



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA DE TELEINFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

IÁLIS CAVALCANTE DE PAULA JÚNIOR

**DETECÇÃO DE CANTOS EM FORMAS BINÁRIAS PLANARES E APLICAÇÃO EM
RECUPERAÇÃO DE FORMAS**

FORTALEZA

2013

IÁLIS CAVALCANTE DE PAULA JÚNIOR

DETECÇÃO DE CANTOS EM FORMAS BINÁRIAS PLANARES E APLICAÇÃO EM
RECUPERAÇÃO DE FORMAS

Tese apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática da Universidade Federal do Ceará como parte dos requisitos para obtenção do Título de Doutor em Engenharia de Teleinformática.

FORTALEZA

2013

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca de Pós-Graduação em Engenharia - BPGE

-
- P347d Paula Júnior, Iális Cavalcante de.
Detecção de cantos em formas binárias planares e aplicação em recuperação de formas / Iális Cavalcante de Paula Júnior. – 2013
70 f. : il. , enc. ; 30 cm.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2013.
Área de concentração: Sinais e Sistemas
Orientação: Profa. Dra. Fátima Nelsizelma Sombra de Medeiros.
Coorientação: Prof. Dr. Francisco Nivandro Bezerra.
1. Teleinformática. 2. Imagens - Recuperação. I. Título.

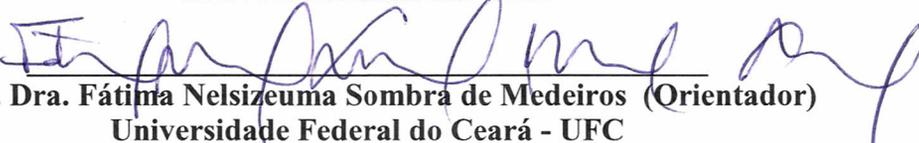
IÁLIS CAVALCANTE DE PAULA JÚNIOR

**DETECÇÃO DE CANTOS EM FORMAS BINÁRIAS PLANARES E APLICAÇÃO
EM RECUPERAÇÃO DE FORMAS**

Tese submetida à Coordenação do Curso de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Doutor em Engenharia de Teleinformática, área de concentração Sinais e Sistemas.

Aprovada em 06/06/2013.

BANCA EXAMINADORA



Prof. Dra. Fátima Nelsizeuma Sombra de Medeiros (Orientador)
Universidade Federal do Ceará - UFC



Prof. Dr. George André Pereira Thé
Universidade Federal do Ceará - UFC



Prof. Dra. Célia Aparecida Zorzo Barcelos
Universidade Federal de Uberlândia- UFU



Prof. Dra. Renata Maria Cardoso Rodrigues de Souza
Universidade Federal de Pernambuco- UFPE



Prof. Dr. Francisco Nivando Bezerra
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCe



Prof. Dr. Régis Cristiano Pinheiro Marques
Instituto Federal de Educação, Ciência e Tecnologia do Ceará - IFCe

À minha família.

AGRADECIMENTOS

Agradeço a Deus, por tudo que ele me ofereceu nesta vida, principalmente por eu sempre ter chegado com segurança nestas viagens entre Fortaleza e Sobral. Só assim pude aproveitar melhor os momentos com os homenageados que seguem.

À minha querida mãe e meus irmãos, por sempre estarem ao meu lado e apresentarem a mim o verdadeiro sentimento de amor e amizade que se deve ter com as pessoas.

À minha orientadora Profa. Fátima Sombra, pelo acompanhamento, atenção e dedicação no decorrer deste trabalho.

À minha amada noiva Anne Araújo, por sempre oferecer companhia, carinho e conforto nos momentos mais conturbados desta caminhada. Aos meus amigos de sempre, em especial a Ricardo Magno e Aline Araújo, por se fazerem presentes, multiplicando as alegrias, compartilhando comigo os sofrimentos e ajudando no meu amadurecimento.

Aos membros do Grupo de Processamento de Imagens (em ordem alfabética: Daniel Ferreira, Geraldo Ramalho, Gilson Gonçalves, Gladeston Leite, Janaina Cruz, Marcelo Souza, Mariana Bento, Otília Nogueira, Régis Marques e Rodrigo Veras) pelo bom convívio e amizade adquirida por todo este tempo.

Aos meus colegas de trabalho, professores e técnicos-administrativos, no Curso de Engenharia da Computação e *Campus* de Sobral da UFC, pela paciência e apoio para que eu dedicasse mais tempo nesta pesquisa.

E por fim, agradeço a meu pai, onde quer que esteja, pela presença constante em meu coração e em minha vida.

“Ou escreves algo que valha a pena ler, ou fazes
algo acerca do qual valha a pena escrever.”
(Benjamim Franklin)

RESUMO

Sistemas de recuperação de imagens baseada em conteúdo (do termo em inglês, *Content-Based Image Retrieval* - CBIR) que operam em bases com grande volume de dados constituem um problema relevante e desafiador em diferentes áreas do conhecimento, a saber, medicina, biologia, computação, catalogação em geral, etc. A indexação das imagens nestas bases pode ser realizada através de conteúdo visual como cor, textura e forma, sendo esta última característica a tradução visual dos objetos em uma cena. Tarefas automatizadas em inspeção industrial, registro de marca, biometria e descrição de imagens utilizam atributos da forma, como os cantos, na geração de descritores para representação, análise e reconhecimento da mesma, possibilitando ainda que estes descritores se adequem ao uso em sistemas de recuperação.

Esta tese aborda o problema da extração de características de formas planares binárias a partir de cantos, na proposta de um detector multiescala de cantos e sua aplicação em um sistema CBIR. O método de detecção de cantos proposto combina uma função de angulação do contorno da forma, a sua decomposição não decimada por transformada *wavelet* Chapéu Mexicano e a correlação espacial entre as escalas do sinal de angulação decomposto. A partir dos resultados de detecção de cantos, foi realizado um experimento com o sistema CBIR proposto, em que informações locais e globais extraídas dos cantos detectados da forma foram combinadas à técnica Deformação Espacial Dinâmica (do termo em inglês, *Dynamic Space Warping*), para fins de análise de similaridade formas de tamanhos distintos. Ainda com este experimento foi traçada uma estratégia de busca e ajuste dos parâmetros multiescala de detectores de cantos, segundo a maximização de uma função de custo.

Na avaliação de desempenho da metodologia proposta, e outras técnicas de detecção de cantos, foram empregadas as medidas *Precisão* e *Revocação*. Estas medidas atestaram o bom desempenho da metodologia proposta na detecção de cantos verdadeiros das formas, em uma base pública de imagens cujas verdades terrestres estão disponíveis. Para a avaliação do experimento de recuperação de imagens, utilizamos a taxa *Bull's eye* em três bases públicas. Os valores alcançados desta taxa mostraram que o experimento proposto foi bem sucedido na descrição e recuperação das formas, dentre os demais métodos avaliados.

Palavras-chave: Análise de Formas, Detecção de Cantos, Descrição de Formas, Recuperação de Imagens Baseada em Conteúdo, Similaridade.

ABSTRACT

Content-based image retrieval (CBIR) applied to large scale datasets is a relevant and challenging problem present in medicine, biology, computer science, general cataloging etc. Image indexing can be done using visual information such as colors, textures and shapes (the visual translation of objects in a scene). Automated tasks in industrial inspection, trademark registration, biostatistics and image description use shape attributes, e.g. corners, to generate descriptors for representation, analysis and recognition; allowing those descriptors to be used in image retrieval systems.

This thesis explores the problem of extracting information from binary planar shapes from corners, by proposing a multiscale corner detector and its use in a CBIR system. The proposed corner detection method combines an angulation function of the shape contour, its non-decimated decomposition using the Mexican hat wavelet and the spatial correlation among scales of the decomposed angulation signal. Using the information provided by our corner detection algorithm, we made experiments with the proposed CBIR. Local and global information extracted from the corners detected on shapes was used in a Dynamic Space Warping technique in order to analyze the similarity among shapes of different sizes. We also devised a strategy for searching and refining the multiscale parameters of the corner detector by maximizing an objective function.

For performance evaluation of the proposed methodology and other techniques, we employed the *Precision* and *Recall* measures. These measures proved the good performance of our method in detecting true corners on shapes from a public image dataset with ground truth information. To assess the image retrieval experiments, we used the Bull's eye score in three public databases. Our experiments showed our method performed well when compared to the existing approaches in the literature.

Keywords: Shape analysis, Corner detection, Shape description, Content-Based Image Retrieval, Similarity Analysis.

LISTA DE FIGURAS

1	A metodologia proposta para detecção de cantos.	12
2	Análise de correlação entre os coeficientes <i>wavelets</i> de diferentes escalas do contorno de forma.	14
3	(<i>Continuação</i>) Análise de correlação entre os coeficientes <i>wavelet</i> de diferentes escalas de um contorno de forma.	15
4	Efeitos da correlação entre os coeficientes <i>wavelet</i> de escalas consecutivas da forma apresentada na Figura 2(a).	16
5	Análise comparativa dos métodos para diferentes valores de desvio-padrão e números de cantos. (a),(b),(e),(f) formas binárias e (c),(d),(g),(h) suas respectivas análises comparativas.	23
6	Cantos detectados (círculos) a partir de uma forma com 1328 pontos de contorno e 12 cantos verdadeiros.	25
7	Cantos detectados (círculos) a partir de uma forma com 1193 pontos de contorno e 03 cantos verdadeiros.	26
8	Ocorrências dos (a) valores de <i>Precisão</i> e (b) <i>Revocação</i> na base de imagens.	27
9	As principais componentes de todos os sistemas de recuperação de imagens baseada em conteúdo. Adaptado de (MULLER <i>et al.</i> , 2004).	32
10	Principais técnicas para recuperação de imagens. Adaptado de Silva <i>et al.</i> (2011) e de Smeulders <i>et al.</i> (2000).	34
11	Busca por similaridade. Adaptado de Silva <i>et al.</i> (2011) e de Smeulders <i>et al.</i> (2000).	35
12	Arquitetura de um sistema de recuperação de imagens por conteúdo. Adaptado de Torres e Falcão (2006) e de Smeulders <i>et al.</i> (2000).	36
13	Metodologia do sistema CBIR proposto.	38
14	Exemplo de aplicação da função de curvatura.	39
15	Definição do descritor extraído de uma forma binária. A partir da (a) forma original e seu centróide calculado, detectam-se (b) os seus cantos e demarca-se um ponto arbitrário no contorno da forma. Com essas informações destacadas, determina-se o (c) descritor da forma segundo os cantos detectados.	41

16	Exemplo de aplicação da técnica <i>DSW</i> sobre uma matriz de diferenças r_{ij} : (1) nova matriz M_{ij} contendo a primeira iteração da Equação 4.7; (2) M_{ij} atualizada com a segunda e terceira iteração da recursividade de <i>DSW</i> e (3) M_{ij} completa após o processamento da quarta iteração de Δ e identificação do custo mínimo no valor de $M(n, m)$ (destacado em vermelho).	45
17	Conjuntos de formas geradas com a inserção/remoção de uma região com canto côncavo ou convexo. Adaptado de Barenholtz <i>et al.</i> (2003).	46
18	A relação entre a distância <i>DSW</i> e a seleção do ponto de partida para duas formas obtida a partir do (a) algoritmo <i>DSW</i> clássico, e (b) a busca grosseira e (c) o estágio refinado do algoritmo <i>HopDSW</i> . Adaptado de Alajlan (2011). . . .	49
19	Curvas <i>Precisão-Revocação</i> - Base de dados Tari 1000. Resultados da recuperação de imagens das três abordagens propostas.	53
20	Curvas <i>Precisão-Revocação</i> - Base de dados MPEG-7 Part B. Resultados da recuperação de imagens das três abordagens propostas.	53
21	Curvas <i>Precisão-Revocação</i> - Base de dados Tari 1000. Resultados da recuperação de imagens entre o método proposto e outras abordagens presentes na literatura.	55
22	Curvas <i>Precisão-Revocação</i> - Base de dados MPEG-7 Part B. Resultados da recuperação de imagens entre o método proposto e outras abordagens presentes na literatura.	55

LISTA DE TABELAS

1	Complexidade temporal dos detectores de cantos.	18
2	Avaliação quantitativa entre os métodos.	24
3	Definição do descritor segundo a forma Figura 14(a). Campo (1) corresponde à posição relativa angular e (2) indica o valor de curvatura no canto.	40
4	Complexidade temporal das técnicas de comparação de descritores de formas. .	50
5	Taxas de <i>Bull's eye</i> nas bases de dados Tari.	56
6	Taxas de <i>Bull's eye</i> na base de dados MPEG-7 Parte B.	56
7	Amostras de formas dentre as bases de dados adotadas.	57
8	Análise comparativa entre sistemas de recuperação avaliados.	57

LISTA DE SIGLAS

<i>CBIR</i>	Recuperação de Imagens baseado em Conteúdo (<i>Content-Based Image Retrieval</i>)
<i>HCP</i>	Pontos de Alta Curvatura (<i>High Curvature Points</i>)
<i>CSS</i>	Curvatura Espaço-Escala (<i>Curvature Space-Scale</i>)
<i>DCSS</i>	Curvatura Direta Espaço-Escala (<i>Direct Curvature Space-Scale</i>)
<i>CR</i>	Proporção de Canto (<i>Corner Ratio</i>)
<i>DoG</i>	Diferença de Gaussianas (<i>Difference of Gaussians</i>)
<i>SVM</i>	Máquinas de Vetor de Suporte (<i>Support Vector Machines</i>)
<i>TAR</i>	Representação de Área-Triângulo (<i>Triangle-Area Representation</i>)
<i>MTAR</i>	Representação de Área-Triângulo Multiescala (<i>Multiscale Triangle-Area Representation</i>)
<i>DP</i>	Programação Dinâmica (<i>Dynamic Programming</i>)
<i>DSW</i>	Deformação Espacial Dinâmica (<i>Dynamic Space Warping</i>)
<i>AR</i>	relação de aspecto entre formas
<i>E</i>	excentricidade entre formas
<i>S</i>	solidez entre formas
<i>HopDSW</i>	Deformação Espacial Dinâmica com passo de salto (<i>Hopping step Dynamic Space Warping</i>)

LISTA DE SÍMBOLOS

τ	parâmetro de suavização da curvatura
Γ	curva do contorno parametrizado
x	coordenada x do ponto do contorno - domínio do espaço
y	coordenada y do ponto do contorno - domínio do espaço
X	coordenada x do ponto do contorno - domínio da frequência
Y	coordenada y do ponto do contorno - domínio da frequência
$g(t, \sigma)$	função gaussiana de largura σ no ponto t do contorno
k	função de curvatura no domínio do espaço
K	função de curvatura no domínio da frequência
$\phi(t)$	sinal de angulação na posição t
q	parâmetro de suavização no sinal de angulação
$\psi(t)$	função chapéu mexicano normalizada na posição t
$c_s(t)$	coeficientes <i>wavelet</i> de aproximação na escala s na posição t
$w_s(t)$	coeficientes <i>wavelet</i> de detalhes na escala s na posição t
$corr_s(t)$	correlação espacial entre as escalas adjacentes s e $s + 1$
$corr_s(t)$	correlação espacial entre as escalas adjacentes s e $s + 1$
$Q(p_1, p_2)$	função de custo para métricas de detecção de cantos
h	limiar de remoção de cantos falsos
\Re	conjunto dos números reais
$(s_{1,t}, s_{2,t})$	posição relativa angular e valor de curvatura em t do contorno

c_m	centróide da forma
P_a	ponto arbitrário do contorno da forma
F_A	forma A
F_B	forma B
A	descritor da forma A
B	descritor da forma B
$d_s(A, B)$	distância entre os descritores A e B
ρ	número de deslocamentos no descritor de forma
φ	função de correspondência no cálculo de distância entre descritores de dimensões distintas
r_{ij}	matriz de diferenças entre descritores
M_{ij}	matriz de memorização das diferenças entre descritores
$d_g(F_A, F_B)$	distância global entre as formas A e B
$c_f(A, B)$	complexidade entre os descritores A e B
n_{conc}	número de cantos côncavos detectados
n_{all}	quantidade total de cantos identificados
h_s	passo de salto no método <i>HopDSW</i>

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Motivação e Objetivos	3
1.2	Contribuições	4
1.3	Organização da Tese	4
	PARTE I - DETECÇÃO DE CANTOS EM IMAGENS BINÁRIAS	6
2	ABORDAGEM MULTIESCALA PARA DETECÇÃO DE CANTOS	7
2.1	Métodos Avaliados	9
2.1.1	Técnica de Pedrosa e Barcelos	10
2.1.2	Técnica de Curvatura Espaço-Escala	10
2.1.3	Técnica de Lee <i>et al.</i>	11
2.2	Detector de Cantos Proposto	12
2.3	Complexidade Computacional Temporal	17
2.4	Considerações Finais	19
3	RESULTADOS EXPERIMENTAIS DE DETECÇÃO DE CANTOS	20
3.1	Metodologia de Avaliação de Desempenho	20
3.2	Experimento I - Ajustes de Parâmetros	21
3.3	Experimento II - Análise dos Resultados	24
3.4	Considerações Finais	28
	PARTE II - RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO	30
4	RECUPERAÇÃO DE FORMAS ATRAVÉS DE CANTOS	31
4.1	Propriedades Visuais	32
4.1.1	Características Locais e Globais	32
4.1.2	Características de Forma	33
4.2	Técnicas de Comparação de Formas	33
4.3	Definição do Sistema CBIR Adotado	36

4.4	Descrição de Formas	37
4.4.1	Sinal de Curvatura	37
4.4.2	Descritor de Formas por Cantos	39
4.5	Recuperação de Formas	40
4.5.1	Deformação Espacial Dinâmica	42
4.5.2	Medida de Distância	44
4.5.3	Aperfeiçoamento da Técnica <i>DSW</i>	46
4.6	Complexidade Computacional Temporal	50
4.7	Considerações Finais	51
5	RESULTADOS EXPERIMENTAIS DE RECUPERAÇÃO DE FORMAS	52
5.1	Análise da Recuperação de Formas	52
5.2	Considerações Finais	59
6	CONCLUSÕES	60
6.1	Produção Bibliográfica	61
6.2	Perspectivas de Trabalhos	62
	REFERÊNCIAS	63

1 INTRODUÇÃO

Com o advento e o crescimento da internet e a popularização dos dispositivos para captura de imagens como câmeras digitais e *scanners* nos últimos anos, há uma maior necessidade de armazenar e recuperar grandes coleções de imagens (SILVA *et al.*, 2011; QUELLEC *et al.*, 2010). Os usuários necessitam cada vez mais de ferramentas eficientes para navegar, pesquisar, organizar e recuperar essas informações oriundas de imagens em domínios distintos. Entre estes domínios destacam-se áreas como sensoriamento remoto, medicina, jornalismo, meteorologia, publicidade, moda, indústria, ciências ambientais e outras que também necessitam de sistemas atuantes na recuperação de imagens.

No entanto, não há ainda disponível um sistema computacional capaz de realizar buscas de imagens em grandes bases com eficiência próxima a do sistema humano (SILVA *et al.*, 2010). O ser humano possui a visão como um dos sentidos mais aprimorados, dado que o cérebro dispõe de uma grande capacidade de interpretação e generalização das informações visuais. Qualquer ser humano é capaz de categorizar, de maneira rápida e simples, uma imagem ou parte dela, além de identificar faces ou objetos previamente conhecidos (SILVA *et al.*, 2011).

Desenvolver um sistema computacional de busca de imagens que apresente um desempenho similar ao sistema visual humano continua sendo um desafio, porque não existe um consenso (CHURCHLAND *et al.*, 1994) ou um entendimento completo a respeito de como o cérebro processa as imagens. Nos campos da psicologia e da neurociência, existem artigos que reportam simulações computacionais de determinadas áreas do cérebro envolvidas no processamento da informação visual (SHASTRI, 2001; PALMERI; GAUTHIER, 2004; RANGANATH, 2006; KAY *et al.*, 2008). Outros trabalhos reproduzem ou explicam fenômenos particulares através de sistemas computacionais (DRAPER *et al.*, 2004; FELSEN; DAN, 2005; DAVIES *et al.*, 2009). Esses artigos utilizam conceitos psicofísicos e neurocientíficos para simular computacionalmente regiões do cérebro (giro para-hipocampal, por exemplo) e testar a validade de modelos teóricos sobre alguma funcionalidade cerebral. Alguns artigos visam contribuir na área de visão computacional (GORDER, 2008; SERRE; POGGIO, 2010) como o trabalho seminal de Marr (1982) (MARR, 1982) que inspirou conceitos da área de neurociência e de neuroanatomia.

Devido ao tipo de informação que pode ser obtida através de uma imagem, as bases de

dados de imagens não podem ser tratadas em um estilo similar a outros tipos de bibliotecas digitais. A dependência contextual apresentada pelas imagens, e a natureza complexa das imagens bidimensionais (que divergem das formas geométricas mais simples) tornam a questão da representação dos dados mais difícil para bases de dados de imagens. Uma representação de uma imagem invariante às transformações geométricas (rotação, translação e escala) é ainda uma questão aberta na pesquisa. Dois paradigmas principais são definidos nos sistemas computacionais para recuperação de imagens: baseado em texto e em conteúdo. As abordagens para recuperação baseadas em texto começaram a ser utilizadas na década de 70. Nesses sistemas, o processo de recuperação compara os termos de uma consulta textual, definida por um usuário, com as anotações associadas às imagens por palavras-chave e, a partir dessa comparação, retorna um conjunto de imagens (SILVA *et al.*, 2010). Esta abordagem dispõe duas desvantagens principais que são: a necessidade do trabalho humano na realização de anotações e a imprecisão dessas mesmas devido à ambiguidade e imprecisão das palavras utilizadas nas anotações definidas pelos usuários. O conteúdo de uma imagem normalmente é absorvido de maneira mais intuitiva pelo usuário do que um conjunto de palavras é possível descrever (DATTA *et al.*, 2008; WANG *et al.*, 2010).

Sistemas de recuperação de imagens baseados em conteúdo (do termo em inglês *Content-Based Image Retrieval*, CBIR) (SMEULDERS *et al.*, 2000; LIU *et al.*, 2007; DATTA *et al.*, 2008; VASCONCELOS; KUNT, 2001) têm sido estudados e propostos na tentativa de superar desvantagens dos sistemas de recuperação de imagens baseados em texto. Na abordagem CBIR, as imagens são indexadas por características relacionadas ao seu conteúdo visual, tal como cor, textura e forma, tornando desnecessária a anotação manual (SILVA *et al.*, 2010). Os métodos de extração de características de imagens utilizam descritores, que representam o seu conjunto de pixels através de um vetor de características, que expressa componentes da imagem relacionada. O processo de busca consiste basicamente em, dado um padrão de consulta (normalmente uma imagem), calcular a similaridade da mesma em relação aos demais objetos armazenados em uma base de imagens e recuperar as mais similares. Esta similaridade é obtida comparando-se os descritores do padrão de consulta com os descritores associados às imagens da base de interesse. De resultados apresentados pela neurociência, sabe-se que a chave para o reconhecimento de objetos é o sistema ser capaz de discriminar objetos sendo tolerante a transformações de rotação, escala, translação, iluminação, mudança de ponto de vista e organização (DATTA *et al.*, 2008; VALLE; CORD, 2009). Por isso, os descritores que buscam traduzir as propriedades visuais utilizadas para descrição das imagens devem ser invariantes a essas transformações.

Ao longo das últimas décadas, o estudo dos sistemas CBIR tem permitido o desenvolvimento de novos métodos com resultados mais eficientes no processo de recuperação de imagens. Existem algumas abordagens assumidas por esses sistemas em relação ao objetivo

do usuário na recuperação de imagens por conteúdo. Em uma dessas abordagens, o usuário visa recuperar uma imagem específica. Neste caso, o usuário sabe exatamente como e qual é a imagem que ele está buscando e a pesquisa termina, quando esta imagem é encontrada. Outra análise da recuperação de imagens emprega a busca de todas as imagens que estão a uma certa distância da imagem de consulta, relacionando-as através de suas características. A abordagem mais comum é a recuperação por similaridade, na qual é especificado o número de imagens que devem ser retornadas as quais são mais próximas do padrão de consulta (SILVA *et al.*, 2010, 2011).

1.1 Motivação e Objetivos

Existe hoje na comunidade científica, um interesse crescente no desenvolvimento de sistemas de recuperação de imagens baseados em suas características. Técnicas de visão computacional e processamento de imagens são usadas para extrair automaticamente estas características em um processo de geração de descritores. Para o reconhecimento das imagens semelhantes à imagem submetida na solicitação do sistema, são requeridos descritores dessas imagens os quais são avaliados segundo medidas de similaridade. Estas medidas também são fontes de inúmeras pesquisas em sistemas de recuperação de imagens. Ao longo das últimas décadas, diversos produtos comerciais e protótipos experimentais para recuperação de imagens por conteúdo foram desenvolvidos, como IBM QBIC (FALOUTSOS *et al.*, 1994), MIT Photobook (PENTLAND *et al.*, 1996), Virage (GUPTA; JAIN, 1997), Columbia VisualSEEK e Web-SEEK (SMITH; CHANG, 1996), Berkeley Chabot (OGLE; STONEBRAKER, 1995), UIUC MARS (MEHROTRA *et al.*, 1997), Stanford WBIIS (WANG *et al.*, 1998), PicHunter (COX *et al.*, 2000), SIMPLIcity (WANG *et al.*, 2001), PicToSeek (GEVERS; SMEULDERS, 2000), Blobworld (CARSON *et al.*, 2002), CIRES (IQBAL; AGGARWAL, 2002), Fids (LI *et al.*, 2005) e Windsurf (BARTOLINI *et al.*, 2010). É importante que as técnicas desenvolvidas para este fim sejam eficientes na descrição da forma mesmo que esta esteja sob o efeito de transformações geométricas (rotação, translação, escala, etc.). A redução de informação e o menor custo de processamento na recuperação das formas tem particular importância, uma vez que estes são grandes desafios na definição dos sistemas CBIR (KHAN; AHMAD, 2007). A adequação dos detectores de cantos à essa probabilidade tem favorecido o surgimento de novos estudos e linhas de pesquisa nesse tema.

Esta tese tem como objetivo propor uma nova abordagem para recuperação de formas que permita ao usuário buscar e encontrar imagens de seu interesse seguindo um padrão de escolha. Esta abordagem se baseia em características morfométricas, mais especificamente os cantos de uma forma. Em (ALAJLAN *et al.*, 2007), os autores assumem que uma forma possui uma descrição mais eficiente, quando esta utiliza características locais e globais da mesma. Esta

tese propõe que um sistema CBIR que associe estas características na descrição do padrão de consulta e de todas as presentes nas bases de imagens avaliadas.

