importante mencionar ainda que nos experimentos realizados, exceto no terceiro, os vetores de características dos métodos HTS, HTSn, HTSh e BAS (*Beam Angle Statistics*) (ARICA; VURAL, 2003) apresentam tamanho de 50 posições. Este tamanho ótimo em termos de taxa de acerto e tempo de processamento foi obtido para o HTS, HTSn e HTSh

(propostos neste trabalho) por meio de testes realizados variando-se o tamanho dos vetores de características. Apenas como um breve exemplo destes testes, a Figura 33 ilustra as curvas *Precision-Recall* obtidas para o HTS sobre a base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999) ao se variar o tamanho dos vetores de características empregados no reconhecimento dos objetos nas imagens desta base. Quanto mais elevada a curva, melhor o desempenho. Percebe-se que o aumento do tamanho dos vetores de características a partir de 40 ou 50 posições não ocasiona ganho de desempenho expressivo ao descritor ao passo que o tempo de processamento aumenta. Comprova-se deste modo que vetores com tamanho de 50 posições são uma boa opção.

A fim de poder comparar de forma justa os resultados dos descritores propostos com os do BAS (ARICA; VURAL, 2003), os vetores deste descritor também apresentam 50 posições nos experimentos, exceto apenas no terceiro.

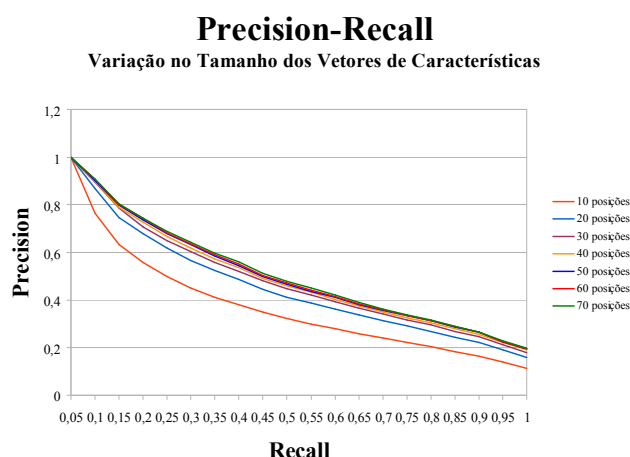


Figura 33. Curvas *Precision-Recall* do método HTS obtidas para a base de imagens MPEG-7 Part B (JEANNIN; BOBER, 1999) ao se variar o tamanho (número de posições) dos vetores de características, conforme indicado na legenda, no reconhecimento dos objetos nas imagens da base. Quanto mais elevada a curva, melhor o desempenho. Percebe-se que os resultados são muito próximos quando se usou vetores com tamanhos a partir de 40 posições.

5.1 Primeiro Experimento - Base de Imagens Kimia-216

Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) é uma base pública tradicional de imagens, disponível para testes de descritores de formas. Nela existem 216 imagens: 18 classes de objetos com 12 imagens por classe. Em cada imagem há a silhueta de um objeto em preto e um fundo branco. O tamanho das imagens na base é pouco variável e suas dimensões não excedem poucas centenas de pixels. A Figura 34 mostra algumas destas imagens. Pode-se

observar que alguns dos objetos (silhuetas) apresentam-se rotacionados e um pouco “deformados”. Há pouca variação em termos de translação e escala.

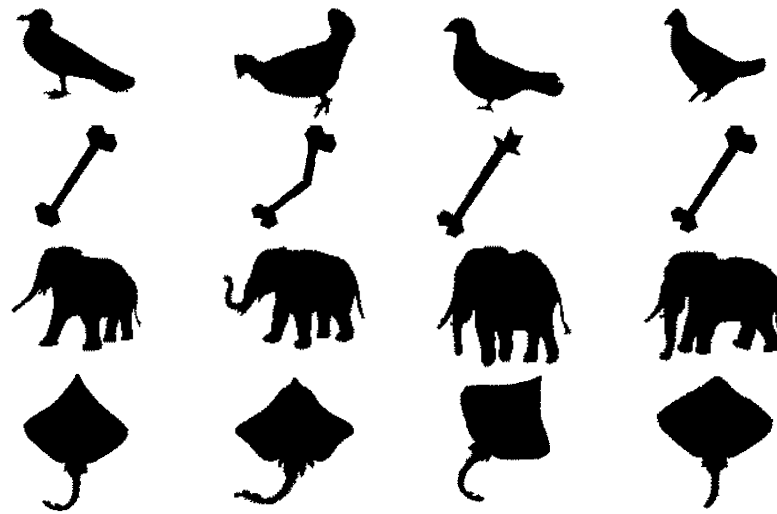


Figura 34. Exemplos de imagens (de 4 classes) presentes na base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).

Para analisar o desempenho dos descritores testados sobre a base de imagens em questão gerou-se as curvas *Precision-Recall* e *Multiscale Separability* dos mesmos e verificou-se suas matrizes de confusão. As curvas citadas são muito utilizadas na análise de desempenho de descritores de formas. A curva *Precision-Recall* mede a taxa *Precision* (razão da quantidade de imagens da classe correta retornadas em uma consulta sobre o total de imagens retornadas na consulta) em função da taxa *Recall* (razão da quantidade de imagens da classe correta retornadas na consulta sobre o total de imagens da classe correta existentes na base de imagens). Por sua vez a curva *Multiscale Separability* indica quão separados no espaço de características estão os agrupamentos de vetores de características de classes distintas. Quanto mais separados, melhor o descritor.

A Figura 35 ilustra as curvas *Precision-Recall* para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). Quanto mais elevada a curva, melhor o descritor. Pode-se perceber que os quatro melhores métodos foram o BAS (ARICA; VURAL, 2003) e os três descritores propostos neste trabalho: HTS, HTSn e HTSh.

Por sua vez, a Figura 36 mostra as curvas *Multiscale Separability* destes quatro melhores descritores para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). De forma similar, quanto mais elevada a curva, melhor o método. Pode-se observar que os três descritores propostos neste trabalho apresentaram resultados bastante superiores ao do BAS (ARICA; VURAL, 2003).

Precision-Recall

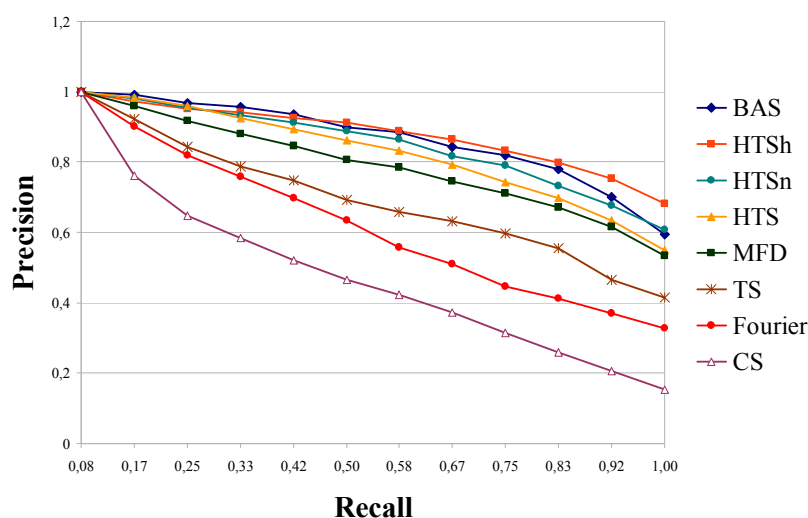


Figura 35. Curvas *Precision-Recall* para os métodos testados na base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004): *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003), HTSh, HTSn, HTS, *Multiscale Fractal Dimension* (MFD) (TORRES; FALCÃO; COSTA, 2004), *Tensor Scale* (TS) (MIRANDA; TORRES; FALCÃO, 2005), Fourier (ZHANG; LU, 2001) e *Contour Saliency* (CS) (TORRES; FALCÃO, 2007). Quanto mais elevada a curva, melhor o método. Os descritores BAS (ARICA; VURAL, 2003), HTSh, HTSn, e HTS apresentaram os melhores resultados.

Multiscale Separability

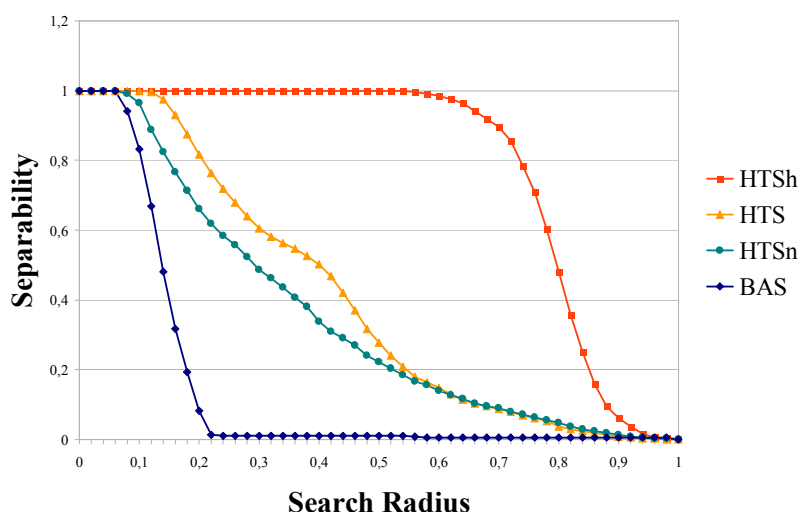


Figura 36. Curvas *Multiscale Separability* para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) dos quatro melhores métodos em termos de resultados *Precision-Recall*: HTSh, HTS, HTSn e *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003). Quanto mais elevada a curva, melhor o descritor. Os métodos propostos neste trabalho obtiveram desempenhos muito superiores ao do BAS (ARICA; VURAL, 2003).

