



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

MARCELO MARQUES SIMÕES DE SOUZA

**UM MÉTODO PARA OTIMIZAÇÃO DE DESCRITORES MULTIESCALA EM
ANÁLISE DE FORMAS**

FORTALEZA

2016

MARCELO MARQUES SIMÕES DE SOUZA

UM MÉTODO PARA OTIMIZAÇÃO DE DESCRITORES MULTIESCALA EM ANÁLISE
DE FORMAS

Tese apresentada ao Curso de Doutorado em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Orientadora: Prof. Dr. Fátima Nelsizeuma Sombra de Medeiros

FORTALEZA

2016

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- S239m Souza, Marcelo Marques Simões de.
Um método para otimização de descritores multiescala em análise de formas / Marcelo Marques Simões de Souza. – 2016.
76 f.: il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2016.
Orientação: Profª. Dra. Fátima Nelsizeuma Sombra de Medeiros.
1. Visão computacional. 2. Reconhecimento de padrões. 3. Análise de formas. 4. Descritores multiescala. 5. Otimização evolucionária. I. Título.

CDD 621.38

MARCELO MARQUES SIMÕES DE SOUZA

UM MÉTODO PARA OTIMIZAÇÃO DE DESCRITORES MULTIESCALA EM ANÁLISE
DE FORMAS

Tese apresentada ao Curso de Doutorado em Engenharia de Teleinformática do Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Engenharia de Teleinformática. Área de Concentração: Engenharia de Teleinformática

Aprovada em: 29 de Julho de 2016

BANCA EXAMINADORA

Prof. Dr. Fátima Nelsizeuma Sombra de
Medeiros (Orientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Victor Hugo Costa de Albuquerque
Universidade de Fortaleza (UNIFOR)

Prof. Dr. Auzuir Ripardo Alexandria
Instituto Federal do Ceará (IFCE)

Prof. Dr. Francisco Nivando Bezerra
Instituto Federal do Ceará (IFCE)

Prof. Dr. Walter da Cruz Freitas Júnior
Universidade Federal do Ceará (UFC)

Dedico este trabalho a meu pai, Márcio, minha mãe, Neusa, meu irmão, Fábio, minha irmã, Simone, minha esposa, Isaura, meu filho, Pedro, minha filha, Sophia e minha orientadora, Profa. Fátima.

AGRADECIMENTOS

Aos colegas do LABVIS pelo companherismo e pela ajuda no desenvolvimento deste trabalho.

Ao Renato Vasconcelos, secretário da Pós-Graduação em Engenharia de Teleinformática.

Ao Colegiado do curso de Pós-Graduação em Engenharia de Teleinformática, pelo desafio imposto à realização deste trabalho.

À FUNCAP e ao CNPq pelo apoio financeiro.

“No words can describe it
No example can point to it
Samsara does not make it worse
Nirvana does not make it better
It has never been born
It has never ceased
It has never been liberated
It has never been deluded
It has never existed
It has never been nonexistent
It has no limits at all
It does not fall into any kind of category”

(Dudjom Rinpoche)

RESUMO

A forma é um importante atributo do sistema visual dos primatas que tem sido amplamente explorado em visão computacional em aplicações de classificação, reconhecimento e recuperação de imagens pelo conteúdo. Um sistema de visão computacional para o reconhecimento de objetos realiza a análise da forma, que consiste na descrição ou representação da mesma e na análise de similaridade entre formas. Um aspecto importante em análise de forma é a adequação do descritor ao problema de reconhecimento de padrões de interesse, porém há uma carência de métodos que sistematizem essa adequação. Este trabalho propõe um método automático para o ajuste dos parâmetros de descritores de formas com otimização evolutiva. A aplicabilidade do referido método é investigada na adequação de um descritor multiescala de forma ao problema de identificação das espécies de plantas a partir das suas folhas, sendo seu desempenho avaliado por técnicas de visualização de dados, medidas de avaliação de agrupamentos e experimentos de classificação e recuperação de formas pelo conteúdo. A análise visual exploratória dos agrupamentos mostrou que a metodologia de otimização melhora os resultados de agrupamento e recuperação de formas. Já os experimentos de classificação com os descritores otimizados alcançaram elevadas taxas de precisão e revocação, assim como da medida bulls-eye. Tendo como função objetivo métricas de qualidade de agrupamentos, a otimização de descritores de forma melhora a representação das formas nos aspectos de coesão intraclasse e separação entre classes, o que reflete positivamente no desempenho em experimentos de classificação e recuperação de formas.

Palavras-chave: Visão computacional. Reconhecimento de padrões. Análise de formas. Descritores multiescala. Otimização evolucionária.

ABSTRACT

Shape is an important attribute of the primate visual system that has been widely explored in computer vision applications, such as object classification, recognition and content-based image retrieval. A computer vision system for object recognition performs shape analysis, which encompasses shape description or representation and shape similarity analysis. A relevant aspect in shape analysis is to adjust the descriptor to the pattern recognition problem of interest, even though there is a lack of consistent methods for doing so. This work introduces an automatic method to setup multiscale shape descriptor through evolutionary optimization. The method was applied to adjust a multiscale shape descriptor to the problem of leaf-based plant specimen identification, where data visualization techniques, clustering quality metrics and shape classification and retrieval experiments were used to assess its performance. The visual exploratory data analysis techniques showed that the proposed methodology improved shape clustering and retrieval. Moreover, supervised and unsupervised classification experiments accomplished high precision and recall rates as well as Bulls-eye scores with the optimized parameters. Thus, taking as objective function clustering quality metrics, shape descriptors optimization leads to improvement in shapes representation in terms of intraclass coesion and inter-class separation, which reflects positively in shape classification experiments performance.

Keywords: Computational vision. Pattern recognition. Shape analysis. Multiscale descriptors. Evolutionary optimization.

LISTA DE ILUSTRAÇÕES

Figura 1 – Arquitetura de um sistema de CBIR.	24
Figura 2 – Obtenção da forma e do contorno da imagem de uma folha.	27
Figura 3 – Processo de obtenção da representação paramétrica do contorno da folha da Figura 2.	28
Figura 4 – Efeito do ruído na estimativa computacional da curvatura para uma forma circular.	30
Figura 5 – Curvatura estimada do contorno da folha da Figura 2 através de método computacional.	30
Figura 6 – Método de análise multiescala a partir da assinatura do contorno de uma forma.	32
Figura 7 – Efeito da suavização sobre o contorno da forma para diferentes valores de escala.	33
Figura 8 – Curvatura multiescala do contorno da forma da folha da Figura 2.	35
Figura 9 – Representação da Matriz-U como imagem em níveis de cinza. As células brancas rotuladas correspondem aos neurônios da rede SOM. Células escuras indicam maior separação entre neurônios e células claras indicam maior proximidade entre neurônios.	41
Figura 10 – Proposta de uma metodologia para otimização evolucionária de um descritor multiescala de forma.	45
Figura 11 – Formas da base de imagens Kimia-99.	47
Figura 12 – Amostras de formas da base de imagens Kimia-216.	47
Figura 13 – Amostras de formas da base de imagens MPEG-7.	47
Figura 14 – Trinta e duas amostras extraídas da base Flavia, que contém 1907 imagens de folhas de 32 espécies diferentes. Cada amostra corresponde a uma espécie.	48
Figura 15 – Matriz-U para as formas da base MPEG7 CE-Shape-1.	49
Figura 16 – Projeções do escalonamento multidimensional das formas da base Kimia-99 Sebastian <i>et al.</i> (2004). As imagens mostram como os agrupamentos evoluem ao longo do processo de otimização (DE), assim como os valores de MAD e R^2 .	50
Figura 17 – Metodologia empregada para os experimentos de recuperação de formas pelo conteúdo.	51
Figura 18 – Metodologia de classificação para avaliação de desempenho do descritor otimizado pelo método proposto exibido na Figura 10.	53

Figura 19 – Amostras de formas de folhas da base de imagens Flavia.	56
Figura 20 – Convergência dos métodos de otimização para o problema de descrição de folhas de plantas: (a) SA, (b) DE, (c) PSO. As curvas em vermelho destacam o valor médio da função MAD para 30 realizações de cada método (curvas em preto).	57
Figura 21 – Resultados de classificação das espécies vegetais da base de imagens Flavia para diferentes estratégias de escolha das escalas do descritor NMBE.	59
Figura 22 – Resultados dos descritores NMBE e DFM para a base Kimia-99. (a) Matriz-U obtida com o descritor NMBE e respectiva <i>silhouette</i> média por classe. (b) Matriz-U obtida com o descritor DFM e respectiva <i>silhouette</i> média por classe.	60
Figura 23 – Resultados dos descritores NMBE e DFM para a base Kimia-216. (a) Matriz-U obtida com o descritor NMBE e respectiva <i>silhouette</i> média por classe. (b) Matriz-U obtida com o descritor DFM e respectiva <i>silhouette</i> média por classe.	61
Figura 24 – Matrizes-U e <i>silhouettes</i> média para as descrições das folhas da base Flavia com o descritor NMBE: (a) otimizado e (b) não otimizado.	62
Figura 25 – Matrizes-U e valores MADs obtidos para a base de imagens Flavia por uso da NMBE otimizada pelos algoritmos: (a) SA, (b) DE e (c) PSO.	63
Figura 26 – Projeções MDS dos descritores NMBE para formas da base de folhas Flavia. (a) Descritor NMBE otimizado pelo algoritmo SA com valor de função $MAD = 0,762$ (b) Descritor NMBE não otimizado com valor de função $MAD = 0,969$ e (c) Descritor NMBE não otimizado com valor de função $MAD = 1,044$	64
Figura 27 – Experimentos realizados com formas da base de imagens de folhas Flavia. (a) taxa de recuperação obtida com os descritores NMBE e IDSC originais e suas versões otimizadas, (b) e (c) recuperação de duas amostras de folhas utilizando os descritores NMBE e IDSC otimizadas e não otimizadas, respectivamente.	66

LISTA DE TABELAS

Tabela 2 – Propriedades da curvatura e as características geométricas que essas representam.	29
Tabela 3 – Características das bases de imagens utilizadas nos experimentos.	46
Tabela 4 – Valores de MAD alcançados, pelos diferentes métodos de ajuste das escalas do descritor NMBE, para representação das folhas da base Flavia.	58
Tabela 5 – Taxa Bulls-eye para a base de imagens Flavia.	66
Tabela 6 – Complexidade computacional dos métodos de otimização.	66

LISTA DE ABREVIATURAS E SIGLAS

CBIR	recuperação de imagens por conteúdo
CCD	<i>charge-coupled device</i>
CDSC	<i>coherent distance shape context</i>
CI	inteligência computacional
CSS	curvatura espaço-escala
d-TAS	<i>differential turning angle scalogram</i>
DE	evolução diferencial
DFM	dimensão fractal multiscala
DTW	<i>dynamic time warping</i>
EC	computação evolucionária
GA	algoritmo genético
IDSC	<i>inner distance shape context</i>
kNN	k-vizinhos próximos
LDA	análise por discriminante linear
MAD	desvio mediano absoluto
MDS	escalonamento multidimensional
NB	<i>naive Bayes</i>
NMBE	energia de dobramento multiescala normalizada
PCA	análise das componentes principais
PSO	otimização por enxame de partículas
QDA	análise por discriminante quadrático
SA	resfriamento simulado
SC	<i>shape context</i>
SOM	mapa auto-organizável do Kohonen
TAR	representação por área de triângulos

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Motivação e objetivos	18
1.2	Produção Científica	20
1.3	Organização da Tese	21
2	FUNDAMENTOS TEÓRICOS	22
2.1	Sistemas de recuperação de imagens pelo conteúdo	22
<i>2.1.1</i>	<i>Arquitetura dos sistemas de recuperação de imagens pelo conteúdo</i>	<i>23</i>
<i>2.1.2</i>	<i>Extração de características</i>	<i>24</i>
2.2	Descritores do contorno das formas	26
<i>2.2.1</i>	<i>Representação paramétrica do contorno e curvatura</i>	<i>27</i>
<i>2.2.2</i>	<i>Representações multiescala</i>	<i>31</i>
2.3	Técnicas de visualização de dados	39
<i>2.3.1</i>	<i>Escalonamento multimensional</i>	<i>39</i>
<i>2.3.2</i>	<i>Mapa auto-organizável de Kohonen</i>	<i>40</i>
2.4	Métodos de otimização	41
<i>2.4.1</i>	<i>Resfriamento simulado</i>	<i>42</i>
<i>2.4.2</i>	<i>Evolução diferencial</i>	<i>42</i>
<i>2.4.3</i>	<i>Otimização por enxame de partículas</i>	<i>43</i>
3	METODOLOGIA	44
3.1	Customização dos descritores	44
3.2	Bases de formas binárias	46
3.3	Avaliação do descritor	48
<i>3.3.1</i>	<i>Visualização dos dados</i>	<i>48</i>
<i>3.3.2</i>	<i>Recuperação de formas</i>	<i>50</i>
<i>3.3.3</i>	<i>Classificação supervisionada</i>	<i>52</i>
4	RESULTADOS E DISCUSSÕES	56
4.1	Classificação de formas e sua relação com a função objetivo	56
4.2	Análise exploratória visual de agrupamentos	58
4.3	Recuperação de imagens baseada em conteúdo	65
4.4	Custo computacional da otimização	66

5	CONCLUSÕES E TRABALHOS FUTUROS	68
	REFERÊNCIAS	70

1 INTRODUÇÃO

A visão é um sentido importante para os primatas que têm inspirado toda uma área de pesquisa denominada visão computacional, cujas origens e evolução, intimamente relacionadas com a história da computação, tem recentemente motivado o desenvolvimento de tecnologias com uma vasta gama de aplicações (COSTA; CESAR JR., 2009). A visão é um processo complexo, pois envolve a análise de várias informações visuais, como a cor, profundidade, movimento, forma e textura dos objetos, com o propósito de locomoção, reconhecimento, classificação e manipulação dos objetos (ULLMAN, 1995).

Dentre as informações visuais, a forma desempenha um papel importante no sistema de percepção visual humano pela riqueza de informação que esta propicia. A forma advém da imagem resultante da projeção dos objetos existentes no espaço tridimensional em estruturas bidimensionais, como a retina do olho humano ou um sensor do tipo *charge-coupled device* (CCD) de uma câmera (COSTA; CESAR JR., 2009).

Em visão computacional, o processo de análise e reconhecimento de formas envolve a descrição computacional e a avaliação da similaridade das formas. Esse processo consiste em detectar e representar atributos visuais das imagens que sejam relevantes para melhoria do desempenho dos algoritmos que realizam as tarefas de classificação e reconhecimento dos objetos (ESCOLANO *et al.*, 2009). Quanto à avaliação de similaridade, procura-se por medidas que sejam semelhantes ao sistema de percepção da visão humana, ainda que não esteja totalmente claro como tal sistema opera (BAI *et al.*, 2010).

Em várias aplicações (ZHANG *et al.*, 2016; ZHAO *et al.*, 2015), informações sobre a forma são suficientes para caracterizar os objetos existentes em uma cena, independente dos demais atributos da imagem. Na medicina, descritores de forma foram aplicados no diagnóstico de doenças da coluna vertebral (LEE *et al.*, 2009), na avaliação clínica de tumores de mama a partir de imagens de ultrassom (YANG *et al.*, 2009) e na identificação automática de pílulas (USHIZIMA *et al.*, 2015). Descritores de forma também foram aplicados, em microscopia, na análise de materiais (ZHANG *et al.*, 2016) e em recuperação de informação, na busca de marcas registradas por similaridade de conteúdo (ANUAR *et al.*, 2013; QI *et al.*, 2010).

Há dois tipos de métodos de descrição de formas: os baseados em região e os baseados em contorno (ZHANG; LU, 2004). Os baseados em região extraem características de toda a área interior à forma. Alguns métodos desta natureza são os momentos de Zernike (KIM; KIM, 2000) e de Legendre (YANG *et al.*, 2006) que, apesar de apresentarem excelente desempenho,

são inadequados para reconhecimento de objetos na presença de oclusões. Utilizando as propriedades da transformada de Fourier, o descritor genérico de Fourier (ZHANG; LU, 2002) permite a análise multiescala de formas. Já os métodos baseados em contorno exploram a informação contida na fronteira entre o interior da forma e o fundo da imagem. Embora mais complexos, pois requerem implementações mais sofisticadas, esses métodos são mais adequados para o reconhecimento de objetos com oclusões. Nesta categoria, a codificação em cadeia (COSTA; CESAR JR., 2009, p 337) mapeia segmentos de linhas em uma grade fixa com um número finito de possíveis orientações, as aproximações poligonais (COSTA; CESAR JR., 2009, p 340–351) e os métodos de esqueletização (COSTA; CESAR JR., 2009, p 394–400).

Na abordagem por aproximações poligonais, a forma é decomposta em segmentos de linha. Os vértices dos polígonos são utilizados como primitivas de onde algumas características são extraídas. A transformação do eixo mediano, ou esqueletização, foi introduzido por Blum (1967). Esta consiste em reduzir regiões a curvas que seguem a forma global de um objeto. Sebastian *et al.* (2004) utilizaram este descritor no reconhecimento de formas. Milios e Petrakis (2000) introduziram a representação as formas como uma coleção de segmentos entre dois pontos de inflexão consecutivos. Os segmentos obtidos são considerados em diferentes níveis de resolução.