Neste trabalho, os experimentos foram realizados em recuperação por similaridade para encontrar um conjunto de imagens semelhantes ao padrão apresentado como consulta. Esse conjunto de imagens se assemelhará à imagem de consulta com relação a pontos relevantes e o vínculo de concavidade/convexidade entre os referidos pontos.

O objetivo desta tese encontra-se na caracterização eficiente de formas a partir de seus cantos e a sua aplicação em um sistema de recuperação de imagens. Será proposto aqui um método multiescala de detecção de cantos em formas binárias cujos parâmetros serão ajustados à base de dados e que aumentem a eficácia na localização dos pontos relevantes da forma. Esse método proporciona uma nova descrição de formas baseada nas características locais e globais das mesmas e que resultará em uma maior discriminação das formas entre os dados avaliados da base. Por último, este trabalho analisa a otimização da etapa de comparação entre os descritores de formas com relação à complexidade temporal, permitindo que esse processo se realize entre vetores de descritores de formas com dimensões diferentes. Essa análise servirá, sobretudo, na implementação de técnicas utilizadas em sistemas CBIR disponíveis na literatura e que são aplicadas em formas, com o objetivo de avaliar comparativamente a técnica proposta.

1.2 Contribuições

As contribuições desta tese são relacionadas como segue:

- desenvolvimento de uma proposta de detecção de cantos segundo uma metodologia de busca de parâmetros que minimiza a detecção de cantos falsos e perdidos;
- desenvolvimento de uma nova descrição de forma que considera a proporção de cantos côncavos como característica relevante para a mesma;
- proposta de um sistema CBIR que utiliza os cantos detectados em uma forma;
- redução de complexidade temporal na etapa de comparações entre formas sob efeito de transformações geométricas.

1.3 Organização da Tese

Os capítulos desta tese seguem a estrutura que segue:

Parte I: Detecção de Cantos em Imagens Binárias

Capítulo 2: proposta de uma metodologia de detecção de cantos em formas, que inclui uma abordagem *wavelet* não-ortogonal e multiescala, a partir de uma representação unidimensional da forma e correlação espacial entre escalas. Ademais, é proposta uma nova metodologia de avaliação de desempenho dos detectores de cantos;

Capítulo 3: apresentação dos resultados obtidos na detecção de cantos e avaliação dos resultados por meio das medidas apresentadas;

Parte II: Recuperação de Imagens Baseada em Conteúdo

Capítulo 4: são descritos conceitos básicos de sistemas de recuperação de formas, com uma discussão sobre a abordagem direcionada a cantos e uma revisão dos métodos relacionados a essa temática; e proposta de uma abordagem para descrição e recuperação de formas através dos cantos, com base na detecção dos mesmos em determinadas bases de imagens;

Capítulo 5: apresentação dos resultados obtidos na recuperação das formas; e avaliação dos resultados por meio das medidas apresentadas;

Capítulo 6: discussão dos resultados alcançados e suas contribuições, bem como perspectivas de trabalhos futuros.

PARTE I - DETECÇÃO DE CANTOS EM IMAGENS BINÁRIAS

2 ABORDAGEM MULTIESCALA PARA DETECÇÃO DE CANTOS

De acordo com Mokhtarian e Mohanna (2006), os cantos de uma forma ou um objeto são características importantes em imagens e os pontos mais relevantes em uma forma. Também conhecidos como pontos de alta curvatura, do termo inglês *High Curvature Points (HCP)*, os cantos são úteis para a descrição de forma e, além disso, representação de um objeto de maneira compacta, invariante à rotação e translação (PEDROSA *et al.*, 2011a; TORRES; FALCÃO, 2007).

Diversos trabalhos de análise e reconhecimento de formas utilizam algoritmos de detecção de cantos e pontos dominantes, que podem ser utilizados para representar objetos (COSTA; CESAR, 2009; PAULA JR. *et al.*, 2006; ZHANG *et al.*, 2011). Aplicações distintas dependem das informações dos cantos nos quais são incluídas análise da cena, aproximação poligonal, correspondência de características, navegação de robôs, similaridade de formas e rastreamento de objetos, entre muitos outros.

Um modelo de detecção de cantos registra que os pontos de interesse correspondem aos pontos de altas curvaturas de um determinado objeto. Os pontos de alta curvatura de formas podem ser associados a cantos, o que indica a presença de pontos de referências a partir das quais é possível reconstruir a forma inicial. Um grande número de detectores de cantos foi proposto na literatura (AWRANGJEB; LU, 2008; GAO *et al.*, 2007; HUA; LIAO, 2000; MOKHTARIAN; SUOMELA, 1998). Os métodos de detecção de cantos podem ser divididos em dois grupos que se baseiam em intensidade (AWRANGJEB; LU, 2008; GAO *et al.*, 2007; ROSTEN *et al.*, 2010; WANG; LEE, 1998) e no contorno da forma (GAO *et al.*, 2007; LEE *et al.*, 1995; ZHANG *et al.*, 2010, 2009). O primeiro grupo inclui algoritmos que indicam a presença de cantos a partir da intensidade do nível de cinza da imagem. O segundo grupo é composto dos métodos que recuperam o contorno do objeto da imagem e identificam os pontos com valor máximo de curvatura ou pontos de inflexão na representação deste contorno (MOKHTARIAN; MOHANNA, 2006). O método proposto nesta tese pertence a este último grupo, uma vez que a base de dados é composto de formas binárias. Detectores de cantos também podem ser classificados em detectores monoescala (HARRIS; STEPHENS,

1998) e detectores multiescala (AWRANGJEB; LU, 2008; LEE *et al.*, 1995; PAULA JR. *et al.*, 2006). Detectores monoescala funcionam bem se as características oriundas dessa imagem possuem dimensões similares; caso contrário, características de escalas distintas não terão todos os seus detalhes detectados. Os detectores de canto multiescala têm suporte na teoria espaço-escala clássica (MOKHTARIAN; MACKWORTH, 1986) e foram propostos a fim de melhorar a eficácia nas diferentes definições das funções de representação da forma em escalas distintas. Rattarangsi e Chin (RATTARANGSI; CHIN, 1992), Mokhtarian e Suomela (MOKHTARIAN; SUOMELA, 1998), e Mohanna e Mokhtarian (MOKHTARIAN; MOHANNA, 2006) introduziram a técnica de curvatura espaço-escala, do termo em inglês *Curvature Space-Scale (CSS)*, para detectar cantos. Extensões desta mesma técnica têm sido propostas como a curvatura espaço-escala direta, do termo em inglês *Direct Curvature Space-Scale (DCSS)*, e uma versão híbrida para reduzir a sensibilidade do *DCSS* ao ruído (ZHONG; LIAO, 2007). O *CSS/DCSS* híbrido pode ajudar a traçar um canto da escala mais grosseira até a mais refinada para que este seja localizado da maneira mais precisa possível (ZHONG; LIAO, 2007). Além disso, o *DCSS* apresenta um menor custo computacional, quando comparado ao *CSS*.

O contorno da forma pode ser representado em termos da direção do gradiente da borda, tal como definido em (LIU; SRINATH, 1990; ROSENFELD; WESZKA, 1975). Em (LEE *et al.*, 1995), os autores introduziram uma estrutura para a detecção de cantos baseada em *wavelets* usando diferentes escalas, e mais tarde outros artigos estenderam o escopo desse trabalho (HUA; LIAO, 2000; ROSTEN *et al.*, 2010; WANG; LEE, 1998). Lee *et al.* relataram uma análise segundo o limiar Proporção de Canto, do termo em inglês *Corner Ratio (CR)*, que é definido de acordo com o conjunto de imagens. Os autores sugerem valores distintos de parâmetros de acordo com a imagem processada. Em geral, a definição dos parâmetros a serem aplicados em detectores de canto não é uma tarefa simples, já que esses valores de parâmetros normalmente não são ajustáveis para todas as imagens de uma base de dados (PEDROSA; BARCELOS, 2010). Em relação à decomposição *wavelet*, o algoritmo introduzido em (LEE *et al.*, 1995) utiliza a transformada *wavelet modulus maxima* que geralmente adota uma *wavelet*-mãe gaussiana ou chapéu mexicano.

Recentemente, Pedrosa e Barcelos (PEDROSA; BARCELOS, 2010) desenvolveram um detector de cantos multiescala robusto ao ruído e simples em termos de complexidade computacional. Segundo esta abordagem, o único parâmetro a ser definido pelo usuário é o número de iterações para a filtragem anisotrópica, o qual independe da base de imagens. A filtragem anisotrópica elimina a influência de ruído e remove detalhes irrelevantes no sinal de curvatura das formas, embora requeira maior esforço computacional, especialmente quando incrementa o número de iterações. Este detector apresenta algumas vantagens em relação a outros métodos disponíveis na literatura, particularmente porque incorpora um processo

de filtragem que reduz a detecção de cantos falsos. No entanto, é necessário determinar o melhor número de iterações no passo de suavização e o parâmetro apropriado τ para calcular a curvatura. Os autores relatam que o algoritmo alcança bons resultados ao fim de 500 iterações na etapa de filtragem. Na verdade, ambos os parâmetros são sensíveis a resolução, rotação e ruído de contorno da forma. Os autores avaliaram o algoritmo utilizando imagens binárias do banco de imagens MPEG7 Parte B (LATECKI *et al.*, 2000) e medidas baseadas na verdade-terrestre. Os cantos da verdade-terrestre destas imagens foram manualmente marcados e aplicados na avaliação dos resultados de simulação.

Neste capítulo, introduzimos um algoritmo multiescala para detecção de cantos que foi apresentado no evento *24th Conference on Graphics, Patterns and Images (Sibgrapi 2011)* (PAULA JR. *et al.*, 2011) e uma versão estendida foi publicada no periódico *Journal of Mathematical Imaging and Vision* (PAULA JR. *et al.*, 2013). O primeiro artigo abordou a detecção de cantos em escalas redundantes obtidas da decomposição do sinal de angulação, destinadas a obtenção de uma reconstrução adequada do contorno original. Da mesma forma, o algoritmo proposto ainda decompõe o sinal de angulação do contorno da forma em múltiplas escalas e identifica máximos e mínimos locais de escalas consecutivas deste sinal através de uma decomposição *wavelet* não-decimada (GAO *et al.*, 2007; HUA; LIAO, 2000; LEE *et al.*, 1995), utilizada na identificação dos pontos dominantes da forma. Essa metodologia foi modificada no artigo estendido com a inclusão de um novo procedimento que busca parâmetros do algoritmo multiescala, em um determinado banco de dados, combinando a ideia subjacente da correlação inter-escala para detectar pontos dominantes em formas planas. O método proposto é também avaliado e comparado com os métodos introduzidos em (LEE *et al.*, 1995; MOKHTARIAN; SUOMELA, 1998; PEDROSA; BARCELOS, 2010) incluindo ainda a técnica CSS.

Este capítulo está organizado em cinco seções. Na Seção 2.1 é apresentada uma breve revisão dos trabalhos relacionados. A Seção 2.2 descreve a técnica de detecção de cantos proposta. Na Seção 2.3 é discutida a complexidade computacional do método proposto e os demais relacionados e as considerações finais são apresentadas na Seção 2.4.

2.1 Métodos Avaliados

Nesta seção, apresentamos três métodos multiescala disponíveis na literatura, nomeados nesta tese como métodos de Pedrosa e Barcelos (PEDROSA; BARCELOS, 2010), CSS (MOKHTARIAN; MACKWORTH, 1986) e Lee *et al.* (LEE *et al.*, 1995). A natureza multiescala é a principal razão para a escolha dessas abordagens, para que a avaliação de desempenho seja justa e compatível com o detector de cantos proposto.

2.1.1 Técnica de Pedrosa e Barcelos

Pedrosa e Barcelos (2010) introduziram um método baseado na detecção de pontos de alta curvatura ao longo do contorno, em que para cada ponto de curvatura k_t considera-se uma curva ou os segmentos de linha que ligam k_t aos pontos $k_{t-\tau}$ e $k_{t+\tau}$ no contorno, tal que τ é um parâmetro de suavização. É sabido que a parametrização ideal é uma tarefa difícil de alcançar, uma vez que pode depender de vários fatores, tais como resolução e orientação da forma (PEDROSA; BARCELOS, 2010). Esta técnica adota $\tau = 1$, o que resulta em um sinal de curvatura detalhado e sensível ao ruído. Este sinal é então suavizado através de um filtro anisotrópico não-linear que elimina a maior parte dos pontos de alta curvatura não relacionados aos cantos verdadeiros. O número de iterações deste processo de filtragem controla o nível de detalhes do sinal e portanto, à medida que este parâmetro diminui maior é a quantidade de pontos relevantes. Após a suavização, a detecção dos cantos verdadeiros é realizada por meio da análise dos pontos de máximos e mínimos locais presentes no sinal de curvatura suavizada. Vale ressaltar que nem todos os pontos de máximos e mínimos restantes constituem cantos reais.

2.1.2 Técnica de Curvatura Espaço-Escala

Uma das representações por curvatura multiescala mais populares de formas planas 2D é a curvatura espaço-escala ou CSS (MOKHTARIAN; MACKWORTH, 1986), a qual foi aperfeiçoada e aplicada em trabalhos posteriores (ALMEIDA *et al.*, 2006; AWRANGJEB; LU, 2008; KINDRATENKO, 2003; MOKHTARIAN; SUOMELA, 1998; MOKHTARIAN; MACKWORTH, 1992; MOKHTARIAN; MOHANNA, 2006; ROH *et al.*, 2006; ZHONG; LIAO, 2007; ZHONG *et al.*, 2009; ZHANG *et al.*, 2010).

A técnica CSS é adequada para a recuperação de características geométricas invariantes, pontos de cruzamento de zeros (MOKHTARIAN; MACKWORTH, 1986) e/ou pontos extremos (ZHONG; LIAO, 2007) da curvatura de uma curva plana em múltiplas escalas. O procedimento assume que a curva Γ é inicialmente parametrizada t :

$$\Gamma(t) = (x(t), y(t)), \quad (2.1)$$

em que $x(t)$ e $y(t)$ referem-se às coordenadas do ponto t do contorno. A versão multiescala de Γ é definida em (MOKHTARIAN; MACKWORTH, 1992) por:

$$\Gamma_\sigma = (X(t, \sigma), Y(t, \sigma)), \quad (2.2)$$

em que

$$X(t, \sigma) = x(t) \otimes g(t, \sigma); Y(t, \sigma) = y(t) \otimes g(t, \sigma). \quad (2.3)$$

O símbolo \otimes refere-se ao operador de convolução e $g(t, \sigma)$ indica a gaussiana de largura σ . Vale destacar que σ descreve o parâmetro de escala e que o processo de geração das versões multiescala de Γ para valores crescentes de σ a partir de zero representa a evolução da curva Γ . A fim de localizar cruzamentos de zeros ou pontos extremos da curvatura nas múltiplas versões do contorno, é necessário calcular a curvatura, K , em uma versão escalonada Γ_σ . A curvatura K em Γ_σ é dada por (MOKHTARIAN; MACKWORTH, 1992):

$$K(t, \sigma) = \frac{\dot{X}(t, \sigma)\dot{Y}(t, \sigma) - \ddot{X}(t, \sigma)\dot{Y}(t, \sigma)}{\sqrt[3]{(\dot{X}(t, \sigma)^2 + \dot{Y}(t, \sigma)^2)}}, \quad (2.4)$$

em que

$$\dot{X}(t, \sigma) = \frac{\partial}{\partial t}(x(t) \otimes g(t, \sigma)) = x(t) \otimes \dot{g}(t, \sigma), \quad (2.5)$$

$$\ddot{X}(t, \sigma) = \frac{\partial^2}{\partial t^2}(x(t) \otimes g(t, \sigma)) = x(t) \otimes \ddot{g}(t, \sigma). \quad (2.6)$$

Similarmente, $\dot{Y}(t, \sigma)$ e $\ddot{Y}(t, \sigma)$ podem ser definidas em correspondência com as expressões das Equações (2.5) e (2.6) para $y(t)$. O algoritmo desenvolvido por Mokhtarian e Suomela (1998) requer o detector de bordas de Canny (1986) aplicado a uma imagem em nível de cinza para obter uma imagem binária com as bordas destacadas.

2.1.3 Técnica de Lee *et al.*

Este método inclui a função de angulação, $\phi(t)$, que é obtida da equação:

$$\phi(t) = \tan^{-1} \left(\frac{y(t+q) - y(t-q)}{x(t+q) - x(t-q)} \right), \quad (2.7)$$

em que o parâmetro q é chamado de parâmetro de suavização. De acordo com Lee *et al.* (1995), a determinação do parâmetro q na Equação (2.7) depende da resolução da orientação e define a capacidade de discriminação dos cantos.

Trabalhos anteriores em análise e representação de curvas digitais (ROSENFELD; JOHNSTON, 1973; ROSENFELD; WESZKA, 1975) foram os pioneiros na utilização de direção do gradiente de borda para representar o contorno de forma. Rosenfeld e Johnston (1973) reportaram que a escolha do parâmetro de suavização q pode se tornar uma etapa desafiadora, e para tanto sugeriram $q > 1$ para proporcionar uma medição de inclinação suavizada no perfil do canto representado pelo sinal de angulação. No entanto, observou-se que a melhor escolha deste parâmetro depende da escala de interesse. Analisando estas observações, o trabalho de Lee *et al.* (1995) apresentou o sinal de angulação para valor de $q = 3$, de modo que esta escolha modificou o perfil de orientação de um canto assemelhando-se a uma rampa no sinal.

A próxima seção descreve a abordagem proposta, além da descrição das principais etapas do algoritmo de Lee *et al.* (1995) que foram incorporados na referida proposta.

2.2 Detector de Cantos Proposto

Os cantos de uma forma se apresentam como características locais no contorno, sendo as *wavelets* ferramentas adequadas para processamento e análise dessas características importantes que persistem ao longo de diversas escalas de contornos suavizados. Com o objetivo de identificar os cantos candidatos, utiliza-se a correlação entre escalas oriundas da decomposição *wavelet* Chapéu Mexicano. A busca dos cantos candidatos ocorre nos pontos de máximo e mínimo locais presentes nos coeficientes de detalhes pelas escalas de decomposição adjacentes. A decomposição utilizando Chapéu Mexicano se mostra eficiente na análise pontual dos dados multiescala, em particular na detecção de singularidades locais (contornos, cantos) (ANTOINE *et al.*, 1993). As demais propriedades desta função *wavelet* podem ser exploradas para detectar características em diferentes escalas (KUTTER *et al.*, 1999). Nesta proposta incorporamos esta função *wavelet* pela sua capacidade de identificação de alterações nos sinais de angulação não-estacionários, a qual pode ser estendida para as abordagens multidimensionais, ao aproximar esta *wavelet* por uma diferença de funções gaussianas (DoGs).

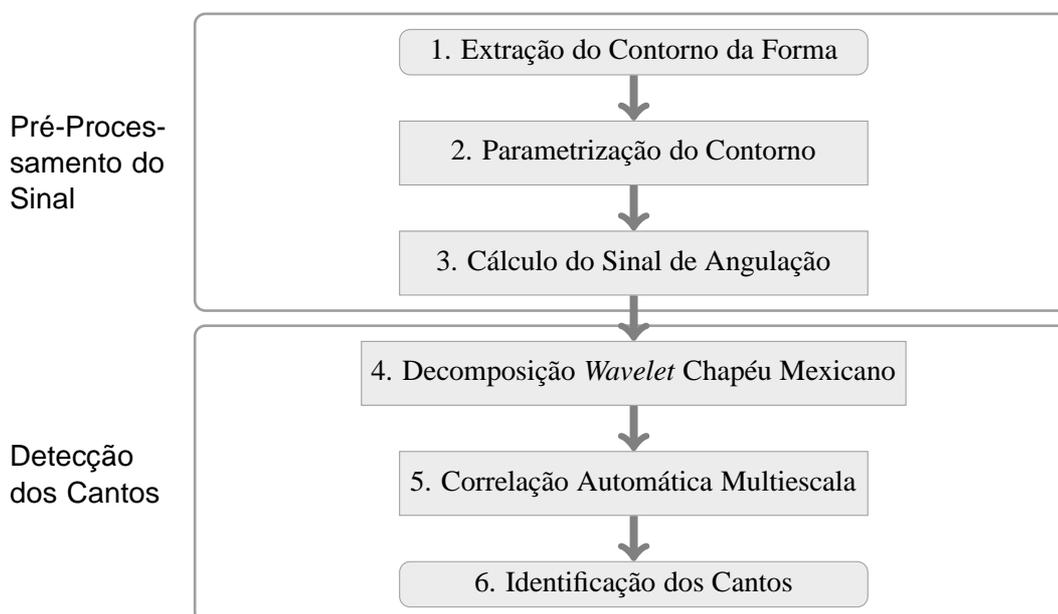


Figura 1: A metodologia proposta para detecção de cantos.

A Figura 1 resume a metodologia proposta para detecção de cantos multiescala. O sinal de representação de contorno é adquirido na extração do mesmo na forma binária. Existem várias maneiras de representar o contorno de uma forma, e nesta tese adotou-se a representação do sinal de contorno pelo código da cadeia 4-direcional (COSTA; CESAR, 2009). Este método segue um ponto inicial do contorno e o percorre no sentido horário. Para cada ponto, o código

da cadeia recupera a sua localização e descreve a direção a ser seguida no contorno para encontrar o ponto seguinte. O código da cadeia é parametrizado em um sinal de angulação (LEE *et al.*, 1995) antes de ser aplicado como entrada no detector de cantos. As principais diferenças entre o método proposto e o de Lee *et al.* são as seguintes: utilização da função Chapéu Mexicano normalizada (MALLAT, 1998) como *wavelet*-mãe; a fase de avaliação dos candidatos; a metodologia de busca dos parâmetros adequados do algoritmo que resulta no melhor desempenho e comparações com métodos publicados mais recentemente.

Como mencionado anteriormente, a entrada do detector de cantos é a função de angulação do contorno da forma. Este sinal unidimensional é gerado de acordo com a Equação 2.7 e corresponde à etapa 3 da metodologia proposta. No passo seguinte, aplica-se a operação de convolução ao sinal de angulação e a uma versão escalonada da *wavelet*-mãe nas diversas escalas de decomposição

$$c_s(t) = \phi(t) \otimes \psi_s(t), \quad (2.8)$$

em que $c_s(t)$ corresponde à versão suavizada do sinal de angulação na escala s e $\psi_s(t)$ é uma versão escalonada da *wavelet*-mãe. A função $\psi_s(t)$ é tal que

$$\psi_s(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t}{s}\right) \quad (2.9)$$

e a *wavelet*-mãe $\psi(t)$ é a segunda derivada de uma gaussiana, ou seja, uma função Chapéu Mexicano normalizada (MALLAT, 1998) dada por:

$$\psi(t) = \frac{2}{\pi^{1/4} (\sqrt{3}\sigma)} \left(\frac{t^2}{\sigma^2} - 1 \right) \exp\left(\frac{-t^2}{2\sigma^2} \right), \quad (2.10)$$

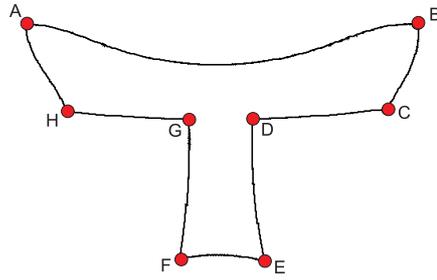
em que σ refere-se ao desvio-padrão.

A Equação (2.8) aplicada ao sinal de angulação produz os coeficientes *wavelets* de aproximação (sinal suavizado) em cada escala s . Com isso, a diferença entre dois sinais suavizados sucessivos gera os coeficientes *wavelets* de detalhes, $w_s(t)$, como segue:

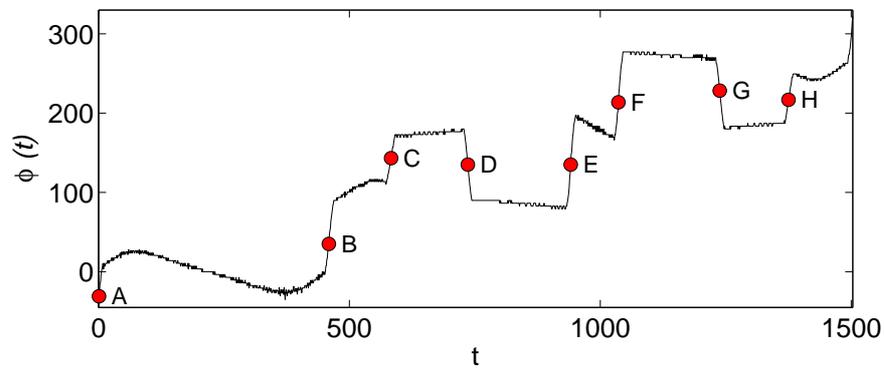
$$w_s(t) = c_{s-1}(t) - c_s(t); \quad s \geq 1, \quad (2.11)$$

em que $c_0(t)$ corresponde ao sinal de angulação original.

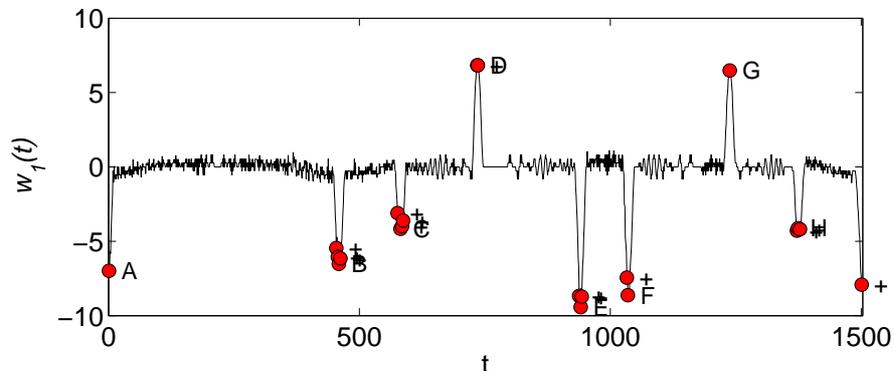
A metodologia proposta mostra que os cantos ocorrem onde existem valores extremos do sinal, isto é, valores de máximos e mínimos nos coeficientes *wavelets* não-ortogonais $w_s(t)$ em duas ou mais escalas adjacentes. A novidade desta estrutura multiescala consiste em inspecionar esses valores de picos, usando sinais de correlação obtidos a partir das escalas redundantes do sinal de decomposição *wavelet*, sem a utilização de um limiar empírico. Assim,



(a) Forma original e seus cantos da verdade-terrestre (círculos vermelhos).