Tabela 7. Taxas de acerto e erro por classe da matriz de confusão do HTSn sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). As taxas médias de acerto e erro bem como o desvio padrão destas taxas são exibidos nas duas últimas linhas.

Classe	Acerto (%)	Erro (%)
C1	91,67	8,33
C2	100,00	0,00
C3	100,00	0,00
C4	83,33	16,67
C5	100,00	0,00
C6	100,00	0,00
C7	91,67	8,33
C8	100,00	0,00
C9	100,00	0,00
C10	100,00	0,00
C11	100,00	0,00
C12	100,00	0,00
C13	91,67	8,33
C14	100,00	0,00
C15	100,00	0,00
C16	91,67	8,33
C17	100,00	0,00
C18	100,00	0,00
Média:	97,22	2,78
Desvio Padrão:	4,95	

Tabela 8. Matriz de confusão do HTSh para a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004).

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18
C1	10	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
C2	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C3	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C4	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0
C6	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	0
C7	0	1	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0
C8	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0
C9	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0
C10	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0
C11	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0
C12	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0
C13	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0
C14	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0
C15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0
C17	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	11	0
C18	0	0	2	0	0	0	1	0	0	0	0	0	1	0	0	0	0	8

Tabela 9. Taxas de acerto e erro por classe da matriz de confusão do HTSh sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004). As taxas médias de acerto e erro bem como o desvio padrão destas taxas são exibidos nas duas últimas linhas.

Classe	Acerto (%)	Erro (%)
C1	83,33	16,67
C2	100,00	0,00
C3	100,00	0,00
C4	100,00	0,00
C5	100,00	0,00
C6	100,00	0,00
C7	91,67	8,33
C8	100,00	0,00
C9	100,00	0,00
C10	100,00	0,00
C11	100,00	0,00
C12	100,00	0,00
C13	100,00	0,00
C14	100,00	0,00
C15	100,00	0,00
C16	100,00	0,00
C17	91,67	8,33
C18	66,67	33,33
Média:	96,30	3,70
Desvio Padrão:	8,68	

Em relação aos tempos de processamento utilizados por cada um dos quatro métodos em questão nas etapas de Extração de Características e Reconhecimento sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004), a Tabela 10 ilustra os resultados obtidos. Vale lembrar que na etapa de Extração de Características mediu-se o tempo gasto por cada método para realizar a extração de características dos objetos das 216 imagens da base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e na etapa de Reconhecimento mediu-se o tempo usado por cada método para realizar as 216-216 comparações de vetores de características. O HTS foi o método mais rápido.

A partir dos resultados obtidos, pode-se verificar que os descritores HTS, HTSn e HTSh apresentaram resultados muito bons, similares aos do BAS (ARICA; VURAL, 2003) e melhores do que os dos outros todos descritores testados.

Em relação às curvas *Precision-Recall*, os descritores HTS, HTSn e HTSh superaram, junto com o BAS (ARICA; VURAL, 2003), os resultados dos outros descritores testados. A abordagem HTSh foi a melhor das três propostas neste trabalho, inclusive superando a curva do BAS (ARICA; VURAL, 2003) para um amplo intervalo de valores de *Recall*.

Tabela 10. Tempos aproximados em segundos tomados pelos quatro métodos em questão sobre a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) nas etapas de Extração de Características e Reconhecimento.

Método	Tempo em segundos (s) para a Extração de Características	Tempo em segundos (s) para o Reconhecimento
HTS	9	6
HTSn	11	6
BAS	11	72
HTSh	11	901

As curvas *Precision-Recall* do descritor HTS e do HTSn ficaram ligeiramente abaixo da curva do BAS (ARICA; VURAL, 2003) mas muito próximas. A curva do HTSn se apresentou um pouco melhor que a do HTS e chegou até a superar a curva do BAS (ARICA; VURAL, 2003) mas apenas para valores de *Recall* elevados. Além de tudo isto, para os valores de *Recall* até 0,25, isto é, no início das curvas, os desempenhos do BAS (ARICA; VURAL, 2003), do HTS, do HTSn e do HTSh foram praticamente iguais. Isto é um fato importante porque um bom número de aplicações que se valem de descritores de formas como sistemas de busca de imagens por conteúdo (*Content-Based Image Retrieval* - CBIR) operam majoritariamente nesta região da curva: com valores de *Precision* elevados mesmo que com valores baixos de *Recall*. Em geral, nestes sistemas, o usuário prioriza mais o fato de as imagens trazidas na busca serem da mesma classe que a imagem apresentada do que o fato de se ter retornado todas as imagens da classe de consulta existentes na base de dados ou não.

Em se tratando dos resultados *Multiscale Separability*, os três métodos propostos superaram muito o resultado do BAS (ARICA; VURAL, 2003). O HTSh foi o melhor método, o HTS o segundo melhor e o HTSn o terceiro.

Com relação às matrizes de confusão, a matriz do BAS (ARICA; VURAL, 2003) foi ligeiramente superior às dos descritores propostos, mas a diferença não foi muito grande. A matriz do HTSh foi ligeiramente pior que as matrizes do HTS e HTSn.

Em termos de tempo de processamento percebe-se que o HTS e o HTSn são métodos mais rápidos que o BAS (ARICA; VURAL, 2003), em especial na etapa de Reconhecimento, por apresentarem menor complexidade que este em seus algoritmos, conforme explicado no Capítulo 4.

O HTSn é mais lento que o HTS na etapa de Extração de Características por nesse método ser necessário verificar os valores das posições vizinhas no espaço de Hough ao se percorrer a senóide de um ponto de borda (no HTS verifica-se apenas os valores nas posições da matriz

sob a senóide). Em relação à etapa de Reconhecimento, os dois métodos utilizam o mesmo algoritmo (casamento usando a distância L_1) e portanto apresentaram tempos iguais.

O HTSh, apesar de também possuir complexidade linear, realiza muitas comparações nos casamentos de vetores de características uma vez que trabalha com distâncias entre diversos histogramas (presentes nos vetores) e por isto foi bastante lento na etapa de Reconhecimento.

5.2 Segundo Experimento - Base de Imagens MPEG-7 Part B

Outra base pública de imagens disponível é a MPEG-7 CE-Shape-1 Part B (JEANNIN; BOBER, 1999). Nesta base existem 1400 imagens (divididas em 70 classes, cada classe com 20 imagens) onde o fundo é preto e a silhueta do objeto é branca.

A base MPEG-7 Part B (JEANNIN; BOBER, 1999) é mais complexa do que a base Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) uma vez que, além de ser maior e de possuir mais classes e imagens, o tamanho das imagens varia muito, há rotação, translação e mudança de escala nos objetos, além de deformações e ruídos. Também, há grande variabilidade intraclasse e em alguns casos grande similaridade interclasses. A Figura 37 ilustra casos destas transformações e características da base.

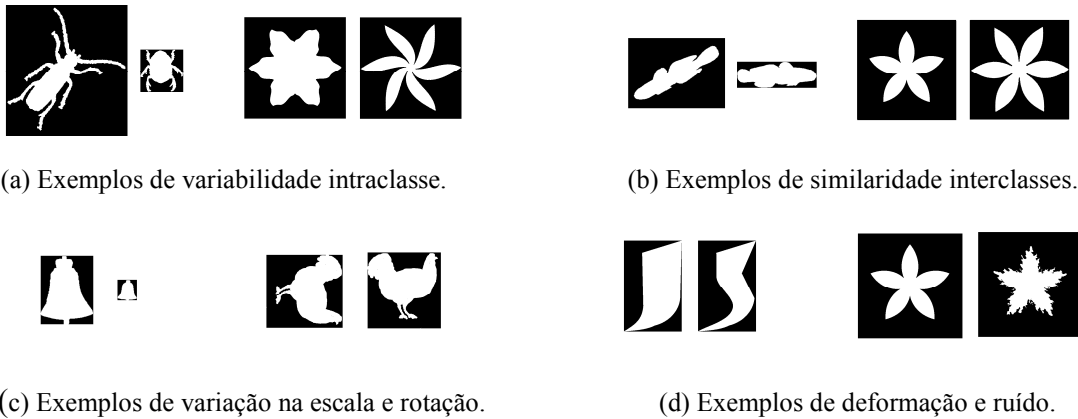


Figura 37. Exemplos das imagens da base MPEG-7 Part B (JEANNIN; BOBER, 1999). Em (a), (c) e (d) cada par de imagens corresponde a imagens de uma mesma classe. Em (b) em cada par há imagens de duas classes diferentes. As imagens estão representadas proporcionalmente aos seus tamanhos reais (exceto as dos peixes e sinos que estão aumentadas para melhorar a visualização).

Os descritores HTS, HTSn e HTSh, bem como os demais métodos encontrados na literatura e descritos nesta dissertação, também foram testados nesta base de imagens e suas

curvas *Precision-Recall* e *Multiscale Separability* foram encontradas. As matrizes de confusão dos métodos também foram verificadas.

A Figura 38 mostra as curvas *Precision-Recall* obtidas e a Figura 39 mostra as curvas *Multiscale Separability*. As matrizes de confusão (e medidas extraídas destas) do HTS, HTSn, HTSh e BAS (ARICA; VURAL, 2003) encontram-se na Figura 40. Como complemento aos dados da Figura 40, a taxa média de acerto, a de erro e o desvio padrão da matriz de confusão do descritor BAS (ARICA; VURAL, 2003) correspondem a 96%, 4% e 8,15%, respectivamente. A matriz de confusão do HTS apresenta taxa média de acerto, de erro e desvio padrão com valores de 85,14%, 14,86% e 17,4%, respectivamente. A taxa média de acerto, a de erro e o desvio padrão da matriz de confusão do HTSn correspondem a 89,21%, 10,79% e 14,86%, respectivamente. Por fim, a taxa média de acerto, a de erro e o desvio padrão da matriz de confusão do HTSh correspondem a 89,71%, 10,29% e 16,61%, respectivamente.