A descrição multiescala de formas (MOKHTARIAN; MACKWORTH, 1986; CESAR JR.; COSTA, 1997; SHU *et al.*, 2015; LI *et al.*, 2016) é uma abordagem promissora em reconhecimento de padrões em imagens (DIREKOGLU; NIXON, 2011). Na descrição multiescala os atributos das formas são representados em vários níveis de detalhes, variando de escalas de baixa resolução, aonde os detalhes que diferenciam as formas de uma mesma classe não são levados em consideração, até escalas de alta resolução aonde esses detalhes são preservados (ULLMAN, 1995; COSTA; CESAR JR., 2009). Desta forma, a combinação desses atributos, de baixa e alta resolução, aumentam o poder de discriminação do descritor, melhorando o desempenho na tarefa de classificação (DIREKOGLU; NIXON, 2011).

Mokhtarian e Mackworth (1986) propuseram o descritor curvatura espaço-escala (CSS), que é baseado na representação multiescala da curvatura das formas. O CSS é invariante às transformações afins, mas é sensível a oclusão e convexidade das formas. A representação por área de triângulos (TAR) (ALAJLAN, 2011) é outro tipo de representação multiescala baseado nas áreas de triângulos formados pelos pontos das borda, calculados para diferentes escalas. Yang *et al.* (2016) também apresentam uma proposta robusta de descritor multiescala

para a recuperação de formas pelo conteúdo, com propriedades de invariância à rotação, escala, variação intraclasse, deformação, oclusão parcial e ruídos. Esta proposta utiliza três assinaturas invariantes de contorno para capturar características locais e globais das formas em múltiplas escalas, sendo a avaliação da similaridade entre as assinaturas feita através de programação dinâmica.

Outras técnicas de representação do contorno das formas são *b-splines* (PAGLIERONI; JAIN, 1985), o *differential turning angle scalogram* (d-TAS) (KPALMA; RONSIN, 2007) e as funções peso, ou *height functions* (WANG *et al.*, 2012). A função peso de um ponto amostral do contorno é definida como sendo pelas distâncias de todos os outros pontos amostrais a sua linha tangente. A função peso obtida é então suavizada para representar e reconhecer objetos bidimensionais

Um método clássico de encontrar correspondência entre conjuntos de pontos, que utiliza a distância euclidiana como métrica para construção da matriz de distâncias, é o *shape context* (SC) (BELONGIE *et al.*, 2002). Já Ling e Jacobs (2005) propuseram o *inner distance shape context* (IDSC) substituindo a distância euclidiana do SC pela distância do produto escalar. Embora ambos os métodos tenham a habilidade de extrair características bastante discriminativas das formas, lidando também com o problema da inexistência de correspondência, estes são sensíveis a diferentes deformações e poses de uma mesma forma. Ademais, por utilizarem programação dinâmica, esses métodos apresentam um custo computacional elevado na avaliação da similaridade entre as formas (FREITAS *et al.*, 2016).

Variantes do SC e do IDSC são encontrados em Nanni *et al.* (2012). O método desses autores transforma os descritores obtidos a partir da IDSC, SC e funções de peso em um descritor matricial por quantização local de fase. Os descritores matriciais das formas são então comparados através do divergente de Jeffrey (ULLAH, 1996). Hu *et al.* (2012) propuseram um descritor baseado em contorno para o reconhecimento de formas de mão, o qual foi nomeado de distância coerente do contexto da forma, do inglês *coherent distance shape context* (CDSC). O método CDSC é robusto a diferentes poses da mão e pode ser utilizado tanto para o reconhecimento das formas como das palmas das mãos.

A representação das formas por suas partes constituintes tem desempenhado um papel relevante no reconhecimento de objetos (ULLMAN, 1995). Tal abordagem resulta em uma representação robusta das formas, pois, além de representar suas partes constituintes, representa as relações espaciais existentes entre as partes.

Em Shoji (1992), as formas são decompostas em diferentes retângulos, sendo a localização e as dimensões dos retângulos selecionadas por programação dinâmica. Pala *et al.* (2000) propuseram o uso dos pontos de cruzamento de zero da curvatura de um contorno suavizado para obtenção de partes cujos autores denominaram de fichas ou *tokens*. As orientações e os pontos de máxima curvatura das partes obtidas são levadas em consideração para representar e comparar as formas. Esse método não é invariante a rotação devido a orientação das fichas ou *tokens* (RUBERTO; CINQUE, 2009).

Usando programação dinâmica, Latecki *et al.* (2007) propuseram um método para comparação parcial de formas aonde tangentes locais ao contorno são utilizadas na descrição de formas. Cui *et al.* (2009) propuseram o uso da integral da curvatura absoluta como descritor de forma. Para comparar as partes das curvas eles utilizaram a correlação cruzada, sendo o método invariante a rotação, escala e translação.

1.1 Motivação e objetivos

Descritores de forma (BELONGIE *et al.*, 2002; LING; JACOBS, 2005; LATECKI *et al.*, 2007; NANNI *et al.*, 2012; HU *et al.*, 2012; WANG *et al.*, 2012), em sua maioria, são utilizados em sistemas de recuperação de imagens por conteúdo (CBIR). Sistemas dessa natureza realizam buscas em bases multimídia utilizando o conteúdo visual das imagens para recuperar aquelas que sejam do interesse do usuário mediante um padrão de consulta por ele especificado (FENG *et al.*, 2003).

Em termos das aplicações de descritores de formas em CBIR, este trabalho enfatiza a identificação de espécies vegetais através das folhas (WANG *et al.*, 2000; NAM *et al.*, 2008; FOTOPOULOU *et al.*, 2011; ZHAO *et al.*, 2015; SOUZA *et al.*, 2016). A abordagem tradicional para a classificação de espécies vegetais é a taxonomia, ou seja, o treinamento de especialistas, denominados taxonomistas, para a identificação das espécies através das características observadas nas plantas (COPE *et al.*, 2012). A deficiência dessa abordagem é o aspecto subjetivo do taxonomista (GOMES *et al.*, 2012) e o pequeno número de taxonomistas disponíveis (COPE *et al.*, 2012). Ademais, os taxonomistas são especialistas em identificar um pequeno grupo de espécies, sendo assim impossível aos mesmos classificarem diversas espécies (COPE *et al.*, 2012).

Historicamente, as amostras de espécies são armazenadas em arquivos físicos, denominados de herbários (COPE *et al.*, 2012), para posterior classificação, tornando o acesso

às amostras difícil e lento. Com a criação das bases de dados digitais, a informação tornou-se facilmente acessível aos pesquisadores em qualquer parte do mundo, o que permite o desenvolvimento de métodos computacionais automatizados para classificação das espécies a partir de imagens. Desta forma, sistemas dessa natureza permitiriam que pessoas com um conhecimento limitado de botânica desempenhem satisfatoriamente em campo tarefas de identificação de espécies que requereriam o trabalho de um taxonomista (COPE *et al.*, 2012).

Um aspecto importante em análise de formas é a adequação do descritor ao problema de reconhecimento de padrões de interesse. Tal tarefa envolve ajuste de parâmetros, assim como a investigação da estrutura que o descritor impõe aos dados. Desta forma, busca-se ajustar a descrição ao domínio de conhecimento do problema, ou seja, colocá-la em conformidade com a percepção dos especialistas da área.

Muitas vezes, no projeto de descritores de forma, atribui-se valores de parâmetros que independem da base de dados representada. Em outras situações, esse ajuste é obtido através de um processo empírico ou mesmo por busca exaustiva (MOKHTARIAN; SUOMELA, 1998; LING; JACOBS, 2007; WANG *et al.*, 2012).

Paula Jr. *et al.* (2012) utilizou um esquema de otimização por busca exaustiva e reportou resultados satisfatórios no ajuste de parâmetros de um descritor a uma base de imagens de formas. Bai *et al.* (2010) propuseram um método para ajuste dos parâmetros de descritores de forma por aprendizagem supervisionada. Com isso a similaridade entre padrões de formas pré-estabelecidos foi utilizada em um processo de aprendizagem sensível ao contexto e que busca os parâmetros mais adequados ao cálculo da matriz de similaridade de um conjunto de formas (BAI *et al.*, 2010).

O ajuste de parâmetros dos descritores pode ser considerado um problema de otimização e inteligência computacional (CI) (ENGELBRECHT, 2007), sendo esta uma subárea da inteligência artificial. Estudos nessa área envolvem o desenvolvimento de algoritmos bio-inspirados, que apresentam comportamento inteligente, para resolução de problemas complexos. Em otimização, os algoritmos de computação evolucionária (EC) imitam a evolução natural dos processos de populações. Desta forma, cada solução candidata é análoga a um indivíduo em uma população e sua qualidade, à aptidão desse indivíduo (EIBEN; SMITH, 2015).

Este trabalho tem como objetivo geral introduzir uma metodologia para o ajuste de parâmetros dos descritores multiescala de formas (COSTA; CESAR JR., 2009), através de otimização evolutiva (ENGELBRECHT, 2007), para adequação dos mesmos a um problema

de visão computacional que seja de interesse. Assim, é investigada a aplicabilidade do referido método ao problema de descrição multiescala da forma das folhas para a identificação das espécies de plantas.

Os objetivos específicos deste trabalho são os seguintes:

- Aplicar a metodologia de otimização ao descritor de formas energia de dobramento multiescala normalizada (NMBE) (COSTA; CESAR JR., 2009) e IDSC (LING; JACOBS, 2005);
- Propor uma função objetivo, a ser empregada nos algoritmos de otimização, para busca de soluções ótimas para os conjuntos de parâmetros multiescala dos descritores de formas;
- Avaliar qualitativamente e quantitativamente a capacidade discriminativa dos descritores multiescala otimizados e não otimizados;
- Empregar técnicas de visualização de dados e medidas de avaliação da qualidade de agrupamentos para quantificar a capacidade discriminativa dos descritores estudados;
- Realizar experimentos de classificação e recuperação de formas pelo conteúdo, em bases públicas de imagens de uso geral e de folhas de plantas, para quantificar a qualidade dos descritores;

1.2 Produção Científica

- SOUZA, M.M.S.; MEDEIROS, F.N.S.; RAMALHO, G.L.B.; DE PAULA JR., I.C.; OLIVEIRA, I.N. Evolutionary optimization of a multiscale descriptor for leaf shape analysis. *Expert Systems with Applications*, p. –, 2016. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S095741741630361X>>.
- USHIZIMA, D.; CARNEIRO, A.; SOUZA, M.; MEDEIROS, F. Investigating pill recognition methods for a new national library of medicine image dataset. In: *Advances in Visual Computing*. Springer Science Business Media, 2015. p 410-419. Disponível em: <http://dx.doi.org/10.1007/978-3-319-27863-6_38>
- NOGUEIRA, O.C.; SOUZA, M.M.S. ; MEDEIROS, F.N.S. ; OLIVEIRA, I.N.S. Dimensão Fractal em Recuperação de Imagens Baseadas em Conteúdo (CBIR). XIII Escola Regional de Informática da SBC, v. 1, p. 1-9, 2013.

1.3 Organização da Tese

A organização dos capítulos desta tese segue a seguinte estrutura:

Capítulo 2: contém a descrição de conceitos teóricos e ferramentas utilizadas no desenvolvimento, avaliação e testes dos descritores, assim como no desenvolvimento da metodologia proposta para análise de formas.

Capítulo 3: introduz a metodologia proposta para ajuste automático de parâmetros multiescala de formas por algoritmos de otimização evolucionária, bem como para a avaliação do descritor resultante desse processo de otimização. Apresenta ainda a função objetivo a ser minimizada na busca de soluções ótimas para o problema.

Capítulo 4: expõe e discute os resultados obtidos pela aplicação do método proposto ao problema de classificação de espécies vegetais com os descritores de forma NMBE e IDSC.

Capítulo 5: traz as conclusões do trabalho e as sugestões de pesquisas futuras.

2 FUNDAMENTOS TEÓRICOS

Este capítulo apresenta os conceitos necessários para a compreensão das técnicas utilizadas na metodologia proposta neste trabalho, destacando os métodos de extração de características do contorno das formas, as técnicas de visualização de dados e os métodos de otimização evolutivos.

Inicialmente são apresentados os métodos de extração de características do contorno de formas baseados em assinaturas, em particular a curvatura (KINDRATENKO, 2003), a curvatura multiescala (MOKHTARIAN; MACKWORTH, 1992) e o descritor NMBE (CESAR JR.; COSTA, 1997). Este último, cujos parâmetros são ajustados pelo método proposto, é obtido a partir da curvatura multiescala que, por sua vez, é obtida através da curvatura.

São apresentadas também duas técnicas de visualização de dados, as quais são empregadas na avaliação do método de otimização de descritores proposto: o escalonamento multidimensional (MDS) (COX; COX, 2000) e o mapa auto-organizável do Kohonen (SOM) (KOHONEN *et al.*, 2001). O capítulo finaliza apresentando três métodos de otimização evolutivos utilizados pelo método proposto: resfriamento simulado (SA), evolução diferencial (DE) e otimização por enxame de partículas (PSO).

2.1 Sistemas de recuperação de imagens pelo conteúdo

A disseminação dos sistemas de *hardware* embarcado com capacidade de coletar, armazenar e disponibilizar imagens digitais em redes de computadores, ampliou a demanda por aplicações de visão computacional em áreas como a medicina, indústria, segurança, biologia, entre outras. Dentre estas demandas destacam-se as ferramentas efetivas de gerenciamento de informação multimídia que facilitam a organização e a busca automática desse tipo de conteúdo.

Busca de imagens em grandes bases de dados multimídia é um dos serviços importantes disponibilizados pelos sistemas de gerenciamento de informação na atualidade. O método clássico para se prover tal serviço emprega a rotulação textual por palavras-chave. Nessa abordagem, o usuário realiza buscas fornecendo informações textuais ao sistema. Esse último recupera a informação multimídia empregando métodos tradicionais de buscas textuais por palavras-chave. Atualmente, essa abordagem tem se tornado inviável devido ao grande volume de informação multimídia disponível na internet, o que torna o processo de anotação textual demasiadamente dispendioso. Ademais o processo de descrição textual é impreciso e sujeito a

erros, uma vez que diferentes indivíduos tendem a interpretar e descrever uma mesma imagem subjetivamente e, portanto, empregando diferentes palavras-chave na descrição. Os sistemas de recuperação de imagens por conteúdo (CBIR) foram propostos em resposta a essas dificuldades. Nesses sistemas, o processo de busca utiliza o conteúdo visual das imagens ao invés da rotulação textual.

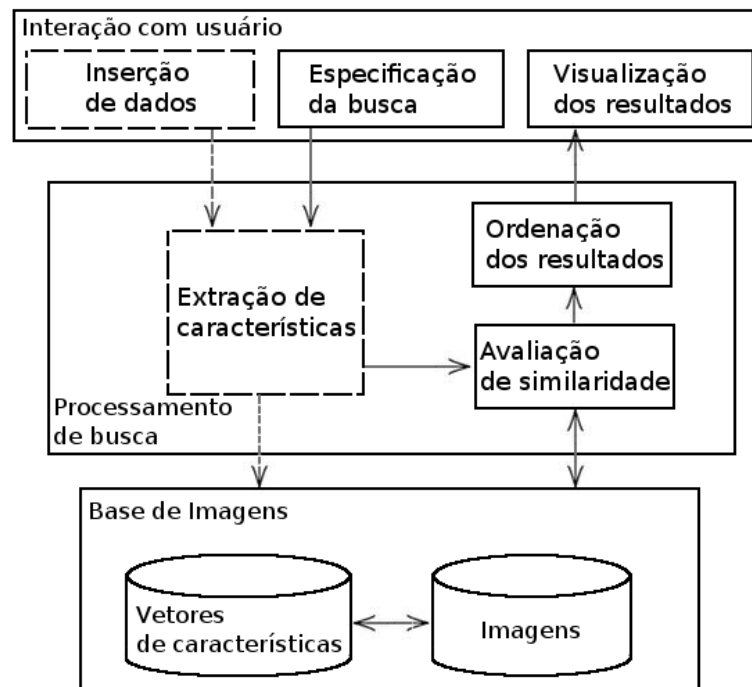
2.1.1 Arquitetura dos sistemas de recuperação de imagens pelo conteúdo

Sistemas de CBIR empregam o conteúdo visual das imagens para formar uma base de dados de representações na forma de vetores de características. A busca é então realizada com o usuário fornecendo ao sistema uma imagem, ou figura de consulta, a qual também é representada vetorialmente. Assim, através da avaliação da similaridade existente entre as representações da imagem de consulta e das imagens da base, o sistema procura retornar ao usuário as imagens que sejam do seu interesse. Dependendo das particularidades da aplicação, esse padrão de consulta, que corresponde também a uma imagem, pode ser um esboço, um modelo ou uma cópia exata do que se deseja procurar (SMEULDERS *et al.*, 2000).