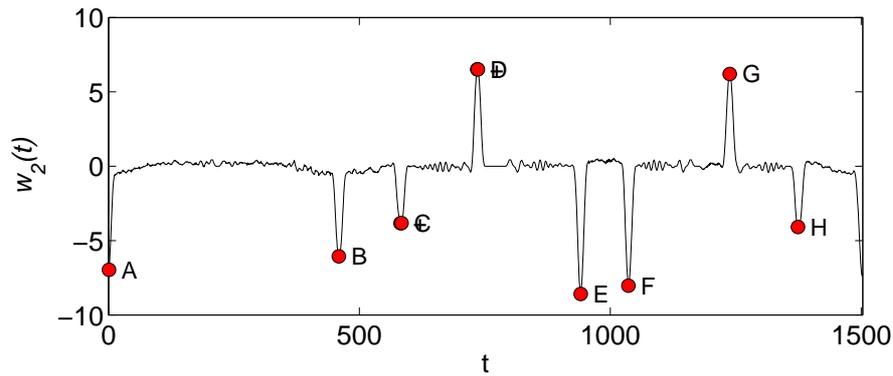
(b) Sinal de angulação com seus cantos destacados (círculos vermelhos).



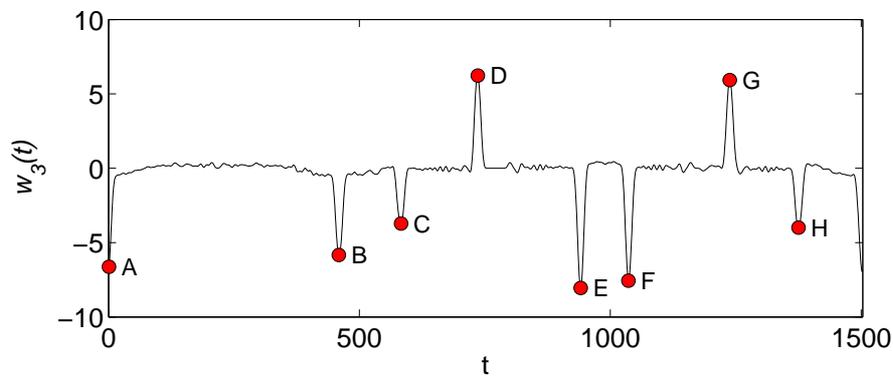
(c) Cantos destacados no 1º nível de coeficientes de detalhes. As marcações em cruz indicam cantos falsos.

Figura 2: Análise de correlação entre os coeficientes *wavelets* de diferentes escalas do contorno de forma.

escalas redundantes diferentes são geradas e correlacionadas para detectar picos persistentes que são potenciais cantos candidatos nos sinais de decomposição. Os resultados experimentais indicam que a análise de correlação multiescala nas primeiras três escalas é capaz de revelar a informação redundante que persiste ao longo das escalas e isso propicia a identificação dos cantos candidatos. A correlação espacial $corr_s(t)$ entre as escalas adjacentes s e $s + 1$ é dada por (SITA; RAMAKRISHNAN, 2000):



(a) Cantos destacados no 2º nível de coeficientes de detalhes. As marcações em cruz indicam cantos falsos.



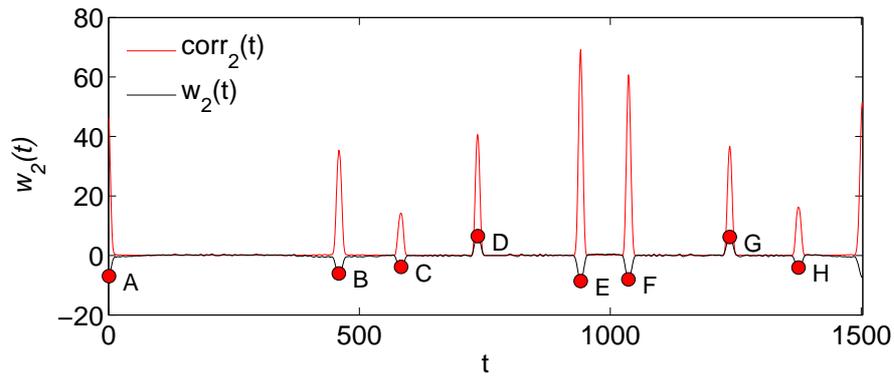
(b) Cantos destacados no 3º nível de coeficientes de detalhes.

Figura 3: (Continuação) Análise de correlação entre os coeficientes *wavelet* de diferentes escalas de um contorno de forma.

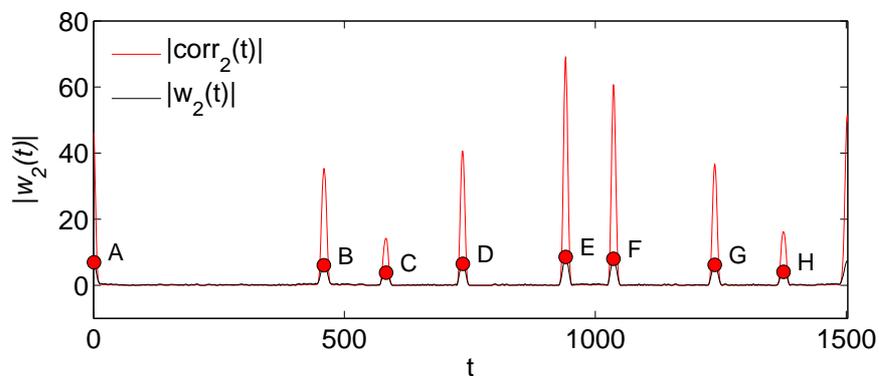
$$corr_s(t) = w_s(t) \cdot w_{s+1}(t), \quad s \geq 1. \quad (2.12)$$

Neste trabalho, o sinal de angulação é decomposto em três escalas espaciais que são correlacionadas e assim, verificam-se informações redundantes entre escalas de decomposição. De acordo com os testes realizados, esta redundância pode ser investigada por mais níveis de decomposição. Na verdade, o número de escalas em que cantos verdadeiros são preservados varia de acordo com a imagem. Além disso, observa-se que em regiões suaves no contorno, os cantos candidatos não persistem, quando se tem sucessivas escalas de decomposição. Quanto aos pontos de alta curvatura, estes se mantêm em escalas mais altas.

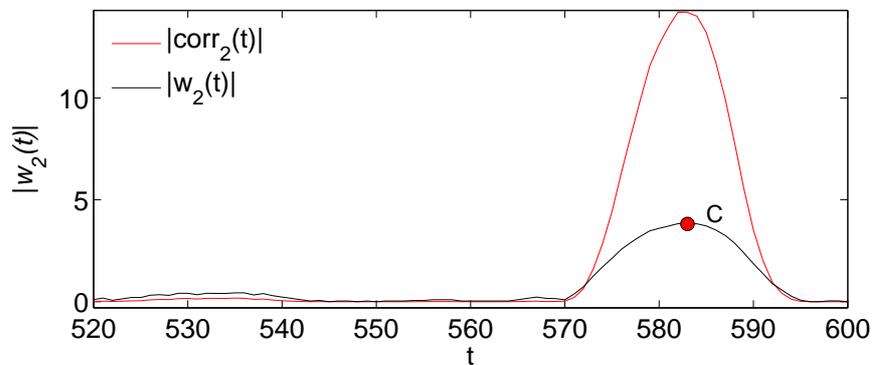
A abordagem proposta é sintetizada nas Figura 2 e Figura 3, em que os cantos reais são exibidos com maior amplitude ao longo de múltiplas escalas de decomposição *wavelet*. Ao mesmo tempo, as magnitudes dos cantos falsos decaem com o incremento das escalas. Os círculos vermelhos nas Figura 2 e Figura 3 ilustram os cantos candidatos selecionados nos coeficientes. O sinal de angulação é decomposto em diferentes escalas, de tal maneira que um sinal contendo os coeficientes de detalhes é comparado elemento a elemento com o sinal de



(a) Cantos destacados (marcações em vermelho) no 2º nível de coeficientes de detalhes e o sinal de correlação.



(b) Cantos destacados no 2º nível de coeficientes de detalhe e os valores absolutos de correlação.



(c) Detalhamento do sinal de correlação sobre o 2º nível de valores absolutos de coeficientes de detalhe.

Figura 4: Efeitos da correlação entre os coeficientes *wavelet* de escalas consecutivas da forma apresentada na Figura 2(a).

coeficientes do próximo nível. Este resultado mostra que os cantos correspondem aos valores de maiores amplitudes que persistem no sinal de correlação nas várias escalas de decomposição *wavelet*.

Admitindo que cantos falsos desaparecem rapidamente com a escala crescente, o algoritmo identifica no sinal de correlação se um canto candidato é verdadeiro ou falso, testando a seguinte

desigualdade:

$$|corr_s(t)| > |w_s(t)|, \quad s \geq 1, \quad (2.13)$$

em que $|\cdot|$ representa o valor absoluto. Se a Equação (2.13) é confirmada (para a escala s e todos os n pontos de contorno), então os valores de máximo da correlação espacial entre duas escalas consecutivas apontam para os cantos localizados. Posteriormente, estes pontos constituem um vetor de cantos candidatos. O algoritmo investiga se o canto candidato apresenta o maior valor em amplitude na vizinhança no sinal de correlação, a fim de distinguir um ponto de alta curvatura de um falso canto detectado. A ausência de cantos ou outras características significativas na vizinhança de uma região de contorno permite que os cantos falsos sejam removidos deste vetor.

Pontos detectados no sinal de correlação são destacados na Figura 4 e estão sobrepostos aos sinais dos coeficientes de detalhes *wavelet*. A Figura 4(c) apresenta uma ampliação de uma parte da Figura 4(b) para mostrar os valores absolutos dos sinais de correlação e coeficientes *wavelet*. Observa-se que a validade da Equação (2.13) ocorre quando o canto é identificado em locais onde a magnitude da correlação é maior que o valor absoluto do coeficiente de detalhe.

2.3 Complexidade Computacional Temporal

Em uma análise mais detalhada do algoritmo proposto, os passos relacionados à parametrização de contorno e extração do sinal de angulação do mesmo são menos críticos, em termos de complexidade computacional, do que o esquema de decomposição *wavelet*. Enquanto a etapa de pré-processamento resulta em uma complexidade da ordem de $\Theta(n)$ (passos 1, 2 e 3 da metodologia), o módulo relacionado com a decomposição *wavelet* chapéu mexicano é executado com uma complexidade temporal de $\Theta(n^2)$ (passos 4, 5 e 6), em que n é o número de pontos de contorno da forma. Na verdade, o custo computacional pode ser reduzido com a realização da convolução espacial no domínio da frequência. No entanto, o procedimento restante referente à correlação multiescala atinge uma complexidade $\Theta(n^2)$. Com isso, a complexidade computacional total do algoritmo é: $\Theta(n) + \Theta(n^2) + \Theta(n^2) = \Theta(n^2)$. No detector de cantos proposto, o sinal de angulação é decomposto em apenas três escalas, o que portanto, não impacta sobre a complexidade assintótica global do algoritmo.

A despeito da natureza multiescala do algoritmo desenvolvido, não foi definida uma faixa de variação das escalas (número mínimo e máximo) para realizar a etapa de correlação do detector de cantos. Certamente a definição desta faixa de valores afetaria a complexidade computacional do algoritmo. A complexidade computacional do método proposto e das demais técnicas de detecção de cantos abordadas nesse trabalho estão dispostas na Tabela 1. Cada um

dos métodos possui a sua representação de complexidade baseada na informação mais influente no desempenho do algoritmo. Para a maioria dos métodos relacionados, o número de pontos do contorno da forma (representado por n) é a mais relevante.

Para efeitos de comparação, o método de detecção de cantos introduzido por Pedrosa e Barcelos (2010) consiste de três etapas, que incluem a representação da forma através do sinal de τ -curvatura, filtragem anisotrópica não-linear e a detecção de cantos no sinal de curvatura suavizado. A complexidade computacional geral deste método é relatada como $O(nm)$, em que n é o número de pontos do contorno da forma e m corresponde ao número de iterações na etapa de suavização utilizando um filtro de difusão anisotrópica não-linear. Em caso de uma base com grande número de imagens a processar, esse número cresceria da ordem da quantidade de imagens a serem processadas, e portanto, resultou-se em um tempo de execução maior do algoritmo. Assim, quanto maior a base de imagens e o número de iterações da etapa de filtragem, mais lento será o processamento da detecção dos cantos.

A abordagem CSS apresenta uma complexidade computacional mais elevada do que o método proposto. Esse custo é decorrente da etapa de filtragem gaussiana, que é aplicada ao contorno da forma diversas vezes, e as respectivas mudanças do desvio-padrão aplicado a cada iteração, tal como definido em (MOKHTARIAN; MACKWORTH, 1986; MOKHTARIAN; SUOMELA, 1998). Este processo é realizado para gerar o mapa de espaço-escala a partir do método original. Na sequência, realiza-se a detecção de cantos com o parâmetro igual a σ_{high} (MOKHTARIAN; SUOMELA, 1998). O método resulta em uma complexidade computacional $O(n^3)$, em que n é o número de pontos presentes em um contorno.

O detector de cantos de Lee *et al.* (LEE *et al.*, 1995) realiza processamento semelhante ao método proposto, diferindo na função *wavelet*-mãe utilizada na etapa de decomposição multiescala. Devido à sua característica não-ortogonal e pelo realce nas regiões de pico, optou-se pela *wavelet* Chapéu Mexicano normalizada no método proposto. Outra diferença se refere à localização dos cantos, dado que o método de Lee *et al.* avalia uma escala específica apenas e busca por valores máximos e mínimos de acordo com um determinado limiar. Sua complexidade computacional é caracterizada por $O(n^2)$, em que n é o número de pontos de um contorno da forma. A Tabela 1 resume a complexidade computacional de todos os métodos apresentados nesta seção.

Tabela 1: Complexidade temporal dos detectores de cantos.

Método	Complexidade
Proposto	$\Theta(n^2)$
Pedrosa e Barcelos	$O(nm)$
CSS	$O(n^3)$
Lee <i>et al.</i>	$O(n^2)$

2.4 Considerações Finais

Este capítulo introduziu uma nova abordagem para detecção de cantos que utiliza correlação entre escalas redundantes de uma decomposição *wavelet* Chapéu Mexicano normalizado. A escolha desta *wavelet* se deve ao fato da sua habilidade em detectar singularidades locais como cantos. Além disso, esta função *wavelet* pode ser estendida para abordagens multidimensionais de uma maneira eficiente ao aproximá-la por diferença de funções gaussianas. A abordagem proposta investiga alterações em sinais de angulação não-estacionários para detecção de cantos multiescala. A representação multiescala auxilia na análise de cantos candidatos relacionados aos pontos de mínimo e máximo que persistem no sinal de correlação. Com isso, o método descarta cantos falsos que desaparecem ao longo das escalas.

Com relação à complexidade computacional, o método proposto alcançou uma menor complexidade computacional em comparação às abordagens *CSS* e *Lee et al.* Em particular, o método de Pedrosa e Barcelos alcançou a menor complexidade computacional, quando o número de pontos de contorno é maior do que o número de iterações da filtragem anisotrópica.

No próximo capítulo são apresentados os resultados experimentais e discussões oriundas da aplicação da técnica proposta e demais métodos.

3 RESULTADOS EXPERIMENTAIS DE DETECÇÃO DE CANTOS

Este capítulo apresenta os experimentos e a avaliação de desempenho da técnica proposta para a detecção de pontos dominantes em comparação com outros métodos baseados em contorno, como as seguintes abordagens: detecção de pontos de cantos da forma (PEDROSA; BARCELOS, 2010), análise por curvatura espaço-escala (MOKHTARIAN; SUOMELA, 1998), e transformada *wavelet* módulo máxima (LEE *et al.*, 1995). Os resultados foram obtidos executando os métodos para um conjunto de formas de diferentes tamanhos e características, que pertencem a um conjunto de 104 formas binárias da base de imagens MPEG7 Parte B (LATECKI *et al.*, 2000). Deve-se destacar que 10% das imagens, cujos contornos foram processados, apresentam ruído de digitalização.

3.1 Metodologia de Avaliação de Desempenho

A inspeção visual é recurso usual na avaliação de desempenho dos métodos de detecção de cantos, entretanto esta é uma tarefa tediosa e muitas vezes subjetiva. Além disso, os pontos de alta curvatura não necessariamente correspondem a pontos visualmente significativos (BALMASHNOVA; FLORACK, 2008; COSTA; CESAR, 2009; MOKHTARIAN; MOHANNA, 2006; MOKHTARIAN; SUOMELA, 1998). Estas razões motivam o uso de medidas quantitativas que relacionem cantos detectados com a informação da verdade-terrestre na avaliação de desempenho dos detectores de cantos. Vale ressaltar que a avaliação de desempenho dos métodos de detecção de cantos também pode ser realizada em termos de erro de reconstrução e proporção de compactação (PAULA JR. *et al.*, 2011).

Esta subseção apresenta a metodologia de avaliação de desempenho que considera medidas baseadas na verdade-terrestre e que são nomeadas na literatura como *Precisão* e *Revocação*.

A *Precisão* mede a capacidade dos detectores multiescala em recuperar os cantos que são relevantes para representar uma forma, ou seja, ausência de cantos falsos. Por outro lado, a *Revocação* quantifica os cantos relevantes que são realmente detectados, indicando que não houve perda de cantos verdadeiros no resultado final. Estes cantos relevantes correspondem à

verdade-terrestre da forma. Formalmente, estas medidas são definidas como:

$$Precisão = \frac{\#(\{\text{cantos relevantes}\} \cap \{\text{cantos detectados}\})}{\#(\{\text{cantos detectados}\})}, \quad (3.1)$$

$$Revocação = \frac{\#(\{\text{cantos relevantes}\} \cap \{\text{cantos detectados}\})}{\#(\{\text{cantos relevantes}\})}, \quad (3.2)$$

em que $\#(\bullet)$ é a cardinalidade do conjunto \bullet e \cap denota a intersecção dos conjuntos. Quando cantos falsos não são detectados, *Precisão* retorna um valor igual a 1,0. Por sua vez, se todos os cantos detectados correspondem aos verdadeiros, *Revocação* retorna valor igual a 1,0.

3.2 Experimento I - Ajustes de Parâmetros

Esta subsecção apresenta o procedimento proposto para busca dos melhores valores de parâmetros para os métodos de detecção multiescala de cantos que se baseiam nas verdades-terrestres de um banco de imagens. O objetivo deste procedimento é maximizar as medidas *Precisão* e *Revocação* para um banco de dados específico, em termos de uma função de custo.

Aqui, introduz-se a função de custo Q como a soma dos valores de *Precisão* e *Revocação* obtidos nos experimentos. Estes valores são recuperados quando os parâmetros p_1 e p_2 dos detectores multiescala são aplicados sobre um banco de imagens I , considerando estes parâmetros como valores relevantes na execução de um determinado método, de acordo com:

$$Q(p_1, p_2) = \sum_{\forall i \in I} Precisão(p_1, p_2, i) + Revocação(p_1, p_2, i), \quad (3.3)$$

em que i refere-se a uma das imagens pertencentes à I . A função Q é maximizada quando os parâmetros p_1 e p_2 retornam valores máximos para *Precisão* e *Revocação*, como segue:

$$\overline{Q(p_1, p_2)} = \operatorname{argmax} Q(p_1, p_2). \quad (3.4)$$

Inicialmente, define-se um intervalo de valores para os parâmetros p_1 e p_2 . Em seguida, aplica-se um esquema de otimização por força bruta, em que o detector de cantos é executado por todos os pares de valores dos parâmetros na faixa de $p_1 = [p_{1_{\min}} \cdot p_{1_{\max}}]$ e $p_2 = [p_{2_{\min}} \cdot p_{2_{\max}}]$ para todas as imagens no banco de dados. Finalmente, o melhor par de valores dos parâmetros é escolhido. No que diz respeito ao método proposto, o par (p_1, p_2) consiste dos parâmetros q e σ , ou seja, passo de suavização e desvio-padrão da função *wavelet*-mãe.

Neste experimento adotou-se as seguintes faixas de valores para os parâmetros do detector,

a saber: $q_{min} = 2$, $q_{max} = 15$ e $\sigma_{min} = 1$, $\sigma_{max} = 15$. O algoritmo proposto foi executado 210 vezes para cada forma da base de dados. Os valores *Precisão*= 1 e *Revocação*= 1 foram alcançados por diferentes pares de (q, σ) e ocorrências dentro desta faixa, ou seja, a partir de $(q_{min}, \sigma_{min}) = (2, 2)$ para o máximo $(q_{max}, \sigma_{max}) = (15, 15)$. Este experimento resultou no par $(q, \sigma) = (8, 4)$ o qual alcançou valor máximo na função custo Q . Muitas vezes, é possível obter o valor máximo da função Q em valores distintos de q e σ . Dentre essas ocorrências, optou-se pelos menores valores dos parâmetros avaliados para definir os seus valores finais na análise comparativa entre os métodos. Realizou-se 224 experimentos para identificar o melhor conjunto de parâmetros para o método proposto.

Os experimentos mostraram que os melhores resultados foram alcançados, quando esses parâmetros eram ajustados seguindo a base de dados utilizada. Em particular, Mokhtarian e Suomela (1998) sugeriram $\sigma_{high} = 4$ para o rastreamento de cantos e um limiar $h = 0,02$ para remover cantos falsos na abordagem CSS, sendo σ_{high} (desvio-padrão limite) e h (intervalo entre escalas) parâmetros definidos em (MOKHTARIAN; SUOMELA, 1998). Da mesma maneira, Lee *et al.* (1995) conduziram seus experimentos com $q = 3$ para alcançar o mesmo objetivo. Redefinimos o método de Lee *et al.* original aplicando a função *wavelet* Chapéu Mexicano com $\sigma = 3$ na decomposição do sinal de angulação o qual foi extraído com $q = 3$, como sugerido pelos autores.

Seguindo a função de custo Q proposta para ajuste dos parâmetros, avaliou-se o melhor par (p_1, p_2) para cada método, com (σ_{high}, h) na abordagem CSS e (q, σ) para o método de Lee *et al.* No que diz respeito aos parâmetros do método CSS, foram testados os valores $\sigma_{high} = [1..15]$, $h_{min} = 0,01$, $h_{max} = 15,0$, com uma passo incremental de 0,01 para h . Portanto, a função Q foi maximizada com o par $(\sigma_{high}, h) = (6, 10)$ e esse par de valores foi alcançado após 24.000 testes de valores de parâmetros para o método CSS. Da mesma maneira, o experimento com o método de Lee *et al.* resultou no par $(q, \sigma) = (4, 3)$ dentro da faixa de $q = [2..15]$ e $\sigma = [1..15]$. Com esta última abordagem, encontrou-se os novos valores dos seus parâmetros após a realização de 224 experimentos.

A Figura 5 exibe o número de cantos detectados pelos métodos discutidos segundo os experimentos realizados, apresentados aqui em escala logarítmica para melhor visualização. Efetivamente, estes resultados foram obtidos com a base de dados mencionada e utilizando diferentes valores para σ . Como esperado, observou-se para todos os métodos multiescala que o número de cantos diminui à medida que σ aumenta. Além disso, valores menores deste parâmetro implicam em detecção de maior número de cantos, incluindo a este falsos positivos. No método proposto, após alcançar um certo valor de σ , ou seja, σ igual a 3, o número de cantos detectados se estabilizou independentemente do incremento deste parâmetro. Os resultados de simulação mostraram que valores de q próximos a 1 implicam em uma grande quantidade de detalhes presentes no sinal de angulação, o que resulta na detecção de maior número de cantos,

incluindo a estes falsos positivos. Por outro lado, valores elevados de q implicam em uma grande redução de detalhes, que resulta na perda de cantos verdadeiros.

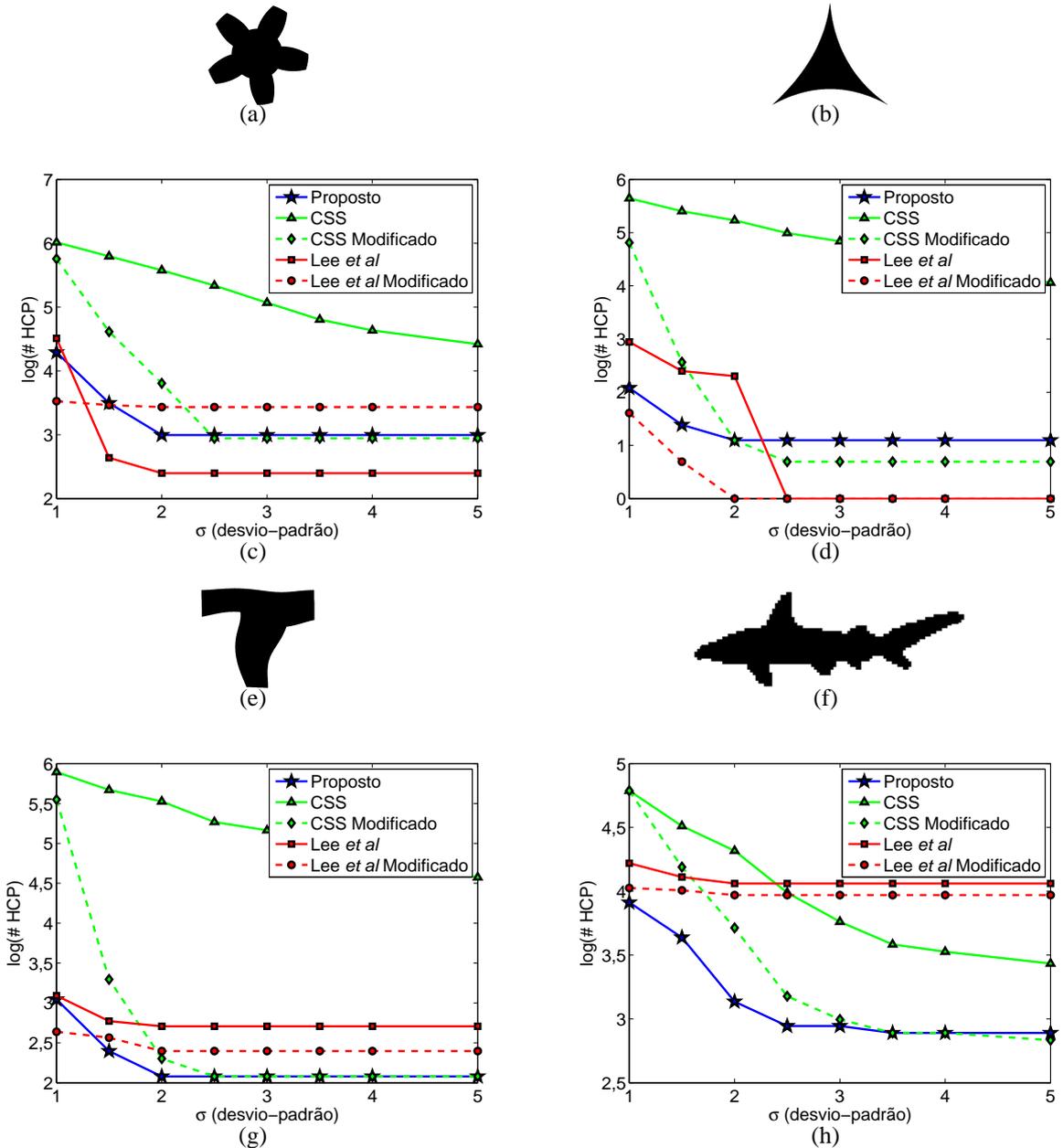


Figura 5: Análise comparativa dos métodos para diferentes valores de desvio-padrão e números de cantos. (a),(b),(e),(f) formas binárias e (c),(d),(g),(h) suas respectivas análises comparativas.