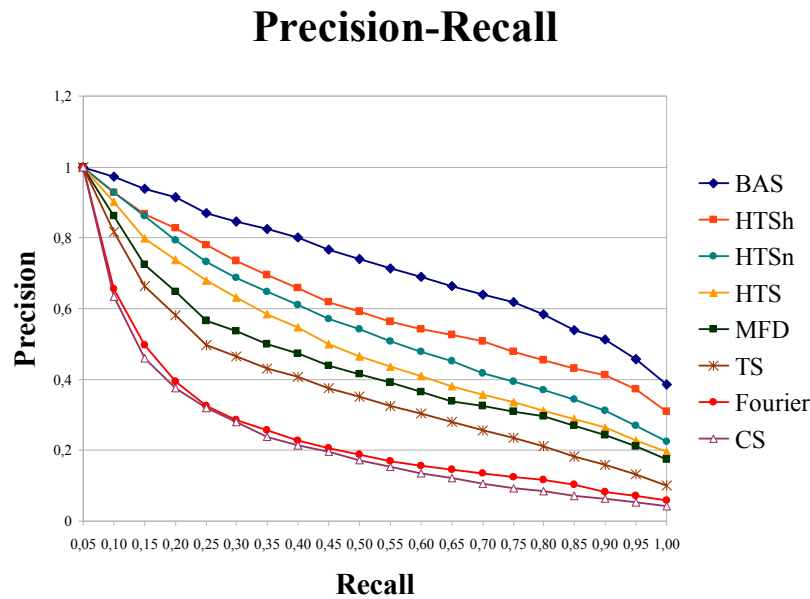


Figura 38. Curvas *Precision-Recall* para os métodos testados na base MPEG-7 Part B (JEANNIN; BOBER, 1999): *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003), HTSh, HTSn, HTS, *Multiscale Fractal Dimension* (MFD) (TORRES; FALCÃO; COSTA, 2004), *Tensor Scale* (TS) (MIRANDA; TORRES; FALCÃO, 2005), Fourier (ZHANG; LU, 2001) e *Contour Saliency* (CS) (TORRES; FALCÃO, 2007). Quanto mais elevada a curva, melhor o método. Mais uma vez o BAS (ARICA; VURAL, 2003) e os métodos propostos (HTS, HTSn e HTSh) apresentaram os melhores resultados.

Multiscale Separability

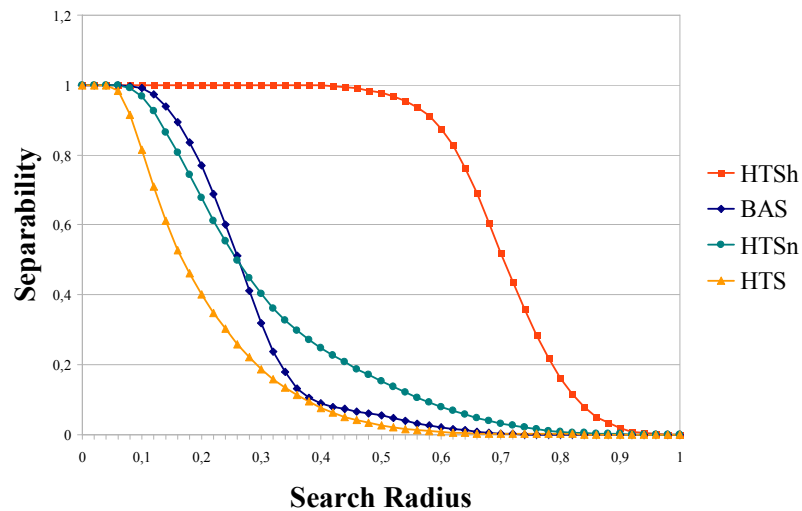


Figura 39. Curvas *Multiscale Separability* para a base MPEG-7 Part B (JEANNIN; BOBER, 1999) dos quatro melhores métodos em termos de resultados *Precision-Recall*: HTSh, BAS (ARICA; VURAL, 2003), HTSn e HTS. Quanto mais elevada a curva, melhor o descritor. Observa-se que o HTSh obteve o melhor desempenho o qual foi muito melhor que o do BAS (ARICA; VURAL, 2003). Os desempenhos do HTSn e HTS foram próximos ao desempenho do BAS (ARICA; VURAL, 2003).

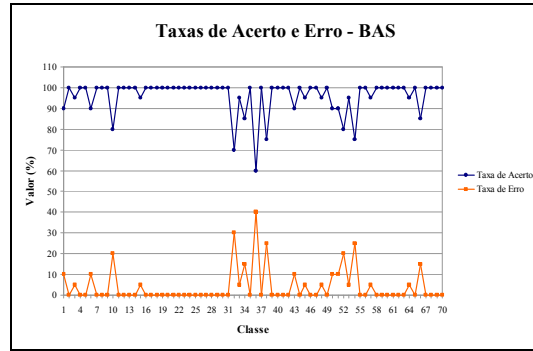
Em relação às curvas *Precision-Recall*, as do BAS (ARICA; VURAL, 2003), HTSh, HTSn e HTS foram as quatro melhores (nesta ordem).

Nas curvas *Multiscale Separability* o melhor resultado foi obtido pelo HTSh com desempenho muito superior em relação aos dos demais métodos. A curva do HTSn se apresentou equivalente à do BAS (ARICA; VURAL, 2003) e a do HTS apenas um pouco pior que esta.

Com relação às matrizes de confusão, a do BAS (ARICA; VURAL, 2003) foi superior às dos outros métodos testados. Vale observar que na construção das matrizes empregou-se a técnica de *cross validation k-folds* (com $k=10$) bem com um classificador 1-*Nearest Neighbor* (1-NN) no reconhecimento das imagens de teste. Também é importante ressaltar que dada a grande quantidade de classes nesta base de imagens, as matrizes de confusão foram representadas por imagens em tons de cinza (na Figura 40). Cada “pixel” da imagem representa uma posição da matriz. Quanto mais claro o “pixel” da imagem, maior o valor na respectiva posição da matriz. As taxas de acerto e erro (por classe) dos métodos nas matrizes de confusão são exibidas na forma de gráficos na própria Figura 40, também devido ao grande número de classes na base de imagens em questão.



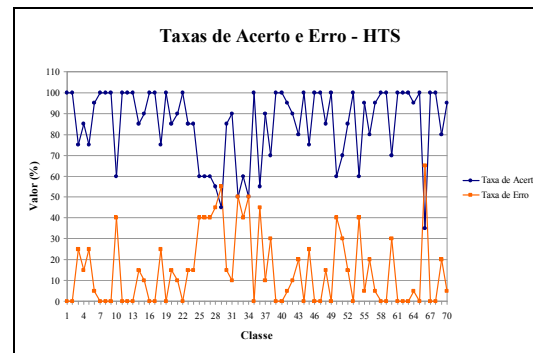
(a)



(b)



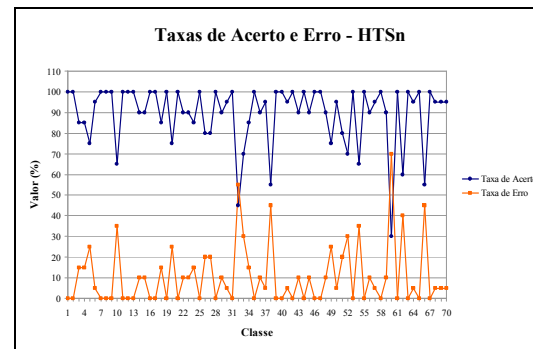
(c)



(d)



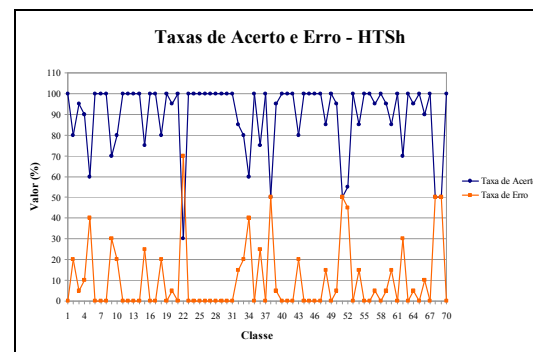
(e)



(f)



(g)



(h)

Figura 40. Matrizes de confusão e gráficos com as taxas de acerto e erro por classe (obtidas a partir destas matrizes) para a base MPEG-7 Part B (JEANNIN; BOBER, 1999). Em (a) e (b) são apresentados os resultados do BAS (ARICA; VURAL, 2003). Em (c) e (d) são apresentados os resultados do HTS. Em (e) e (f) são apresentados os resultados do HTSn. Por fim, em (g) e (h) são apresentados os resultados do HTSh.

A Tabela 11 ilustra os tempos de processamento utilizados por cada um dos quatro métodos em questão nas etapas de Extração de Características e Reconhecimento sobre a base MPEG-7 Part B (JEANNIN; BOBER, 1999). Na etapa de Extração de Características mediu-se o tempo gasto por cada método para realizar a extração de características dos objetos das 1400 imagens da base MPEG-7 Part B (JEANNIN; BOBER, 1999) e na etapa de Reconhecimento mediu-se o tempo usado por cada método para realizar as 1400·1400 comparações de vetores de características. O HTS foi o método mais rápido.