A Figura 1 apresenta a arquitetura de um sistema de CBIR clássico, sendo este composto pelos seguintes componentes (TORRES; FALCÃO, 2006):

- a) Módulo de interação com usuário: através desse módulo o usuário fornece como entrada ao sistema uma imagem de consulta e visualiza, como saída, os resultados das buscas. É também através desse módulo que são realizadas a inserção e a remoção de imagens da base de imagens, sendo esse último processo, destacado em linhas tracejadas, realizado pelo administrador do sistema;
- b) Módulo de processamento de busca: esse módulo engloba grande parte da funcionalidade do sistema CBIR, sendo esse composto pelos seguintes submódulos:
 - Extração de características: contempla os algoritmos computacionais responsáveis em obter uma representação matemática, na forma de vetor de características, dos atributos visuais das imagens como cor, textura e forma;
 - Avaliação de similaridade: contempla o método empregado para a avaliação do grau de similaridade existente entre as imagens a partir de suas respectivas representações vetoriais;
 - Ordenação dos resultados: método para seleção das imagens que serão apresentadas ao usuário conforme o grau de similaridade ao padrão fornecido;

Figura 1 – Arquitetura de um sistema de CBIR.



Fonte: Torres e Falcão (2006).

- c) Base de imagens: repositórios onde estão contidas as imagens, bem como suas correspondentes representações vetoriais.

Este trabalho enfatiza os processos de extração de características e de avaliação de similaridade em experimentos CBIR, sendo esses processos destacados nas subseções a seguir.

2.1.2 Extração de características

O processo de extração de características das imagens é fundamental para os sistemas CBIR, e se dá através de operações de processamento de imagens para extrair informação que seja semanticamente relevante dos atributos da imagem para a realização das buscas por similaridade de conteúdo. Os atributos comumente empregados nesse processo são a cor, a textura, a forma e a localização espacial dos elementos constituintes de uma imagem. Alguns métodos de representação de formas por transformação empregam transformada de Fourier (CESAR JR.; COSTA, 1997; SHU *et al.*, 2015). Em CBIR, o custo computacional das buscas deve levar em consideração o custo de se aplicar a transformada à forma de consulta.

A cor é o atributo mais empregado em extração de características em CBIR, pois a mesma permite discriminar uma ampla gama de imagens. Quando se utiliza esse atributo deve-se considerar que o registro da cor em uma imagem varia consideravelmente com a orientação

da superfície imageada, o posicionamento da câmera, a posição da fonte de iluminação e a maneira como a luz interage com os objetos imageados (SMEULDERS *et al.*, 2000). Ademais, a percepção humana da cor é um assunto complexo que ainda não está completamente elucidado (SMEULDERS *et al.*, 2000). De acordo com Torres e Falcão (2006), as técnicas de descrição por cores podem ser agrupadas em duas grandes classes dependendo se a informação de cor é codificada correlacionada com a sua distribuição espacial ou não.

Atributos de textura são úteis na recuperação de imagens de satélites, bem como na busca de documentos digitalizados (SMEULDERS *et al.*, 2000). A textura é definida em termos de estruturas formadas por grupos de *pixels* que aparecem na imagem com certa periodicidade ou aleatoriedade (GONZALEZ; WOODS, 2006). O método já consagrado para descrição de características de texturas é a matriz de co-ocorrência proposta por Haralick *et al.* (1973). No entanto, métodos baseados em transformações wavelet têm sido utilizados, mais notoriamente a transformada wavelet de Gabor (MANJUNATH; MA, 1996; LIU *et al.*, 2009). Outra técnica que tem sido explorada na atualidade para descrição de texturas é a dimensão fractal multiescala (FLORINDO; BRUNO, 2013).

Informação relativa à localização espacial dos elementos de uma imagem permite, por exemplo, recuperar imagens fotográficas de paisagens externas, aonde distingue-se a localização do céu na região superior da imagem. Ademais, imagens fotográficas também tendem a apresentar em seu centro alguma informação significativa que difere do entorno da imagem, sendo essa informação útil no processo de recuperação de imagens de fotografias de pessoas, animais, etc (LIU *et al.*, 2007).

A forma é uma característica intrínseca das imagens que é amplamente explorada pelo sistema de visão dos primatas para o reconhecimento de objetos. Progressos significativos têm sido alcançados no desenvolvimentos de sistemas CBIR a partir desse atributo (SMEULDERS *et al.*, 2000; LIU *et al.*, 2007). Dentre as abordagens empregadas na obtenção de representações a partir das formas, a mais popular é a que as representa através de vetores de características invariantes (ULLMAN, 1995). Tal representação permite que a similaridade entre as formas seja avaliada através de medidas de distância entre vetores. Além de depender da qualidade da segmentação dos objetos, por si só, a representação por vetores de características é pouco discriminativa (ULLMAN, 1995), restringindo a sua aplicação em CBIR a problemas específicos.

Na representação estrutural, as formas são divididas em um conjunto de partes constituintes, sendo cada parte representada individualmente através de vetores de características.

O descritor é composto a partir desses vetores e da relação existente entre as partes, representadas através de estruturas de dados como grafos, árvores ou cadeias de caracteres. Uma vez que esses métodos comparam as formas a partir das suas estruturas, esses conseguem uma boa acurácia, porém a um custo computacional elevado.

O descritor do contorno da forma *shape context* (SC) (BELONGIE *et al.*, 2002) avalia a similaridade entre formas por correspondência de pontos. Esse descritor caracteriza a distribuição dos demais pontos do contorno em relação a um ponto de referência para cada ponto do contorno tomado como referência. Os métodos que usam correspondência, ou alinhamento, representam as formas diretamente no domínio espacial. Esses métodos determinam o grau de similaridade entre duas formas tentando alinhá-las de modo a medir a diferença residual existente entre as mesmas.

2.2 Descritores do contorno das formas

Os estímulos visuais, ao qual os sistemas de visão biológicos estão submetidos, apresentam um elevado grau de redundância de informação. No processamento dessa informação, tais sistemas buscam eliminar essas redundâncias, retendo apenas a quantidade de informação necessária para o desempenho da tarefa requerida.

Em visão computacional e reconhecimento de padrões, dá-se o nome a esse processo de eliminação das redundâncias de extração de características. Esta etapa de baixo nível busca encontrar informação relevante em uma imagem e representá-la convenientemente para a realização de uma tarefa específica de visão computacional. Uma vez que o processo de extração de características determina o desempenho de um sistema de visão computacional, este recebe especial atenção no desenvolvimento de tais sistemas.

Dentre os atributos dos quais se realizam a extração de características, a forma é considerada a mais relevante em diversas aplicações de visão artificial pela riqueza de informações que esta possui. Uma forma é obtida quando um objeto de interesse é identificado e segmentado em uma imagem.

De acordo com Zhang e Lu (2004), a representação das formas envolve a busca por informações, tanto em sua borda como em sua região interior, que sejam perceptualmente relevantes. Obter uma representação, ou descrição dessa natureza é uma tarefa complexa porque, quando projetamos os objetos tridimensionais do mundo real em duas dimensões, perdemos as informações de uma das dimensões. Como resultado, temos uma representação bidimensional

Figura 2 – Obtenção da forma e do contorno da imagem de uma folha.



parcial do objeto projetado. O problema torna-se ainda mais complexo se levarmos em conta que a forma é frequentemente corrompida por ruídos, defeitos, distorções arbitrárias e oclusões.

Zhang e Lu (2004) classificam as técnicas de representação de formas em dois grupos: métodos baseados em contorno e em região. No primeiro grupo, as características das formas são extraídas de toda sua região, enquanto que no segundo grupo, as características são extraídas apenas da borda. Os referidos autores ainda subdividem os métodos em cada um dos grupos supracitados em estruturais e globais. Essa subdivisão é baseada em se a forma como um todo é utilizada na representação ou se apenas partes, segmentos e seções das formas são utilizadas.

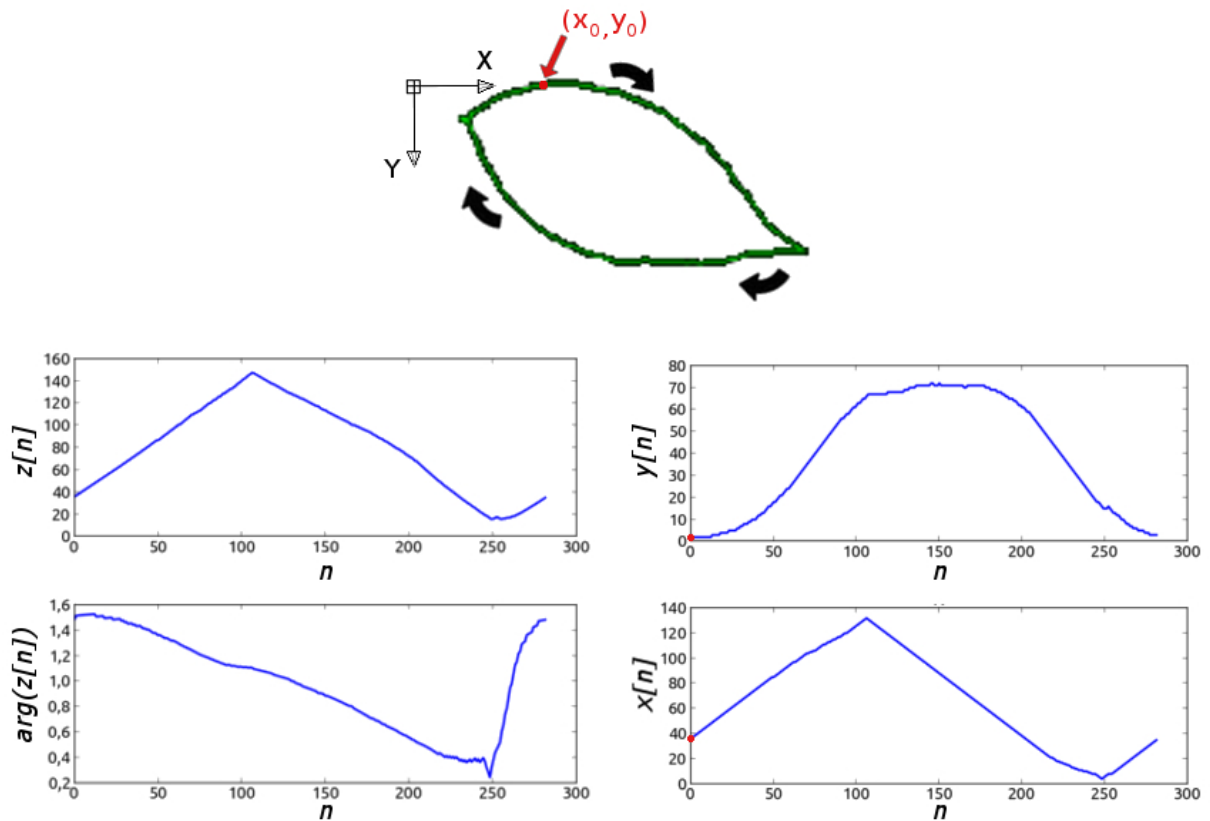
A Figura 2 ilustra as etapas envolvidas no processo de representação de uma forma. Temos nesse caso a segmentação da imagem por limiar seguida da extração do contorno. Com base no contorno, obtém-se uma variedade de representações, na forma de sinais, adequadas para as tarefas de comparação, classificação e reconhecimento de formas.

Este trabalho enfatiza a aplicação de técnicas de processamento de sinais para a obtenção de representações globais do contorno das formas, uma vez que tais técnicas têm sido empregadas com sucesso para esse propósito (COSTA; CESAR JR., 2009).

2.2.1 Representação paramétrica do contorno e curvatura

Costa e Cesar Jr. (2009) definem assinatura de uma forma como sendo um sinal discreto unidimensional que descreve algumas das características do seu contorno ou da sua região. Devido a redução de dimensionalidade, as assinaturas do contorno são representações compactas das formas. Estas assinaturas podem ser utilizadas diretamente como descritores, porém o custo computacional envolvido em sua comparação direta é elevado. Isso porque as assinaturas, em geral, não são totalmente insensíveis a rotação, translação e escalamento das formas, bem como variam no total de amostras conforme varia a resolução das imagens. Algumas assinaturas são mais sensíveis a ruído e a pequenas distorções dos contornos, tornando necessário realizar filtragens. Embora melhore a robustez, tal processo acarreta em alguma perda

Figura 3 – Processo de obtenção da representação paramétrica do contorno da folha da Figura 2.



de informação.

A representação por coordenadas paramétricas consiste na obtenção das coordenadas dos pontos amostrados do contorno, representados a partir de um sistema de coordenadas preestabelecido, percorrendo-o sequencialmente em sentido horário ou anti-horário. Para o contorno discreto Γ , com N amostras, representado num sistema de coordenadas cartesianas, tal processo origina um conjunto de tuplas $\Gamma = (\mathbf{x}[n], \mathbf{y}[n])$, $n \in \{0, 1, \dots, N-1\}$, cujas componentes $\mathbf{x}[n]$ e $\mathbf{y}[n]$ são sinais discretos vetoriais das coordenadas amostradas do contorno.

Na Figura 3, estão representados os sinais obtidos para o contorno da folha da Figura 2 com a amostragem $N = 280$ pontos. Em vermelho, está destacado o ponto de origem aonde a varredura, em sentido horário, se inicia. Os dois pontos observados aonde a evolução do sinal $\mathbf{x}[n]$ inverte sua tendência (de crescente para decrescente e de decrescente para crescente) correspondem aos pontos mais salientes da folha. Já o platô observado no sinal $\mathbf{y}[n]$ corresponde a região da parte inferior da folha, em que quase não se observa variações do contorno ao longo do eixo Y .

Outra representação paramétrica para o contorno é obtida compondo-se um sinal complexo $\mathbf{z}[n] = \mathbf{x}[n] + j\mathbf{y}[n]$, $j = \sqrt{-1}$, $n \in \{0, 1, \dots, N-1\}$. Essa representação é conveniente

Tabela 2 – Propriedades da curvatura e as características geométricas que essas representam.

Propriedade	Característica da forma
Máximo valor absoluto	Ponto saliente
Máximo valor positivo	Saliência convexa
Mínimo valor negativo	Saliência côncava
Valores constantes e nulos	Segmentos retilíneos
Valores constantes e não nulos	Segmentos circulares
Cruzamentos de zero	Pontos de inflexão

Fonte: Costa e Cesar Jr. (2009)

quando se deseja realizar extração de características do contorno através de operações de processamento de sinais. A Figura 3 ilustra o módulo e a fase da representação complexa do contorno da folha da Figura 2.

Embora sejam descritivas, as representações por coordenadas paramétricas apresentam o inconveniente de não serem invariantes a translação, rotação ou escalamento das formas, porém, outras assinaturas com tais propriedades podem ser derivadas a partir dessas mesmas (KINDRATENKO, 2003).

A curvatura é uma assinatura do contorno da forma com importantes propriedades geométricas, o que motiva sua utilização para obtenção de descritores. Há evidências biológicas de que as propriedades desta assinatura sejam exploradas pelo sistema de visão dos primatas nas tarefas de reconhecimento de formas (COSTA; CESAR JR., 2009). Na Tabela 2, estão destacadas as principais propriedades que a curvatura apresenta.

Para uma curva contínua fechada $\Gamma = (x(l), y(l))$, cujo perímetro é L e que se encontra parametrizada em $l \in [0, L]$, a curvatura ($\kappa(l)$) é definida como sendo (KINDRATENKO, 2003):

$$\kappa(l) = \frac{x'(l)y''(l) - x''(l)y'(l)}{((x''(l))^2 + (y''(l))^2)^{\frac{3}{2}}}, \quad (2.1)$$

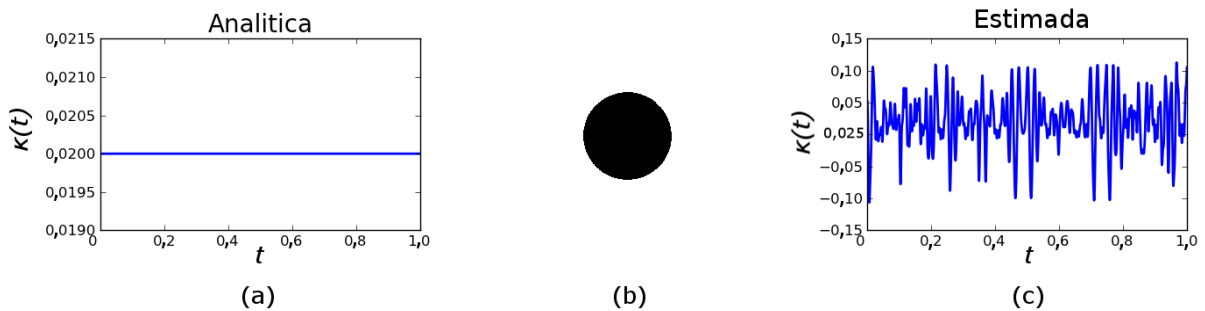
sendo $(x'(l), y'(l))$ e $(x''(l), y''(l))$ as derivadas primeira e segunda das coordenadas paramétricas da curva, respectivamente.

Sob o aspecto computacional, o cálculo da curvatura do contorno de uma forma requer que o mesmo seja espacialmente amostrado e discretizado. Tal processo torna o cálculo das derivadas da Equação 2.1 sensível ao ruído e inviável, o que limita a aplicação direta da curvatura para a obtenção de descritores.