Vale destacar que essa análise comparativa exclui o método Pedrosa e Barcelos, dado que seus parâmetros multiescala não fazem correspondência com as variáveis σ e q adotadas na metodologia proposta.

3.3 Experimento II - Análise dos Resultados

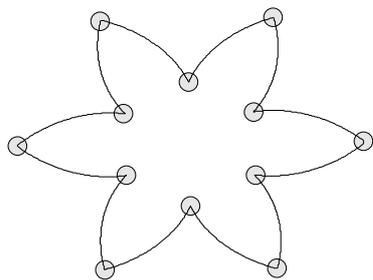
Esta subsecção apresenta os experimentos realizados com parâmetros gerados pela metodologia apresentada na Subsecção 3.2. As Figura 6 e Figura 7 mostram que o algoritmo de ajuste de parâmetros obteve melhores resultados para os métodos *CSS* e *Lee et al.* Os métodos multiescala proposto, Pedrosa e Barcelos e *CSS* modificado produziram resultados próximos, quando se trata de formas de contorno suave (ver Figura 6). O método proposto detectou todos os cantos verdadeiros sem recuperar qualquer canto falso, enquanto o método de Pedrosa e Barcelos perdeu dois cantos verdadeiros e produziu dois falsos positivos para a mesma imagem em destaque. O método de *Lee et al.* modificado alcançou um valor intermediário de *Precisão* uma vez que o mesmo foi penalizado por retornar maior número de cantos falsos. Os métodos originais *CSS* e *Lee et al.* produziram grande quantidade de cantos falsos e consequentemente alcançaram os piores resultados para esta forma.

A Figura 7 apresenta o resultado da aplicação de cada detector de cantos em uma amostra ruidosa presente na base de dados. Apenas o método proposto detectou todos os cantos que compõem a verdade-terrestre da imagem. A etapa de filtragem do método de Pedrosa e Barcelos resultou em uma detecção com menos cantos falsos, no entanto, esta suavização causou na perda de cantos verdadeiros. Os outros métodos foram realizados de maneira semelhante ao proposto para esta imagem ruidosa.

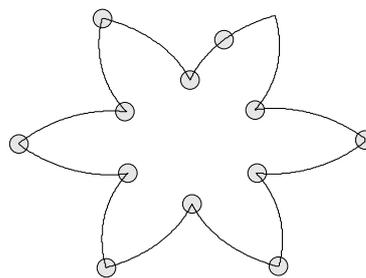
Tabela 2: Avaliação quantitativa entre os métodos.

Método	<i>Precisão</i>		<i>Revocação</i>	
	Média	Desvio-Padrão	Média	Desvio-Padrão
Proposto	0,87	0,29	0,97	0,08
Pedrosa e Barcelos	0,87	0,19	0,68	0,18
<i>CSS</i>	0,72	0,33	0,61	0,18
<i>CSS</i> Modificado	0,84	0,29	0,78	0,22
<i>Lee et al.</i>	0,51	0,33	0,71	0,36
<i>Lee et al.</i> Modificado	0,61	0,31	0,75	0,35

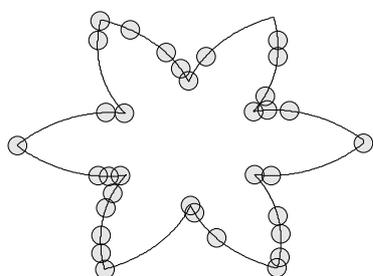
Com relação à avaliação quantitativa de desempenho, os valores médios das medidas *Precisão* e *Revocação* foram calculados para todos os métodos e a Tabela 2 resume os resultados obtidos. O método de Pedrosa e Barcelos e a abordagem proposta alcançaram os melhores valores de *Precisão* e, consequentemente, uma menor quantidade de cantos falsos foram detectados. As versões modificadas de *CSS* e *Lee et al.* alcançaram valores mais elevados para essa medida do que suas versões originais. Quanto ao desvio-padrão nesta mesma medida, o método Pedrosa e Barcelos superou as outras abordagens devido à filtragem inerente ao seu processo. Por outro lado, a abordagem proposta alcançou a melhor taxa de detecção, com *Revocação*=0,97. A partir desses resultados, concluímos que o desempenho dos



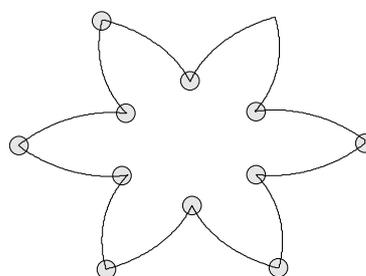
(a) Método proposto: 12 cantos (nenhum perdido, nenhum falso positivo).



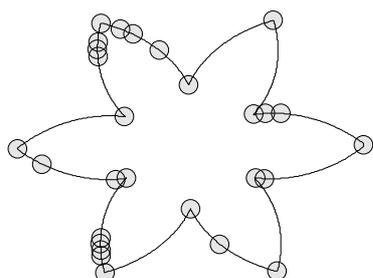
(b) Método de Pedrosa e Barcelos: 12 cantos (01 perdido, 01 falso positivo).



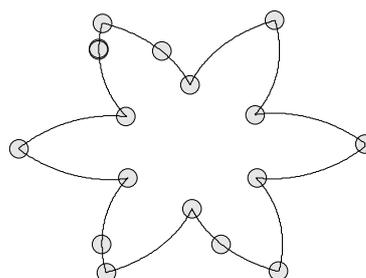
(c) Método CSS: 33 cantos (01 perdido, 22 falsos positivos).



(d) Método CSS modificado: 11 cantos (01 perdido, nenhum falso positivo).



(e) Método de Lee *et al.*: 28 cantos (nenhum perdido, 16 falsos positivos).

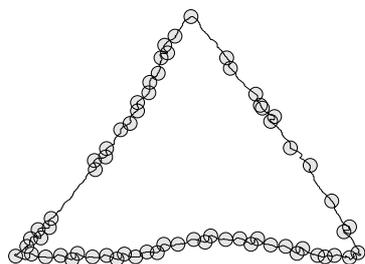


(f) Método de Lee *et al.* modificado: 17 cantos (nenhum perdido, 05 falsos positivos).

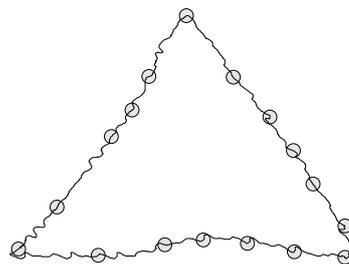
Figura 6: Cantos detectados (círculos) a partir de uma forma com 1328 pontos de contorno e 12 cantos verdadeiros.

detectores de cantos modificados foi melhorado com a redefinição dos parâmetros alcançados e, conseqüentemente, as medidas de *Precisão* e *Revocação*.

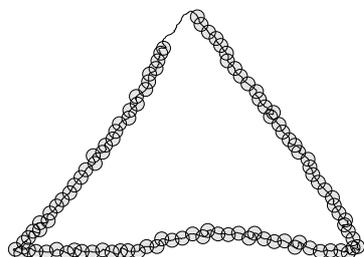
O método de Pedrosa e Barcelos detectou um número de cantos mais próximo às



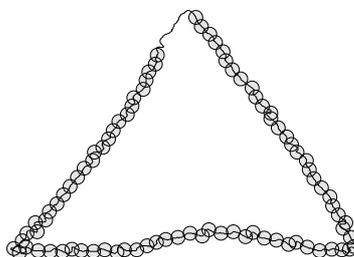
(a) Método proposto: 61 cantos (nenhum perdido, 58 falsos positivos).



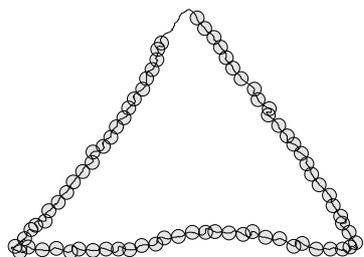
(b) Método de Pedrosa e Barcelos: 17 cantos (02 perdidos, 16 falsos positivos).



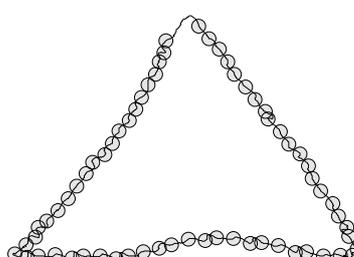
(c) Método CSS: 56 cantos (03 perdidos, 56 falsos positivos).



(d) Método CSS modificado: 31 cantos (03 perdidos, 31 falsos positivos).



(e) Método de Lee *et al.*: 76 cantos (02 perdidos, 75 falsos positivos).



(f) Método de Lee *et al.* modificado: 64 cantos (02 perdidos, 63 falsos positivos).

Figura 7: Cantos detectados (círculos) a partir de uma forma com 1193 pontos de contorno e 03 cantos verdadeiros.

verdades-terrestres das formas mencionadas. No entanto, as coordenadas espaciais dos cantos detectados não correspondem aos cantos verdadeiros e isso é confirmado pelo menor valor de *Revocação*. Como este detector de cantos realiza uma filtragem anisotrópica iterativa nas imagens ruidosas e não-ruidosas, provavelmente o número fixo de iterações causou efeitos de suavização indesejáveis sobre os contornos não-ruidosos. É importante mencionar que estas duas classes de imagens requerem um número diferente de iterações para esse filtro. O

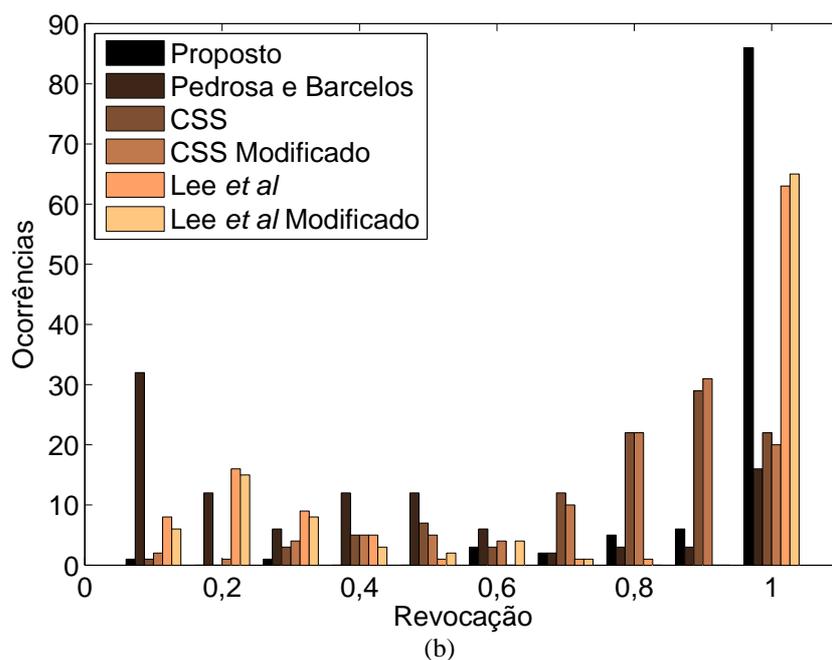
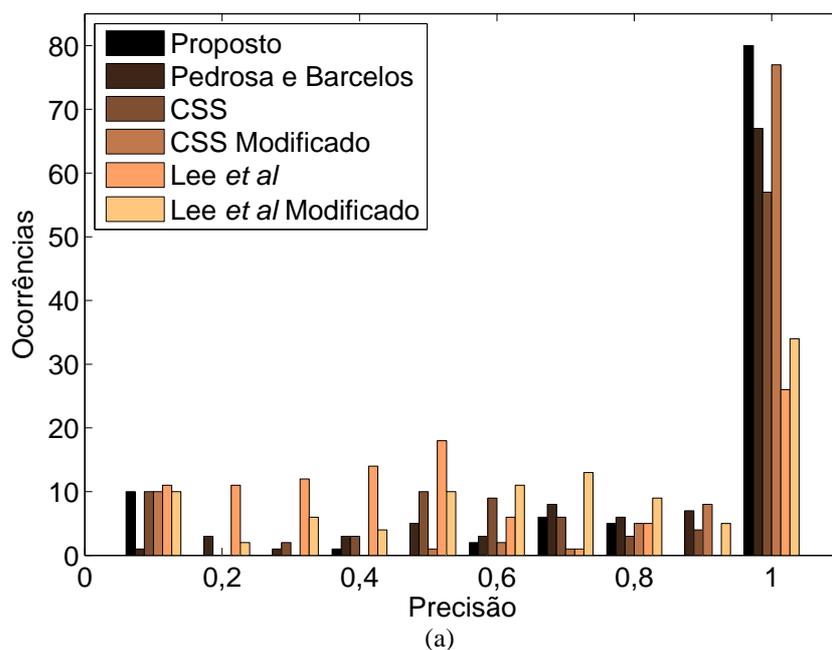


Figura 8: Ocorrências dos (a) valores de *Precisão* e (b) *Revocação* na base de imagens.

método modificado *CSS* alcançou um menor número de cantos perdidos, ou seja, alto valor de *Revocação*, enquanto o detector de *Lee et al.* atingiu valores crescentes de *Precisão* em sua versão modificada.

Com relação ao método proposto, observou-se que o mesmo alcançou o menor número de cantos falsos e o maior número de cantos verdadeiros entre os métodos discutidos. Outra

vantagem importante deste método é o fato que o mesmo requer apenas três escalas de decomposição *wavelet*, que permite extrair a informação relevante a partir da correlação entre as escalas subsequentes. A avaliação comparativa atestou que o método proposto alcançou maiores valores de medida *Revocação* e isso proporcionou as menores taxas de cantos perdidos, dentre as técnicas avaliadas.

De acordo com a Figura 8, todos os histogramas são assimetricamente distribuídos. Com relação ao método proposto, a medida *Precisão* apresenta maior número de ocorrência com valores próximos a 1. No entanto, todos os métodos exibem mais de uma moda em seus histogramas, exceto o método de Pedrosa e Barcelos, que atingiu os valores mais elevados de *Precisão*, encontrando menos cantos falsos devido à filtragem anisotrópica. Para a medida *Revocação*, o método proposto alcançou a maior ocorrência de valores próximos de 1. Merece destaque o fato de que quase todas as ocorrências desta medida estão próximas ao valor máximo. Esta informação confirma que o método proposto apresenta o melhor desempenho dentre os demais na localização dos cantos verdadeiros das formas.

O hardware utilizado nos experimentos foi um notebook com as seguintes especificações: Processador Intel Core 2 Duo T8100, Adaptador Gráfico Intel Graphics Media Accelerator (GMA) X3100, Windows Vista Home Basic 32-bit e 4 GB de Memória RAM. Além disso, as simulações foram desenvolvidas no ambiente Matlab 2010b.

3.4 Considerações Finais

Os resultados apresentados neste capítulo para o detector de cantos proposto permitem concluir que a decomposição sucessiva do sinal de angulação favorece a identificação e retenção de características relevantes sem distorcer a verdadeira localização dos cantos no contorno da forma. Observamos ainda que o processo de correlação entre as escalas recuperou cantos candidatos nas três escalas iniciais. Escalas subsequentes não contribuem de forma significativa com informações relevantes para a detecção dos cantos candidatos, sendo que o algoritmo descartou informações fracamente correlacionadas nessas escalas.

Uma contribuição relevante deste trabalho diz respeito à busca pelos parâmetros dos algoritmos de detecção de cantos multiescala que se baseiam na verdade-terrestre e nas medidas *Precisão* e *Revocação*. Este procedimento de ajuste dos parâmetros melhorou o desempenho dos dois métodos multiescala que foram projetados para trabalhar com parâmetros constantes, independentemente da base de dados. Concluiu-se a partir dos experimentos realizados, que o detector de cantos proposto e o método de Pedrosa e Barcelos apresentaram desempenho semelhante em se tratando da medida *Precisão*. Em relação à medida *Revocação*, o detector de cantos proposto superou os demais e, conseqüentemente, foi capaz de recuperar cantos verdadeiros e eliminar falsos positivos. Finalmente, a metodologia proposta apresentou uma

alternativa viável para a busca e ajuste dos parâmetros dos algoritmos de detecção de cantos. Por se tratar de uma metodologia adaptável à base de imagens trabalhada, observou-se a viabilidade de sua utilização em sistemas de recuperação de imagens, entretanto a função custo proposta requer medidas baseadas na verdade-terrestre.

Alguns sistemas de recuperação de imagens possuem uma etapa de extração de características das imagens. Esse processamento reduz a dimensionalidade do problema ao comparar duas imagens (sinais bidimensionais) através de sinais unidimensionais parametrizados que descrevem as imagens analisadas. Várias características podem ser obtidas a partir de uma imagem e estas mesmas podem ser agrupadas na estrutura de um descritor. Os cantos de uma forma servem de base para a definição de novos descritores de formas. A partir destes descritores é possível analisar a recuperação de formas presentes em uma base de dados. Os conceitos relacionados a descritores e sistemas de recuperação de formas baseadas em cantos serão apresentados e discutidos no próximo capítulo.

PARTE II - RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

4 RECUPERAÇÃO DE FORMAS ATRAVÉS DE CANTOS

Os primeiros sistemas de recuperação de imagens datam do início da década de 80 (CHANG; FU, 1980), e a maioria destes remete a sistemas como o QBIC da IBM (Consulta por conteúdo de imagem, do original em inglês *Query by Image Content*) como sendo o início da recuperação de imagens baseada em conteúdo (FLICKNER *et al.*, 1995). Por isso, o software QBIC é definitivamente o sistema mais difundido na área comercial. No entanto, a maioria dos sistemas disponíveis na atualidade são acadêmicos. Alguns exemplos bem conhecidos incluem Candid (KELLY *et al.*, 1995), Photobook (PENTLAND *et al.*, 1996) e Netra (MA *et al.*, 1997). Ambos usam características simples de cores e de textura para descrever o conteúdo das imagens. O sistema Blobworld (BELONGIE *et al.*, 1998) introduziu o uso de informações de mais alto nível, tais como partes segmentadas da imagem para consultas. Esse é um conceito importante para o estudo de CBIRs: as propriedades visuais de alto nível de uma imagem são aquelas em que o usuário tem a percepção da informação visual; enquanto que a descrição de baixo nível (vetores de características) é utilizada para a representação das imagens. A plataforma Pichunter (COX *et al.*, 2000) é um navegador de imagem que auxilia o usuário na busca de uma específica em um banco de dados, exibindo as imagens do usuário na tela que maximizam o ganho de informação em cada etapa de resposta. Um sistema que está disponível gratuitamente é a ferramenta GNU Image Finding (GIFT - <http://www.gnu.org/software/gift/>) (SQUIRE *et al.*, 2000).

A maioria desses sistemas têm arquiteturas semelhantes, quanto à navegação e arquivamento/indexação de imagens que compreendem ferramentas para a extração de recursos visuais, armazenamento e recuperação eficiente dessas características, medidas de distância ou de cálculo de similaridade e um tipo de interface gráfica do usuário. A configuração de um sistema genérico é mostrada na Figura 9. Os componentes mencionados são descritos com mais detalhes no decorrer do capítulo.

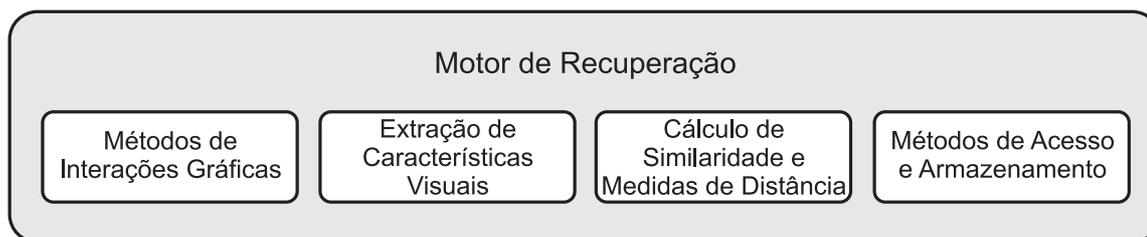


Figura 9: As principais componentes de todos os sistemas de recuperação de imagens baseada em conteúdo. Adaptado de (MULLER *et al.*, 2004).

4.1 Propriedades Visuais

As características visuais foram classificados em (PFUND; MARCHAND-MAILLET, 2001) como primitivas, nas quais estão cor ou a forma; recursos lógicos, tais como identidade dos objetos mostrados e características abstratas, como significado de cenas retratadas. Os sistemas atualmente disponíveis apenas usam recursos primitivos, menos anotação manual que é acoplada com as características visuais. Mesmo os sistemas que usam segmentos e características locais ainda estão longe de identificar objetos de maneira confiável. O usuário de sistemas CBIR pode se deparar com uma quantidade excessiva de dados irrelevantes na recuperação de imagem. Esse problema pode surgir por dificuldade na escolha da imagem adequada para consulta ou pela falta de descritores apropriados que permitam agregar a subjetividade da percepção visual. Essa questão está relacionada com a lacuna semântica, que representa a diferença entre características de baixo nível extraídas da imagem e o alto nível da informação que o usuário necessita (SMEULDERS *et al.*, 2000). Quanto mais um pedido de recuperação é especializado para um determinado domínio, então menor será a lacuna que pode ser identificada usando o conhecimento deste domínio. Outra diferença é a lacuna sensorial que descreve a perda real entre a estrutura e a representação da imagem (digital) (MULLER *et al.*, 2004).

4.1.1 Características Locais e Globais

A fusão de descritores tem sido usada para melhorar o desempenho de sistemas de recuperação de imagens. Algumas características inerentes a esses descritores podem ser globais ou locais, dependendo da análise do conteúdo visual. As características de cor, textura e de forma podem ser utilizadas em um nível global da imagem ou localmente, sobre determinadas regiões da imagem processada (SMEULDERS *et al.*, 2000; SQUIRE *et al.*, 2000). O sistema CBIR pode se concentrar na união de características locais e globais em um determinado descritor que permitam uma maior separabilidade entre os dados analisados (COMANICIU *et al.*, 1998). Esta tese aproveita esta possibilidade de associar características

em níveis distintos e analisa o impacto da combinação de descritores locais e globais em bases de imagens binárias.

4.1.2 Características de Forma

Em sistemas CBIR, adotar o conteúdo da forma de uma imagem como meta do processo de recuperação pode tornar essa abordagem mais difícil do que utilizar cor ou textura, especialmente quando o sistema requer a segmentação e o reconhecimento dos objetos contidos na imagem. Costa e Cesar (2009) apresentam diversas técnicas de análise e classificação de formas em que o processo de extração de suas características é definido com base em três tarefas: (a) organizar e visualizar as características; (b) extraí-las; e (c) mensurar as que foram extraídas dos objetos.

Existem duas abordagens para caracterizar a imagem pela forma (CASTAÑÓN; TRAINA, 2002). A primeira abrange técnicas que incluem uma descrição total do contorno do objeto (número de pontos de alta curvatura) e a outra utiliza informações sobre as características morfológicas da região. Esta tese se insere na abordagem que adota as características baseadas em cantos do contorno presentes na forma.

4.2 Técnicas de Comparação de Formas

Basicamente todos os sistemas CBIR fazem uso da premissa de equivalência de uma imagem e sua representação no espaço de características. Esses sistemas geralmente usam medidas de distância tais como o modelo de espaço vetorial euclidiano para mensurar distâncias entre uma imagem de consulta (descrita por suas características) e seus possíveis resultados (NIBLACK *et al.*, 1993; JAIN; VAILAYA, 1996). Estes sistemas representam todas as imagens como vetores de características em um espaço vetorial n -dimensional. O fato é que a despeito de seu uso, essas métricas de comparação entre imagens ainda não simulam a percepção visual humana (TVERSKY, 1977) com precisão. Outras medidas de distância também são definidas para um modelo de espaço vetorial, tais como a distância *city-block* (GONZALEZ; WOODS, 2007) ou a distância Mahalanobis (NIBLACK *et al.*, 1993). Ainda assim, o uso de espaços de características de elevadas dimensões não garante bons resultados, o que reforça o cuidado a ser tomado na escolha da medida de distância a ser adotada no processo de recuperação (HINNEBURG *et al.*, 2000). Estes problemas associados com a definição de similaridade em espaços de características de altas dimensões são conhecidos como o problema da dimensionalidade.

A relação entre a probabilidade da recuperação da imagem e as medidas de distância em um espaço vetorial foi introduzido por Vasconcelos e Lippmann (2000). Os autores argumentaram

que as medidas de distância em um espaço vetorial descritas na literatura se adequam, em princípio, à recuperação probabilística sob determinados pressupostos das distribuições de características. Outra técnica de recuperação probabilística é o uso de Máquinas de Vetor de Suporte (*Support Vector Machines - SVMs*) para a classificação das imagens em relevantes e não relevantes (GOH *et al.*, 2001).

Alguns métodos inerentes a um sistema CBIR são bem na literatura relacionada de recuperação de texto. Esses sistemas de recuperação são aplicadas a características visuais em que os elementos de alto nível correspondem aproximadamente às palavras de um texto (SQUIRE *et al.*, 2000; ZHU *et al.*, 2001). Este tipo de sistema se baseia em dois princípios: uma característica frequente em uma imagem descreve bem esta mesma imagem; e uma característica frequente na coleção de dados é um fraco indicador de dissimilaridade entre as imagens.