Tabela 11. Tempos aproximados em segundos tomados pelos quatro métodos em questão sobre a base MPEG-7 Part B (JEANNIN; BOBER, 1999) nas etapas de Extração de Características e Reconhecimento.

Método	Tempo em segundos (s) para a Extração de Características	Tempo em segundos (s) para o Reconhecimento
HTS	296	221
HTSn	315	221
BAS	1157	2997
HTSh	302	38238

Em resumo, os resultados apresentados mostram que os descritores HTS, HTSn e HTSh também demonstraram desempenhos muito bons na a base MPEG-7 Part B (JEANNIN; BOBER, 1999). Suas curvas *Precision-Recall* foram as melhores dentre as curvas de todos os descritores testados, exceto em relação à do BAS (ARICA; VURAL, 2003), descritor cuja curva se apresentou um pouco melhor que as dos três métodos propostos neste trabalho. Observa-se que a abordagem HTSh obteve curva consideravelmente superior às curvas dos métodos HTS e HTSn e bastante próxima à do BAS (ARICA; VURAL, 2003).

A curva *Multiscale Separability* do HTSh foi a melhor de todas superando em muito a do BAS (ARICA; VURAL, 2003). As curvas do HTS e HTSn ficaram muito próximas da do BAS (ARICA; VURAL, 2003): a curva do HTSn foi praticamente equivalente à do BAS (ARICA; VURAL, 2003) e a do HTS apenas um pouco pior que esta.

Em relação às matrizes de confusão, o BAS (ARICA; VURAL, 2003) apresentou desempenho um pouco superior aos do HTS, HTSn e HTSh.

Com relação ao tempo de processamento, percebe-se que o HTS e o HTSn foram muito mais rápidos que o BAS (ARICA; VURAL, 2003). Conforme explicado no Capítulo 4, isto se justifica pelo fato da complexidade dos dois primeiros métodos ser linear enquanto que a do BAS (ARICA; VURAL, 2003) é quadrática.

Apesar de também possuir complexidade linear em todo seu algoritmo, o HTSh na etapa de Reconhecimento faz muitas comparações no casamento de dois dados vetores de características e acaba se tornando bastante lento, mais lento até que o BAS (ARICA; VURAL, 2003). No entanto, vale observar que, com o aumento do tamanho dos vetores de características, o BAS (ARICA; VURAL, 2003) tende a se tornar mais lento que o HTSh na etapa de Reconhecimento devido à sua maior complexidade.

5.3 Terceiro Experimento - Tempo de Processamento

O HTS e o HTSn foram muito mais rápidos que o BAS (ARICA; VURAL, 2003) nos experimentos com as bases Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER, 1999). O HTSh foi mais rápido na etapa de Extração de Características mas muito mais lento na etapa de Reconhecimento que o BAS (ARICA; VURAL, 2003).

Para se obter uma melhor comparação dos tempos de processamento nas etapas de Extração de Características e Reconhecimento dos três descritores de formas propostos nesta dissertação com os do BAS (ARICA; VURAL, 2003), um terceiro experimento foi realizado. Para o teste da etapa de Extração de Características foram criadas oito imagens binárias contendo em cada uma delas um quadrado com dimensões variando de 20 pixels por lado a 2560 pixels por lado (de uma imagem para outra dobra-se o tamanho dos lados do quadrado). Mediu-se o tempo utilizado por cada método para, a partir dos pontos de borda segmentados dos quadrados (dobrando em quantidade de uma imagem para outra), gerar os vetores de características para as formas.

As Figuras 41, 42 e 43 ilustram as comparações dos tempos de processamento na etapa de Extração de Características dos métodos propostos (HTS, HTSn e HTSh) com os do BAS (ARICA; VURAL, 2003). Os tempos do BAS (ARICA; VURAL, 2003) são iguais nas três imagens. Não se representou os tempos dos quatro métodos em uma só imagem para não dificultar a visualização.

Percebe-se que o HTS é mais rápido que o BAS (ARICA; VURAL, 2003) e conforme se aumenta o tamanho do objeto, esta diferença cresce muito. O HTSn e o HTSh nos testes com quadrados pequenos (até 80 pontos por lado) foram um pouco mais lentos que o BAS (ARICA; VURAL, 2003), mas ao se aumentar o tamanho dos quadrados o BAS (ARICA; VURAL, 2003) se torna muito mais lento que eles. Esta maior lentidão do BAS (ARICA;

VURAL, 2003) pode ser explicada, conforme descrito no Capítulo 4, por este método apresentar complexidade quadrática em seu algoritmo enquanto que os três métodos propostos neste trabalho (HTS, HTSn e HTSh) apresentam complexidade linear.

Tempo Utilizado - Extração de Características

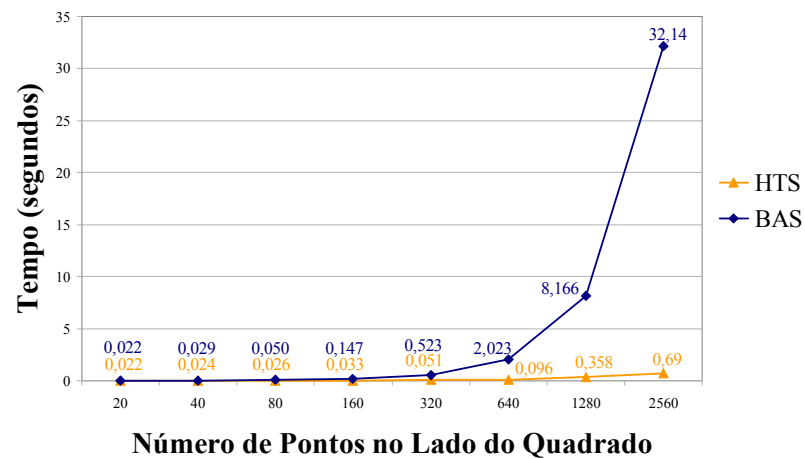


Figura 41. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTS (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens.

Tempo Utilizado - Extração de Características

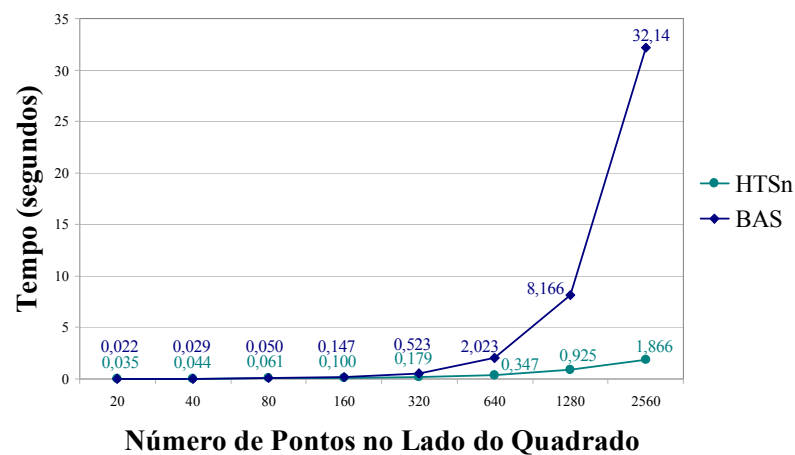


Figura 42. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSn (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens.

Tempo Utilizado - Extração de Características

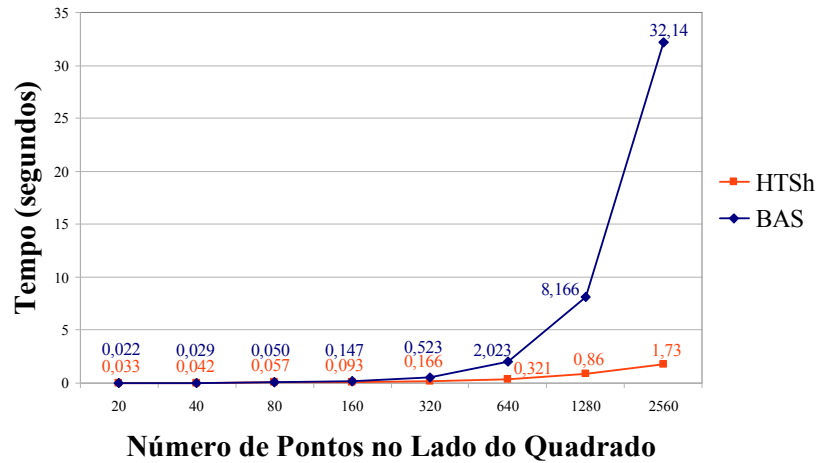


Figura 43. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSh (etapa de Extração de Características) obtidos ao se aumentar o tamanho do quadrado nas imagens.

Para testar o tempo de processamento da etapa de Reconhecimento do BAS (ARICA; VURAL, 2003) e dos métodos propostos (HTS, HTSn, HTSh) utilizou-se apenas a imagem do maior quadrado criada e a partir dela foram gerados vetores de características com dimensões cada vez maiores (cada método gerou seus vetores amostrando as funções unidimensionais do quadrado em quantidade crescente de posições): o primeiro vetor continha tamanho de 20 posições e o último 2560 posições (de um vetor para outro dobra-se o tamanho). Os tempos gastos por cada método na comparação de dois vetores de características com cada um dos comprimentos testados foram verificados. As Figuras 44, 45 e 46 ilustram os resultados. Novamente o HTS foi muito mais rápido que o BAS (ARICA; VURAL, 2003). Por sua vez, o HTSn apresentou tempos muito próximos do HTS visto que ambos utilizam a distância L_1 no casamento de vetores de características. Por fim, o HTSh apresentou tempos muito maiores que os do BAS (ARICA; VURAL, 2003) ao se analisar vetores de características pequenos, mas ao aumentar a dimensão destes vetores, o BAS (ARICA; VURAL, 2003) tornou-se muito mais lento devido à maior complexidade em seu algoritmo (conforme exposto no Capítulo 4).