A Figura 4 ilustra tal efeito para uma imagem de forma circular. Em particular, a

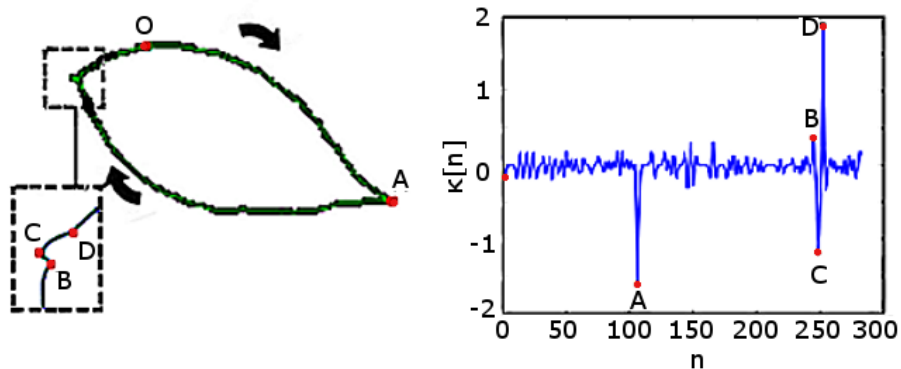
Figura 4a mostra o gráfico da curvatura teórica, ao longo do perímetro t do contorno, obtida analiticamente pela aplicação da Equação 2.1 ao contorno parametrizado do círculo de mesmo raio ao do círculo da Figura 4b. A Figura 4c ilustra a curvatura obtida computacionalmente a partir do contorno discreto extraído do círculo da Figura 4b. Nota-se que a curvatura da Figura 4a (analítica) tem um valor constante $\kappa(l) = \frac{1}{r}$, sendo r o raio do círculo central, enquanto que a curvatura Figura 4c (computacional) varia significativamente em torno do valor analítico esperado, como mostra a Figura 4.

Figura 4 – Efeito do ruído na estimativa computacional da curvatura para uma forma circular.



Diversas estratégias foram propostas na literatura para contornar o problema da sensibilidade ao ruído do cálculo computacional da curvatura. Trabalhos clássicos, como os de Rosenfeld e Johnston (1973) e Beus e Tiu (1987) estimam a curvatura através do ângulo formado entre vetores obtidos dos pontos do contorno. Mokhtarian e Mackworth (1992) introduziram um método que suaviza o contorno, através da convolução do mesmo com um filtro gaussiano, antes de calcular a curvatura pela Equação 2.1. Cazals e Pouget (2005) e Shi *et al.* (2002) apresentaram trabalhos que utilizavam métodos para estimação da curvatura por interpolação de pontos.

Figura 5 – Curvatura estimada do contorno da folha da Figura 2 através de método computacional.



A Figura 5 mostra a curvatura estimada, para o contorno da folha da Figura 2, através

do método proposto por Mokhtarian e Mackworth (1992). O contorno foi previamente suavizado com um filtro gaussiano com desvio padrão $\sigma = 20$. Os pontos aonde a curvatura apresenta os picos em destaque correspondem aos pontos salientes da folha.

2.2.2 Representações multiescala

Qualquer método que se proponha a descrever características das formas com base em contornos deve ser capaz de representá-las de forma confiável e precisa.

Embora carregue grande parte da informação a respeito das formas, o contorno é sensível a ruídos, oclusões e variações das formas. Em outros casos, o contorno não se apresenta completamente disponível, com regiões disjuntas e descontinuidades. Tais aspectos comprometem a confiabilidade e a precisão das assinaturas baseadas em contorno.

A descrição multiescala do contorno das formas vem se mostrando uma alternativa viável para superação de tais problemas. Esta técnica se baseia no método proposto por Witkin (1987) e Koenderink (1984) para análise de sinais em vários níveis de resolução. Esses autores introduziram o conceito de fator de escala, cujo ajuste determina o grau de resolução na qual ocorre a análise do sinal.

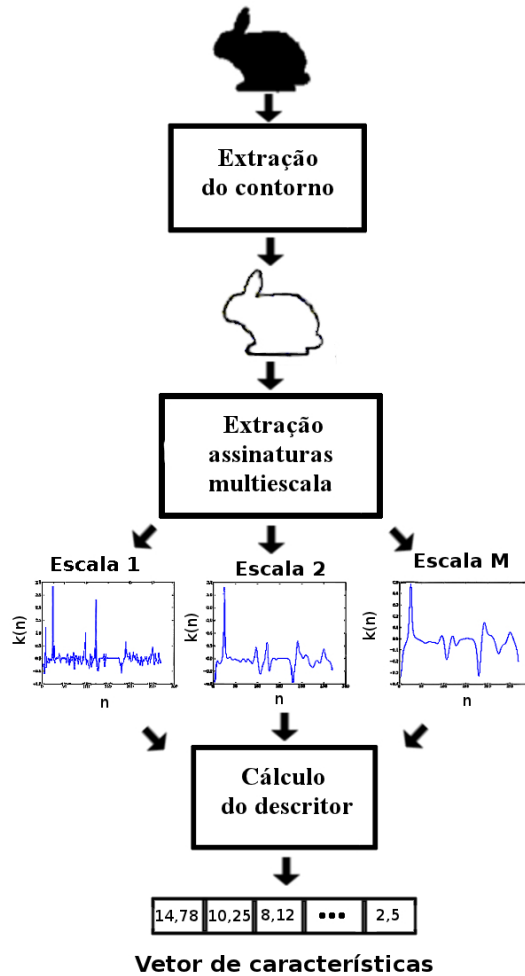
Perante outros métodos, a grande vantagem da representação multiescala é a sua habilidade em representar os atributos das formas em vários níveis de detalhes, variando de escalas de baixa resolução, aonde os detalhes que diferenciam as formas de uma mesma classe não são levados em consideração, até escalas de alta resolução aonde esses detalhes são preservados (ULLMAN, 1995).

A Figura 6 ilustra o processo para obtenção de uma representação multiescala de um sinal unidimensional discreto $k[n]$ que represente o contorno de uma forma, tal como as assinaturas apresentadas na Subseção 2.2.1. Emprega-se nesse processo uma função de transformação $F[n, \sigma]$, geralmente com características de filtragem passa-baixas e com frequência de corte σ . Na função de transformação, a referida frequência de corte corresponde ao fator de escala na qual o sinal $k[n]$ será analisado.

Para um vetor de escalas $\sigma = (\sigma_1 \sigma_2 \dots \sigma_M)$, a representação multiescala consiste em um vetor de sinais $k[n, \sigma] = (k[n, \sigma_1] k[n, \sigma_2] \dots k[n, \sigma_N])$, sendo cada termo $k[n, \sigma_i]$ obtido através de

$$k[n, \sigma_i] = F[n, \sigma_i] * k[n], \quad (2.2)$$

Figura 6 – Método de análise multiescala a partir da assinatura do contorno de uma forma.



em que o operador $*$ denota a convolução entre o sinal $k[n]$ e a transformação $F[n, \sigma_i]$. Assim, cada elemento de $k[n, \sigma]$ corresponde ao sinal $k[n]$ analisado na escala σ_i .

No caso de curvas planas, a análise multiescala permite descrevê-las em vários níveis de detalhes e de abstração a partir de suas assinaturas, tornando a representação mais discriminativa que os métodos que empregam medidas quantitativas globais, como área, perímetro e compacidade, devido a sua robustez e estabilidade (DIREKOGLU; NIXON, 2008).

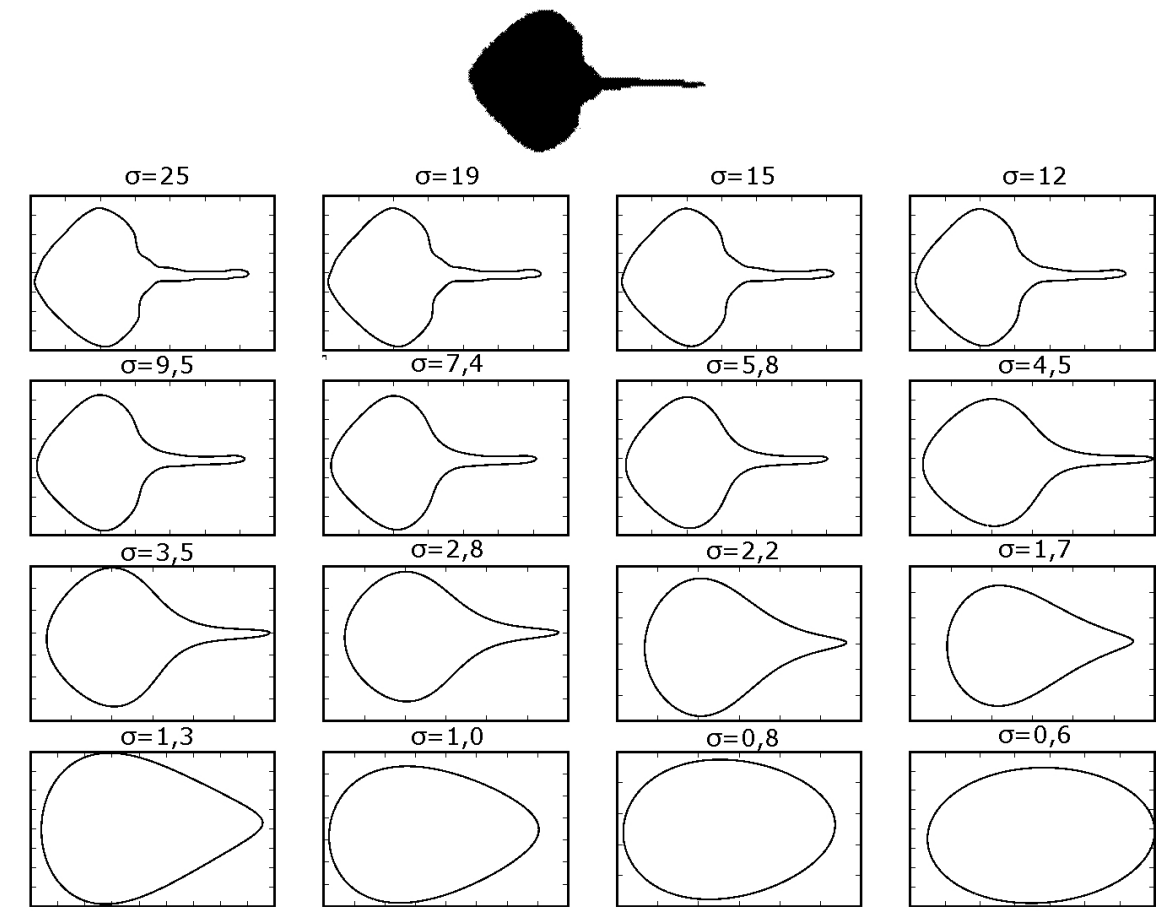
Curvatura Multiescala

Inspirados na técnica de Witkin (1987) e Koenderink (1984), Mokhtarian e Mackworth (1992) propuseram um método para análise multiescala do contorno através da assinatura da curvatura. O referido método emprega como função de transformação um filtro passa-baixas gaussiano, $g_{\sigma}(l) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{l^2}{2\sigma^2}}$, que suaviza o contorno antes do cálculo de sua curvatura. Nesse caso, o ajuste do desvio padrão da função gaussiana (σ^2) atua como fator de escala, regulando a

largura de banda do filtro e o nível de suavização do contorno.

A Figura 7 ilustra o efeito da suavização do contorno para diferentes valores do fator de escala. Observa-se nesta figura que, à medida que o fator de escala aumenta, o grau de suavização também aumenta e, conseqüentemente, mais informações de detalhes são eliminadas do contorno suavizado. Esse efeito permite obter uma representação do contorno em diferentes níveis de detalhes.

Figura 7 – Efeito da suavização sobre o contorno da forma para diferentes valores de escala.



No caso contínuo, temos as coordenadas do contorno suavizado realizando a convolução entre as coordenadas do contorno parametrizado $C(l) = (x(l), y(l))$ e o filtro gaussiano:

$$x_{\sigma}(l) = x(l) * g_{\sigma}(l) = \int_{-\infty}^{\infty} x(v)g_{\sigma}(l-v)dv \quad (2.3)$$

$$y_{\sigma}(l) = y(l) * g_{\sigma}(l) = \int_{-\infty}^{\infty} y(v)g_{\sigma}(l-v)dv \quad (2.4)$$

Para o cálculo das derivadas $x'_\sigma(l)$, $y'_\sigma(l)$, $x''_\sigma(l)$ e $y''_\sigma(l)$, necessárias para o cálculo da curvatura do contorno suavizado, temos, pelas propriedades da convolução, $x'_\sigma(l) = x(l) * g'_\sigma(l)$, $y'_\sigma(l) = y(l) * g'_\sigma(l)$, $x''_\sigma(l) = x(l) * g''_\sigma(l)$ e $y''_\sigma(l) = y(l) * g''_\sigma(l)$.

O cálculo da curvatura do contorno suavizado $\kappa_\sigma(l)$ se dá através da equação 2.1, substituindo-se $x'(l)$, $y'(l)$, $x''(l)$ e $y''(l)$ por $x'_\sigma(l)$, $y'_\sigma(l)$, $x''_\sigma(l)$ e $y''_\sigma(l)$, respectivamente, ou seja,

$$\kappa_\sigma(l) = \frac{x'_\sigma(l)y''_\sigma(l) - x''_\sigma(l)y'_\sigma(l)}{((x''_\sigma(l))^2 + (y''_\sigma(l))^2)^{\frac{3}{2}}}. \quad (2.5)$$

Uma outra abordagem utilizada para se calcular a curvatura multiescala, que foi proposta por Cesar Jr. e Costa (1996) e adotada nesta tese, opera com a representação do contorno no domínio da frequência. A partir da transformada de Fourier das coordenadas do contorno suavizado $X_\sigma(f) = F\{x_\sigma(l)\}$ e $Y_\sigma(f) = F\{y_\sigma(l)\}$. Os referidos autores calculam derivadas utilizando a propriedade da derivada da transformada de Fourier, ou seja:

$$x'_\sigma(l) = F^{-1}\{2\pi jfX_\sigma(f)\}, \quad (2.6)$$

$$y'_\sigma(l) = F^{-1}\{2\pi jfY_\sigma(f)\}, \quad (2.7)$$

$$x''_\sigma(l) = F^{-1}\{-(2\pi f)^2X_\sigma(f)\} \text{ e} \quad (2.8)$$

$$y''_\sigma(l) = F^{-1}\{-(2\pi f)^2Y_\sigma(f)\}, \quad (2.9)$$

em que $F^{-1}\{X(f)\}$ denota a transformada de Fourier inversa do sinal $X(f)$.

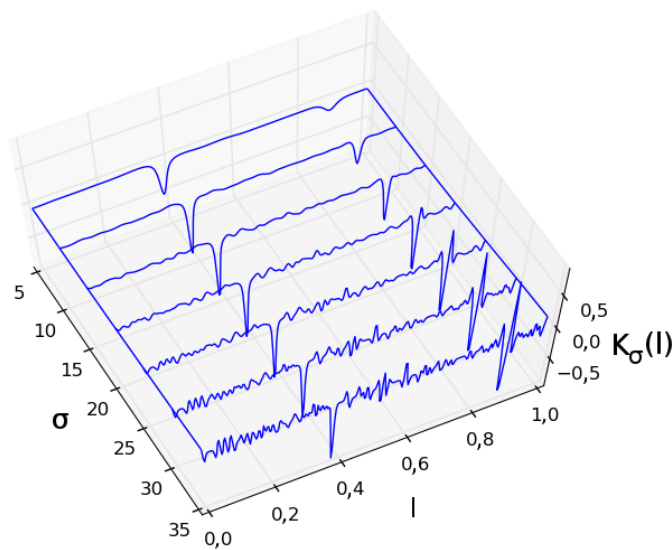
Na suavização do contorno no domínio da frequência, ao invés da convolução, realiza-se o produto dos sinais $X(f)$ e $Y(f)$ com a transformada de Fourier da expressão do filtro gaussiano:

$$X_\sigma(f) = X(f)G_\sigma(f), \quad (2.10)$$

$$Y_{\sigma}(f) = Y(f)G_{\sigma}(f) \quad (2.11)$$

$$G_{\sigma}(f) = F\{g_{\sigma}(l)\} = e^{-2\pi^2 f^2 \sigma^2}. \quad (2.12)$$

Figura 8 – Curvatura multiescala do contorno da forma da folha da Figura 2.



A Figura 8 ilustra a evolução do sinal de curvatura do contorno da folha da Figura 2 para diferentes níveis de suavização. Os picos da curvatura que se preservam nas escalas de baixa resolução correspondem às informações mais salientes do contorno, que o caracterizam globalmente. As informações de detalhes, que tendem a desaparecer nas escalas de baixa resolução e se preservam nas escalas de alta resolução, representam as características mais específicas do contorno. Essa imagem com diferentes curvaturas multiescalas também é conhecida na literatura como curvograma (CESAR JR.; COSTA, 1996).