As diversas técnicas para recuperação de imagens, podem ser categorizadas em: busca por similaridade, por palavras-chave, aprendizagem de máquina, rastreamento por *Web Crawler* e realimentação de relevância. O diagrama apresentado na Figura 10 resume as diferentes técnicas de recuperação de imagens e a caixa em cinza destaca o enfoque adotado nesta tese. Estas técnicas têm por objetivo reduzir a lacuna semântica entre as características de baixo nível (descritores) e a expectativa do usuário.

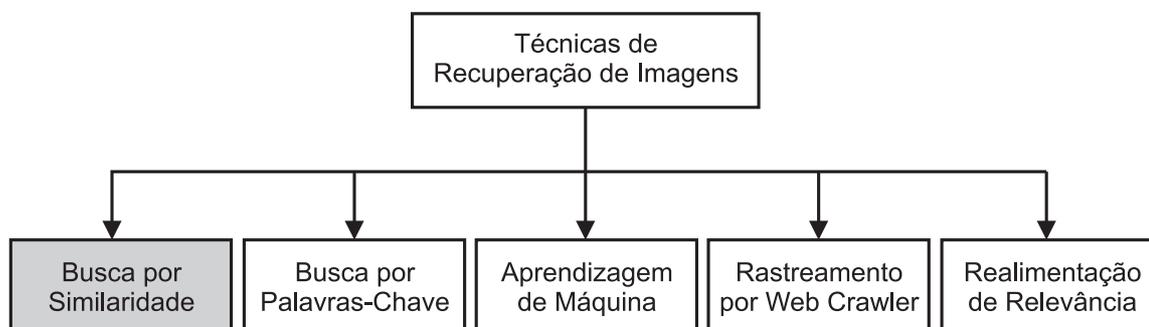


Figura 10: Principais técnicas para recuperação de imagens. Adaptado de Silva *et al.* (2011) e de Smeulders *et al.* (2000).

Na busca por similaridade (*query by example*), uma imagem inicial é utilizada como entrada para o sistema, servindo de base para a consulta. O algoritmo de pesquisa varia conforme a aplicação, e o objetivo é que as imagens resultantes compartilhem elementos ou características comuns com a imagem original. Neste tipo de pesquisa, o usuário apresenta uma imagem de consulta com uma característica desejada, dentre os exemplos as cores predominantes ou um desenho com uma forma aproximada do objeto desejado. Na sequência do processamento são extraídas as características da imagem de consulta para que sejam comparadas com aquelas armazenadas na base de imagens. O sistema retorna então a imagem ou o conjunto de imagens mais similares ao padrão inicial de consulta, dependendo de como este foi projetado

(Figura 11). Esses sistemas geralmente são projetados e desenvolvidos por cientistas e usados exclusivamente por especialistas.

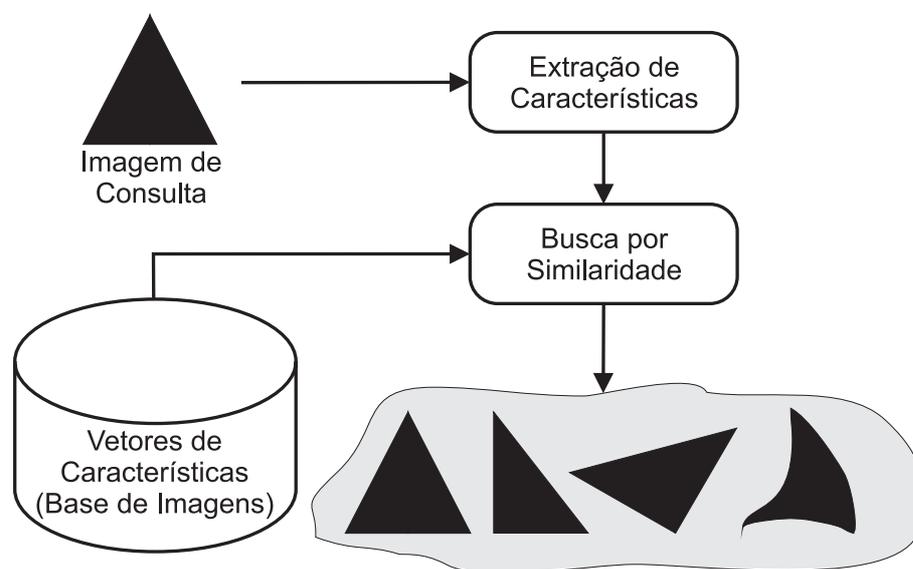


Figura 11: Busca por similaridade. Adaptado de Silva *et al.* (2011) e de Smeulders *et al.* (2000).

Na literatura, tem-se observado trabalhos (BELONGIE *et al.*, 1998; DORAIKARAJ; NAMUDURI, 2004; GEVERS; SMEULDERS, 2000) que utilizam mais de um descritor para recuperação de imagens, representando diferentes características da imagem (como por exemplo cor e textura). Para isso, os descritores podem ser combinados pela aplicação de diferentes pesos a cada um dos descritores utilizados (DORAIKARAJ; NAMUDURI, 2004; RUI *et al.*, 1998). Caso o resultado da busca não seja satisfatório, é possível avaliar qual é o descritor mais adequado, definir uma métrica diferente para cálculo de similaridade entre os vetores de característica ou ainda criar uma maneira de combinar diferentes descritores. Especialistas modificam o sistema com o objetivo de otimizá-lo para resolução do problema em questão. A técnica de busca por similaridade é normalmente utilizada em problemas específicos como reconhecimento de faces, impressões digitais, placas de veículos, entre outros. Uma arquitetura típica de um sistema de recuperação de imagens por conteúdo é mostrada na Figura 12 (TORRES; FALCÃO, 2006).

As buscas por similaridade podem ser úteis em diversos contextos, como verificar se um logotipo similar já foi registrado, buscar rostos suspeitos para prevenção de crime e detectar nudez em imagens ou vídeos. Um dos maiores desafios na área de recuperação de imagens é a criação de modelos computacionais que respondem questões como “mostre imagens de pessoas andando de bicicleta na praia”, reproduzindo assim a necessidade de um usuário comum. Como citado anteriormente, não existe ainda um sistema computacional que reproduza o processo de recuperação e busca visual realizada pelo cérebro humano. Por isso, diferentemente dos sistemas usados para resolver problemas de domínios específicos, alguns sistemas de

recuperação de imagens realizam buscas utilizando-se da experiência do usuário, mimetizando assim o conhecimento humano para busca de similaridade.

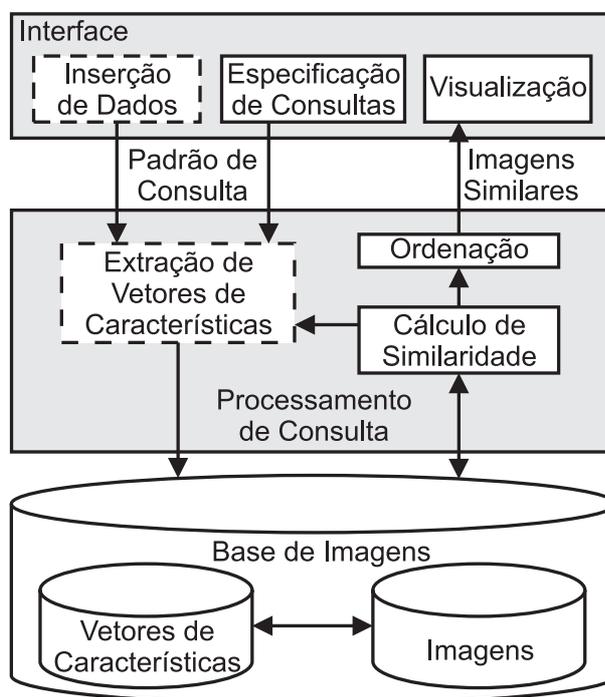


Figura 12: Arquitetura de um sistema de recuperação de imagens por conteúdo. Adaptado de Torres e Falcão (2006) e de Smeulders *et al.* (2000).

4.3 Definição do Sistema CBIR Adotado

O sistema CBIR proposto nesta tese se enquadra na classe de sistemas de recuperação da busca por similaridade. Não serão abordados métodos de interface gráfica, nem métodos de acesso e armazenamento, uma vez que dentre os objetivos deste trabalho consta o desempenho computacional do sistema proposto. Este mesmo desempenho é muito mais crítico nos componentes de extração de características visuais e no cálculo de similaridade e medidas de distância. Conseqüentemente, os trabalhos avaliados e abordados em termos de comparação também descrevem apenas estes componentes dos seus motores e utilizam a mesma técnica de recuperação. Como exposto anteriormente, as coleções de imagens digitais têm crescido continuamente nos últimos anos e a maneira mais frequente para recuperação de imagens pela Web ainda utiliza as informações textuais obtidas conjuntamente com as imagens (CAI *et al.*, 2004; FENG *et al.*, 2004).

A percepção dos objetos pela visão humana é subjetiva e dificulta a representação de uma imagem através de uma única característica. A cor é provavelmente a característica mais utilizada para recuperação visual, sendo relativamente robusta por apresentar independência do tamanho e da orientação da imagem. Textura é outra propriedade presente em praticamente

todas as estruturas, como nuvens, vegetação, paredes, cabelo entre outras, contendo informação importante sobre o arranjo estrutural da superfície. A forma é também uma importante informação utilizada pela visão humana no reconhecimento de objetos. Essa característica é aplicada em reconhecimento de caracteres, detecção e reconhecimento de pessoas em sistemas de segurança e rastreamento de objetos em vídeo. Assim, para a recuperação de imagens, atributos de cor, textura e forma podem ser extraídos de uma imagem e expressos através de um vetor de características. Esse vetor pode ser interpretado como um ponto no espaço \mathcal{R}^n , em que n é o tamanho do vetor de características e a busca das imagens poderá ser realizada, com base na distância entre o vetor de características da imagem de consulta e os vetores de características de imagens de uma base de dados.

O trabalho desenvolvido nesta tese admite que a recuperação de imagens consiste da combinação de um algoritmo de extração de vetores de características de uma forma e da função de distância entre vetores. A extração de características codifica o conteúdo de uma imagem no vetor de características, enquanto a função de distância define a similaridade entre dois vetores (ou dois descritores de forma) e, conseqüentemente, entre duas imagens. Essa similaridade é dada pelo inverso da função de distância, ou seja, quanto menor a distância entre os vetores de características das imagens, maior a similaridade entre elas.

O descritor global captura a informação diretamente da imagem e caso as características da imagem descrevam diferentes partes da mesma (regiões, bordas ou pontos de interesse), este descritor é denominado de local. Os vetores de características dos descritores locais são calculados sobre as partes, como por exemplo as regiões em torno de pontos de interesse. Tanto descritores locais quanto globais são aplicados no sistema de recuperação proposto.

4.4 Descrição de Formas

A metodologia adotada no sistema de recuperação de formas proposto está disposta na Figura 13. A partir de uma imagem de consulta, o sistema inicia com a detecção dos cantos da forma de entrada. Com esses cantos identificados, é possível gerar uma descrição da forma relacionando características obtidas a partir dos mesmos. Em seguida, o descritor produzido é comparado a todos os outros descritores definidos previamente a partir das bases de imagens. Por fim, o sistema retorna as imagens mais semelhantes de consulta através dos cálculos de medidas de similaridade realizados entre os descritores mencionados.

4.4.1 Sinal de Curvatura

A curvatura é uma característica que pode ser extraída, a partir de contornos de formas, e seu estudo tem sido motivado por apresentar forte relação biológica com o processamento da

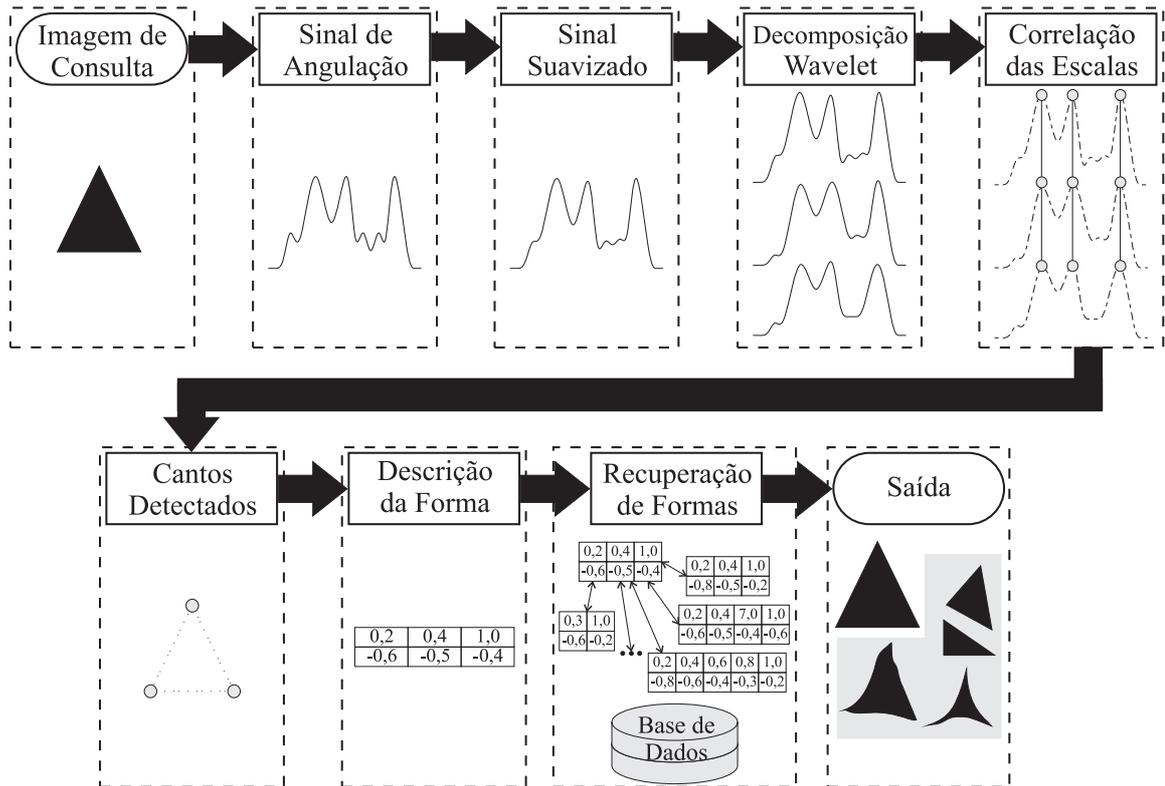


Figura 13: Metodologia do sistema CBIR proposto.

informação e percepção visual humana (CÉSAR; COSTA, 1996; PASSARINHO, 2005). Esta tese utiliza o método de estimação de curvatura digital proposto por César e Costa (1996), que se baseia nas propriedades da transformada de Fourier para obter a função de curvatura do contorno parametrizado da forma a ser analisada.

Seja P_t a forma parametrizada de uma dada curva regular P , em que $t \in \Gamma \subset \mathfrak{R}$, isto é, o parâmetro t assume valores sobre um intervalo Γ da curva \mathfrak{R} . A curvatura em qualquer ponto $P_t = (x(t), y(t))$, $t \in \Gamma$ é dada pela equação

$$k(t) = \frac{\dot{x}\ddot{y} - \ddot{x}y}{(\dot{x}^2 + \dot{y}^2)^{3/2}}, \quad (4.1)$$

em que $\dot{x} = dx/dt$, $\ddot{x} = d^2x/dt^2$, $\dot{y} = dy/dt$, $\ddot{y} = d^2y/dt^2$ são as derivadas primeira (\dot{x}, \dot{y}) e segunda (\ddot{x}, \ddot{y}) com relação ao parâmetro t .

A Figura 14(c) mostra a curvatura da forma apresentada na Figura 14(a). Os picos na Figura 14(c) correspondem às regiões em que o contorno na Figura 14(b) muda bruscamente de direção. Essa mudança caracteriza a presença dos pontos de alta curvatura, ou cantos da forma, no sinal de representação. Todo intervalo nulo corresponde às linhas retas presentes na forma. Os intervalos não-nulos representam arcos e semi-círculos nos contornos a serem analisados.

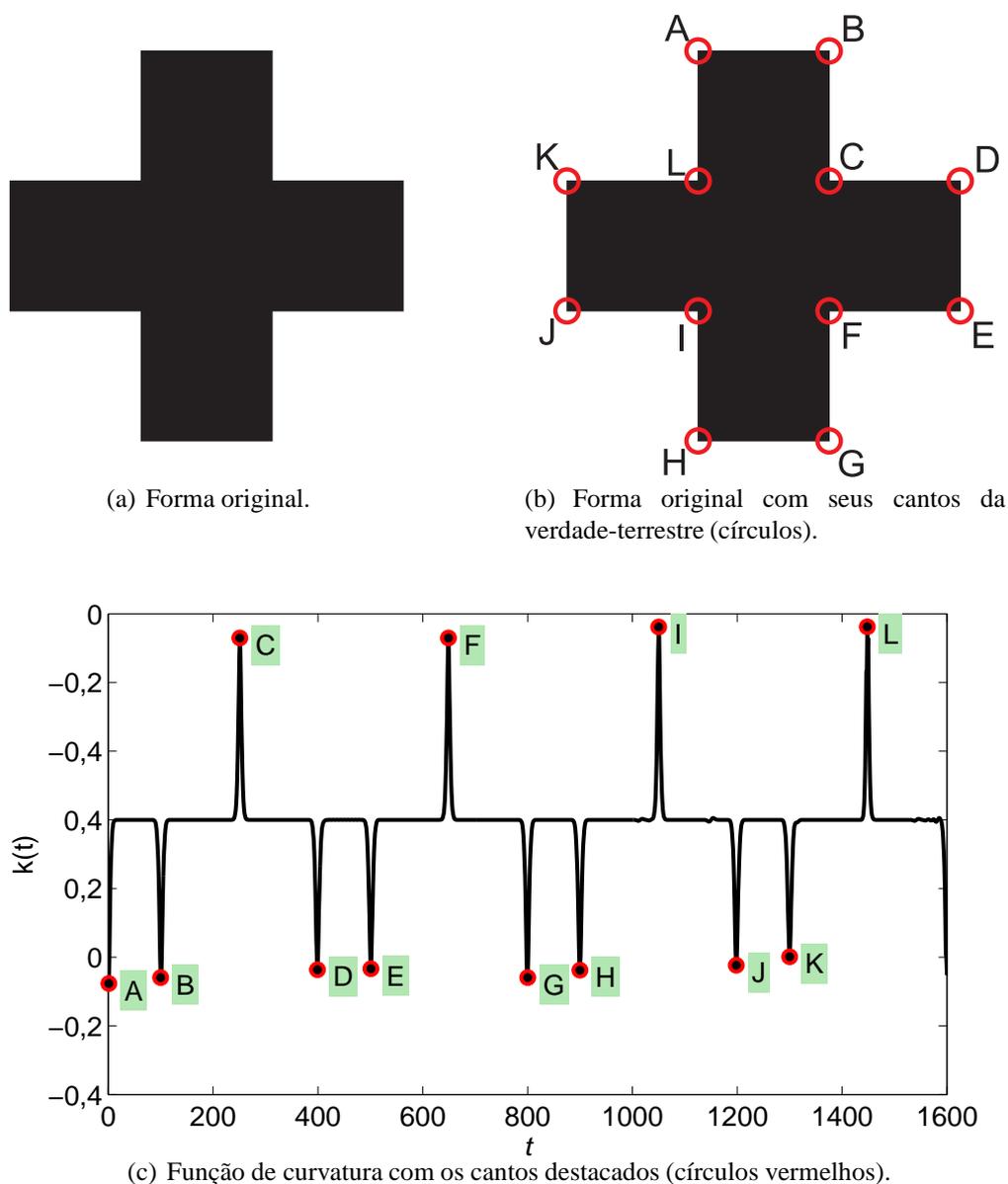


Figura 14: Exemplo de aplicação da função de curvatura.

4.4.2 Descritor de Formas por Cantos

Neste trabalho, os cantos da forma fornecem informações relevantes da mesma e a partir destes é possível derivar outras características, de modo que o poder de discriminação entre formas aumente. A técnica proposta utiliza a descrição da forma que relaciona posição angular e valor de curvatura dos cantos detectados.

Este trabalho aborda a descrição de cantos seguindo a proposta de Pedrosa, Barcelos e Batista (PEDROSA *et al.*, 2011a) que define cada canto detectado s_t em um vetor por um par $(s_{1,t}, s_{2,t})$, em que $s_{1,t}$ representa a posição relativa angular e $s_{2,t}$ o valor de curvatura. A posição

relativa angular de um canto é definida a partir da acumulação dos ângulos $c' \hat{c}_m s_t$, calculados com a posição do canto, o centróide da forma c_m e qualquer outro ponto arbitrário c' do contorno da forma. Para facilitar o entendimento do cálculo da posição relativa angular e sua relação com esse descritor, deve-se observar a disposição do centróide da forma e a localização dos cantos no contorno da mesma, como exibidos nas Figuras 15(a) e 15(b). O descritor de forma segue a definição disposta na Figura 15(c) de acordo com as informações oriundas dos cantos detectados, relacionadas as características da posição relativa angular e a curvatura oriunda dessa imagem. Estas informações são agrupadas em uma estrutura que será utilizada em um procedimento de correspondência entre descritores para alcançar a distância entre duas formas, como apresentada na Tabela 3.

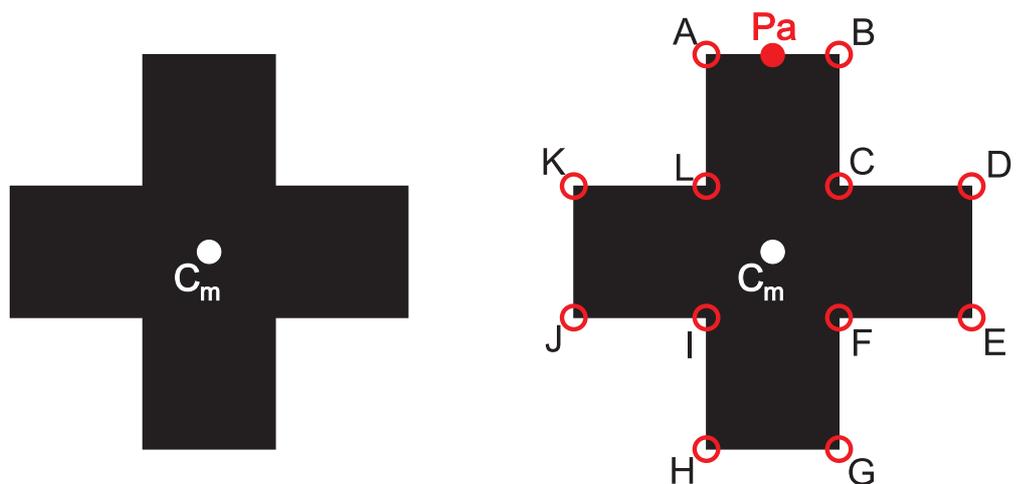
Tabela 3: Definição do descritor segundo a forma Figura 14(a). Campo (1) corresponde à posição relativa angular e (2) indica o valor de curvatura no canto.

(1)	0,0128	0,0254	0,0663	0,1358	0,2319	0,3567	0,5101	0,6646	0,7910	0,8880	0,9586	1,0000
(2)	-0,2581	-0,2295	0,2648	-0,2052	-0,2060	0,2503	-0,2295	-0,2132	0,2807	-0,2116	-0,1919	0,3810

4.5 Recuperação de Formas

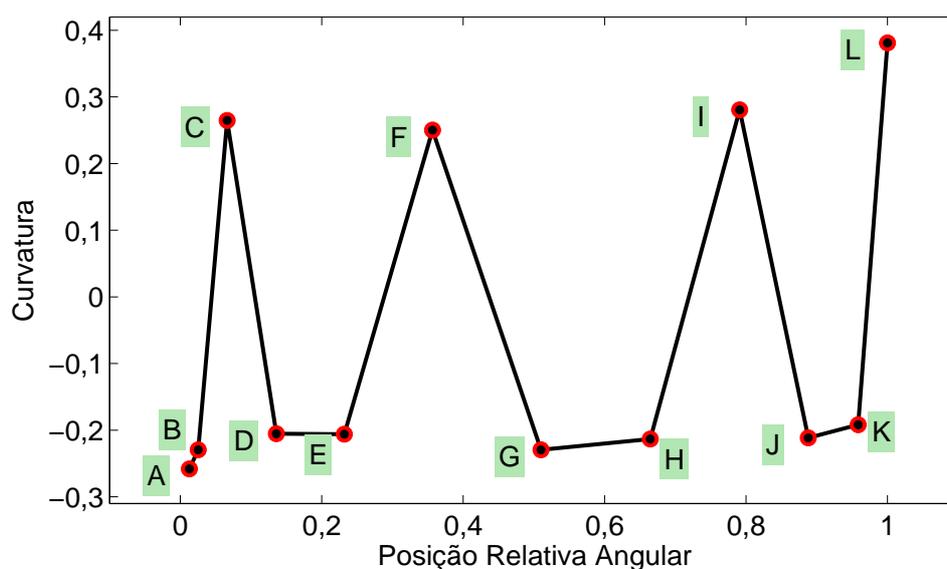
A análise de similaridade de formas mostra que mesmo que duas formas sejam representadas por contornos fechados ou funções paramétricas, estas podem não ter correspondência entre si (DATTA *et al.*, 2008). Por outro lado, duas formas que apresentam o mesmo contorno, descrito por diferentes pontos iniciais, nem sempre podem ser representadas da mesma maneira. Portanto, o cálculo de similaridade entre duas formas pode ser realizado de duas maneiras: através do cálculo da diferença entre características extraídas dos contornos das formas, que são invariantes ao ponto inicial e robusto a algumas transformações geométricas; ou ainda, através do cálculo de casamento entre as formas a fim de encontrar correspondências pontuais que identifiquem a distância entre os objetos. Esta segunda alternativa de medição de similaridade tornou-se a mais adotada recentemente, devido à sua capacidade de capturar propriedades intrínsecas às imagens e sem perder informação ao permitir comparar representação de imagens com dimensões distintas, conduzindo assim a medidas mais precisas.