Tempo Utilizado - Reconhecimento

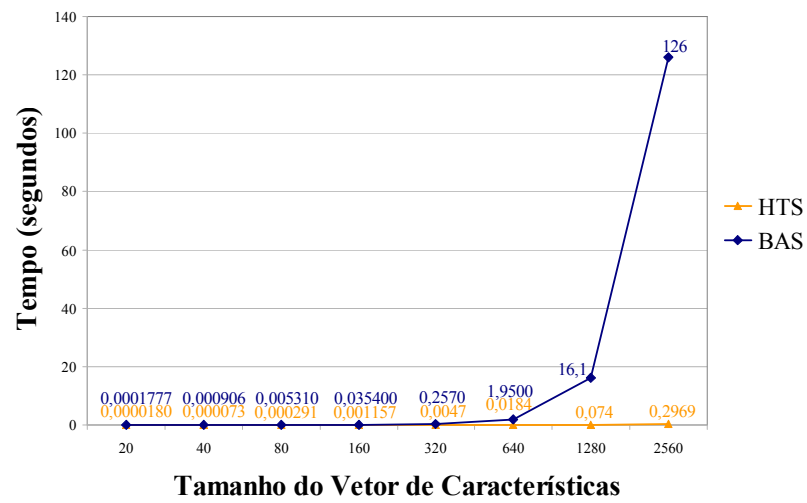


Figura 44. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTS (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.

Tempo Utilizado - Reconhecimento

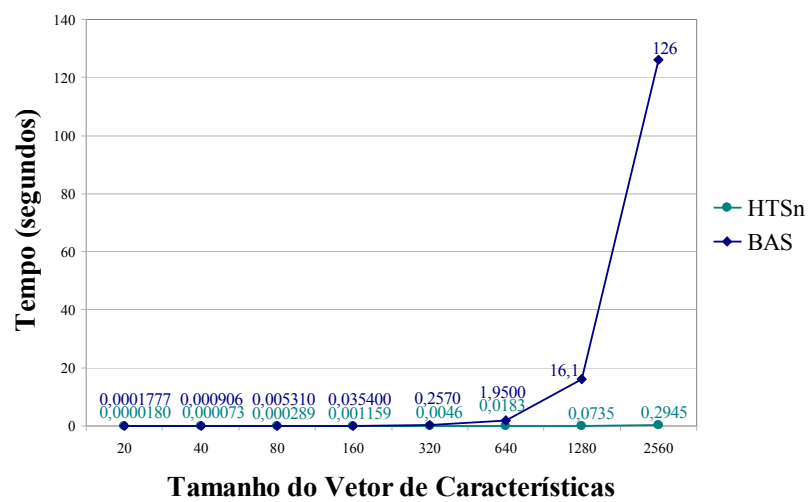


Figura 45. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSn (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.

Tempo Utilizado - Reconhecimento

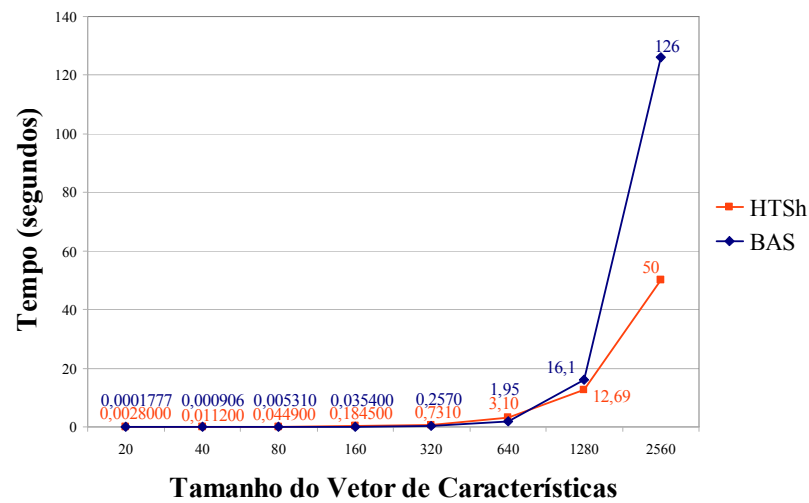


Figura 46. Tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTSh (etapa Reconhecimento) obtidos ao se aumentar o tamanho dos vetores de características sendo comparados.

Em síntese, conforme pode-se observar nos gráficos dos tempos de processamento do BAS (ARICA; VURAL, 2003) e do HTS, este último método foi muito mais rápido que aquele. Em relação à etapa de Extração de Características para quadrados com lados pequenos a diferença não é muito expressiva porém conforme se aumenta o tamanho dos lados do quadrado na imagens de teste ela cresce acentuadamente. Para o maior quadrado com 2560 pontos em cada lado o tempo tomado pelo BAS (ARICA; VURAL, 2003) foi mais de trinta vezes o valor gasto pelo HTS. Os tempos obtidos eram esperados visto que o BAS (ARICA; VURAL, 2003) apresenta complexidade quadrática em sua fase de Extração de Características e o HTS complexidade linear (conforme explicado no Capítulo 4).

Na fase de Reconhecimento mesmo ao se trabalhar com vetores de características pequenos (20 ou 40 posições) a diferença entre o tempo dos dois métodos já é grande: o BAS (ARICA; VURAL, 2003) se mostra muito mais lento que o HTS. Ao se aumentar o tamanho dos vetores de características os tempos do BAS (ARICA; VURAL, 2003) crescem vertiginosamente chegando a demorar aproximadamente 126 segundos no casamento dos vetores com 2560 posições enquanto que o HTS utilizou apenas 0,2969 segundos para realizar a mesma tarefa. Estes tempos também podem ser explicados pelo fato da complexidade do BAS (ARICA; VURAL, 2003) na etapa de Reconhecimento ser quadrática e do HTS linear. O BAS (ARICA; VURAL, 2003) emprega o método OCS (WANG; PAVLIDIS, 1990) no

casamento dos vetores o qual apesar de melhorar muito os resultados do método o deixa bastante lento.

Em relação aos gráficos comparativos do HTSn e do BAS (ARICA; VURAL, 2003), na etapa de Extração de características o HTSn foi um pouco mais lento que BAS (ARICA; VURAL, 2003) quando se trabalhou com quadrados de pequenas dimensões, mas pela complexidade quadrática do BAS (ARICA; VURAL, 2003), este método logo se tornou mais lento que o HTSn ao aumentar o tamanho dos quadrados. Na fase de Reconhecimento o HTSn foi muito mais rápido que o BAS (ARICA; VURAL, 2003), com tempos similares ao do HTS por ambos usarem a mesmo algoritmo (estruturas de dados, função de distância, etc.).

O HTSh também foi um pouco mais lento que o BAS (ARICA; VURAL, 2003) na etapa de Extração de Características quando se trabalhou com quadrados pequenos, mas ao aumentar os lados dos quadrados nas imagens esta situação se inverteu. Apesar de o HTSh ser muito lento na etapa de Reconhecimento como verificado nos experimentos anteriores nas bases Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER, 1999), quando o tamanho dos vetores de características cresce, o BAS (ARICA; VURAL, 2003) acaba ficando mais lento que o método proposto. Como mencionado, isto também se justifica pela complexidade linear do HTSh e quadrática do BAS (ARICA; VURAL, 2003).

5.4 Quarto Experimento - Base de Imagens de Sinus Frontais

Um último experimento utilizando uma base de imagens radiográficas (na qual não há transformações de rotação, translação e escala) foi realizado para comprovar se os bons resultados do descritor HTSh sobre as bases Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER, 1999), visto que o mesmo não é invariante à rotação e há imagens rotacionadas nestas duas bases, poderiam ser repetidos em uma aplicação real, na área de biometria.

A base de imagens utilizada neste último experimento corresponde a uma base de imagens radiográficas anteroposteriores do crânio (frente do crânio) de alguns indivíduos. As imagens foram disponibilizadas por professores da Universidade de Michigan, Estados Unidos, e são parte do projeto *Denver Growth Study*. A base contém 90 imagens radiográficas do crânio de 29 indivíduos maiores de 20 anos. Existem pelo menos duas radiografias por indivíduo. O descritor HTSh foi utilizado para reconhecer os indivíduos presentes nas imagens a partir das

formas de seus sinus frontais (reconhecimento biométrico). Como dito, apesar de ele não ser invariante à rotação, nesta base de imagens não há mudança em termos de escala, rotação e translação de uma imagem para outra.

Os sinus frontais (ou seios frontais) da face estão localizados no osso frontal do crânio e consistem em um par de cavidades loculadas e irregulares que se comunicam com a fossa nasal através do “infundibulum” (BROGDON, 1998; FALGUERA; MARANA, 2008). A Figura 47(a) exibe uma imagem radiográfica pertencente à base de imagens em questão na qual é possível visualizar os sinus frontais de um indivíduo. A mesma imagem com a borda dos sinus frontais demarcada manualmente com o auxílio de uma médica radiologista é exibida na Figura 47(b).

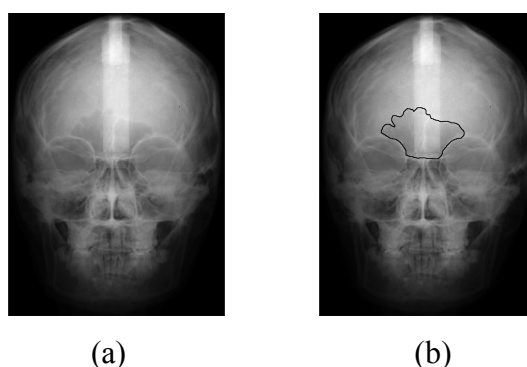


Figura 47. (a) Imagem radiográfica original do crânio de um indivíduo e (b) a mesma imagem com a borda dos sinus frontais demarcada manualmente com o auxílio de uma médica radiologista.