Cesar Jr. e Costa (1997) propuseram um método para o ajuste das escalas através das expressões

$$oct_l = (\sqrt{2})^l; l = 1, \dots, S \quad (2.13)$$

e

$$\sigma_l = \left(\frac{\sigma_{max} - \sigma_{min}}{\sigma_{max} - \sqrt{2}} \right) (\sigma_l - \sqrt{2}) + \sigma_{min}, \quad (2.14)$$

em que S é o número de escalas, l a l -ésima escala calculada $\sigma_{min} = \sigma_1$ a menor escala e $\sigma_{max} = \sigma_S$ a maior escala.

Energia de dobramento multiescala normalizada

Young *et al.* (1974) propuseram a energia de dobramento como uma medida de complexidade para análise de formas biológicas. Conceitualmente, esta é definida como sendo a energia necessária para se modificar uma forma, através de deformações, ao seu estado de menor energia, ou seja, um círculo com o mesmo perímetro da forma.

A maneira mais direta de se obter a energia de dobramento de um contorno fechado é a partir de sua curvatura pela seguinte expressão:

$$E = \frac{1}{L} \int_l \kappa^2(l) dl, \quad (2.15)$$

sendo L o perímetro do contorno e a integral calculada ao longo do comprimento de seu arco. O resultado da equação 2.15 é um escalar que representa a energia média do sinal da curvatura.

A energia de dobramento multiescala normalizada (NMBE) foi introduzida por Cesar Jr. e Costa (1997) para a análise de formas de neurônios. Nesta tese, investigamos sua utilização como um descritor de propósito geral em recuperação de formas pelo conteúdo. Para um contorno discreto, com N pontos, representado na forma complexa $z[n] = x[n] + jy[n]$, $n \in [0, 1, \dots, N-1]$, a NMBE é dada por:

$$E_{be} = \frac{L^2}{N} \sum_{n=0}^{N-1} \kappa^2[n], \quad (2.16)$$

sendo o perímetro do contorno elevado ao quadrado (L^2) uma constante de normalização para que o descritor tenha invariância a escala. A curvatura discreta $\kappa[n]$ é calculada a partir de $z[n]$ através da seguinte expressão:

$$\kappa[n] = \frac{-Im(z'[n](z''[n])^*)}{|z'[n]|^3}, \quad (2.17)$$

onde $z'[n]$ e $z''[n]$ correspondem as derivadas primeira e segunda, $z^*[n]$ o conjugado de $z[n]$ e $Im(z)$ a coordenada imaginária de z .

A NMBE resulta da versão discreta da curvatura multiescala. Podemos calcular as derivadas primeira e segunda do contorno discreto suavizado ($z'_\sigma[n]$ e $z''_\sigma[n]$) através das propriedades da derivada da convolução ou da transformada de Fourier (CESAR JR.; COSTA, 1997). No caso discreto, a transformada de Fourier de $z[n]$ é dada por:

$$Z[s] = F\{z[n]\} = \sum_{n=0}^{N-1} z[n] e^{-j2\pi ns/N}, \quad (2.18)$$

em que $s = -N_2 \dots N - N_2 - 1$ e $N_2 = \text{floor}(\frac{N}{2})$ (maior valor inteiro $\leq \frac{N}{2}$). A transformada inversa é dada por

$$z[n] = F^{-1}\{Z[s]\} = \sum_{s=-N_2}^{N-N_2-1} Z[s] e^{j2\pi ns/N}, \quad n = 0 \dots N - 1. \quad (2.19)$$

No domínio s o contorno é suavizado multiplicando-se, elemento a elemento, $Z[s]$ pela transformada de Fourier da versão discreta do filtro passa-baixas gaussiano $g_\sigma[n] = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{n^2}{2\sigma^2}}$, $n = 0 \dots N - 1$:

$$Z_\sigma[s] = Z[s] \cdot F\{g_\sigma[n]\}, \quad (2.20)$$

sendo as referidas derivadas do contorno suavizado dadas por:

$$z'_\sigma[n] = F^{-1}\{j2\pi s Z_\sigma[s]\} \quad (2.21)$$

$$z''_\sigma[n] = F^{-1}\{-(2\pi s)^2 Z_\sigma[s]\}. \quad (2.22)$$

O processo de filtragem passa-baixas reduz a energia espectral da representação complexa do contorno causando portanto, o encolhimento do seu perímetro. Uma estratégia para compensar tal efeito é normalizar o contorno suavizado com a razão entre o perímetro do contorno não suavizado (L) e o seu perímetro (L_σ) (CESAR JR.; COSTA, 1996; CESAR JR.; COSTA, 1997):

$$\check{z}_\sigma[n] = \frac{L}{L_\sigma} z_\sigma[n]. \quad (2.23)$$

Substituindo $\kappa[n]$ por $\kappa_\sigma[n]$ e $z[n]$ por $\check{z}_\sigma[n]$ nas Equações 2.17 e 2.16, e realizando estes cálculos para M escalas distintas ($\sigma_1, \sigma_2, \dots, \sigma_M$) resulta na seguinte representação multiescala da energia de dobramento:

$$NMBE = (\log E_{\sigma_1}, \log E_{\sigma_2}, \dots, \log E_{\sigma_M}). \quad (2.24)$$

Dimensão fractal multiescala

As raízes conceituais da geometria fractal foram introduzidas por Mandelbrot (2000) para descrever os objetos que as definições tradicionais da geometria euclidiana não consegue descrever. A dimensão fractal é definida a partir de um número não inteiro que quantifica a complexidade de um objeto, sob os aspectos de fragmentação e auto-similaridade (MANDELBROT, 2000). Complexidade é uma propriedade importante dos objetos que informa quanto espaço este ocupa (COSTA; CESAR JR., 2009).

A dimensão fractal multiescala (DFM) estima a complexidade das formas através de uma curva que representa as mudanças na complexidade à medida que a escala de visualização da forma varia (FLORINDO *et al.*, 2012). O método empregado para estimar a dimensão fractal e a DFM nesta tese é o método de Minkowski-Bouligand (COSTA; CESAR JR., 2009). Este método dilata a forma sob análise utilizando como elemento estruturante um disco de raio $r > 0$, sucessivamente. A inclinação da interpolação linear da curva $\log A(r)$ versus $\log r$, sendo $A(r)$ a área total dilatada para o elemento estruturante de raio r , determina a estimativa da dimensão fractal D_f como sendo:

$$D_f = 2 - \lim_{r \rightarrow 0} \frac{\log A(r)}{\log r}. \quad (2.25)$$

A derivada da curva log-log, representada através de N valores discretos de raios $r_i > 0$, determina a DFM:

$$DFM = (D_f(t_1), D_f(t_2), \dots, D_f(t_N)), \quad (2.26)$$

sendo $D_f(t) = 2 - \frac{du(t)}{dt}$, $t = \log r$ e $u(t) = \log A(t)$.

2.3 Técnicas de visualização de dados

Um problema dos algoritmos empregados na extração de características que foram abordados nesse capítulo é a elevada dimensionalidade da representação vetorial obtida. Além de dificultar compreender a organização dos dados, essa elevada dimensionalidade acarreta em maior complexidade do sistema computacional que deve lidar com essa informação.

Através de técnicas de redução de dimensionalidade é possível gerar visualizações gráficas dos dados que permitem compreender melhor sua estrutura. Este capítulo apresenta duas técnicas de visualização de dados que foram empregadas neste trabalho na avaliação da capacidade discriminativa dos descritores do contorno de formas: escalonamento multidimensional (MDS) e mapa auto-organizável do Kohonen (SOM).

2.3.1 Escalonamento multidimensional

O MDS (COX; COX, 2000) consiste em um método para representação de dados de um espaço inicial multidimensional para um espaço de dimensão reduzida, preservando as relações de distâncias existentes no espaço inicial. Esse algoritmo modela a similaridade ou dissimilaridade dos dados como distâncias no espaço geométrico.

Há dois tipos de algoritmos MDS: o métrico e o não métrico. No primeiro, a matriz de similaridade de entrada é construída a partir de uma métrica de distância que respeita a desigualdade triangular. Assim, as distâncias entre dois pontos no espaço de dimensão reduzida são ajustadas para serem as mais próximas possível daquelas encontradas no espaço de dimensão inicial. Em sua versão não métrica, o algoritmo tenta preservar a ordem das distâncias, procurando uma relação monotônica não paramétrica entre as proximidades das distâncias do espaço de dimensão reduzida e as similaridades no espaço de dimensão inicial.

O coeficiente de determinação R^2 mede o ajuste da representação dos dados em um espaço de menor dimensão. Este coeficiente indica, em porcentagem, o ajuste do modelo aos dados observados. Assim, quanto mais próximo de 1 o valor de R^2 , melhor é o ajuste do modelo aos dados observados. Sejam $\mathbf{D} = (d_{ij})_{N \times N}$ e $\hat{\mathbf{D}} = (\hat{d}_{ij})_{N \times N}$ as matrizes de distância simétricas entre N vetores nos espaços de dimensão M e O ($M > O$) respectivamente. O coeficiente de determinação é dado por

$$R^2 = 1 - \frac{\sum_{i=1}^N \sum_{j=i}^N (\hat{d}_{ij} - d_{ij})^2}{\sum_{i=1}^N \sum_{j=i}^N (d_{ij} - \bar{d})^2}, R^2 \in [0, 1], \quad (2.27)$$

em que d_{ij} é a distância entre as amostras i e j no espaço de dimensão M , \hat{d}_{ij} a distância entre as amostras i e j no espaço de dimensão O e \bar{d} é a distância média no espaço de dimensão M .

2.3.2 *Mapa auto-organizável de Kohonen*

O SOM (KOHONEN *et al.*, 2001) consiste em um tipo de rede neural de aprendizagem não supervisionada. Desenvolvida por Kohonen (1982), a SOM projeta os vetores apresentados em sua entrada de um espaço N-dimensional para um espaço bidimensional, preservando a estrutura topográfica do espaço vetorial de origem. Em outras palavras, se dois vetores encontram-se próximos no espaço de entrada, estes preservarão essa relação de proximidade no espaço de projeção.

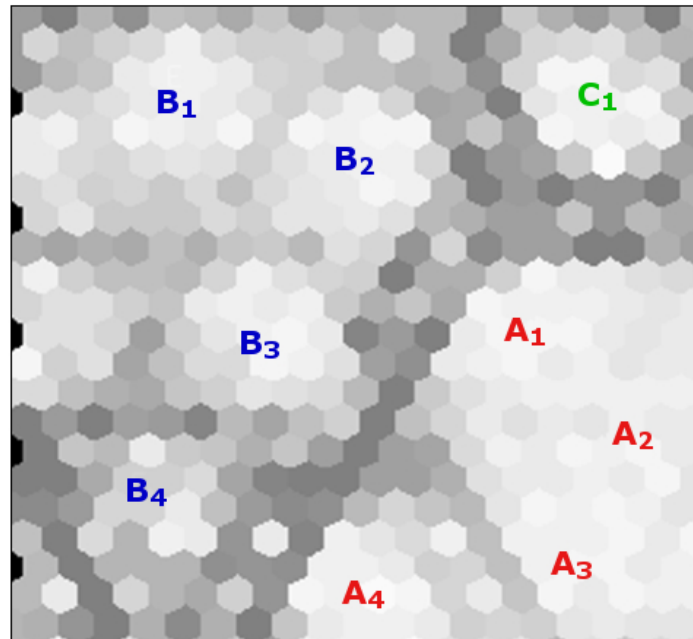
Sendo uma ferramenta para análise exploratória de dados, esse tipo de rede neural tem sido empregada para visualização de imagens (STRONG; GONG, 2011), identificação de agrupamentos (KUROIWA *et al.*, 2000), classificação de texturas (MARANA *et al.*, 1997), bem como outras aplicações.

Ultsch e Siemon (1990) demonstraram que, embora a SOM organize os vetores em agrupamentos, esta não representa as distâncias entre os mesmos de maneira fidedigna. Isso torna a análise direta do mapa de projeções não adequada para a avaliação dos agrupamentos estabelecidos. Para contornar esse problema, os referidos autores desenvolveram um método bidimensional de representação conhecido como matriz unificada de distâncias ou matriz-U. Obtida a partir da SOM essa matriz mostra, preservando a topologia, a relação de distância entre as estruturas mapeadas (ULTSCH; SIEMON, 1990).

A Figura 9 tem por objetivo auxiliar na compreensão e interpretação da informação contida na matriz-U. Neste exemplo a matriz-U é representada por um conjunto de células em níveis de cinza. As estruturas rotuladas representam os neurônios da SOM com maior ativação para um dado padrão apresentado à rede durante o treinamento. Estruturas cujos rótulos são de mesma cor correspondem aos padrões pertencentes a uma mesma classe, enquanto as demais células, em níveis de cinza, representam o grau de separação existente entre os padrões. Assim, células mais escuras implicam em maior separação entre as estruturas mapeadas, enquanto células mais claras indicam maior grau de proximidade.

Nesta matriz, a rede SOM mapeou as três estruturas da classe vermelha (A_1 , A_2 e A_3) com pouca separação inter-classes, uma vez que não há regiões escuras no entorno das mesmas. A estrutura A_4 , que também pertence a classe vermelha, encontra-se mais afastada das estruturas

Figura 9 – Representação da Matriz-U como imagem em níveis de cinza. As células brancas rotuladas correspondem aos neurônios da rede SOM. Células escuras indicam maior separação entre neurônios e células claras indicam maior proximidade entre neurônios.



referidas anteriormente, pois há uma barreira mais escura a separando das demais. Há também uma grande barreira que separa as classes entre si. As quatro estruturas da classe vermelha, por exemplo, estão bem separadas das demais, pertencentes a classe azul (B_1 , B_2 , B_3 e B_4) e verde (C_1).

A mesma análise realizada para classe vermelha vale para a classe azul, embora haja maior separação intra-classe entre as estruturas desta última, como no caso da estrutura B_4 , que encontra-se bem afastada das demais de sua classe. A estrutura C_1 , da classe verde, encontra-se bem separada de todas as demais, uma vez que há uma barreira escura em todo seu entorno.

2.4 Métodos de otimização

Problemas de otimização, que consistem em encontrar mínimo e máximo de uma função objetivo de múltiplas variáveis, ocorrem com frequência em diversas áreas do conhecimento, como engenharia, economia e ciências. Métodos clássicos como Newton-Rapson e suas variantes, funcionam muito bem no caso de funções convexas, em que a função objetivo é diferenciável e com um único ótimo global, porém tais condições não são encontradas em diversos problemas.

Os métodos de otimização evolutivos foram desenvolvidos para lidar com o problema

de otimização de funções não convexas e não diferenciáveis. Inspirados nos processos biológicos evolutivos, esses métodos utilizam de metaheurísticas e uma população de soluções candidatas para realizar buscas no espaço de pesquisa pela solução ótima. Esta seção apresenta três algoritmos metaheurísticos de otimização que foram utilizados nesta tese no ajuste de parâmetros dos descritores multiescala: resfriamento simulado (SA) (ENGELBRECHT, 2007), evolução diferencial (DE) (STORN; PRICE, 1997) e otimização por enxame de partículas (PSO) (SHI; EBERHART, 1998).

2.4.1 *Resfriamento simulado*

O método de otimização SA simula o processo termodinâmico de aquecimento e resfriamento de um metal. O mecanismo desse método para exploração do espaço de busca consiste numa variável de temperatura que determina a probabilidade de aceitação de uma solução pior àquela encontrada até o momento, sendo que quanto maior for a temperatura, maior será essa probabilidade. Soluções melhores que a solução atual são sempre aceitas pelo algoritmo, enquanto que a probabilidade de aceitar soluções piores decresce exponencialmente à medida que o algoritmo interage.

Na exploração do espaço de busca, esse algoritmo também requer um mecanismo de perturbação da solução atual, que consiste em adicionar ou subtrair, com uma dada probabilidade P_r , um ruído de natureza estocástica a cada coordenada da solução atual. Na implementação utilizada neste trabalho esse mecanismo corresponde a um ruído branco de variância 0,2.

2.4.2 *Evolução diferencial*

Analogamente ao algoritmo genético (GA) de otimização, o algoritmo DE (STORN, 1996) emprega mecanismos de cruzamento e mutação de espécies para evoluir uma população de N indivíduos. No entanto, a forma como tais mecanismos são implementados no DE difere significativamente em relação ao GA.

O cruzamento consiste em produzir um novo indivíduo (\mathbf{d}) a partir de três indivíduos distintos sorteados da população σ_1 , σ_2 e σ_3 que sejam diferentes do indivíduo atualmente em evolução (σ_k). Para a variante do algoritmo *DE/rand/1* (STORN, 1996), a regra de cruzamento é

$$d = \sigma_1 + \beta \cdot (\sigma_2 - \sigma_3), \quad (2.28)$$

sendo o parâmetro β um fator de amplificação da diferença entre os vetores σ_2 e σ_3 .

O mecanismo de mutação do DE utiliza o indivíduo d , produzido pela Equação 2.28, para perturbar o vetor de solução candidata atualmente em evolução σ_k . Assim, com probabilidade de mutação P_r , substitui-se cada coordenada de σ_k pela coordenada de d correspondente.