Na literatura é possível encontrar trabalhos (ALAJLAN *et al.*, 2008, 2007) que adotam os cruzamentos de zero da representação área do triângulo (*Triangle-Area Representation, TAR*) na caracterização das formas de uma base de dados. Há ainda versões que abordam uma descrição *wavelet* multiescala para construir uma imagem de representação área do triângulo multiescala (*Multiscale Triangle-Area Representation, MTAR*) (ALAJLAN *et al.*, 2007; ALAJLAN, 2011). A correspondência *MTAR* entre imagens considera os picos e locais de máxima concavidade



(a) Forma original e seu centróide c_m (círculo branco).

(b) Forma original com um ponto arbitrário P_a e cantos correspondentes à verdade-terrestre (círculos vermelhos).



(c) Descritor com seus cantos destacados (círculos vermelhos).

Figura 15: Definição do descritor extraído de uma forma binária. A partir da (a) forma original e seu centróide calculado, detectam-se (b) os seus cantos e demarca-se um ponto arbitrário no contorno da forma. Com essas informações destacadas, determina-se o (c) descritor da forma segundo os cantos detectados.

nas imagens. A representação *TAR* é utilizada para descritores multiescala invariantes a transformações afins em contornos fechados, tal que as áreas triangulares em cada ponto de fronteira são aplicados na correspondência via programação dinâmica (*Dynamic Programming*, *DP*). A mesma representação fornece informações sobre características de forma, como a convexidade/concavidade em cada ponto de fronteira, portanto, *TAR* melhora a discriminação. Para a correspondência de formas, um algoritmo *DP* chamado de deformação espacial dinâmico

(*Dynamic Space Warping, DSW*), é empregado para encontrar o melhor alinhamento entre duas representações de forma.

Pedrosa *et al.* (2011a) descreveram formas planares usando pontos de canto, que são definidos como os pontos de alta curvatura ao longo do contorno da forma. Estes pontos foram adotados por representar a forma de uma maneira compacta e serem invariantes à translação. Neste trabalho mencionado, a forma é representada por um sinal definido pelas informações dos cantos relacionados à posição angular e valor de curvatura. Além disso, utilizou-se também um algoritmo de *DSW* para medir a similaridade entre duas formas caracterizadas por seus cantos. O trabalho de Pedrosa *et al.* (2011b) expandiu esta técnica agregando características globais e complexidade na descrição da forma.

A abordagem proposta nesta tese utiliza uma representação baseada em cantos que é sensível à seleção do ponto inicial ou à rotação da forma. Essa representação passa a ser descrita com informações relacionadas aos cantos identificados e aplicando uma análise de similaridade igualmente baseada em *DSW*. Na descrição da forma é adicionada uma característica global para destacar a relevância dos cantos côncavos em sua representação. A principal contribuição da técnica é a sua capacidade de fornecer uma maior acurácia na recuperação de formas que os demais métodos avaliados com base em experimentos sobre as bases de dados Kimia (SEBASTIAN *et al.*, 2004), MPEG-7 (CE-shape-1 parte B) (BOBER, 2001) e Tari (ASLAN *et al.*, 2008; BAI *et al.*, 2012).

A etapa de recuperação de formas é a mais complexa dentre as definidas na metodologia proposta. Os descritores de formas permitem comparar formas e calcular a distância elas. Para um conjunto de formas conhecidas, recuperam-se aquelas mais similares à imagem de consulta por apresentarem as menores distâncias em relação às demais calculadas na base de imagens. A técnica *DSW* fornece uma medida de distância que possibilita comparar descritores de tamanhos distintos. Associado a esse processamento, calculam-se ainda características de aspecto global para cada forma e uma medida de complexidade que permita que as distâncias entre os descritores se encontrem na mesma ordem de grandeza.

4.5.1 Deformação Espacial Dinâmica

Nos últimos anos, a modificação da deformação temporal dinâmica para correspondência e recuperação de formas recebeu relevante destaque em estudos e aplicações no domínio espacial. A abordagem adotada por essas metodologias é motivada pela eficácia do algoritmo *DSW* em realizar correspondências intuitivas na comparação entre os pontos de contorno de duas formas. Várias técnicas de correspondência de formas que alcançaram as maiores taxas de recuperação são baseadas em programação dinâmica (ALAJLAN *et al.*, 2007; ALAJLAN, 2011; PEDROSA *et al.*, 2011a, 2011b).

Na análise de similaridade, o algoritmo *DSW* efetua a correspondência por alinhamento não-linear, em que um ponto na forma F_A pode ser combinado com um ou mais pontos na forma F_B . Tal alinhamento compensa deformações não-rígidas em imagens de formas. O algoritmo *DSW* encontra o menor custo de deformação para alinhar os pontos de dois contornos. Esta deformação deve ser monotonicamente crescente, contínua e limitada (iniciando e terminando com o alinhamento dos pontos finais dos dois descritores) (ALAJLAN, 2011). Tais restrições asseguram que o algoritmo seja executado de maneira recursiva e eficiente, além de um alinhamento intuitivo. Assim, todos os pontos nos dois contornos são usados no processo de deformação.

Para o descritor adotado, os vetores correspondentes às formas F_A e F_B estão ordenados pelos seus respectivos valores de posição relativa angular. Sendo que n corresponde ao tamanho do descritor A , m o tamanho do vetor B e que $m \leq n$, a similaridade entre as formas representa a melhor correspondência (ou menor distância) entre os cantos de F_A e F_B . Assim, a distância entre as formas é calculada por:

$$D^{A,B}(\varphi) = \sum_{i=1}^n \|a_i - b_{\varphi(i)}\|. \quad (4.2)$$

em que $\|\bullet\|$ define a distância entre os descritores pela correspondência entre os atributos dos descritores de forma A e B , respectivamente. Nesta distância há um mapeamento dos atributos a_i , da forma A , sobre os atributos $b_{\varphi(i)}$ de B .

Para tornar a função de distância invariante à rotação, esta é calculada sobre todos os possíveis deslocamentos entre os vetores A e B . O valor do deslocamento é definido como a diferença entre o primeiro ponto do vetor B sobre todos os pontos do vetor A . Para que a medida de similaridade se confirme como invariante à rotação, deve-se encontrar o melhor deslocamento ρ tal que a distância $D^{A\rho,B}(\varphi)$ seja mínima. Portanto, a função de distância para esse valor deslocado é definida como segue:

$$D^{A\rho,B}(\varphi) = \sum_{i=1}^n \|a_i^\rho - b_{\varphi(i)}\|. \quad (4.3)$$

em que a_i^ρ representa os atributos deslocados do descritor de forma A .

Para que a função de distância seja invariante ao espelhamento da imagem, basta inverter a posição dos elementos associados à curvatura presentes no descritor A . Desse modo, a distância final entre A e B é dada pela menor distância obtida ao utilizar os valores originais e invertidos do descritor A , considerando todos os seus possíveis deslocamentos. Formalmente, a distância final entre os cantos de A e B é definida por:

$$d_s(A, B) = \min\left\{\min_{\rho}(D^{A\rho, B}(\varphi)), \min_{\rho}(D^{A\rho, B}(\varphi))\right\}. \quad (4.4)$$

A partir dos descritores das duas formas, é calculada uma matriz de diferenças r que relaciona a posição relativa angular aos valores de curvatura de cada descritor que funcionará como a implementação computacional da Equação 4.2. Essa matriz definirá o cálculo da função de correspondência φ o qual pode ser realizado através de DP . A matriz de diferenças r é dada por:

$$r = (r_{ij}) = ||a_i - b_j||. \quad (4.5)$$

Calcular a similaridade entre as formas F_A e F_B torna-se um problema de minimização da matriz r . A solução é encontrar o caminho de custo mínimo sobre essa matriz de diferenças. Assim, a minimização da Equação 4.5 pode ser obtida através da função recursiva que segue:

$$M(i, j) = r(i, j) + \Delta, \quad (4.6)$$

$$\Delta = \begin{cases} 0, & \text{if}(i = 1) \text{ and } (j = 1); & \mathbf{(1)} \\ M(i, j - 1), & \text{if}(i = 1) \text{ and } (j > 1); & \mathbf{(2)} \\ M(i - 1, j), & \text{if}(i > 1) \text{ and } (j = 1); & \mathbf{(2)} \\ \min(M(i - 1, j), M(i - 1, j - 1)), & & \\ M(i, j - 1)), & \text{if}(1 < i < n) \text{ and } (1 < j < m). & \mathbf{(3)} \end{cases} \quad (4.7)$$

O valor calculado na Equação 4.6 indica a distância entre os descritores das duas formas. A otimização computacional deste processamento está na utilização de DP para o cálculo de $M(n, m)$, como pode ser observado na Figura 16. A fim de tornar o método invariante à seleção do ponto inicial do contorno, repete-se este mesmo cálculo de distância com a rotação do descritor B . Armazena-se ao fim do processamento a menor distância entre todas possibilidades de seleção do ponto inicial do contorno, sendo a distância final ($d_s(A, B)$) a soma das distâncias calculadas para as componentes dos dois descritores.

4.5.2 Medida de Distância

Para alcançar a distância entre duas formas no sistema de recuperação proposto, calcula-se a diferença entre as informações dos cantos detectados em duas formas e suas respectivas características globais, como apresentado anteriormente em (ALAJLAN *et al.*, 2007). Considerando duas formas, F_A e F_B , e levando em conta seus respectivos descritores A e B , a distância entre as imagens é dada por:

(1)	$r_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> </table>	1	1	1	1	1	1	1	1	1	1	1	1	$M_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	1	0	0	0	0	0	0	0	0	0	0	0															
1	1	1	1																																						
1	1	1	1																																						
1	1	1	1																																						
1	0	0	0																																						
0	0	0	0																																						
0	0	0	0																																						

(2)	$r_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> </table>	1	1	1	1	1	1	1	1	1	1	1	1	$M_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>2</td><td>3</td><td>4</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td></tr> </table>	1	2	3	4	0	0	0	0	0	0	0	0	$M_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>2</td><td>3</td><td>4</td></tr> <tr><td>2</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>3</td><td>0</td><td>0</td><td>0</td></tr> </table>	1	2	3	4	2	0	0	0	3	0	0	0		
1	1	1	1																																						
1	1	1	1																																						
1	1	1	1																																						
1	2	3	4																																						
0	0	0	0																																						
0	0	0	0																																						
1	2	3	4																																						
2	0	0	0																																						
3	0	0	0																																						

(3)	$r_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>1</td></tr> </table>	1	1	1	1	1	1	1	1	1	1	1	1	$M_{n \times m}$ <table border="1" style="display: inline-table; border-collapse: collapse;"> <tr><td>1</td><td>2</td><td>3</td><td>4</td></tr> <tr><td>2</td><td>2</td><td>3</td><td>4</td></tr> <tr><td>3</td><td>3</td><td>3</td><td style="background-color: #f08080;">4</td></tr> </table>	1	2	3	4	2	2	3	4	3	3	3	4															
1	1	1	1																																						
1	1	1	1																																						
1	1	1	1																																						
1	2	3	4																																						
2	2	3	4																																						
3	3	3	4																																						

Figura 16: Exemplo de aplicação da técnica *DSW* sobre uma matriz de diferenças r_{ij} : (1) nova matriz M_{ij} contendo a primeira iteração da Equação 4.7; (2) M_{ij} atualizada com a segunda e terceira iteração da recursividade de *DSW* e (3) M_{ij} completa após o processamento da quarta iteração de Δ e identificação do custo mínimo no valor de $M(n, m)$ (destacado em vermelho).

$$D_f(F_A, F_B) = d_{g_\alpha}(F_A, F_B) + \frac{d_s(A, B)}{c_f(A, B)}, \quad (4.8)$$

em que $(d_{g_\alpha}(F_A, F_B))$ corresponde à distância entre as características globais das duas formas e a componente $c_f(A, B)$ é definida como a distância de custo mínimo d_s normalizada pela complexidade da forma (*Shape Complexity, SC*) de cada contorno. A motivação por trás desta normalização tem por base a observação de que a sensibilidade da percepção humana para as variações de contorno diminui com o aumento da *SC*. Visando aumentar a capacidade de discriminação da distância, combinam-se as características globais, com propriedades geométricas que favorecem a discriminação de formas, e que incluem relação de aspecto (*AR*), excentricidade (*E*) e solidez (*S*), como descrito na Equação 4.9 que segue (ALAJLAN *et al.*, 2007, 2008):

$$d_{g_1}(F_A, F_B) = |AR_A - AR_B| + |E_A - E_B| + |S_A - S_B|. \quad (4.9)$$

A proporção de concavidade é inserida como uma característica global no conjunto de informações da forma. Esta medida relaciona a quantidade de cantos côncavos detectados, n_{conc} , e a quantidade total de cantos identificados em uma forma, n_{all} , na proporção $\left(\frac{n_{conc}^T}{n_{all}^T}\right)$ relacionada ao descritor da forma T . A inclusão dessa característica incrementa a distância entre formas que pertencem a classes distintas e estende o conceito anteriormente apresentado na definição da medida de distância, sendo redefinido segundo:

$$d_{g_2}(F_A, F_B) = d_{g_1}(F_A, F_B) + \left| \frac{n_{conc}^A}{n_{all}^A} - \frac{n_{conc}^B}{n_{all}^B} \right|. \quad (4.10)$$

Em análises psicofísicas descritas em (BARENHOLTZ *et al.*, 2003; BARENHOLTZ; FELDMAN, 2003), os cantos côncavos alcançam maior vantagem em sua percepção e relevância em relação às mudanças convexas de igual magnitude. Os autores neste artigo concluíram que a percepção dos observadores, quando pontos de concavidade ao longo do contorno foram introduzidos ou removidos, em comparação com pontos de convexidade na mesma situação. Esta diferença de sensibilidade visual sugere que seções côncavas e convexas de um contorno tenham representações perceptivas consideravelmente diferentes, não obstante suas geometrias sejam semelhantes. A Figura 17 dispõe de uma amostra com mudanças em seu contorno de acordo com a adição ou remoção de uma nova região na forma. Esta região pode assumir um ponto de alta curvatura com definição côncava ou convexa. Os cantos côncavos apresentam um valor de curvatura negativo, como mencionado em (BARENHOLTZ *et al.*, 2003).

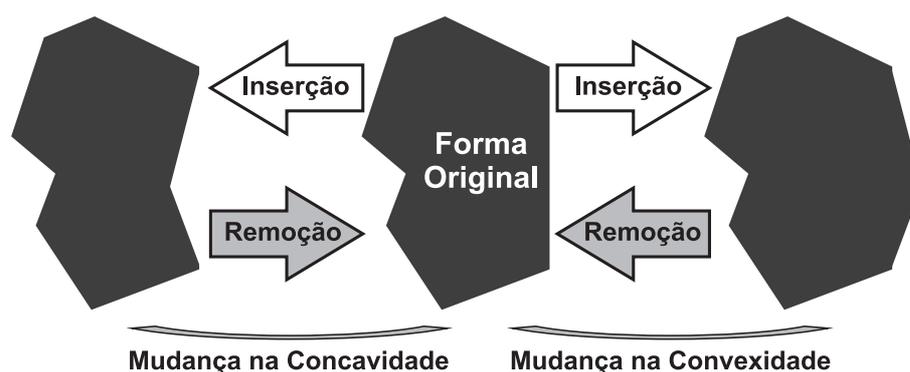


Figura 17: Conjuntos de formas geradas com a inserção/remoção de uma região com canto côncavo ou convexo. Adaptado de Barenholtz *et al.* (2003).

4.5.3 Aperfeiçoamento da Técnica DSW

Tendo por objetivo aprimorar o desempenho computacional da técnica de recuperação proposta, realizou-se neste tese o aperfeiçoamento da técnica *DSW* de acordo com a técnica *HopDSW* definida em (ALAJLAN, 2011). O algoritmo *HopDSW* trata da solução do problema da invariância de rotação ou a seleção do ponto inicial para correspondência de forma baseada em contorno usando a técnica *DSW* básica. Deve-se observar que este algoritmo funciona para qualquer sequência unidimensional, representando formas, e que o mesmo é sensível à seleção do ponto inicial. Ao invés de executar a *DSW* básica para todas as possibilidades de seleção do ponto inicial do contorno, somente pontos com maior probabilidade de corresponder à solução de custo mínimo são pesquisados em duas fases: grosseira e refinada (*coarse-to-fine search*

Algoritmo 1 DSW Aproximado (algoritmo principal): $dist = HopDSW(A, B, h_s)$

Adaptado de (ALAJLAN, 2011).

Entrada: A e B são os descritores das formas comparadas. N é o número de pontos do contorno da primeira forma. h_s é o passo de salto.

Saída: $dist$ é a distância entre os descritores A e B .

```

1:  $A \leftarrow [A \ A \ A]$  // permite deslocamento cíclico
2:  $dist \leftarrow \infty$ 
   // busca grosseira
3: for  $i = 1$  to  $\lceil N/h_s \rceil$  do
4:    $c \leftarrow (i - 1) \times h_s + 1 + N$ 
5:    $d \leftarrow DSW(A(c : c + N - 1), B)$ 
6:   if  $d < dist$  then
7:      $dist \leftarrow d$ 
8:      $t \leftarrow c$ 
9:   end if
10: end for
   // busca refinada
11: for  $i = t - h_s + 1$  to  $t + h_s - 1$  do
12:    $d \leftarrow DSW(A(i : i + N - 1), B)$ 
13:   if  $d < dist$  then
14:      $dist \leftarrow d$ 
15:   end if
16: end for
17: Refletir descritor  $A$  e repetir os passos 3-16
18: return  $dist$ 

```

techniques).

Algoritmo 2 $dist = DSW(A, B)$

Adaptado de (ALAJLAN, 2011).

Entrada: w é a dimensão da matriz avaliada na programação dinâmica, em que $w = 3$ (SAKOE; CHIBA, 1978). D_T é uma matriz de distâncias $N \times N$ inicializada como:

$$D_T(i, j) = \begin{cases} 0 & \max(1, i - w + 1) \leq j \leq \min(N, i + w - 1) \\ \infty & \text{caso contrário} \end{cases}$$

Saída: d é a distância DSW entre os descritores das formas A e B .

```

1: for  $i = 1$  to  $w$  do
2:    $D_T(i, 1) \leftarrow |A(i) - B(1)|$ 
3:    $D_T(1, i) \leftarrow |A(1) - B(i)|$ 
4: end for
5: for  $i = 1$  to  $N$  do
6:   for  $j = \max(1, i - w + 1)$  to  $\min(N, i + w - 1)$  do
7:      $p \leftarrow \min([D_T(i - 1, j) \ D_T(i - 1, j - 1) \ D_T(i, j - 1)])$ 
8:      $D_T(i, j) \leftarrow |A(i) - B(j)| + p$ 
9:   end for
10: end for
11: return  $d \leftarrow D_T(N, N)$ 

```

Um pseudo-código do algoritmo *HopDSW* é dado no Algoritmo 1. O algoritmo aceita como entrada os descritores das formas A e B, em que $n \in \{1, N\}$, e um passo de salto $h_s \geq 1$. O passo de salto (*hopping step*) controla o número de pontos de contorno a serem ignorados durante a busca grosseira do ponto ótimo de partida. Na etapa grosseira, um grupo de pontos candidatos de partida, com cada par consecutivo separado por $h_s - 1$ pontos, é utilizado para processar o algoritmo básico de *DSW* que está disposto em Algoritmo 2. Portanto, o número destes pontos é igual a $\lceil N/h_s \rceil$, em que $\lceil \cdot \rceil$ é a função teto. O ponto t do contorno, que produz a distância mínima é passado para a fase de busca refinada, em que a mesma é retomada dentro dos pontos circundantes até $h_s - 1$ pontos em cada sentido. Portanto, o número de pontos de partida executados nesta fase é de $2(h_s - 1)$.

O algoritmo *DSW* básico é descrito no pseudo-código do Algoritmo 2. Este algoritmo inicia com uma matriz de distâncias, D_T , cujas dimensões são iguais aos comprimentos das duas sequências a serem comparadas. A distância d entre dois pontos do contorno A_i e B_j é definida na Equação 4.8. Os elementos da primeira linha e da coluna de D_T são calculados como a distância entre os pontos correspondentes. Em seguida, os elementos restantes são calculados como definido na Equação 4.6.

A distância entre as formas A e B resulta no valor $D_T(N, N)$, o qual corresponde ao caminho de menor custo entre seus descritores de acordo com o ponto inicial selecionado. Considerando a transformação de reflexão da forma, é suficiente inverter sua assinatura e repetir o algoritmo. Um exemplo ilustrativo do mecanismo de busca do ponto ótimo de partida está representado na Figura 18, em que *HopDSW* representa a técnica *DSW* clássica que inclui todos os possíveis pontos de partida de uma pesquisa grosseira, como mostrado na Figura 18(a). Encontrar o ponto de partida ótimo, $n = 29$, implica na execução do algoritmo básico *DSW* por 94 vezes. Com relação à pesquisa grosseira do algoritmo *HopDSW* é utilizado um passo de salto $h_s = 8$ e retornado $n = 23$ como o ponto de distância mínima, como mostra na Figura 18(b). Em seguida, o estágio de refinamento, mostrado na Figura 18(c), busca em pontos circundantes $n = 23$ e retorna o ponto ideal, que é de $n = 29$. Ao fim, notou-se neste exemplo que apenas 23 execuções do algoritmo básico *DSW* foram necessárias pelo algoritmo *HopDSW*.

Vale ressaltar que a seleção do parâmetro de passo h_s é crítica para o desempenho do algoritmo *HopDSW*. Seja $F(h_s)$ a representação do número de mudanças dos pontos iniciais de contorno calculados pelo algoritmo *HopDSW* em um dado h_s . Então, F é dada por:

$$F(h_s) = \left\lceil \frac{N}{h_s} \right\rceil + 2(h_s - 1), \quad (4.11)$$

em que $h_s \in \{1, \lfloor N/2 \rfloor - 1\}$, e $\lceil \cdot \rceil$ e $\lfloor \cdot \rfloor$ são as funções teto (*ceiling*) e piso (*floor*), respectivamente. Enquanto h_s aumenta, o número de mudanças de pontos iniciais do contorno executadas durante a busca grosseira diminui (representado pelo primeiro termo da Equação

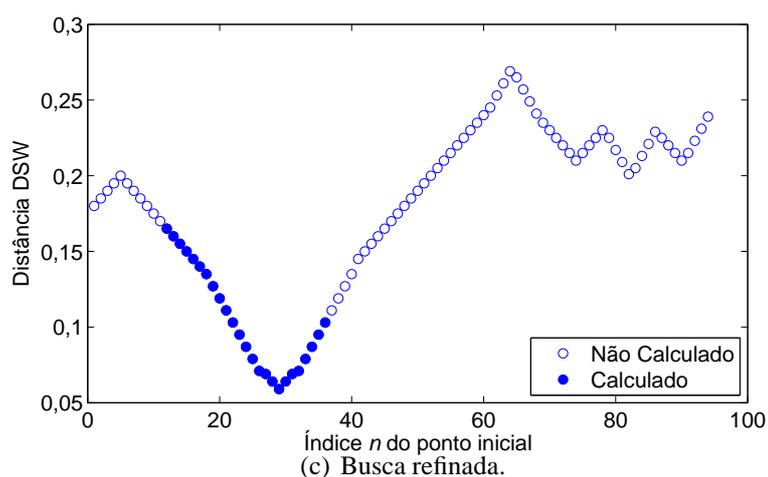
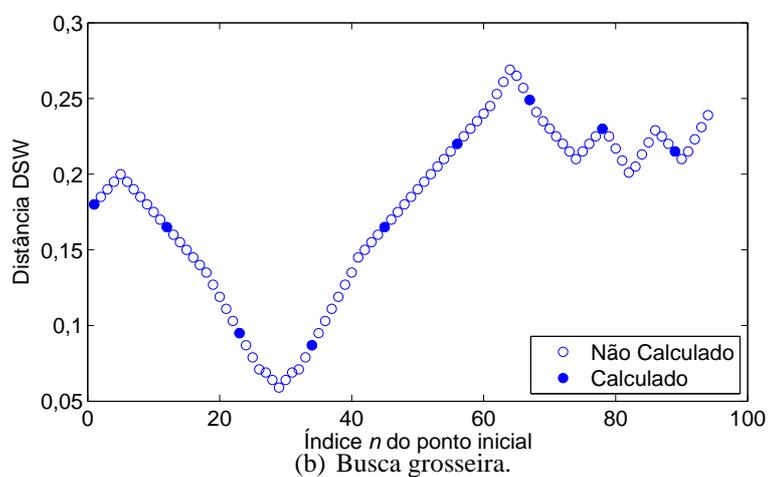
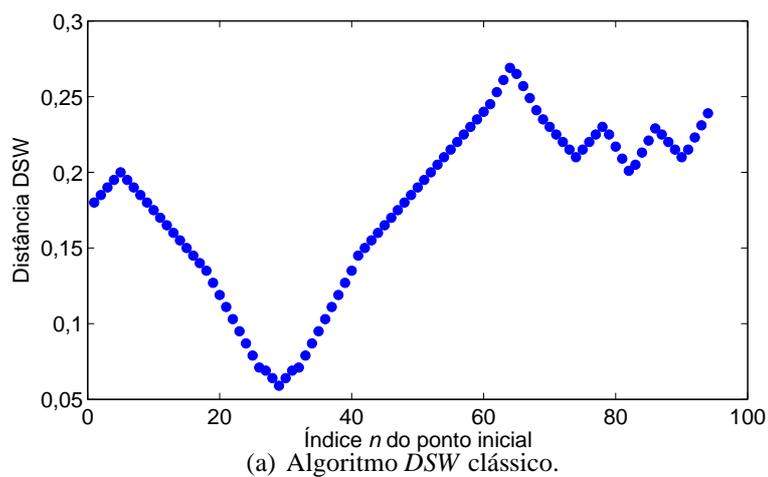


Figura 18: A relação entre a distância *DSW* e a seleção do ponto de partida para duas formas obtida a partir do (a) algoritmo *DSW* clássico, e (b) a busca grosseira e (c) o estágio refinado do algoritmo *HopDSW*. Adaptado de Alajlan (2011).

4.11) e este mesmo número aumenta com a busca refinada (representada pelo segundo termo da Equação 4.11). Observa-se que a função F possui seu valor máximo igual a N (quando $h_s = 1$),

enquanto que $h_s \geq \lfloor N/2 \rfloor$ não deve ser considerado, porque o valor de $F(h_s)$ ultrapassará N , o que significa uma repetição desnecessária de alguns pontos iniciais do contorno. Nota-se também que a função da Equação 4.11 tem um mínimo único uma vez que os primeiro e segundo termos são inversamente monotônicos. Alajlan (2011) determina que o valor de h_s a ser escolhido deve minimizar a função F . O método proposto modifica esta versão original de *HopDSW* adequando a mesma para descritores de dimensões distintas. A metodologia inicial desse aperfeiçoamento da técnica *DSW* foi definida apenas para descritores de mesma dimensão (ALAJLAN, 2011).