Além de apresentarem forma única para cada indivíduo (CAPUTO et al., 2012; FALGUERA; MARANA, 2008; SCHULLER, 1943), os sinus frontais se constituem como uma importante característica biométrica especialmente para sistemas de reconhecimento biométrico forense (identificação de cadáveres) uma vez que eles não se decompõem com a morte da pessoa. Caso o indivíduo não tenha danificado a região frontal do crânio, as imagens radiográficas podem ser obtidas mesmo após sua morte a fim de identificá-lo: compara-se a forma de seus sinus frontais depois de sua morte com as formas dos sinus frontais previamente cadastrados na base de dados.

Neste experimento, após segmentar todos os sinus frontais das imagens radiográficas, tarefa realizada manualmente e com a supervisão de uma médica radiologista, a fim de garantir a qualidade da segmentação, a abordagem HTSh bem como outros descritores de

formas da literatura foram testados na análise das formas dos sinus a fim de reconhecer as pessoas nas imagens.

Cada método foi testado separadamente e seu desempenho foi avaliado em termos de sua curva ROC (*Receiver Operating Characteristic*), de sua taxa EER (*Equal Error Rate*) e de seu valor “Top 1”, na CMC (*Cumulative Matching Curve*). Como na base de dados utilizada cada indivíduo apresenta poucas imagens e em número variável, optamos por utilizar as curvas ROC na comparação dos métodos de análise de formas ao invés das curvas *Precision-Recall* e *Multiscale Separability*.

As curvas ROC são muito utilizadas na avaliação de sistemas biométricos e permitem verificar o comportamento do sistema em termos de Falsa Aceitação (reconhecimento de um indivíduo impostor como sendo legítimo) ou Falsa Rejeição (não reconhecimento de um indivíduo legítimo). A Figura 48 mostra as curvas ROC obtidas pelos descritores testados. Quanto mais próxima dos eixos estiver a curva, melhor é o método de análise de formas. Pode-se perceber que o descritor HTSh obteve resultado equivalente ao do BAS (ARICA; VURAL, 2003) e melhor que de todos os outros descritores de forma avaliados.

Curvas ROC

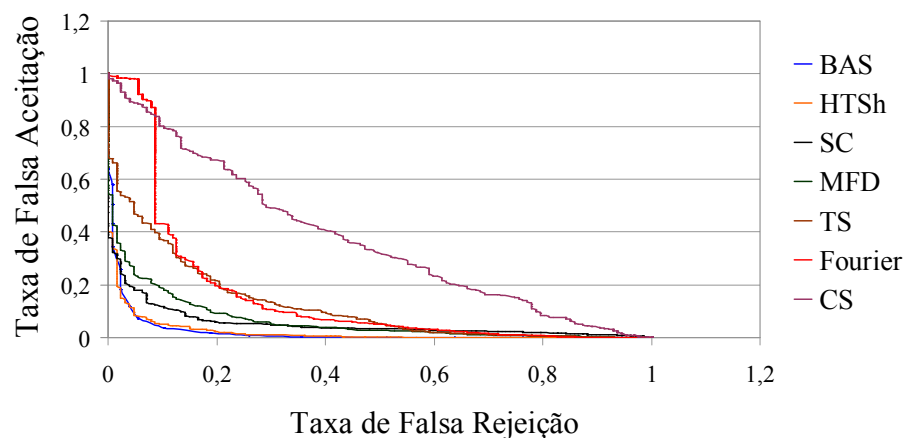


Figura 48. Curvas ROC dos descritores testados no reconhecimento das pessoas da base de sinus frontais: *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003), HTSh, *Shape Context* (SC) (BELONGIE; MALIK, 2000), *Multiscale Fractal Dimension* (MFD) (TORRES; FALCÃO; COSTA, 2004), *Tensor Scale* (TS) (MIRANDA; TORRES; FALCÃO, 2005), Fourier (ZHANG; LU, 2001) e *Contour Saliency* (CS) (TORRES; FALCÃO, 2007). O HTSh e o BAS (ARICA; VURAL, 2003) foram os melhores métodos com resultados muito próximos.

Como mencionado, além da curva ROC, para cada descritor foi encontrado o valor de EER e de “Top 1”. A taxa EER corresponde ao ponto da curva ROC em que a Taxa de Falsa

Aceitação e a Taxa de Falsa Rejeição são iguais (este valor é considerado como o EER da curva). Quanto menor o EER, melhor o descritor. Por sua vez o valor “Top 1” corresponde à quantidade de vezes, dentre os 90 reconhecimentos realizados (uma vez que existem 90 imagens na base usada), que a imagem mais similar à de consulta trazida pelo descritor era da mesma pessoa que estava sendo identificada. Quanto maior o “Top 1”, melhor o descritor. A taxa EER e o valor de “Top 1” são exibidos na Tabela 12. Percebe-se que o BAS (ARICA; VURAL, 2003) foi o melhor método e o HTSh foi o segundo melhor com resultados similares aos do BAS (ARICA; VURAL, 2003).

Tabela 12. Taxa EER e valor “Top 1” (aproximados) para os descritores testados - *Beam Angle Statistics* (BAS) (ARICA; VURAL, 2003), HTSh, *Shape Context* (SC) (BELONGIE; MALIK, 2000), *Multiscale Fractal Dimension* (MFD) (TORRES; FALCÃO; COSTA, 2004), *Tensor Scale* (TS) (MIRANDA; TORRES; FALCÃO, 2005), Fourier (ZHANG; LU, 2001) e *Contour Saliency* (CS) (TORRES; FALCÃO, 2007) - sobre a base de sinus frontais do projeto *Denver Growth Study*, da Universidade de Michigan. O BAS (ARICA; VURAL, 2003) e a abordagem HTSh foram os melhores métodos com resultados muito parecidos.

Descritor	EER (%)	“Top 1” (%)
BAS	6	93
HTSh	7	90
SC	11	30
MFD	13	59
Fourier	19	49
TS	20	59
CS	40	17

As matrizes de confusão dos dois melhores métodos segundo os resultados obtidos em termos de curvas ROC, EER e “Top 1” para a base do projeto *Denver Growth Study* também foram verificadas e são mostradas nas Tabelas 13 e 14. Percebe-se que são relativamente similares e ilustram o bom desempenho dos métodos. Como observação, para esta base de imagens, as matrizes têm 29 linhas por 29 colunas (porque existem imagens radiográficas de 29 pessoas na base de dados). Um método ideal deve apresentar valores não nulos apenas na diagonal principal de sua matriz de confusão, porém não se espera para esta base de imagens que a diagonal principal da matriz apresente o mesmo valor em toda sua extensão visto que cada classe (cada pessoa) possui número variável de imagens na base. Também vale dizer que na construção das matrizes de confusão, dado o limitado número de imagens por classe (por

Tabela 14. Matriz de confusão da abordagem HTSh para a base do projeto *Denver Growth Study*.

	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11	C 12	C 13	C 14	C 15	C 16	C 17	C 18	C 19	C 20	C 21	C 22	C 23	C 24	C 25	C 26	C 27	C 28	C 29
C 1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 2	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
C 3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 4	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 5	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 6	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 7	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 8	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 9	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 10	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 11	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 12	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
C 14	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
C 15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C 16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
C 17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0
C 18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0
C 19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
C 20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
C 21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0
C 22	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
C 23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
C 24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	1	0	0	0
C 25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0
C 26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0
C 27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
C 28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
C 29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Os resultados obtidos para a base de sinus frontais comprovam os bons desempenhos do descritor HTSh verificados de maneira “aproximada” (sem técnica de invariância à rotação) nos experimentos anteriores. As curvas ROC, as taxas EER e “Top 1” do descritor HTSh

foram próximas às dos BAS (ARICA; VURAL, 2003) e melhores que de todos os outros descritores. As matrizes de confusão destes dois métodos também são semelhantes.

5.5 Análise Geral dos Resultados

Em se tratando do HTS, os resultados obtidos sobre as bases Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER, 1999) comprovam seu bom desempenho. O método apresentou desempenho relativamente similar ao do BAS (ARICA; VURAL, 2003) (em um dos testes o seu desempenho foi bastante superior ao deste descritor) e melhor que os de todos os outros descritores da literatura testados. Além disto, o HTS é muito mais rápido que o BAS (ARICA; VURAL, 2003) devido ao seu algoritmo com menor complexidade (conforme descrito no Capítulo 4). Isto é comprovado pelos tempos tomados pelos dois métodos na análise das imagens nos experimentos. No experimento com a base MPEG-7 Part B (JEANNIN; BOBER, 1999), por exemplo, o HTS gastou apenas 296 segundos para obter os 1400 vetores de características a partir das imagens enquanto que o BAS (ARICA; VURAL, 2003) demorou 1157 segundos na mesma atividade.

Por sua vez, o descritor HTSn apresentou resultados ainda melhores que os do HTS (na maioria dos testes) e mais próximos aos do BAS (ARICA; VURAL, 2003), chegando a superar este método em mais casos que o HTS. O HTSn também superou todos os outros descritores de formas da literatura. Em relação ao tempo de processamento, por apresentar o mesmo algoritmo do HTS (com apenas uma pequena modificação no passo de construção dos histogramas dos pontos de borda), o HTSn também apresenta complexidade linear e foi muito mais rápido que o BAS (ARICA; VURAL, 2003) como provam os experimentos realizados.