2.4.3 Otimização por enxame de partículas

A PSO consiste em um algoritmo bio-inspirado que evolui uma população de N partículas que se movimentam no espaço de busca a uma dada velocidade. A cada iteração o PSO: a) registra as melhores posições alcançadas por cada partícula (**pp**) e a melhor posição globalmente encontrada por todas as partículas do enxame (**gp**); b) corrige a velocidade das partículas considerando dois fatores de atração c_1 e c_2 . O primeiro fator controla a tendência da partícula procurar a solução ótima nas redondezas de **pp** e o segundo a tendência da partícula se movimentar em direção a **gp**; c) Atualiza a posição das partículas a partir das suas velocidades.

Para melhorar a convergência a um mínimo global, as partículas são desaceleradas exponencialmente a cada iteração por um fator de inércia ω (SHI; EBERHART, 1998). O ajuste dos parâmetros do algoritmo PSO é um aspecto importante a ser considerado. Um estudo das características de convergência do algoritmo, bem como recomendações de ajuste dos parâmetros com base nessas características é encontrado em (JIANG *et al.*, 2007).

3 METODOLOGIA

O Capítulo 2 apresentou alguns dos métodos encontrados na literatura para a representação computacional de formas a partir do contorno. Em muitos casos, esses métodos apresentam parâmetros que requerem ajustes dependentes da natureza da aplicação, ou seja, das características da base de imagens e o propósito ao qual o sistema de reconhecimento se destina. A energia de dobramento multiescala, por exemplo, requer o ajuste do número e dos valores das escalas utilizadas na representação das formas.

Esse capítulo trata do método proposto neste trabalho, que customizar descritores multiescala à base de imagens de formas melhorando o desempenho em classificação supervisionada e não supervisionada. Inicialmente é apresentando, em linhas gerais, o método em questão para em seguida detalhar cada uma das etapas envolvidas tanto no processo de customização dos descritores como na avaliação dos resultados.

3.1 Customização dos descritores

O objetivo do método proposto para customização de descritores, cujo esquema é apresentado na Figura 10, é melhorar a representação do descritor de formas ajustando seus parâmetros através de algoritmos de otimização evolutivos. Foi utilizada como função objetivo na otimização o desvio mediano absoluto (MAD) da métrica de avaliação de agrupamentos *silhouette* (ROUSSEEUW, 1987). O que motivou a escolha dessa função objetivo é sua robustez a *outliers* (ROUSSEEUW; LEROY, 1987), sendo sua equação definida por

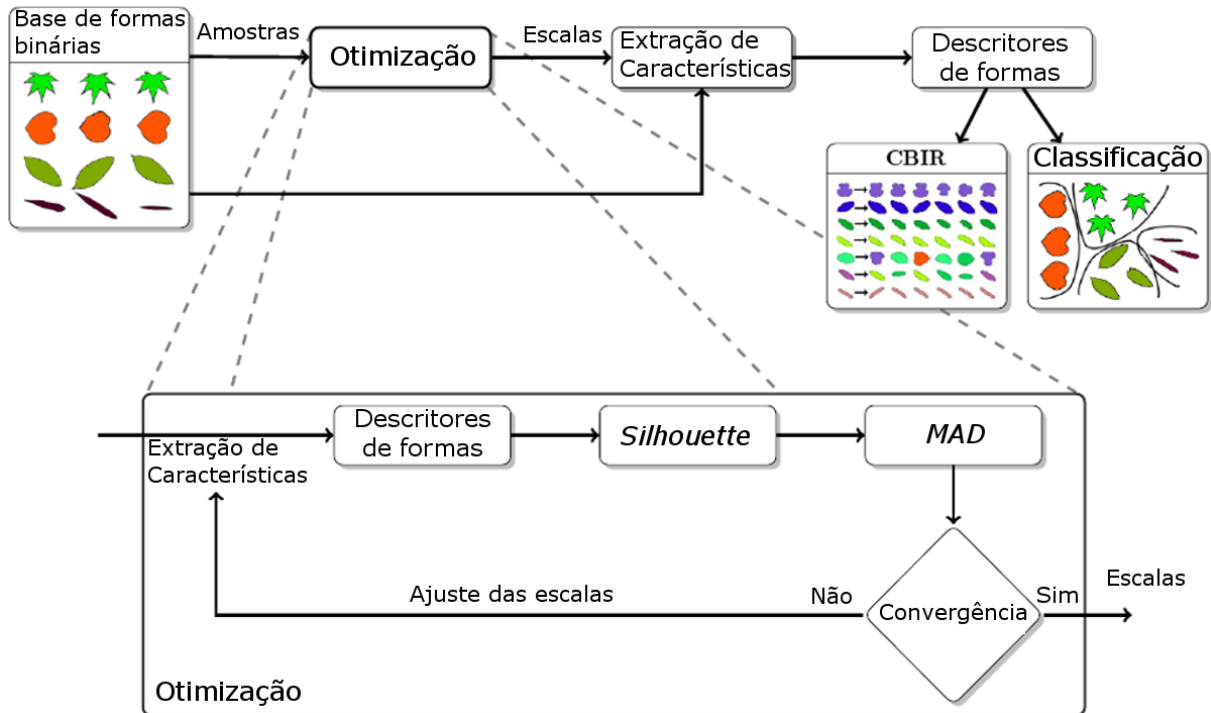
$$MAD = \text{mediana}(|s_i - 1|_{i=1,2,\dots,L}), \quad (3.1)$$

sendo $S = \{s_1, s_2, \dots, s_L\}$ é o conjunto das *silhouettes* calculadas para L descritores de forma. Os operadores $|\cdot|$ e *mediana*() retornam o valor absoluto e a mediana de um conjunto de valores, respectivamente.

A *silhouette* (ROUSSEEUW, 1987) é uma medida de qualidade de agrupamentos que indica o grau de afinidade de uma amostra a um agrupamento, considerando as distâncias médias entre classes e intra classes de um objeto i atribuído a uma dada classe A . Logo, esta métrica é definida como

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \in [-1, 1], \quad (3.2)$$

Figura 10 – Proposta de uma metodologia para otimização evolucionária de um descritor multi-escala de forma.



sendo a_i a dissimilaridade média entre o objeto i e os demais objetos pertencentes a mesma classe de a_i e b_i é a dissimilaridade média do objeto i e a classe vizinha mais próxima de i , excluída sua própria classe.

Essa métrica pode assumir valores no intervalo $[-1, 1]$, sendo que valores negativos indicam que o grau de afinidade de um objeto à classe que este fora atribuído é baixo. Já valores positivos indicam alta afinidade de um objeto à classe que este fora atribuído. Um valor de *silhouette* próximo de zero indica que o objeto está na fronteira entre duas classes e que há, portanto, um grau de incerteza a respeito de qual classes este pertence.

A função objetivo MAD assume valores no intervalo $[0, 2]$. De forma análoga à *silhouette*, um valor igual a zero indica que a estrutura dos agrupamentos é perfeita, enquanto que valores próximos de 2 indicam que a estrutura dos agrupamentos é deficiente, com baixa similaridade entre os objetos de mesma classe ou alta similaridade entre os objetos de classes distintas.

Uma vez definida a função custo, o processo de otimização dos parâmetros do descritor visa ajustá-lo ao problema em estudo. No caso da análise de formas de folhas, a otimização permite que os parâmetros, que minimizam a função objetivo ou função custo (MAD), incorporem nuances e detalhes do contorno das formas de folhas. Assim, tal ajuste reflete em melhoria na acurácia da classificação e recuperação de formas de folhas de plantas.

Vale destacar que a metodologia é versátil pois é ajustável a outras aplicações e portanto, suporta a definição de uma outra função objetivo. A Figura 10 ilustra, em detalhes, como ocorre o ajuste dos parâmetros do descritor multiescala de acordo com a metodologia proposta, aplicada a um problema de análise de formas. Primeiramente, é amostrado na base de folhas um subconjunto das formas para, em seguida, realizar o procedimento de otimização e encontrar o melhor conjunto de parâmetros de escala $\sigma_{otim} = (\sigma_1, \sigma_2, \dots, \sigma_k)$ do descritor NMBE que minimize a função custo da Equação 3.1. Então, utilizando-se as escalas encontradas realiza-se, com o descritor multiescala, a extração de características de toda a base de folhas e, em seguida, a avaliação de desempenho do mesmo.

3.2 Bases de formas binárias

Foram utilizadas nos experimentos três bases públicas de imagens de formas binárias sintéticas e uma base de folhas, cujas características estão apresentadas na Tabela 3. Embora o formato das imagens em cada base seja diferente, estas foram todas convertidas, para fins de padronização, a um padrão comum, no caso o formato PNG.

A Figura 11 apresenta todas as 99 formas da base Kimia-99, enquanto as Figuras 12 e 13 ilustram algumas das formas da base Kimia-216 e MPEG7-CE, respectivamente.

Exemplares extraídos da base pública de imagens de folhas intitulada Flavia (WU *et al.*, 2007) são exibidos na Figura 14. Esta base, composta de 1907 imagens de folhas de 32 espécies distintas, tem sido extensivamente utilizada em trabalhos recentes para testes de métodos de reconhecimento automático de espécies de plantas (WANG *et al.*, 2015; HALL *et al.*, 2015; QUADRI; SIRSHAR, 2015; CHAKI *et al.*, 2015). A base Flavia é bastante desafiadora pois, além de ser desbalanceada, ou seja, conter classes com quantidades distintas de folhas, ela ainda apresenta espécies com grande variabilidade intra classe e pequena variabilidade entre classes.

Tabela 3 – Características das bases de imagens utilizadas nos experimentos.

Base de imagens	Formato	Total de formas	Total de classes
Kimia-99	PNG	99	9
Kimia-216	PGM	216	12
MPEG7-CE	GIF	1400	70
Flavia leaves	JPG	1907	32

Figura 11 – Formas da base de imagens Kimia-99.



Figura 12 – Amostras de formas da base de imagens Kimia-216.

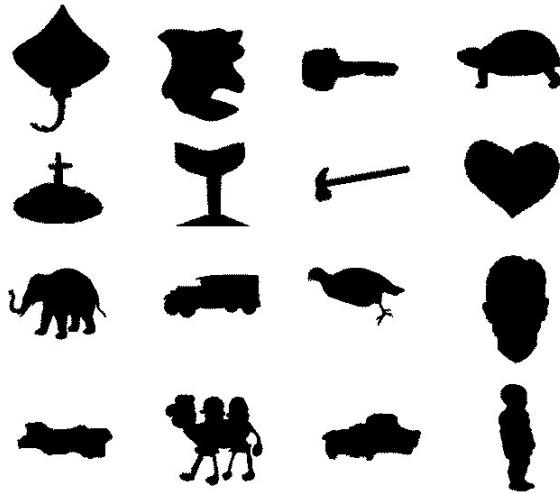


Figura 13 – Amostras de formas da base de imagens MPEG-7.

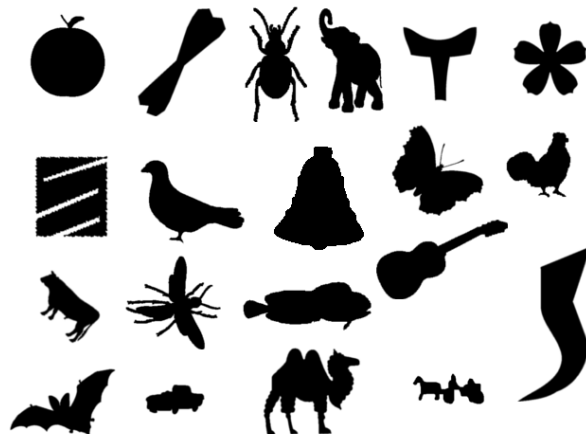


Figura 14 – Trinta e duas amostras extraídas da base Flavia, que contém 1907 imagens de folhas de 32 espécies diferentes. Cada amostra corresponde a uma espécie.



3.3 Avaliação do descritor

Nesta tese, o desempenho do descritor customizado é avaliado qualitativa e quantitativa-mente. Na avaliação qualitativa, dois algoritmos de visualização de dados são utilizados: o mapa auto-organizável de Kohonen (KOHONEN *et al.*, 2001) e o escalonamento multidimensional (COX; COX, 2000). Na avaliação quantitativa, são analisadas as métricas de avaliação Precisão e Revocação (PAULA JR. *et al.*, 2012) obtidas em experimentos de classificação supervisionada, a medida *Bulls-eye* (ALAJLAN, 2011; LING; JACOBS, 2007), classicamente utilizada em experimentos de recuperação de formas (CBIR) e a medida de avaliação de agrupamentos *silhouette* (ROUSSEEUW, 1987).

Para fins de comparação, as avaliações de desempenho qualitativa e quantitativa foram realizadas com as versões dos descritores otimizados pelo método proposto e não otimizados, seja por escolha aleatória dos parâmetros ou por meio de ajuste apresentado em (CESAR JR.; COSTA, 1997).

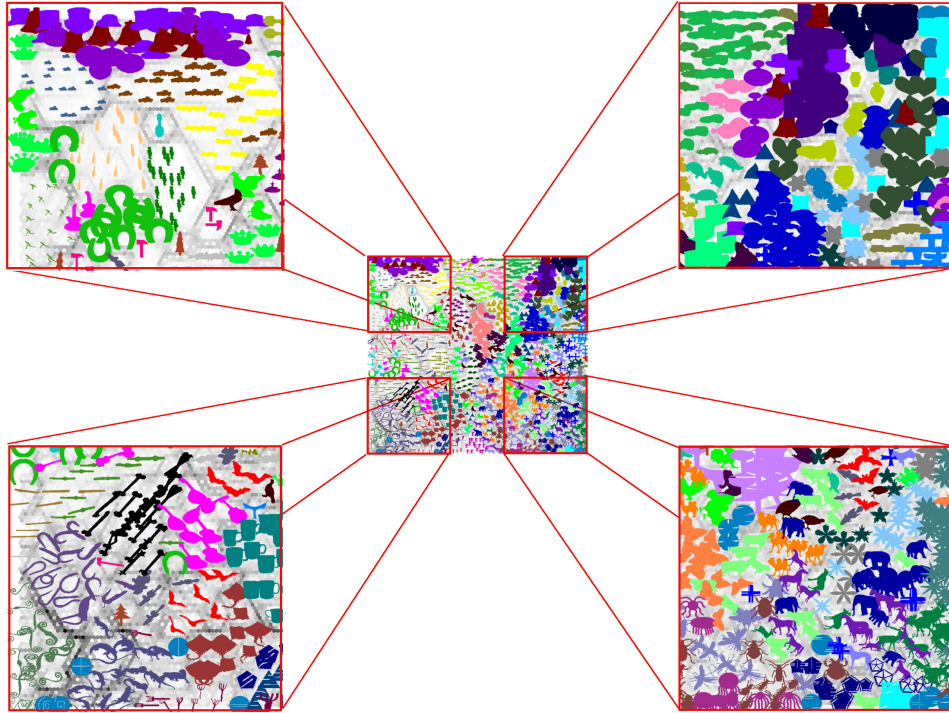
3.3.1 Visualização dos dados

Algoritmos de visualização de dados produzem projeções bidimensionais das descrições das formas da base de folhas, provendo uma representação gráfica que possibilita a análise da qualidade dos agrupamentos. Assim, consegue-se inferir o quão eficaz o descritor é em organizar espacialmente as formas.

A Figura 15 mostra a matriz-U (ULTSCH; SIEMON, 1990) para formas da base MPEG-7 CE-Shape-1 data set (LATECKI *et al.*, 2000). Essa matriz mostra, preservando a topologia do mapa auto-organizável de Kohonen (KOHONEN *et al.*, 2001), a relação de distância entre as estruturas mapeadas. A imagem central é o resultado do processamento de

todas as imagens da base descrita pela NMBE otimizada. As quatro imagens dos cantos são detalhes da imagem central.

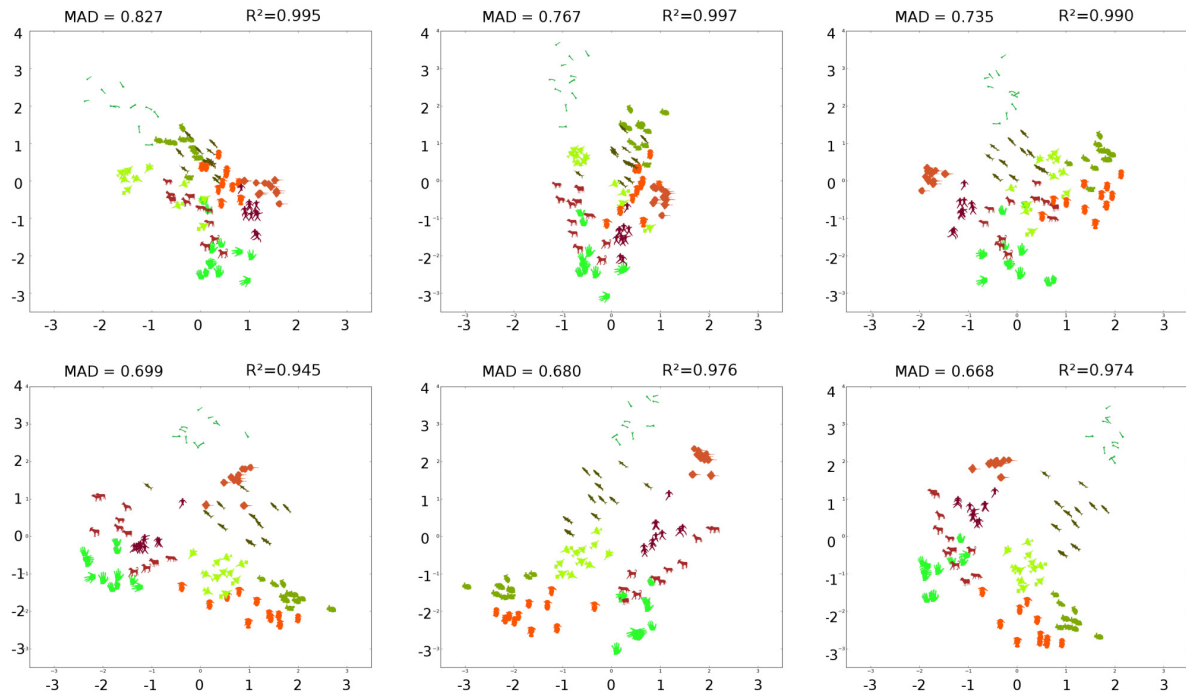
Figura 15 – Matriz-U para as formas da base MPEG7 CE-Shape-1.