4.6 Complexidade Computacional Temporal

A complexidade do algoritmo *HopDSW* merece discussão dado que um dos objetivos deste trabalho é aperfeiçoar o desempenho computacional da metodologia publicada em (ALAJLAN *et al.*, 2007, 2008; ALAJLAN, 2011; PEDROSA *et al.*, 2011b) para o algoritmo proposto. O algoritmo *DSW* básico alcança uma complexidade temporal na ordem de $O(N \log N)$ em uma só comparação (uma posição do ponto inicial) entre dois descritores. Este mesmo algoritmo, ao considerar todas as possibilidades de seleção do ponto inicial do contorno, possui complexidade $O(N^2 \log N)$. Alternativamente, o número de mudanças do ponto inicial do contorno considerado pelo algoritmo *HopDSW* é definido pela Equação 4.11. Como definido em (ALAJLAN, 2011), o número de mudanças do ponto inicial do contorno executadas pelo algoritmo *HopDSW* é menor que N por uma ordem $O(\log N)$. Portanto, a complexidade do algoritmo torna-se $O(N \log^2 N)$.

Tabela 4: Complexidade temporal das técnicas de comparação de descritores de formas.

Método	Complexidade
<i>DSW</i> clássico	$O(N^2 \log N)$
<i>HopDSW</i> (mesma dimensão)	$O(N \log^2 N)$
<i>HopDSW</i> (dimensões distintas)	$O(N \log^2 N)$

No sistema de recuperação de imagens proposto, adaptou-se esse algoritmo *HopDSW* para descritores com dimensões distintas sem interferência nos *loops* do algoritmo que mais oferecem custo computacional. Dessa maneira, a complexidade do algoritmo proposto mantém-se em $O(N \log^2 N)$, como descrito na Tabela 4. Para todos os métodos avaliados, N representa a dimensão do maior descritor utilizado na técnica de correspondência entre formas.

4.7 Considerações Finais

Este capítulo apresentou a metodologia proposta que integra detecção de cantos em formas e a técnica de recuperação *DSW*. Na definição desta metodologia observou-se a necessidade de aprimoramento do desempenho da técnica de recuperação de formas apresentada. Foram abordados conceitos de descrição de formas, proporção de concavidade, deformação espacial dinâmica e sua versão aperfeiçoada.

O próximo capítulo apresenta discussões sobre as simulações que envolvem a técnica de recuperação de formas proposta.

5 RESULTADOS EXPERIMENTAIS DE RECUPERAÇÃO DE FORMAS

Este capítulo apresenta os experimentos e a avaliação de desempenho da técnica proposta para a recuperação de formas em comparação com outros métodos baseados em cantos. Inicialmente, realizou-se um experimento para confirmar que a influência dos cantos convexos é maior que a contribuição dos cantos côncavos nas características globais da recuperação das imagens proposta. O objetivo nesse momento é destacar a relevância da proporção de concavidade para os experimentos seguintes. Na sequência, outros experimentos foram efetuados para ratificar o melhor desempenho do método proposto em relação a outros trabalhos disponíveis na literatura.

5.1 Análise da Recuperação de Formas

A análise comparativa do desempenho dos algoritmos na recuperação de formas foi realizada em termos das medidas *Precisão* e *Revocação*. A *Precisão* representa a razão entre o número de imagens relevantes recuperadas e o número total de imagens recuperadas. A *Revocação* é dada pela razão entre o número de imagens relevantes recuperadas e o número de imagens relevantes na base de dados. Considerando imagem relevante aquela que pertence à mesma classe do objeto avaliado, cada uma das imagens, em cada base de dados, foi utilizada como consulta.

No primeiro experimento, foram propostas três abordagens para um sistema de recuperação de formas. Todos seguiam a metodologia definida no capítulo anterior, associando características locais e globais através das informações dos cantos das formas. A primeira proposta, identificado a seguir como **Proposto 01**, utilizou a proporção de concavidade definida na Subseção 4.5.2 para o cálculo da distância entre duas formas. A segunda abordagem (**Proposto 02**) realizou o mesmo processamento da proposta anterior mas utilizando os cantos convexos no mesmo raciocínio da proporção de concavidade. A terceira e última abordagem (**Proposto 03**) seguiu a mesma técnica envolvida em **Proposto 01**, mas excluindo a proporção de concavidade de suas características globais. As abordagens **Proposto 01**, **Proposto 02**

e **Proposto 03** estão identificadas dessa maneira nas Figuras 19 e 20, que são discutidas na próxima seção.

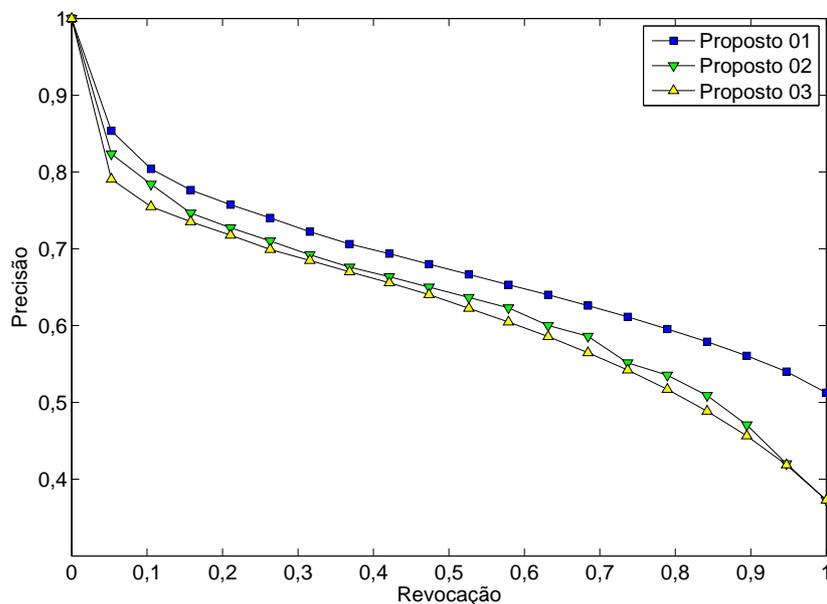


Figura 19: Curvas *Precisão-Revocação* - Base de dados Tari 1000. Resultados da recuperação de imagens das três abordagens propostas.

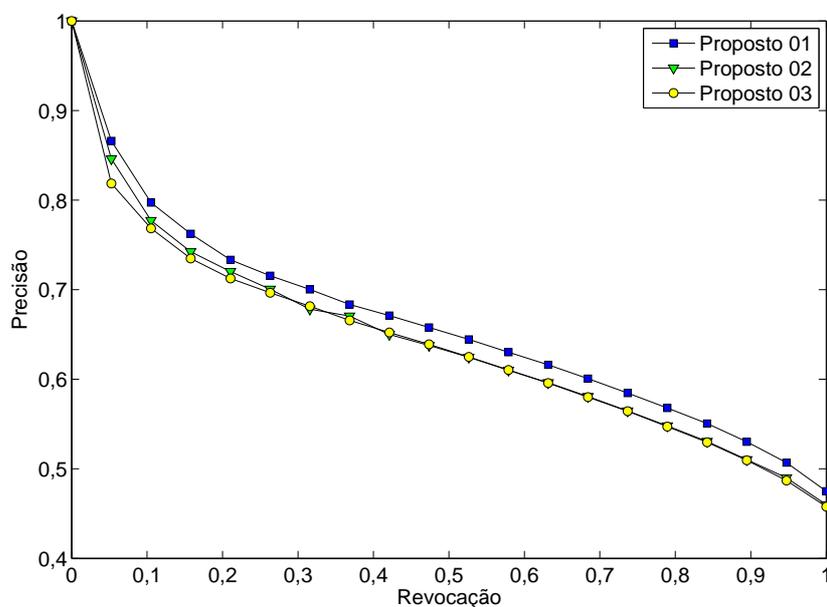


Figura 20: Curvas *Precisão-Revocação* - Base de dados MPEG-7 Part B. Resultados da recuperação de imagens das três abordagens propostas.

Após estes experimentos, procedeu-se uma nova sequência de experimentos envolvendo a melhor abordagem proposta com as demais técnicas de recuperação de imagens presentes na

literatura. Dentre estes últimos, o primeiro método foi introduzido por Alajlan *et al.* (2007) e é definido a partir da representação da área de triângulo (do termo original *Triangle-Area Representation* e nomeada como **TAR**). As duas outras técnicas possuem descritores baseados em cantos da forma (PEDROSA *et al.*, 2011a, 2011b): a primeira delas utiliza apenas a *DSW* nos descritores (identificada como **PedrosaA**) e a segunda agrega informações globais aos descritores definidos por cantos (**PedrosaB**). Todos os experimentos mencionados foram conduzidos utilizando três conjuntos de dados; a saber, base de dados Kimia (SEBASTIAN *et al.*, 2004), com 216 imagens divididas em 18 classes (12 imagens em cada classe); a base de dados MPEG-7 Parte B (BOBER, 2001) que contém 1400 formas distribuídas em 70 classes (20 imagens em cada classe); e a base de dados Tari (ASLAN *et al.*, 2008) que possui três coleções de dados contendo 56, 180 e 1000 imagens com 14, 30 e 50 classes em cada uma, respectivamente.

Para o método proposto, assumiu-se que o mesmo utilizou a proporção de concavidade e foi identificado como **Proposto**. Outros experimentos foram realizados de modo que a característica global de proporção de concavidade (denominado como **PropostoSemConc** nesta seção) foi removida. Por sua vez, no método **PedrosaB** foi adicionada essa mesma característica no cálculo da distância (aqui definida como **PedrosaBConc**). Essas modificações foram aplicadas com o objetivo de analisar a relevância da contribuição dessa característica às duas abordagens. As abordagens **TAR**, **PedrosaA**, **PedrosaB**, **PedrosaBConc**, **Proposto**, **PropostoSemConc** estão identificadas dessa maneira nas Figuras 21 e 22, que são discutidas na próxima seção.

As Figuras 21 e 22 mostram as curvas médias *Precisão-Revocação* para os métodos avaliados nos conjuntos de imagens mais complexos (maior variabilidade de dimensões de descritores de formas), respectivamente relacionados às bases de dados Tari 1000 e MPEG-7.

A técnica proposta apresentou a melhor curva *Precisão-Revocação* dentre os métodos e bases de dados testados. Com o auxílio da proporção de concavidade, este método alcançou o melhor desempenho na relação entre *Precisão* e *Revocação* que as demais abordagens para a base de imagens MPEG-7. Ao mesmo tempo, o incremento da proporção de concavidade no método PedrosaB promoveu uma melhora em seu desempenho, como pode ser observado na Figura 22 e suas respectivas curvas *Precisão-Revocação*. O método proposto também exibe melhor desempenho quando aplicado a base de imagens Tari, e se torna mais bem sucedido na recuperação de formas com o incremento da proporção de concavidade. Da mesma forma, a abordagem PedrosaB sobre a mesma base de dados apresentou melhor desempenho, como destaca a Figura 21.

Outra medida de avaliação de desempenho utilizada neste trabalho foi a taxa de *Bull's eye* (BAI *et al.*, 2010; EGOZI *et al.*, 2010; BAI *et al.*, 2012). Considere um conjunto de N_t formas agrupadas em c_t classes contendo t diferentes formas. Cada forma nessa base de dados

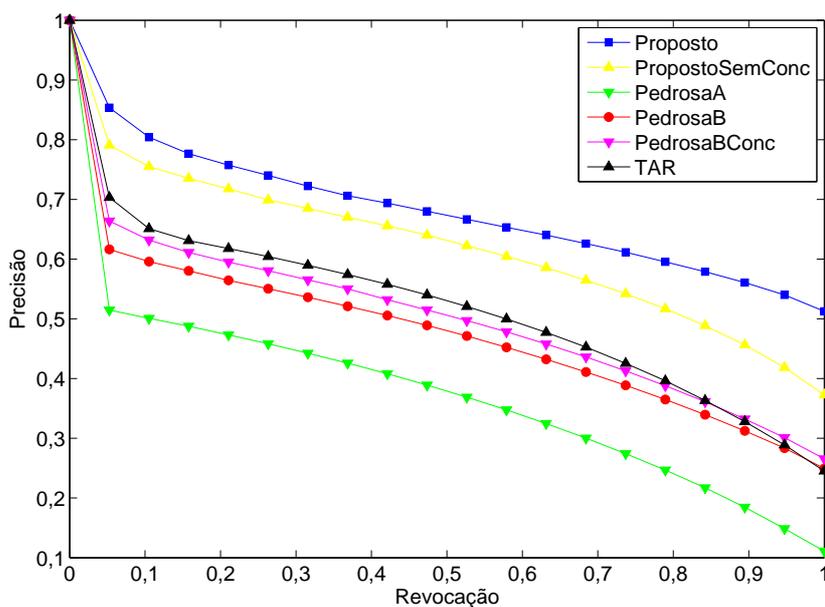


Figura 21: Curvas *Precisão-Revocação* - Base de dados Tari 1000. Resultados da recuperação de imagens entre o método proposto e outras abordagens presentes na literatura.

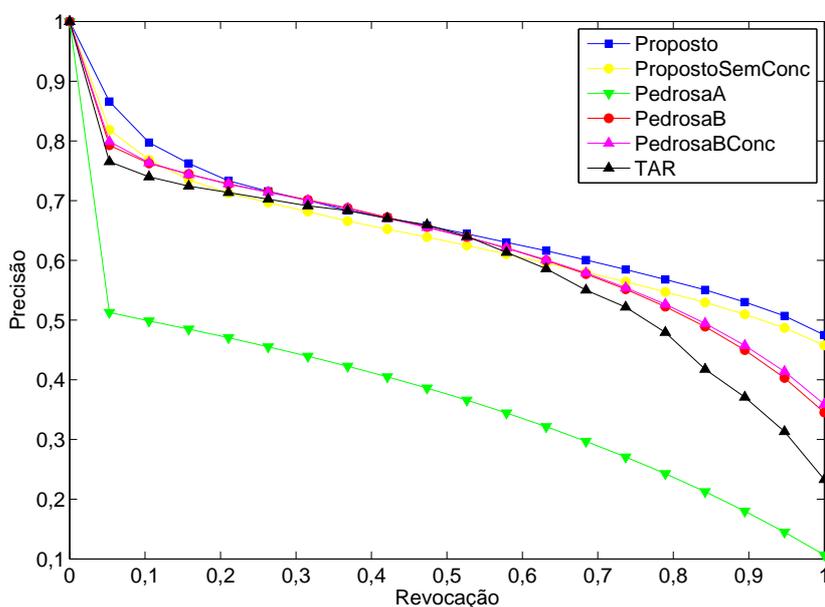


Figura 22: Curvas *Precisão-Revocação* - Base de dados MPEG-7 Part B. Resultados da recuperação de imagens entre o método proposto e outras abordagens presentes na literatura.

é comparada com todas as outras formas, e o número de formas de uma mesma classe, entre as $2 \times t$ formas mais semelhantes são relatadas. A taxa de recuperação *Bull's eye* é a razão entre o número total de formas relatadas a partir de uma mesma classe para o número possível de formas em toda a base de dados ($t \times N_t$). Assim, a melhor taxa possível é 100% nesta medida.

Tabela 5: Taxas de *Bull's eye* nas bases de dados Tari.

Método	Base de Dados Tari 56	Base de Dados Tari 180	Base de Dados Tari 1000
Proposto	99,84%	96,66%	81,60%
PropostoSemConc	93,82%	90,84%	73,96%
PedrosaA	43,25%	26,40%	15,49%
PedrosaB	75,06%	61,46%	53,35%
PedrosaBConc	76,48%	66,13%	57,83%
TAR	86,30%	84,86%	64,27%

Tabela 6: Taxas de *Bull's eye* na base de dados MPEG-7 Parte B.

Método	Base de Dados Kimia	Base de Dados MPEG-7
Proposto	82,45%	82,13%
PropostoSemConc	80,32%	80,38%
PedrosaA	26,50%	13,08%
PedrosaB	50,93%	77,98%
PedrosaBConc	69,06%	79,89%
TAR	71,17%	73,41%

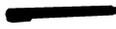
As Tabelas 5 e 6 exibem as taxas de teste *Bull's eye* para diferentes métodos de modo que se observe o desempenho dos mesmos na recuperação com a inclusão da proporção de concavidade.

A técnica de recuperação de formas proposta superou os algoritmos que também utilizam a *DSW*, como pode ser observado nas Tabelas 5 e 6 para todas as bases de dados abordadas. O método proposto foi o mais eficiente na recuperação de formas da própria classe e alcançou os melhores resultados com o incremento da proporção de concavidade. Essa característica também permitiu uma maior eficácia no método PedrosaB, como destacado em negrito nos resultados a cerca de todas as bases de imagens testadas. Tabela 5 também mostra que quanto maior a base de imagens, maior a dificuldade em alcançar o valor máximo de acerto na recuperação das formas. Dessa maneira, bases de dados menores alcançam melhores índices de desempenho com relação à acurácia.

As bases Tari que contêm menor quantidade de imagens (Tari 56 e 180) dispõem de poucas formas com contornos distintos dentro da mesma classe.

Notam-se poucas mudanças na definição dos contornos das formas presentes nas bases Tari com menor quantidade de imagens (Tari 56 e 180). Essas imagens mencionadas possuem uma maior variação de transformações geométricas aplicadas às mesmas, como em classes que apresentam a mesma imagem rotacionada ou com mudança na escala. Estas variações entre as imagens são mais simples de ser identificadas pelo sistema de recuperação do que os detalhes de definição dos contornos das imagens. A base Tari 1000 possui uma quantidade maior de

Tabela 7: Amostras de formas dentre as bases de dados adotadas.

Base de Imagens	Amostras de Formas			
Kimia				
MPEG-7				
Tari 56				
Tari 1000				

imagens e, por consequência, uma maior variedade de efeitos nas mesmas. Nesta base são perceptíveis mudanças de transformações geométricas e de formas dentro da mesma classe. Portanto, o desempenho do sistema de recuperação decai com a diversidade dos padrões dentro de uma mesma classe e o efeito das transformações sobre as mesmas.

As bases MPEG-7 e Kimia apresentam grande diversidade nas formas e consequentemente em seus contornos que compõem mesma classe de objetos. As mudanças de efeito de escala e rotação possuem uma menor ocorrência dentro destas bases de imagens. Amostras das imagens presentes nas bases de imagens mencionadas estão dispostas na Tabela 7.

Tabela 8: Análise comparativa entre sistemas de recuperação avaliados.

Método	Característica Global	Característica Local	Correspondência entre Formas
TAR	Solidez, Excentricidade e Proporção de Aspecto	Assinatura TAR	<i>DSW</i> em descritores de mesma dimensão
PedrosaA	Nenhuma	Posição relativa e Valor de curvatura	<i>DSW</i> em descritores de dimensões distintas
PedrosaB	Solidez, Excentricidade e Proporção de Aspecto	Posição relativa, Valor de curvatura e Complexidade	<i>DSW</i> em descritores de dimensões distintas
Proposto	Solidez, Excentricidade, Proporção de Aspecto e Proporção de Concavidade	Posição relativa, Valor de curvatura e Complexidade	<i>HopDSW</i> em descritores de dimensões distintas

A Tabela 8 exibe a análise comparativa dos métodos de recuperação de formas avaliados, a partir de suas características e métodos de correspondências. O método proposto se diferencia dos demais pela proporção de concavidade e adequação da técnica *HopDSW* (ALAJLAN, 2011) aos descritores com dimensões distintas. Os trabalhos de Alajlan *et al.* (2007) e Alajlan (2011) comparavam apenas entre descritores de mesma dimensão.

Os métodos PedrosaA e *TAR* não apresentam a característica da complexidade (definida na SubSeção 4.5.2) associada ao descritor da forma. Esta característica é sensível às variações de dimensão do contorno e torna o sistema de recuperação de formas invariante à mudança de escala. Reduz assim o desempenho dos sistemas CBIR apresentados. *TAR* exibe melhores resultados para as bases Kimia e Tari porque agrega características globais em sua definição, inclusive superando a técnica PedrosaB nestes casos. Adicionar as propriedades de formas *AR*, *E* e *S* à técnica de comparação entre os objetos das bases de imagens permite superar os métodos que não se associam a essas características.

A base MPEG-7 dispõe de várias imagens que apresentam maiores dimensões e com variações de escala que as contidas nas demais bases de imagens. Para sistemas CBIR que adotam descritores de mesma dimensão, destaca-se um problema na subamostragem desse descritor (ALAJLAN *et al.*, 2007), causando o decaimento no desempenho deste sistema de recuperação. Para esta referida base de dados, o método PedrosaB alcançou os melhores resultados já que o descritor de formas adotado é invariante a escala e com isso opera sobre formas de dimensões variadas. Mesmo com o tratamento aos descritores com dimensões diferentes, o sistema PedrosaB apresenta-se invariante aos efeitos de mudança de escala das formas processadas. Por sua vez, o sistema Proposto emprega a propriedade da complexidade da forma em seu descritor, que pode assumir dimensões distintas ao longo da recuperação das formas, além de associar as mesmas características globais de *TAR* e PedrosaB e inserir uma nova característica que incremente a separabilidade entre classes.

Mais uma vez, o hardware utilizado nos experimentos foi um notebook com as seguintes especificações: Processador Intel Core 2 Duo T8100, Adaptador Gráfico Intel Graphics Media Accelerator (GMA) X3100, Windows Vista Home Basic 32-bit e 4 GB de Memória RAM. Com relação ao software, as simulações foram aplicadas no ambiente Matlab 2010b. Para cada imagem presente nas bases de dados utilizadas foram executados 6 simulações de sistemas de recuperação em comparação com toda a base de imagens associada. Dessa forma, foram realizados 17.112 experimentos para reconhecer o melhor método de recuperação de formas dentre os avaliados.

5.2 Considerações Finais

Neste capítulo, apresentamos um método de recuperação de formas associando detecção de cantos baseada em transformada *wavelet* e a técnica *DSW* para análise de similaridade ou correspondência entre formas. As informações dos cantos permitiram a geração de um descritor adequado ao casamento de formas por meio de programação dinâmica.

Os resultados mostraram que o método proposto apresentou desempenho superior entre todos os métodos avaliados com uso de combinação da característica de proporção de concavidade a outras informações globais. Desse modo, a proposta alcançou os melhores índices de *Precisão*, *Revocação* e taxa de *Bull's eye* com a vantagem de avaliar distâncias entre formas distintas. A proporção de concavidade ressaltou a distância entre formas que pertencem a classes distintas e definiu bem a separabilidade inter-classe dentro da base de imagens. Ademais, a aplicação da programação dinâmica reduziu o custo computacional na recuperação de imagens e permitiu encontrar correspondências entre formas através de seus cantos (concavidade e demais características globais).

No capítulo seguinte serão apresentadas as conclusões finais desta tese, produção científica associada e perspectivas de trabalhos futuros.

6 CONCLUSÕES

Nesta tese foram propostos métodos de análise de formas pertencentes a bases de imagens binárias, com vistas a aplicações em detecção de cantos e recuperação de imagens. A tese inicialmente destaca a relevância dos detectores de cantos em contornos de formas e propõe de um detector de cantos baseado na decomposição *wavelet* Chapéu Mexicano e correlação espacial entre as suas escalas. Esta técnica avalia a representação multiescala dos cantos candidatos e seleciona aqueles que persistem no sinal de correlação entre diferentes escalas de sua decomposição. Assim, a técnica descarta cantos falsos que desaparecem ao longo das escalas de decomposição.

A segunda parte da tese apresenta a arquitetura de um sistema de recuperação de imagens baseado em conteúdo aplicado ao estudo de formas e cantos relevantes. Os métodos de recuperação de formas presentes na literatura normalmente apresentam alta complexidade computacional com relação à comparação entre imagens, sendo que estas são ajustadas de modo que os descritores apresentam as mesmas dimensões. Adaptando estas mesmas técnicas, foi proposta ainda uma redução na complexidade temporal na etapa de correspondência entre os descritores, adequando esse processo às diferentes dimensões dos vetores de descrição analisados. Vale destacar que o descritor de formas adotado foi ajustado para se tornar invariante à rotação, espelhamento, translação e mudança de escala da imagem a ser processada.

O desempenho dos métodos propostos foi analisado e comparado com outras técnicas existentes na literatura. Os testes foram realizados com uma base de imagens contendo uma quantidade reduzida de elementos, que apresentava a descrição da verdade-terrestre de cada imagem de modo a facilitar a compreensão do funcionamento do detector de cantos proposto. Os testes ainda constataram que o método proposto alcançou melhor desempenho em termos das medidas quantitativas denominadas *Precisão* e *Revocação*, com relação às demais técnicas avaliadas. A partir dos experimentos da primeira parte da tese foi possível concluir que:

- a representação do sinal de angulação se mostrou adequada na avaliação dos cantos candidatos que persistem nos coeficientes de detalhes nas suas diferentes escalas de decomposição. Desse modo, são descartados os cantos falsos ao longo das escalas, sem distorção da verdadeira localização dos cantos no contorno da forma;

- a operação de correlação entre as escalas foi capaz de recuperar cantos candidatos nas três primeiras escalas de decomposição. Nas escalas subsequentes, a decomposição exhibe pontos fracamente correlacionados e assim o algoritmo descarta essas informações;
- a metodologia proposta de identificação dos parâmetros mais adequados para a detecção dos cantos na forma se mostrou eficaz, e com isso apresentamos uma alternativa viável de busca e ajuste de parâmetros dos algoritmos de detecção de cantos multiescala. Importante destacar que a metodologia é adaptável à base de imagens, embora dependa da verdade-terrestre das formas;
- o método apresenta limitações, quando aplicado às imagens em nível de cinza.

Na análise de desempenho das técnicas de recuperação, a metodologia proposta alcançou taxas de acerto mais altas nos testes de *Bull's eye* e nos gráficos que relacionam *Precisão* e *Revocação*. Esse desempenho comprovou a eficiência do descritor adotado na recuperação de formas baseada em *DSW* nas três bases de imagens. A partir dos testes realizados destacam-se algumas observações:

- o uso de programação dinâmica e o recurso do passo de salto de execução das comparações entre as formas contribuiu para a redução da complexidade computacional temporal do algoritmo;
- as características globais melhoraram a capacidade de discriminação entre as formas, principalmente após a introdução da proporção de concavidade. Os resultados favoreceram a percepção da influência dessa característica no desempenho da recuperação de formas, dado que a proporção de concavidade ressaltou a distância entre formas de classes distintas e definiu bem a separabilidade entre classes.