O terceiro descritor proposto, o HTSh, apresentou desempenho ainda melhor que o do HTSn e do HTS. Seus resultados também foram melhores que os do BAS (ARICA; VURAL, 2003) em muitos casos. A única desvantagem deste método é que ele pode ser muito demorado na etapa de Reconhecimento uma vez que na comparação entre dois vetores de características são realizados muitos casamentos de histogramas através da distância de Canberra. Em relação ao HTS e ao HTSn, apesar de estes métodos apresentarem resultados um pouco inferiores, eles são muito mais rápidos na etapa de Reconhecimento que o HTSh (na de Extração de Características não há muita diferença de tempo de processamento entre os métodos). O BAS (ARICA; VURAL, 2003) também mostrou nos experimentos com as bases Kimia-216 (SEBASTIAN; KLEIN; KIMIA, 2004) e MPEG-7 Part B (JEANNIN; BOBER,

1999) ser mais rápido que o HTSh. No entanto, vale observar, que por apresentar complexidade linear em todas suas etapas como o HTS e o HTSn, o HTSh tende a se tornar mais rápido que o BAS (ARICA; VURAL, 2003) (descriptor com complexidade quadrática) caso o tamanho dos vetores de características aumente muito. Isto se comprova no gráfico de comparação dos tempos de processamento do HTSh e do BAS (ARICA; VURAL, 2003) no terceiro experimento (Figura 46).

Outro problema ainda com o HTSh corresponde à necessidade de incorporação de algum método a fim de se obter invariância à rotação. O HTS e do HTSn não apresentam esta necessidade por já serem invariantes à rotação: por extraírem medidas estatísticas dos histogramas dos pontos de borda a fim de gerar o vetor de características para uma forma (não usando os próprios histogramas para construir este vetor como faz o HTSh), mesmo com deslocamentos (“rotações”) nos valores dos histogramas dos pontos de borda causados pela rotação do objeto como mostrado na Figura 27, os valores estatísticos associados aos pontos de borda permanecem os mesmos (só há uma mudança na disposição do histograma e não em seus valores ao se rotacionar o objeto) e consequentemente o vetor de características da forma permanece inalterado.

Apesar destes dois aspectos negativos do HTSh (em relação ao tempo e à invariância) é importante ressaltar que nos casos como o apresentado no último experimento, isto é, quando se trabalha com sistemas biométricos forenses baseados em sinus frontais, não é necessário que o descriptor de formas seja invariante às transformações geométricas. Nos sistemas forenses também o tempo de processamento não é um fator crucial (obviamente dentro de limites razoáveis) uma vez que trabalha-se com identificação de pessoas já mortas, em geral em investigações criminais ou em acidentes. Alguns minutos a mais de processamento não representam grande problema. Desta forma o HTSh nestes casos, mesmo que um pouco mais lento e sem invariância à rotação, pode ser a escolha ideal pois apresenta melhores resultados que o HTS, o HTSn e, dependendo do caso, que o BAS (ARICA; VURAL, 2003).

6 Conclusão

Com a realização desta dissertação de mestrado pôde-se perceber que a análise automatizada de imagens pode ser uma poderosa ferramenta em muitas atividades como a busca de imagens baseada em conteúdo e o reconhecimento biométrico de pessoas. Dentre as características que podem ser analisadas em uma imagem encontra-se a forma de objetos e regiões da mesma. A forma em geral é a característica mais significativa e às vezes a única passível de análise.

Diversos descritores de formas encontram-se na literatura e cada um deles se baseia em um princípio a fim de conseguir caracterizar a forma e classificá-la posteriormente com base nas características extraídas. Dependendo de seu modo de funcionamento e das características com que o descritor trabalha ele pode obter melhores ou piores resultados bem como ser mais rápido ou mais lento. Neste contexto, novos métodos de análise de formas são propostos buscando-se obter maior eficácia, extraindo de maneira eficiente, isto é, com baixo custo computacional e consumo de recursos, as informações importantes das formas dos objetos nas imagens.

Neste trabalho é apresentado um novo descritor de formas denominado HTS (*Hough Transform Statistics*). Este descritor associa valores estatísticos aos pontos de borda do objeto sob análise com base em histogramas obtidos do espaço de Hough do mesmo. Através dos resultados obtidos pode-se concluir que o HTS é muito eficaz e eficiente e uma boa escolha caso se necessite analisar as formas planas de objetos em imagens digitais uma vez que apresenta resultados similares aos do BAS (ARICA; VURAL, 2003) e melhores que os dos demais métodos da literatura testados, além de ser muito rápido devido à complexidade linear de seu algoritmo. Outros métodos como o BAS (ARICA; VURAL, 2003) apresentam complexidade quadrática (e muitas vezes até superior) e se tornam muito lentos em especial quando se trabalha com imagens de grandes dimensões, com alta resolução, com grandes bases de imagens ou com grandes vetores de características, uma tendência nos dias atuais dado o barateamento dos sensores de captura e das mídias de armazenamento.

Duas outras versões do HTS com taxas de acerto ainda melhores, denominadas HTSn (*Hough Transform Statistics Neighborhood*) e HTSh (*Hough Transform Statistics Histograms*), também são propostas. Como o HTS, elas possuem algoritmo com complexidade linear e ótimas taxas de acerto. Desta forma, pode-se concluir que também são

muito eficazes e eficientes, superando os desempenhos dos outros descritores de formas da literatura avaliados, inclusive o do BAS (ARICA; VURAL, 2003) em boa parte dos testes.

Por fim, pode-se também concluir que os espaços de Hough são uma ótima opção para a caracterização de formas: além de armazenarem importantes informações sobre elas, suas representações são realizadas através de matrizes acumuladoras cujo acesso é simples, direto e rápido.

7 Trabalhos Futuros

Como trabalho futuro pretende-se propor novas versões do HTS. Uma ideia seria de extrair outros valores dos histogramas dos pontos de borda como energia, amplitude, ângulo de máximo e mínimo (global ou também local), etc. para construir as funções unidimensionais.

Em relação ao HTSn, pretende-se analisar novos sistemas de vizinhança (não só apenas as posições imediatamente adjacentes aos pontos das senóides nos espaços de Hough) na construção dos histogramas dos pontos de borda obviamente sem deixar o método muito mais lento.

Pretende-se também trabalhar na otimização do HTSh, em especial na fase de Reconhecimento, de forma a conseguir representar e comparar os histogramas de forma mais eficiente, isto é, utilizando se possível menor quantidade de memória e em menor período de tempo, sem perder muito o desempenho. Planeja-se ainda estudar métodos de registro de formas para selecionar o mais apropriado a fim de deixar o HTSh invariante à rotação (obviamente apenas nos casos em que isto se fizer necessário). Pretende-se investigar o próprio espaço de Hough a fim de extrair informações sobre o ângulo de rotação dos objetos nas imagens, uma ideia a ser estudada seria verificar, para cada ponto de borda de dois objetos sob análise, o ângulo θ de máximo global do seu respectivo histograma (ângulo correspondente ao maior pico no histograma do ponto). A partir dos ângulos obtidos para os pontos das duas bordas sendo comparadas, pode-se, por exemplo, encontrar a diferença entre estes ângulos para cada par possível de pontos (sempre um de uma borda e outro da outra) e a partir destas diferenças fazer um esquema de votação. O ângulo mais votado corresponderia ao ângulo de diferença de rotação entre os dois objetos sendo comparados nas imagens e a partir deste ângulo poder-se-ia alinhá-los em termos de rotação (rotacionando-se apropriadamente um dos objetos).

Uma vantagem da Transformada de Hough (DUDA; HART, 1972) consiste no fato de que, para se obter o novo espaço de Hough do objeto que foi rotacionado na imagem para se alinhar ao outro, não é necessário recalcular todas senóides de todos os pontos de borda do objeto, isto, é, reconstruir do início o espaço de Hough com base nas novas coordenadas dos pontos de borda na imagem. Através de “deslocamentos” nos valores da matriz acumuladora do objeto antes da rotação pode-se obter o espaço final para o mesmo.

Com a utilização do método de registro de formas não será mais necessário que no casamento de dois vetores de características tenha que se ficar rotacionando um dos vetores k

vezes (e comparando k vezes os dois vetores) como explicado anteriormente, onde k corresponde ao tamanho dos vetores. Além de associar o método ao HTSh, pretende-se ainda estudar a possibilidade de utilizá-lo junto ao HTS e HTSn para o mesmo fim.

8 Publicações Realizadas

SOUZA, G. B.; MARANA, A. N. Shape analysis on digital images: A Hough Transform based approach. In: WORKSHOP DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO (UNESP), 2., 2012, São José do Rio Preto. **Anais...** São José do Rio Preto: UNESP, 2012.

(Apresentação de pôster; Trabalho selecionado para publicação de versão estendida na revista Interciência & Sociedade).

SOUZA, G. B.; MARANA, A. N. Análise de formas planas em imagens digitais. **Interciência & Sociedade**, v. 2, p. 39-50, 2012.

SOUZA, G. B.; MARANA, A. N. HTS: A new shape descriptor based on Hough Transform. In: IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS, 2013, Pequim. **Anais...** Pequim: IEEE, 2013. p. 974-977.

(Apresentação oral; Artigo finalista do *Best Student Paper Contest*).

SOUZA, G. B.; MARANA, A. N. Um novo descritor de formas baseado na Transformada de Hough. In: WORKSHOP DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO DA UNESP, 3., 2013, Rio Claro. **Anais...** Rio Claro: UNESP, 2013.

(Apresentação oral; Trabalho selecionado para publicação de versão estendida na revista Interciência & Sociedade).