Essa ferramenta de visualização identifica quão bem a NMBE descreve detalhes sutis das formas de um mesmo agrupamento e de diferentes agrupamentos. Nas imagens de detalhes se observa que as fronteiras entre os agrupamentos (cor cinza) estão evidentes e separam agrupamentos bem definidos. Nesta tese assumimos que agrupamentos bem definidos são aqueles com menor distância dentro das classes e maior distância entre classes.

As projeções MDS da Figura 16 ilustram como os agrupamentos evoluem à medida que o algoritmo DE busca os parâmetros otimizados do descritor (NMBE). As amostras de formas exibidas pertencem à base Kimia-99 (SEBASTIAN *et al.*, 2004), a qual contém 99 imagens. Neste trabalho, aplicamos técnicas de aprendizagem (*manifold learning*) para produzir as projeções MDS do descritor NMBE otimizado utilizando o algoritmo DE. A Figura 16 demonstra que à medida que os valores de MAD decrescem, as distâncias entre as classes aumentam e conseqüentemente os agrupamentos das classes se tornam mais evidentes. Nesta figura, quando o método de otimização converge, ou seja MAD alcança o menor valor, os únicos agrupamentos que não estão bem separados são aqueles relacionados às formas de animais quadrúpedes e de mãos.

Figura 16 – Projeções do escalonamento multidimensional das formas da base Kimia-99 Sebastian *et al.* (2004). As imagens mostram como os agrupamentos evoluem ao longo do processo de otimização (DE), assim como os valores de MAD e R^2 .



3.3.2 Recuperação de formas

Visando avaliar o desempenho dos descritores de forma, foram realizados experimentos de recuperação de imagens pelo conteúdo em duas bases de imagens de formas binárias (Kimia-99 e MPEG-7 CE-Shape-1) e uma base de folhas (Flavia).

A Figura 17 ilustra a metodologia dos experimentos de recuperação de formas. Primeiramente realiza-se a extração de características das formas da base de imagens com o método de descrição sob avaliação. Desse processo, resulta uma base de dados com os vetores de características associados às formas, que serão utilizados no experimento.

O mesmo processo de extração de características é aplicado à imagem da forma de consulta. Essa última corresponde ao padrão de entrada ao qual as demais formas da base serão comparadas, através de uma medida de similaridade, a fim de se estabelecer o grau de correspondência existente entre as mesmas. Desta forma, tem-se como resultado a lista das imagens recuperadas na ordem decrescente da similaridade que apresentam em relação à forma de consulta. Todo esse processo é realizado tomando-se cada forma da base de imagens como forma de consulta para realização da recuperação das demais.

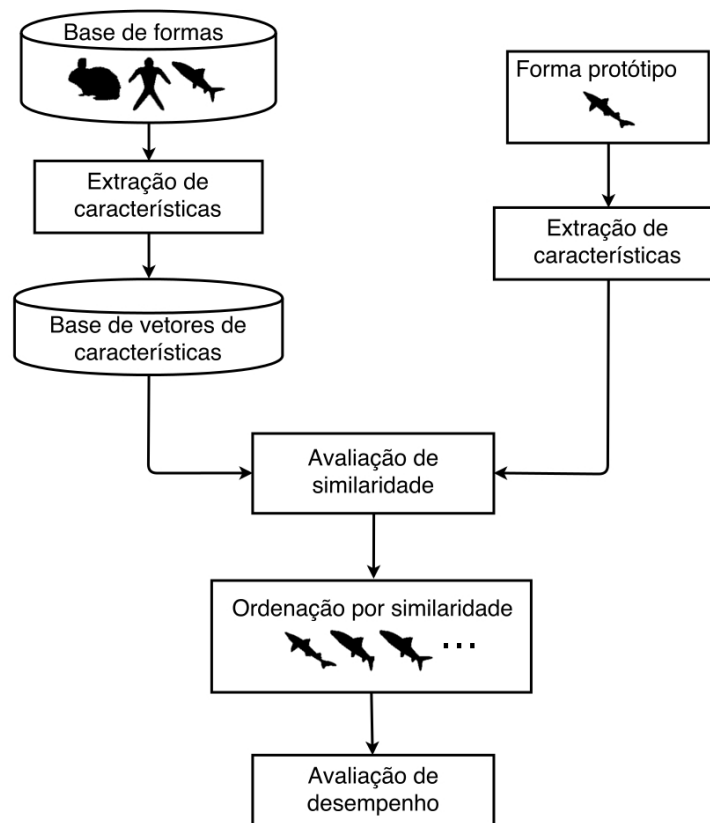
Como medida de dissimilaridade nos experimentos CBIR com descritores multi-

escala, empregamos a distância euclidiana entre os vetores de características das formas. A distância euclidiana entre os vetores a e b de mesma dimensão n é definida como:

$$d_{euclidiana} = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}, \quad (3.3)$$

em a_i e b_i correspondem às i -ésimas componentes dos vetores a e b , respectivamente.

Figura 17 – Metodologia empregada para os experimentos de recuperação de formas pelo conteúdo.



Na avaliação do desempenho destes experimentos, empregamos duas medidas que são o número total de acertos por posição recuperada e a medida *Bulls-eye* (LING; JACOBS, 2007; LATECKI *et al.*, 2000). A primeira medida consiste no número total de ocorrências de formas da mesma classe que a forma de consulta em cada posição recuperada. Em diversos trabalhos de recuperação de formas pelo conteúdo, o número total de acertos por posição recuperada é calculado para a base Kimia-99 (BERNIER; LANDRY, 2003). Para esta base, que contém 99 formas igualmente distribuídas em 9 classes, ou seja é uma base balanceada, são realizadas 99 recuperações das 11 formas mais similares à imagem de consulta. Como resultado espera-se obter um total de 99 formas recuperadas corretamente para cada posição recuperada.

A medida *Bulls-eye* (LATECKI *et al.*, 2000) originalmente foi definida para avaliar o desempenho de descritores de forma em experimentos CBIR com a base de imagens MPEG-7, podendo ser adaptada para outras bases. Esta medida também é utilizada na literatura para a comparação de diferentes métodos de recuperação de formas. Seu cálculo para a base MPEG-7 CE-Shape-1 é obtido da seguinte maneira: tomando-se cada forma dessa base de imagens como elemento de consulta, contabiliza-se o número de recuperações pertencentes à mesma classe da forma de consulta dentre as 40 primeiras posições recuperadas. Como resultado calcula-se a percentagem da quantidade máxima de recuperações corretas possíveis de se alcançar, sendo esta última quantidade $28000 = 1400 \text{ formas} \times 20 \text{ recuperações corretas por forma}$.

Em caso de bases de imagens balanceadas, ou seja, com a mesma quantidade de objetos em todas suas classes, pode-se definir N_B como o número de imagens da base e N_C como o número de elementos de uma determinada classe. Sendo assim, ao realizar o experimento CBIR, recupera-se as $2N_C$ imagens mais semelhantes e contabiliza-se quantas foram recuperadas da mesma classe da imagem usada na consulta, sendo esse processo executado para todas N_B imagens da base. O número de imagens recuperadas da mesma classe das imagens testadas em todas iterações é denotado por N_A , ou seja, a quantidade de acertos. Assim a medida *Bulls-eye* (B_e) é definida por:

$$B_e = \frac{N_A}{N_B N_C}. \quad (3.4)$$

O teste *Bulls-eye* apresenta escores ou valores no intervalo $[0, 1]$. Valores próximos a 0 indicam que os descritores utilizados no experimento *CBIR* não desempenharam bem na caracterização dos elementos da base em estudo, e valores perto de 1 indicam bom desempenho.

3.3.3 Classificação supervisionada

O procedimento para avaliar o desempenho do descritor multiescala em classificação supervisionada está ilustrado na Figura 18. Para fins de clareza, este procedimento foi dividido em três etapas: pré-processamento, extração de características e a classificação propriamente dita. Detalhamos cada dessas etapas, bem como as sub-etapas a elas relacionadas.

Pré-processamento

A etapa de pré-processamento é aplicada às imagens das folhas da base Flavia, uma vez que as imagens das bases MPEG7-CE e Kimia-99 já encontram-se binarizadas. Essa etapa

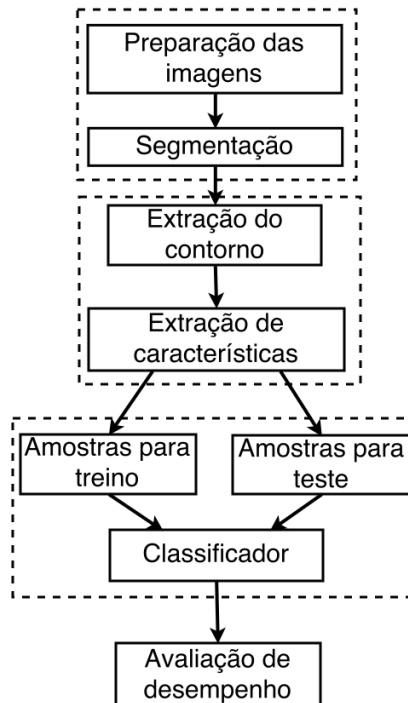
está dividida em duas sub-etapas: preparação das imagens e segmentação.

A preparação das imagens elimina os ruídos inerentes ao processo de aquisição das imagens das folhas. Essa sub-etapa é um tratamento preliminar, que é realizado para garantir um bom resultado de segmentação. Esse tratamento preliminar começa com a extração de um dos três canais de cor RGB da imagem, sendo o canal G o que foi extraído para segmentação. A razão dessa escolha é que este canal incorpora boa parte da informação contida nas imagens de interesse, que correspondem apenas a imagens de folhas verdes.

O próximo passo consiste na aplicação de um filtro bilateral (GONZALEZ; WOODS, 2006) às imagens para homogeneizar as intensidades dos pixels. Além de favorecer a segmentação, este tipo de filtro tem a propriedade de preservar as bordas das formas, que é onde se encontra a informação de interesse para a extração dos atributos.

Finalizando o pré-processamento, a segmentação por limiar (GONZALEZ; WOODS, 2006) é aplicada às imagens tratadas, estando assim as imagens preparadas para a próxima etapa, que é a extração de características.

Figura 18 – Metodologia de classificação para avaliação de desempenho do descritor otimizado pelo método proposto exibido na Figura 10.



Extração de características

Antes da etapa de extração de características, obtém-se as coordenadas paramétricas do contorno das formas segmentadas através do seguidor de bordas proposto por Suzuki e Be (1985). Os descritores são então obtidos, a partir dos contornos das formas, pelos métodos de extração de atributos descritos na Seção 2.2.

Classificação

Para a etapa de classificação, os descritores obtidos na etapa anterior são divididos em grupos de amostras de teste/treino, sendo o método utilizado para geração desses grupos o k-fold ($k = 10$) (WEBB, 2002).

Para cada grupo de testes e treino, os descritores são transformados pelo discriminante linear de Fisher (WEBB, 2002), para melhorar a separação inter classes e a coesão intra classe, e pela análise das componentes principais (PCA), para descorrelacioná-los. Foram então escolhidas as componentes principais que maximizam o desempenho da classificação. Cabe também salientar que as matrizes utilizadas nessas transformações são obtidas apenas considerando as amostras selecionadas para treino dos classificadores.

A última sub-etapa da classificação envolve a realização da validação cruzada, com os 10 grupos de teste e treino gerados (WEBB, 2002), para quatro classificadores estatísticos: *naive Bayes* (NB) (FUKUNAGA, 1990), k-vizinhos próximos (kNN), (FUKUNAGA, 1990; WEBB, 2002), análise por discriminante linear (LDA) (WEBB, 2002) e análise por discriminante quadrático (QDA) (FUKUNAGA, 1990).

O LDA e o QDA são classificadores que assumem modelo Gaussiano multivariado dos dados. Como os nomes sugerem, estes classificadores estabelecem superfícies de decisão lineares e quadráticas, respectivamente. Por serem paramétricos apresentam soluções fechadas, o que os torna simples de serem implementados. Ademais, o LDA e o QDA são inerentemente adequados para resolução de problemas que envolvam a classificação de múltiplas classes.

Quando se assume, no modelo do classificador QDA, que as matrizes de covariância das classes são diagonal e idênticas, que significa assumir classes condicionalmente independentes, o QDA se torna o NB.

Apesar da aparente simplicidade, o classificador NB tem demonstrado desempenho satisfatório em diversas aplicações de reconhecimento de padrões, como a classificação de

documentos e a filtragem de spams. Além de sua eficiência computacional, o NB requer pequenas quantidades de amostras de treino para estimação de seus parâmetros.

































O kNN é um classificador não paramétrico, comumente utilizado em situações onde a superfície de decisão entre classes é irregular. Por não ser paramétrico, este classificador não constrói um modelo de distribuição das classes a partir de amostras de treino, mas infere a classe das amostras de teste diretamente a partir das de treino. Desta forma, a classificação se dá por votação entre os k vizinhos mais próximos à amostra de teste da qual se deseja inferir a classe, sendo atribuído o rótulo da classe que é a mais representativa entre os k pontos vizinhos consultados. Em geral, a métrica de distância utilizada é a euclidiana, porém outras métricas podem ser empregadas. A desvantagem do kNN é que a escolha do parâmetro k é extremamente dependente dos dados. Em geral, valores grandes de k aumentam a robustez do classificador a ruídos, mas torna as fronteiras de decisão entre as classes menos evidente.

A partir dos resultados de 100 execuções da validação cruzada, foram calculados os valores médios e os desvios padrão das medidas clássicas de desempenho Precisão e Revocação.

4 RESULTADOS E DISCUSSÕES

Com o método de otimização proposto, realizamos experimentos de classificação e recuperação de formas pelo conteúdo, para dois descritores, com a base de folhas de plantas Flavia (WU *et al.*, 2007), que possui 1907 imagens de folhas de 32 espécies. A Figura 19 ilustra imagens das folhas desta base. Cada exemplar ilustrado foi segmentado e colorido de acordo com a espécie ao qual pertence. Essa base de folhas é amplamente utilizada na validação de trabalhos de reconhecimento automático de espécies de plantas.

Figura 19 – Amostras de formas de folhas da base de imagens Flavia.

1		9		17		25	
2		10		18		26	
3		11		19		27	
4		12		20		28	
5		13		21		29	
6		14		22		30	
7		15		23		31	
8		16		24		32	

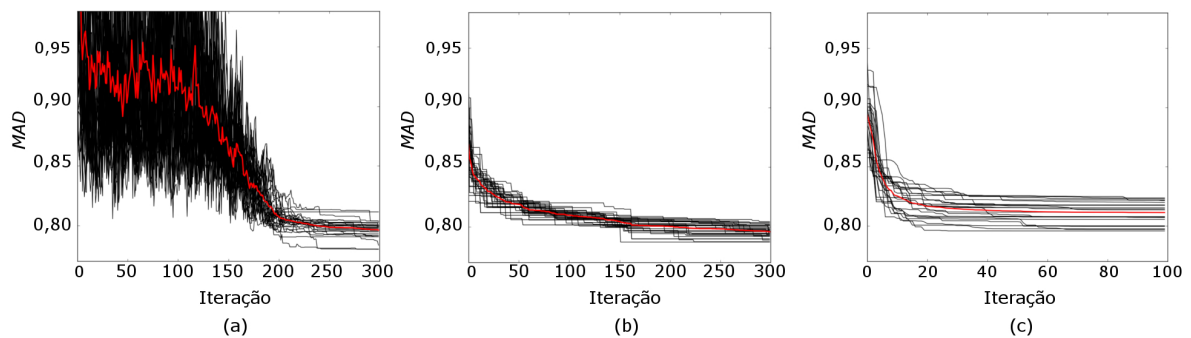
4.1 Classificação de formas e sua relação com a função objetivo

Nos experimentos realizados para comparação dos algoritmos de otimização, avaliamos inicialmente aspectos relacionados à convergência dos algoritmos de otimização e a relação entre os resultados de classificação e as funções MADs obtidas por diferentes estratégias. Além disso investigamos o custo computacional dos mesmos. A Figura 20 ilustra a convergência de cada um dos três algoritmos de otimização para 30 repetições realizadas em um subconjunto da base de formas de folhas de plantas. Os resultados obtidos indicam que a otimização realizada pelo PSO requereu um pequeno número de iterações para convergir a uma solução ótima, quando comparado aos algoritmos SA e DE. Uma vez que a convergência está relacionada com o número de iterações necessárias para se atingir uma solução ótima, um menor número de iterações implica em exploração deficiente do espaço de busca, ou seja, em convergência prematura (ENGELBRECHT, 2007). Nestes casos, o método de otimização apresenta tendência de ficar preso a mínimos locais, encontrando assim soluções sub-ótimas para o problema. Por outro lado, a convergência mais lenta permite ao algoritmo explorar melhor o espaço de busca, aumentando portanto a chance deste convergir para um mínimo global, ou seja, uma solução

ótima (ENGELBRECHT, 2007).