Com os resultados apresentados, concluímos que o detector de cantos proposto localiza com maior eficiência cantos relevantes das formas, sendo que com essa informação de cantos é possível gerar um descritor de características locais e globais destas formas. Em consequência, obtém-se uma maior separabilidade entre as imagens das bases de dados avaliadas.

6.1 Produção Bibliográfica

Artigos publicados em conferências:

- **Iális C. Paula Jr**, Fátima N. S. Medeiros, Francisco N. Bezerra. **Shape Retrieval by Corners and Dynamic Space Warping**, *XVIII International Conference on Digital Signal Processing*, 2013, Santorini, Grécia.

- **Iális C. Paula Jr**, Fátima N. S. Medeiros, Francisco N. Bezerra, Daniela M. Ushizima. **Corner Detection within a Multiscale Framework**. *XXIV Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2011. p. 141-148, Maceió, AL.
- **Iális C. Paula Jr**, Fátima N. S. Medeiros, Thiago P. Serpa. **Abordagem Wavelet para Detecção de Cantos sob Transformações de Similaridade**. *Anais da V Escola Regional de Computação dos Estados do Ceará, Maranhão e Piauí*, 2011, Teresina-PI.
- Francisco N. Bezerra, **Iális C. Paula Jr**, Fátima N. S. Medeiros, Daniela M. Ushizima, Luis H. S. Cintra. **Morphological Segmentation for Sagittal Plane Image Analysis**. *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010. p. 4773-4776, Buenos Aires.

Artigo publicado em periódico:

- **Iális C. Paula Jr**, Fátima N. S. Medeiros, Francisco N. Bezerra, Daniela M. Ushizima. **Multiscale Corner Detection in Planar Shapes**, *Journal of Mathematical Imaging and Vision*, v. 45, n. 3, p. 251-263, Março 2013.

6.2 Perspectivas de Trabalhos

Como perspectivas de trabalhos a serem abordados são recomendadas as seguintes investigações:

- estudo mais aprofundado da detecção de cantos em imagens em níveis de cinza, descartando o processo de segmentação prévia;
- aplicar esta metodologia de recuperação de imagens para descrição e recuperação de formas de pacientes humanos em base de dados clínicos;
- combinação da busca por similaridade utilizada nesta tese à técnica de realimentação de relevância para a recuperação de formas;
- estudo e associação da técnica de recuperação utilizada com outras características como cor e textura em bases distintas de imagens;
- investigar técnicas de processamento massivo paralelo (como GPUs) para a abordagem proposta de recuperação de imagens para acelerar o processo de comparação e classificação de formas;
- realizar experimentos juntamente com usuários, quantificando também os resultados através de testes de usabilidade.

REFERÊNCIAS

- ALAJLAN, N. HopDSW: An approximate dynamic space warping algorithm for fast shape matching and retrieval. *Journal of King Saud University - Computer and Information Sciences*, v. 23, n. 1, p. 7–14, Janeiro 2011.
- ALAJLAN, N.; KAMEL, M. S.; FREEMAN, G. H. Geometry-Based image retrieval in binary image databases. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 6, p. 1003–1013, Junho 2008.
- ALAJLAN, N.; RUBE, I. E.; KAMEL, M. S.; FREEMAN, G. Shape retrieval using triangle-area representation and dynamic space warping. *Pattern Recognition*, v. 40, n. 7, p. 1911–1920, Julho 2007.
- ALMEIDA, C. W. D.; SOUZA, R. M. C. R.; CAVALCANTI, N. L. J. A shape-based image retrieval system using the curvature scale space CSS technique and the self-organizing map SOM model. In: *IEEE. VI International Conference on Hybrid Intelligent Systems*. Auckland, Nova Zelândia, 2006. p. 25–29.
- ANTOINE, J. P.; CARRETTE, P.; MURENZI, R.; PIETTE, B. Image analysis with two-dimensional continuous wavelet transform. *Signal Processing*, v. 31, n. 3, p. 241–272, Abril 1993.
- ASLAN, C.; ERDEM, A.; ERDEM, E.; TARI, S. Disconnected skeleton: shape at its absolute scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 30, n. 12, p. 2188–2203, Dezembro 2008.
- AWRANGJEB, M.; LU, G. An improved curvature scale-space corner detector and a robust corner matching approach for transformed image identification. *IEEE Transactions on Image Processing*, v. 17, n. 12, p. 2425–2441, Dezembro 2008.
- BAI, X.; WANG, B.; YAO, C.; LIU, W.; TU, Z. Co-Transduction for shape retrieval. *IEEE Transactions on Image Processing*, v. 21, n. 5, p. 2747–2757, Maio 2012.
- BAI, X.; YANG, X.; LATECKI, L. J.; LIU, W.; TU, Z. Learning context-sensitive shape similarity by graph transduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 32, n. 5, p. 861–874, Maio 2010.
- BALMASHNOVA, E.; FLORACK, L. M. J. Novel similarity measures for differential invariant descriptors for generic object retrieval. *Journal of Mathematical Imaging and Vision*, v. 31, n. 2, p. 121–132, Julho 2008.
- BARENHOLTZ, E.; COHEN, E. H.; FELDMAN, J.; SINGH, M. Detection of change in shape: an advantage for concavities. *Cognition*, v. 89, n. 1, p. 1–9, Agosto 2003.

- BARENHOLTZ, E.; FELDMAN, J. Visual comparisons within and between object parts: evidence for a single-part superiority effect. *Vision Research*, v. 43, n. 15, p. 1655–1666, Julho 2003.
- BARTOLINI, I.; CIACCIA, P.; PATELLA, M. Query processing issues in region-based image databases. *Knowledge and Information Systems*, v. 25, n. 2, p. 389–420, Novembro 2010.
- BELONGIE, S.; CARSON, C.; GREENSPAN, H.; MALIK, J. Color- and texture-based image segmentation using EM and its application to content-based image retrieval. In: IEEE. *Sixth International Conference on Computer Vision*. Bombay, India, 1998. p. 675–682.
- BOBER, M. Mpeg-7 visual shape descriptors. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 11, n. 6, p. 716–719, Junho 2001.
- CAI, D.; HE, X.; LI, Z.; MA, W.-Y.; WEN, J.-R. Hierarchical clustering of WWW image search results using visual, textual and link information. In: *Proceedings of the 12th annual ACM International Conference on Multimedia*. Nova York, EUA: ACM, 2004. p. 952–959.
- CANNY, J. F. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 8, n. 6, p. 679–698, Novembro 1986.
- CARSON, C.; BELONGIE, S.; GREENSPAN, H.; MALIK, J. Blobworld: Image segmentation using Expectation-Maximization and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 8, p. 1026–1038, Agosto 2002.
- CASTAÑÓN, C. B.; TRAINA, A. J. M. Indexación de imágenes médicas basadas en la transformada de wavelet para estructuras de datos multi-dimensionales. In: COMPUTACIÓN, S. C. de (Ed.). *Anais do I Workshop de Base de Dados das Jornadas Chilenas de Computación*. Copiapó, Chile, 2002. v. 1, p. 1–6.
- CHANG, N.-S.; FU, K.-S. Query-by-pictorial-example. *IEEE Transactions on Software Engineering*, SE-6, n. 6, p. 519–524, Novembro 1980.
- CHURCHLAND, P. S.; RAMACHANDRAN, V. S.; SEJNOWSKI, T. J. A critique of pure vision. *Largescale neuronal theories of the brain*, p. 23–60, 1994.
- COMANICIU, D.; MEER, P.; FORAN, D.; MEDL, A. Bimodal system for interactive indexing and retrieval of pathology images. In: IEEE. *Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision*. Los Alamitos, EUA, 1998. v. 1, n. 1, p. 76–81.
- COSTA, L. F.; CESAR, R. M. J. *Shape Analysis and Classification: Theory and Practice*. 2a. ed. Orlando, EUA: CRC Press, 2009.
- COX, I. J.; MILLER, M. L.; MINKA, T. P.; PAPATHOMAS, T. V.; YIANILOS, P. N. The Bayesian image retrieval system, PicHunter: theory, implementation, and psychophysical experiments. *IEEE Transactions on Image Processing*, v. 9, n. 1, p. 20–37, Janeiro 2000.
- CéSAR, R. M. J.; COSTA, L. F. Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. *Pattern Recognition*, v. 28, n. 9, p. 1559–1569, Setembro 1996.
- DATTA, R.; JOSHI, D.; LI, J.; WANG, J. Z. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, v. 40, n. 2, p. 1–60, Abril 2008.

DAVIES, J.; GOEL, A. K.; NERSESSIAN, N. J. A computational model of visual analogies in design. *Cognitive Systems Research*, v. 10, n. 3, p. 204–215, Setembro 2009.

DORAIRAJ, R.; NAMUDURI, K. R. Compact combination of MPEG-7 color and texture descriptors for image retrieval. In: IEEE. *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004*. Pacific Grove, EUA, 2004. v. 1, n. 1, p. 387–391.

DRAPER, B.; BAEK, K.; BOODY, J. Implementing the expert object recognition pathway. *Machine Vision and Applications*, v. 16, p. 27–32, 2004.

EGOZI, A.; KELLER, Y.; GUTERMAN, H. Improving shape retrieval by spectral matching and meta similarity. *IEEE Transactions on Image Processing*, v. 19, n. 5, p. 1319–1327, Maio 2010.

FALOUTSOS, C.; BARBER, R.; FLICKNER, M.; HAFNER, J.; NIBLACK, W.; PETKOVIC, D.; EQUITZ, W. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, v. 3, n. 3-4, p. 231–262, Julho 1994.

FELSEN, G.; DAN, Y. A natural approach to studying vision. *Nature Neuroscience*, v. 8, n. 12, p. 1643–1646, Novembro 2005.

FENG, H.; SHI, R.; CHUA, T.-S. A bootstrapping framework for annotating and retrieving WWW images. In: *Proceedings of the 12th Annual ACM International Conference on Multimedia*. Nova York, EUA: [s.n.], 2004. v. 1, n. 1, p. 960–967.

FLICKNER, M.; SAWHNEY, H.; NIBLACK, W.; ASHLEY, J.; HUANG, Q.; DOM, B.; GORKANI, M.; HAFNER, J.; LEE, D.; PETKOVIC, D.; STEELE, D.; YANKER, P. Query by image and video content: the QBIC system. *Computer*, v. 28, n. 9, p. 23–32, Setembro 1995.

GAO, X.; SATTAR, F.; QUDDUS, A.; VENKATESWARLU, R. Multiscale contour corner detection based on local natural scale and wavelet transform. *Image and Vision Computing*, v. 25, n. 6, p. 890–898, Junho 2007.

GEVERS, T.; SMEULDERS, A. W. M. PicToSeek: combining color and shape invariant features for image retrieval. *IEEE Transactions on Image Processing*, v. 9, n. 1, p. 102–119, Janeiro 2000.

GOH, K.-S.; CHANG, E.; CHENG, K.-T. Support vector machine pairwise classifiers with error reduction for image classification. In: ACM. *Proceedings of the ACM Multimedia Workshop on Multimedia Information Retrieval*. Ottawa, Canadá, 2001. v. 1, n. 1, p. 32–37.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 3a. ed. Nova York, EUA: Prentice Hall, 2007.

GORDER, P. F. Computer vision, inspired by the human brain. *Computing in Science and Engineering*, v. 10, p. 6–11, Março 2008.

GUPTA, A.; JAIN, R. Visual information retrieval. *Communications of the ACM*, v. 40, n. 5, p. 70–79, Maio 1997.

HARRIS, C. J.; STEPHENS, M. A combined corner and edge detector. In: IED. *IV Alvey Vision Conference*. Manchester, Inglaterra, 1998. v. 1, n. 1, p. 147–151.

HINNEBURG, A.; AGGARWAL, C. C.; KEIM, D. A. What is the nearest neighbor in high-dimensional spaces? In: ACM. *Proceedings of XXVI International Conference on Very Large Databases*. Cairo, Egito, 2000. p. 506–516.

HUA, J.; LIAO, Q. Wavelet-based multiscale corner detection. In: IEEE. *Proceedings of the V International Conference on Signal Processing*. Beijing, China, 2000. v. 1, p. 341–344.

IQBAL, Q.; AGGARWAL, J. K. CIRES: a system for content-based retrieval in digital image libraries. In: IEEE. *7th International Conference on Control Automation, Robotics and Vision (ICARCV)*. Singapura, 2002. v. 1, p. 205–210.

JAIN, A. K.; VAILAYA, A. Image retrieval using color and shape. *Pattern Recognition*, v. 29, n. 8, p. 1233–1244, Agosto 1996.

KAY, K. N.; NASELARIS, T.; PRENGER, R. J.; GALLANT, J. L. Identifying natural images from human brain activity. *Nature*, v. 452, n. 7185, p. 352–355, Março 2008.

KELLY, P. M.; CANNON, M.; HUSH, D. R. Query by image example: The candid approach. In: SPIE. *SPIE Storage and Retrieval for Image and Video Databases III*. Milão, Itália, 1995. v. 2420, p. 238–248.

KHAN, N. M.; AHMAD, I. S. An efficient signature representation for retrieval of spatially similar images. *Signal, Image and Video Processing*, v. 6, n. 1, p. 55–70, Agosto 2007.

KINDRATENKO, V. V. On using functions to describe the shape. *Journal of Mathematical Imaging and Vision*, v. 18, n. 3, p. 225–245, Maio 2003.

KUTTER, M.; BHATTACHARJEE, K. S.; EBRAHIMI, T. Towards second generation watermarking schemes. In: IEEE. *Proceedings of the IEEE Conference on Image Processing*. Kobe, Japão, 1999. v. 1, p. 320–323.

LATECKI, L. J.; LAKAMPER, R.; ECKHARDT, U. Shape descriptors for non-rigid shapes with a single closed contour. In: IEEE. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Hilton Head, EUA, 2000. v. 1, p. 424–429.

LEE, J. S.; SUN, Y. N.; CHEN, C. H. Multiscale corner detection by using wavelet transform. *IEEE Transactions on Image Processing*, v. 4, n. 1, p. 100–104, Janeiro 1995.

LI, Y.; SHAPIRO, L. O.; BILMES, J. A. A generative/discriminative learning algorithm for image classification. In: IEEE. *IEEE International Conference on Computer Vision*. Beijing, China, 2005. v. 2, p. 1605–1612.

LIU, H. C.; SRINATH, M. D. Partial shape classification using contour matching in distance transformation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, n. 11, p. 1072–1079, Novembro 1990.

LIU, Y.; ZHANG, D.; LU, G.; MA, W. Y. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, v. 40, p. 262–282, Janeiro 2007.

MA, W. Y.; DENG, Y.; MANJUNATH, B. S. Tools for texture- and color-based search of images. In: SPIE. *Proc. SPIE Human Vision and Electronic Imaging II*. San Jose, EUA, 1997. v. 3016, p. 496–507.

- MALLAT, S. *A Wavelet Tour of Signal Processing*. 2a. ed. San Diego, EUA: Academic Press, 1998.
- MARR, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York, EUA: The MIT Press, 1982.
- MEHROTRA, S.; RUI, Y.; ORTEGA-BINDERBERGER, M.; HUANG, T. S. Supporting content-based queries over images in MARS. In: IEEE. *IEEE International Conference on Multimedia Computing and Systems 1997*. Ottawa, Canadá, 1997. p. 632–633.
- MOKHTARIAN, F.; MACKWORTH, A. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 8, n. 1, p. 34–43, Janeiro 1986.
- MOKHTARIAN, F.; MACKWORTH, A. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 14, n. 8, p. 789–805, Agosto 1992.
- MOKHTARIAN, F.; MOHANNA, F. Performance evaluation of corner detectors using consistency and accuracy measures. *Computer Vision and Image Understanding*, v. 102, n. 1, p. 81–94, Abril 2006.
- MOKHTARIAN, F.; SUOMELA, R. Robust image corner detection through curvature scale space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 20, n. 12, p. 1376–1381, Dezembro 1998.
- MULLER, H.; MICHOUX, N.; BANDON, D.; GEISSBUHLER, A. A review of content-based image retrieval systems in medical applications – clinical benefits and future directions. *International Journal of Medical Informatics*, v. 73, n. 1, p. 1–23, Fevereiro 2004.
- NIBLACK, C. W.; BARBER, R.; EQUITZ, W.; FLICKNER, M. D.; GLASMAN, E. H.; PETKOVIC, D.; YANKER, P.; FALOUTSOS, C.; TAUBIN, G. QBIC project: querying images by content, using color, texture, and shape. In: SPIE. *Proc. SPIE Storage and Retrieval for Image and Video Databases*. San Jose, EUA, 1993. v. 1908, p. 173–187.
- OGLE, V. E.; STONEBRAKER, M. Chabot: retrieval from a relational database of images. *IEEE Computer*, v. 28, n. 9, p. 40–48, Setembro 1995.
- PALMERI, T. J.; GAUTHIER, I. Visual object understanding. *Nature Reviews Neuroscience*, v. 5, p. 291–304, Abril 2004.
- PASSARINHO, C. J. P. *Avaliação do Tratamento Fisioterapêutico de Desvios Posturais por Reeducação Postural Global Utilizando Técnicas de Processamento Digital de Imagens*. Dissertação (Dissertação de Mestrado em Sinais e Sistemas) — Universidade Federal do Ceará - Dept. de Eng. de Teleinformática, Fortaleza, CE, Janeiro 2005.
- PAULA JR., I. C.; MEDEIROS, F. N. S.; BEZERRA, F. N.; USHIZIMA, D. M. Corner detection within a multiscale framework. In: IEEE. *Anais de XXIV Conference on Graphics, Patterns and Images*. Maceió-AL, 2011. p. 141–148.
- PAULA JR., I. C.; MEDEIROS, F. N. S.; BEZERRA, F. N.; USHIZIMA, D. M. Multiscale corner detection in planar shapes. *Journal of Mathematical Imaging and Vision*, v. 45, n. 3, p. 251–263, Março 2013.

- PAULA JR., I. C.; MEDEIROS, F. N. S.; MENDONCA, G. A.; PASSARINHO, C. J. P.; OLIVEIRA, I. N. S. Correlating multiple redundant scales for corner detection. In: IEEE. *Proceedings of the VI International Telecommunications Symposium*. Fortaleza-CE, 2006. p. 650–655.
- PEDROSA, G. V.; BARCELOS, C. A. Z. Anisotropic diffusion for effective shape corner point detection. *Pattern Recognition Letters*, v. 31, n. 12, p. 1658–1664, Setembro 2010.
- PEDROSA, G. V.; BARCELOS, C. A. Z.; BATISTA, M. A. An image retrieval system using shape salience points. In: IEEE. *Proceedings of the IEEE International Symposium on Circuits and System (ISCAS 2011)*. Rio de Janeiro-RJ, 2011. p. 2797–2800.
- PEDROSA, G. V.; BARCELOS, C. A. Z.; BATISTA, M. A. A shape contour descriptor based on salience points. In: IEEE. *18th IEEE International Conference on Image Processing (ICIP 2011)*. Bruxelas, Bélgica, 2011. p. 1049–1052.
- PENTLAND, A.; PICARD, R. W.; SCLAROFF, S. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, v. 18, n. 3, p. 233–254, Junho 1996.
- PFUND, T.; MARCHAND-MAILLET, S. Dynamic multimedia annotation tool. In: SPIE. *Proceedings of SPIE Internet Imaging III*. San Jose, EUA, 2001. v. 4672, p. 216–224.
- QUELLEC, G.; LAMARD, M.; CAZUGUEL, G.; COCHENER, B.; ROUX, C. Wavelet optimization for content-based image retrieval in medical databases. *Medical Image Analysis*, v. 14, n. 2, p. 227–241, Abril 2010.
- RANGANATH, C. Working memory for visual objects: Complementary roles of inferior temporal, medial temporal, and prefrontal cortex. *Neuroscience*, v. 139, n. 1, p. 277–289, Abril 2006.
- RATTARANGSI, A.; CHIN, R. T. Scale-based detection of corners of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 14, n. 4, p. 430–449, Abril 1992.
- ROH, M. C.; CHRISTMAS, B.; KITTLER, J.; LEE, S. W. Gesture spotting in low-quality video with features based on curvature scale space. In: IEEE. *Proceedings of the VII International Conference on Automatic Face and Gesture Recognition*. Southampton, Inglaterra, 2006. p. 375–380.
- ROSENFELD, A.; JOHNSTON, E. Angle detection on digital curves. *IEEE Transactions on Computers*, v. 22, n. 9, p. 875–878, Setembro 1973.
- ROSENFELD, A.; WESZKA, J. S. An improved method of angle detection on digital curves. *IEEE Transactions on Computers*, v. 24, n. 9, p. 940–941, Setembro 1975.
- ROSTEN, E.; PORTER, R.; DRUMMOND, T. Faster and better: A machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 32, n. 1, p. 105–119, Janeiro 2010.
- RUI, Y.; HUANG, T. S.; ORTEGA, M.; MEHROTRA, S. Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 8, n. 5, p. 644–655, Setembro 1998.

- SAKOE, H.; CHIBA, S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions Acoustics, Speech, and Signal Processing*, v. 26, p. 43–49, Fevereiro 1978.
- SEBASTIAN, T. B.; KLEIN, P. N.; KIMIA, B. B. Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 26, n. 5, p. 550–571, Maio 2004.
- SERRE, T.; POGGIO, T. A neuromorphic approach to computer vision. *Communications of the ACM*, v. 53, p. 54–61, Outubro 2010.
- SHASTRI, L. A computational model of episodic memory formation in the hippocampal system. *Neurocomputing*, v. 38-40, p. 889–897, Junho 2001.
- SILVA, A. T.; FALCÃO, A. X.; MAGALHÃES, L. P. A new CBIR approach based on relevance feedback and optimum-path forest classification. *Journal of WSCG*, v. 18, n. 1-3, p. 73–80, Fevereiro 2010.
- SILVA, A. T.; FALCÃO, A. X.; MAGALHÃES, L. P. Active learning paradigms for CBIR systems based on optimum-path forest classification. *Pattern Recognition*, v. 44, n. 12, p. 2971–2978, Dezembro 2011.
- SITA, G.; RAMAKRISHNAN, A. G. Wavelet domain nonlinear filtering for evoked potential signal enhancement. *Computers and Biomedical Research*, v. 33, n. 6, p. 431–446, Dezembro 2000.
- SMEULDERS, A. W. M.; WORRING, M.; SANTINI, S.; GUPTA, A.; JAIN, R. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, p. 1349–1380, Dezembro 2000.
- SMITH, J. R.; CHANG, S.-F. VisualSEEK: a fully automated content-based image query system. In: *ACM International Conference on Multimedia, MULTIMEDIA'96*. New York, EUA: ACM, 1996. p. 87–98.
- SQUIRE, D. M.; MÜLLER, W.; MÜLLER, H.; PUN, T. Content-based query of image databases: inspirations from text retrieval. *Pattern Recognition Letters*, v. 21, n. 13-14, p. 1193–1198, Dezembro 2000.
- TORRES, R. S.; FALCÃO, A. X. Content-based image retrieval: Theory and applications. *Revista de Informática Teórica e Aplicada*, v. 13, n. 2, p. 161–185, 2006.
- TORRES, R. S.; FALCÃO, A. X. Contour salience descriptors for effective image retrieval and analysis. *Image and Vision Computing*, v. 25, n. 1, p. 3–13, Janeiro 2007.
- TVERSKY, A. Features of similarity. *Psychological Review*, v. 84, n. 4, p. 327–352, Julho 1977.
- VALLE, E.; CORD, M. Advanced techniques in CBIR: Local descriptors, visual dictionaries and bags of features. In: *Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing*. Rio de Janeiro, Brasil: [s.n.], 2009. p. 72–78.
- VASCONCELOS, N.; KUNT, M. Content-based retrieval from image databases: current solutions and future directions. In: *IEEE. Proceedings of the International Conference in Image Processing (ICIP)*. Thessaloniki, Grécia, 2001. v. 3, p. 6–9.

- VASCONCELOS, N.; LIPPMANN, A. A unifying view of image similarity. In: IAPR. *Proceedings of the XV International Conference on Pattern Recognition*. Barcelona, Espanha, 2000. v. 1, p. 38–41.
- WANG, H. H.; MOHAMAD, D.; ISMAIL, N. A. Semantic gap in CBIR: Automatic objects spatial relationships semantic extraction and representation. *International Journal of Image Processing*, v. 4, p. 192–204, Julho 2010.
- WANG, J. Z.; LI, J.; WIEDERHOLD, G. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 23, p. 947–963, Setembro 2001.
- WANG, J. Z.; WIEDERHOLD, G.; FIRSCHEIN, O.; WEI, S. X. Content-based image indexing and searching using Daubechies' wavelets. *International Journal on Digital Libraries*, v. 1, p. 311–328, 1998.
- WANG, Y. P.; LEE, S. L. Scale-space derived from B-splines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 20, n. 10, p. 1040–1055, Outubro 1998.
- ZHANG, C.; MENGYANG, Z.; YUAN, L. An improved algorithm for corner detection. In: IEEE. *Proceedings of the International Conference on Electronic and Mechanical Engineering and Information Technology 2011*. Shanghai, China, 2011. v. 9, p. 4639–4641.
- ZHANG, X.; WANG, H.; SMITH, A. W. B.; XU, L.; LOVELL, B. C.; YANG, D. Corner detection based on gradient correlation matrices of planar curves. *Pattern Recognition*, v. 43, n. 4, p. 1207–1223, Abril 2010.
- ZHANG, X.; WANG, H.; HONG, M.; XU, L.; YANG, D.; LOVELL, B. C. Robust image corner detection based on scale evolution difference of planar curves. *Pattern Recognition Letters*, v. 30, n. 4, p. 449–455, Março 2009.
- ZHONG, B.; LIAO, W. Direct curvature scale space: Theory and corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 3, p. 508–512, Março 2007.
- ZHONG, B.; MA, K. K.; LIAO, W. Scale-space behavior of planar-curve corners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 31, n. 8, p. 1517–1524, Agosto 2009.
- ZHU, L.; TANG, C.; RAO, A.; ZHANG, A. Using thesaurus to model keyblock-based image retrieval. In: *Proceedings of the Second International Conference on Multimedia and Exposition*. Tóquio, Japão: [s.n.], 2001. p. 237–240.