SOUZA, G. B.; MARANA, A. N. Frontal sinus recognition: A comparison among different shape descriptors. In: WORKSHOP DE VISÃO COMPUTACIONAL, 9., 2013, Rio de Janeiro. **Anais...** Rio de Janeiro: [s.n.], 2013.

(Apresentação oral).

Referências

ABBASI, S.; MOKHTARIAN, F.; KITTLER, J. Enhancing CSS-based shape retrieval for objects with shallow concavities. **Image and Vision Computing**, v. 18, n. 3, p. 199-211, 2000.

ANDALÓ, F. A.; MIRANDA, P. A. V.; TORRES, R. S.; FALCÃO, A. X. A new shape descriptor based on Tensor Scale. In: INTERNATIONAL SYMPOSIUM ON MATHEMATICAL MORPHOLOGY, 8., 2007, Rio de Janeiro. **Anais...** Rio de Janeiro: [s.n.], 2007. p. 141-152.

ANDALÓ, F. A.; MIRANDA, P. A. V.; TORRES, R. S.; FALCÃO, A. X. Shape feature extraction and description based on Tensor Scale. **Pattern Recognition**, v. 43, n. 1, p. 26-36, 2010.

ARICA, N.; VURAL, F. T. Y. BAS: A perceptual shape descriptor based on the beam angle statistics. **Pattern Recognition Letters**, v. 24, n. 9-10, p. 1627-1639, 2003.

BACKES, A. R.; BRUNO, O. M. **Implementação e comparação de métodos de estimativa da dimensão fractal e sua aplicação à análise e processamento de imagens**. 2006. Dissertação (Mestrado em Ciência da Computação e Matemática Computacional) - Universidade de São Paulo (USP), São Carlos.

BACKES, A. R.; BRUNO, O. M. **Estudos de métodos de análise de complexidade em imagens**. 2010. Tese (Doutorado em Ciência da Computação e Matemática Computacional) - Universidade de São Paulo (USP), São Carlos.

BACKES, A. R.; FLORINDO, J. B.; BRUNO, O. M. Shape analysis using fractal dimension: A curvature based approach. **Chaos – An Interdisciplinary Journal of Non Linear Sciences**, v. 22, n. 4, p. 043103, 2012.

BELONGIE, S.; MALIK, J. Matching with Shape Contexts. In: IEEE WORKSHOP ON CONTENT-BASED ACCESS OF IMAGE AND VIDEO LIBRARIES, 2000, Hilton Head Island. **Anais...** Estados Unidos: IEEE, 2000. p. 20-26.

BELONGIE, S.; MALIK, J.; PUZICHA, J. Shape matching and object recognition using Shape Context. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 4, p. 509-522, 2002.

BHATTACHARYYA, A. On a measure of divergence between two statistical populations defined by their probability distributions. **Bulletin of the Calcutta Mathematical Society**, v. 35, p. 99-109, 1943.

BROGDON, B. G. **Forensic radiology**. Estados Unidos: CRC Press LLC, 1998.

CANNY, J. F. A computational approach to edge detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 8, n. 6, p. 679-698, 1986.

CAPUTO, I. G. C.; PRADO, F. B.; JÚNIOR, E. D.; MUGLIA, V. F. Frontal sinuses in human identification: Literature review. **Brazilian Journal of Forensic Sciences**, v. 1, n. 2, p. 8-14, 2012.

COSTA, L. F.; JÚNIOR, R. M. C. **Shape analysis and classification – Theory and practice**. Estados Unidos: CRC Press, 2000.

COSTA, L. F.; CAMPOS, A. G.; MANOEL, E. T. M. An integrated approach to shape analysis: Results and perspectives. In: INTERNATIONAL CONFERENCE ON QUALITY CONTROL BY ARTIFICIAL VISION, 2001, [S.l.]. **Anais...** [S.l.: s.n.], 2001. p. 23-34.

DUDA, R. O.; HART, P. E. Use of the Hough Transformation to detect lines and curves in pictures. **Communications of ACM**, v. 15, n. 1, p. 1-15, 1972.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. 2 ed. Estados Unidos: Wiley-Interscience, 2000.

FALCÃO, A. X. **Análise de imagens**. 2008a. Notas de aula - Universidade Estadual de Campinas (UNICAMP), Campinas. Disponível em: <<http://www.ic.unicamp.br/~afalcao/mo445/aula6-old.pdf>>. Acesso em: 20 jun. 2013.

FALCÃO, A. X. **Análise de imagens**. 2008b. Notas de aula - Universidade Estadual de Campinas (UNICAMP), Campinas. Disponível em: <<http://www.ic.unicamp.br/~afalcao/mo445/aula1.pdf>>. Acesso em: 25 jun. 2013.

FALCÃO, A. X. **Fundamentos de imagem digital**. 2011. Notas de aula - Universidade Estadual de Campinas (UNICAMP), Campinas. Disponível em: <<http://www.ic.unicamp.br/~afalcao/mo443/aula1.pdf>>. Acesso em: 25 jun. 2013.

FALCONER, K. **Techniques in fractal geometry**. Inglaterra: John Wiley & Sons, 1997.

FALGUERA, J. R.; MARANA, A. N. **Reconhecimento automático de sinus frontais para identificação humana forense baseada na Transformada Imagem-Floresta e no Contexto da Forma**. 2008. Dissertação (Mestrado em Ciência da Computação) - Universidade Estadual Paulista (UNESP), São José do Rio Preto.

FLORINDO, J. B. BACKES, A. R.; CASTRO, M.; BRUNO, O. M. A comparative study on multiscale fractal dimension descriptors. **Pattern Recognition Letters**, v. 33, n. 6, p. 798-806, 2012.

HOUGH, P. V. C. **A method and means for recognizing complex patterns**. US 3069654, 1962.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **O que é análise estatística?** 1996. Disponível em: <<http://www.dpi.inpe.br/spring/teoria/estatist/estatist.htm>>. Acesso em: 25 jun. 2013.

JEANNIN, S.; BOBER, M. **Description of core experiments for MPEG-7 motion/shape**. 1999. Documento Técnico ISO/IECJTC1/SC29/WG11MPEG99/N2690.

KAUPPINEN, H.; SEPPANEN, T.; PIETIKAINEN, M. An experimental comparison of autoregressive and Fourier-based descriptors in shape classification. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 17, n. 2, p. 201-207, 1995.

LATECKI, L. J.; LAKÄMPER, R. Shape similarity measure based on correspondence of visual parts. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 10, p. 1185-1190, 2000.

MANDELBROT, B. B. **The fractal geometry of nature**. Estados Unidos: W. H. Freeman and Company, 1982.

MIRANDA, P. A. V.; TORRES, R. S.; FALCÃO, A. X. TSD: a shape descriptor based on a distribution of Tensor Scale local orientation. In: BRAZILIAN SYMPOSIUM ON COMPUTER GRAPHICS AND IMAGE PROCESSING, 18., 2005, Natal. **Anais...** Natal: [s.n.], 2005. p. 139-146.

MOKHTARIAN, F.; ABBASI, S. Shape similarity retrieval under affine transforms. **Pattern Recognition**, v. 1, n. 35, p. 31-41, 2002.

MURR, C.; MELO, F. L.; MUSIAL, J. E.; BARA, R. P.; SOUZA, R.; PRADO, S.; SERRA, C. P.; KARAS, E. W.; BILOTI, R. **Fractais: propriedades e construção**. 2005. Monografia (Graduação em Matemática) - Universidade Federal do Paraná (UFPR), Curitiba.

PAPADIMITRIOU, C.; STIEGLITZ, K. **Combinatorial optimization: Algorithms and complexity**. Estados Unidos: Prentice Hall, 1982.

SAHA, P. K. Tensor Scale: A local morphometric parameter with applications to computer vision and image processing. **Computer Vision and Image Understanding**, v. 99, n. 3, p. 384-413, 2005.

SAITO, T.; TORIWAKI, J. New algorithms for Euclidean Distance Transformation of an n-dimensional digitized picture with applications. **Pattern Recognition**, v. 27, n. 11, p. 1551-1565, 1994.

SCHULLER, A. A note on the identification of skulls by x-ray pictures of the frontal sinuses. **Medical Journal of Australia**, v. 1, p. 554-556, 1943.

SEBASTIAN, T.; KLEIN, P.; KIMIA, B. Recognition of shapes by editing their shock graphs. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 26, n. 5, p. 550-571, 2004.

TORRES, R. S.; FALCÃO, A. X.; COSTA, L. F. A graph-based approach for multiscale shape analysis. **Pattern Recognition**, v. 37, p. 1163-1174, 2004.

TORRES, R. S.; FALCÃO, A. X. Contour salience descriptors for effective image retrieval and analysis. **Image and Vision Computing**, v. 25, p. 3-13, 2007.

TRICOT, C. **Curves and fractal dimension**. Estados Unidos: Springer-Verlag, 1995.

WANG, Y. P.; PAVLIDIS, T. Optimal correspondence of string subsequences. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 11, p. 1080-1087, 1990.

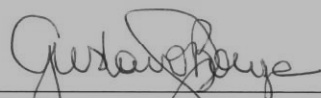
ZAHN, C. T.; ROSKIES, R. Z. Fourier descriptors for plane closed curves. **IEEE Transactions on Computers**, v. C-21, n. 3, p. 269-281, 1972.

ZHANG, D.; LU, G. A comparative study on shape retrieval using Fourier descriptors with different shape signatures. In: INTERNATIONAL CONFERENCE ON INTELLIGENT

MULTIMEDIA AND DISTANCE EDUCATION, 2001, [S.l.]. **Anais...** [S.l.: s.n.], 2001. p. 1-9.

Autorizo a reprodução xerográfica para fins de pesquisa.

São José do Rio Preto, 09/08/2013.



Gustavo Botelho de Souza