Neste contexto, observa-se na Figura 20c que o PSO encontrou soluções sub-ótimas com maior frequência, alcançando valores médios da função MAD igual a $0,805 \pm 0,006$. Com relação aos algoritmos SA e DE, estes alcançaram os menores valores médios de MAD : $0,795 \pm 0,006$ e $0,798 \pm 0,004$, respectivamente. Esses resultados demonstram que os algoritmos SA e DE foram mais eficazes que o PSO em encontrar soluções ótimas. No entanto, o custo para alcançar tais soluções torna-se maior. A Seção 4.4 aborda com mais detalhes as questões relativas ao custo computacional dos algoritmos de otimização utilizados.

Figura 20 – Convergência dos métodos de otimização para o problema de descrição de folhas de plantas: (a) SA, (b) DE, (c) PSO. As curvas em vermelho destacam o valor médio da função MAD para 30 realizações de cada método (curvas em preto).



A análise corrente tem como hipótese o fato de que as escalas otimizadas e o valor ótimo de MAD, que estão inter-relacionadas, implicam em melhorias na taxa de acerto da classificação das espécies vegetais. A Tabela 4 exibe os diversos valores de MAD e os correspondentes métodos de ajustes de parâmetros utilizados. Observa-se que os parâmetros do descritor ajustados pelo método proposto, para diferentes otimizadores, são os que alcançaram os menores valores de MAD, sendo o melhor resultado aquele obtido pelo método de otimização SA. Já os parâmetros ajustados tanto empiricamente como pelo método proposto em (CESAR JR.; COSTA, 1997), foram os que apresentaram os maiores valores de MAD.

Já a Figura 21 ilustra os resultados de classificação para os diferentes métodos de ajuste de parâmetros investigados. Esses resultados foram obtidos para os parâmetros ajustados que alcançaram os valores de MAD reportados na Tabela 4. Os melhores resultados de classificação correspondem aos menores valores de MAD obtidos através da metodologia de otimização de parâmetros. Os resultados de classificação para o descritor NMBE, com as escalas ajustadas conforme sugerido por Cesar Jr. e Costa (1997) ($NMBE_{orig}$), resultou em um desempenho intermediário quando comparado aos resultados para escalas otimizadas e escolhidas arbitrariamente.

Tabela 4 – Valores de MAD alcançados, pelos diferentes métodos de ajuste das escalas do descritor NMBE, para representação das folhas da base Flavia.

Método de ajuste	MAD
SA	0,762
DE	0,783
PSO	0,829
Cesar Jr. e Costa (1997)	0,867
Empírico	0,969

Os resultados mostram que o descritor NMBE otimizado ($NMBE_{opt}$) melhorou o desempenho de todos os classificadores avaliados, uma vez que este alcançou as maiores taxas de Precisão e Revocação para os menores valores de MAD, o que confirma a hipótese inicial.

Apesar do descritor NMBE ter sido originalmente projetado para descrição de formas de neurônios, nesta tese o aplicamos com sucesso em caracterização de folhas de plantas, uma vez que a metodologia de otimização o ajustou ao problema em questão. É importante ressaltar que os diferentes algoritmos de otimização, atingiram valores de MAD distintos, e portanto conjuntos de escalas otimizadas distintos, os resultados de classificação corroboraram a hipótese e a conclusão anterior, pois tais conjuntos de escalas permitiram que os descritores fossem capazes de capturar detalhes dos contornos das formas e as diferenciasse apropriadamente. Com isso, observa-se que os algoritmos de otimização trazem vantagens à representação multiescala com impacto significativo na caracterização das folhas.

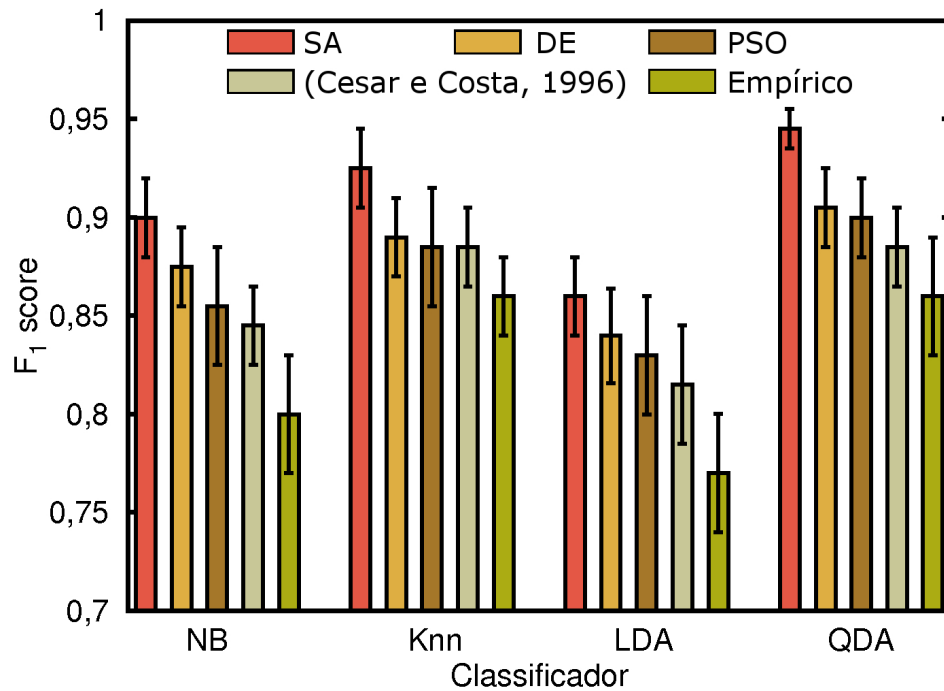
Os resultados da Tabela 4 confirmam que os classificadores que utilizaram descritores gerados a partir de escalas selecionadas arbitrariamente não tiveram desempenho satisfatório, pois alcançaram os menores valores de Precisão e Revocação e os maiores valores da função MAD. Portanto, escalas aleatórias levam a uma representação multiescala menos sensível a variações de características das folhas, conseqüentemente, mais erros de classificação.

4.2 Análise exploratória visual de agrupamentos

Nesta seção utilizamos técnicas de visualização de dados multidimensionais, como MDS e matriz-U, para observar os resultados dos descritores multiescala NMBE e DFM. O uso dessas técnicas suportam a análise do comportamento dos descritores a partir de suas relações com os agrupamentos observados em um espaço bidimensional.

Para ilustrar o poder discriminativo dos descritores de forma multiescala NMBE e DFM, bem como mostrar a relação existente entre a medida de avaliação quantitativa da

Figura 21 – Resultados de classificação das espécies vegetais da base de imagens Flavia para diferentes estratégias de escolha das escalas do descritor NMBE.



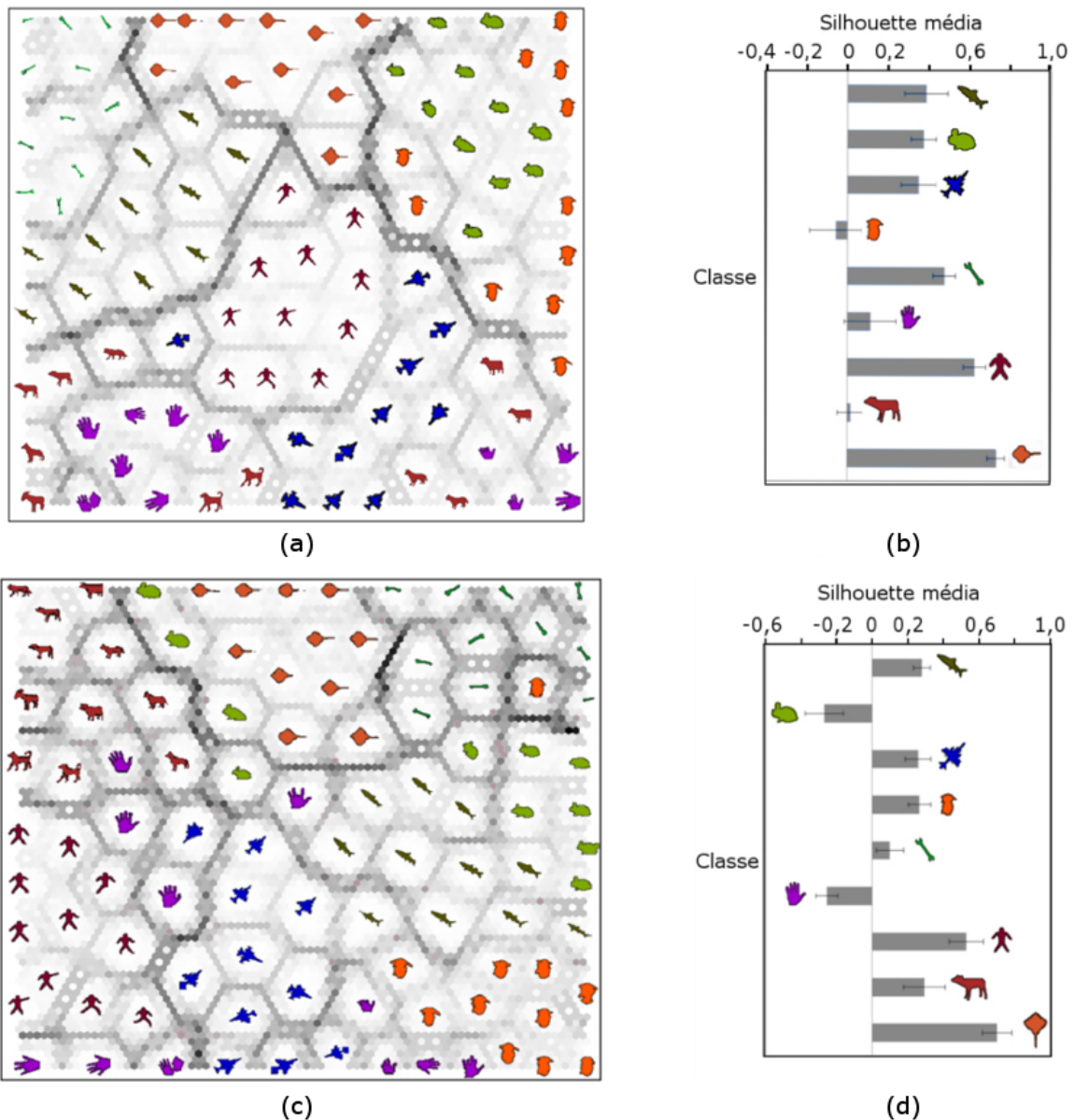
qualidade dos agrupamentos, *silhouette*, e a representação qualitativa obtida através da matriz U, foram realizadas as extrações de características das formas das bases Kimia-99 e Kimia-216 com os referidos descritores. Os resultados estão expostos nas Figuras 22 e 23.

Em ambas as figuras, observa-se que as formas correspondentes às classes com *silhouette* média positiva são visualizadas na matriz-U mantendo forte relação de vizinhança entre si, enquanto que formas correspondentes às classes cuja *silhouette* média é negativa encontram-se dispersas nessas matrizes. Isso denota a inabilidade da descrição em representá-las adequadamente.

Assim, os descritores conseguem capturar características tanto globais como locais, uma vez que classes de formas com características globais semelhantes, como por exemplo as formas alongadas do canto superior esquerdo da Figura 23a, estabelecem uma relação de vizinhança entre si. Apesar disso, formas alongadas de classes distintas não se misturam, o que é um indicativo de que o descritor foi capaz também de representar características locais discriminativas que permitem a separação entre as classes.

Existem no entanto casos, em que ambos os descritores não foram capazes de discriminar formas com características globais semelhantes, como é o caso das classes de formas de elefantes e camelos na Figura 23. Na matriz-U, estas formas aparecem agrupadas como se pertencessem a uma mesma classe e consequentemente, com valores de *silhouette* média

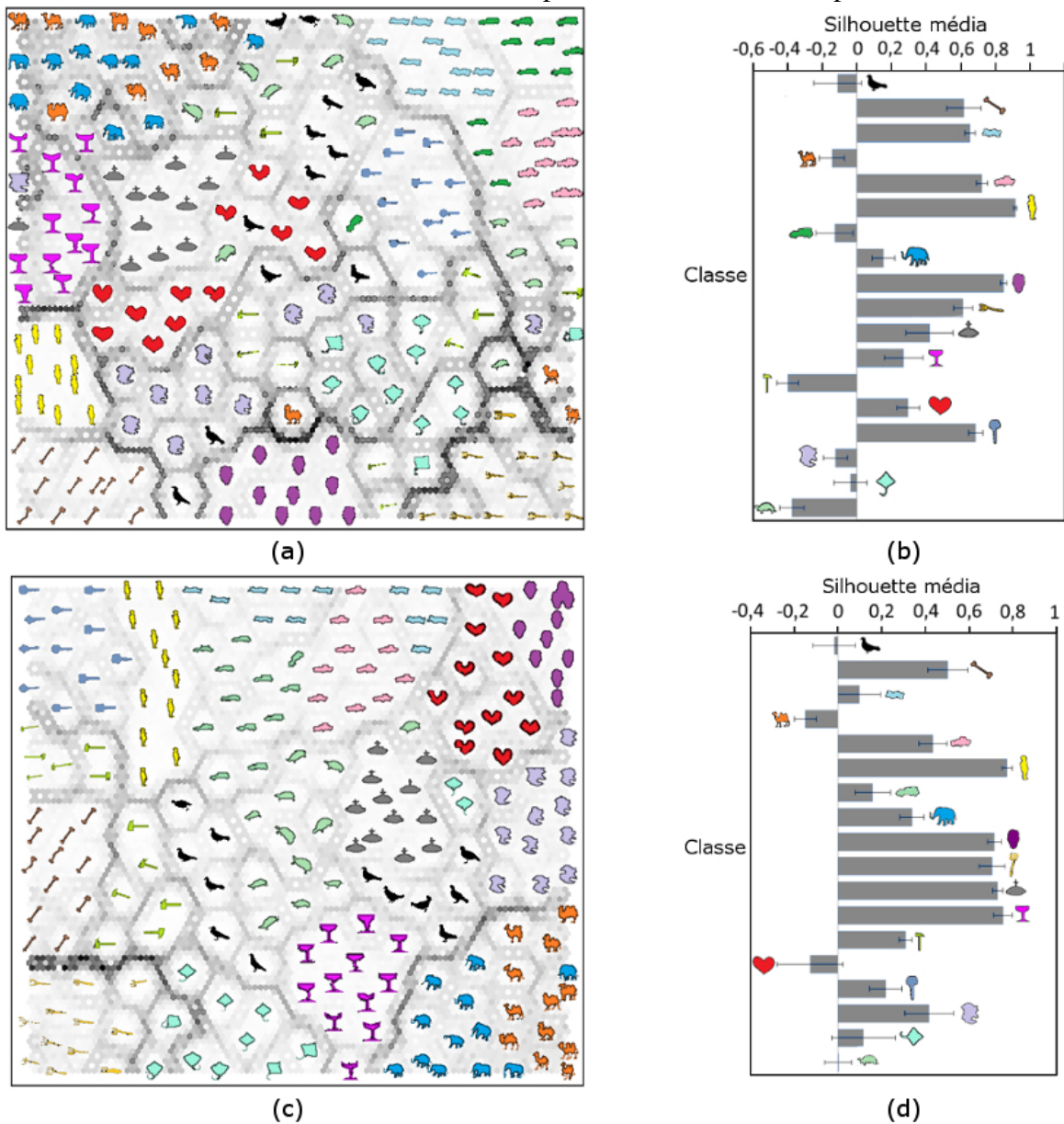
Figura 22 – Resultados dos descritores NMBE e DFM para a base Kimia-99. (a) Matriz-U obtida com o descritor NMBE e respectiva *silhouette* média por classe. (b) Matriz-U obtida com o descritor DFM e respectiva *silhouette* média por classe.



significativamente menores em relação às formas que foram discriminadas corretamente pelos descritores.

A Figura 24a mostra a matriz-U obtida para as formas de folhas de plantas da base Flavia representadas pelo descritor NMBE otimizado e a Figura 24b, pelo descritor NMBE não otimizado. A análise exploratória visual dos agrupamentos produzidos pela descrição NMBE indica que a descrição otimizada melhora a organização dos agrupamentos. Essa evidência então confirma que a função objetivo MAD é adequada para guiar o processo de otimização desse descritor. A visualização das formas nas matrizes-U da Figura 24 aponta diferenças entre os

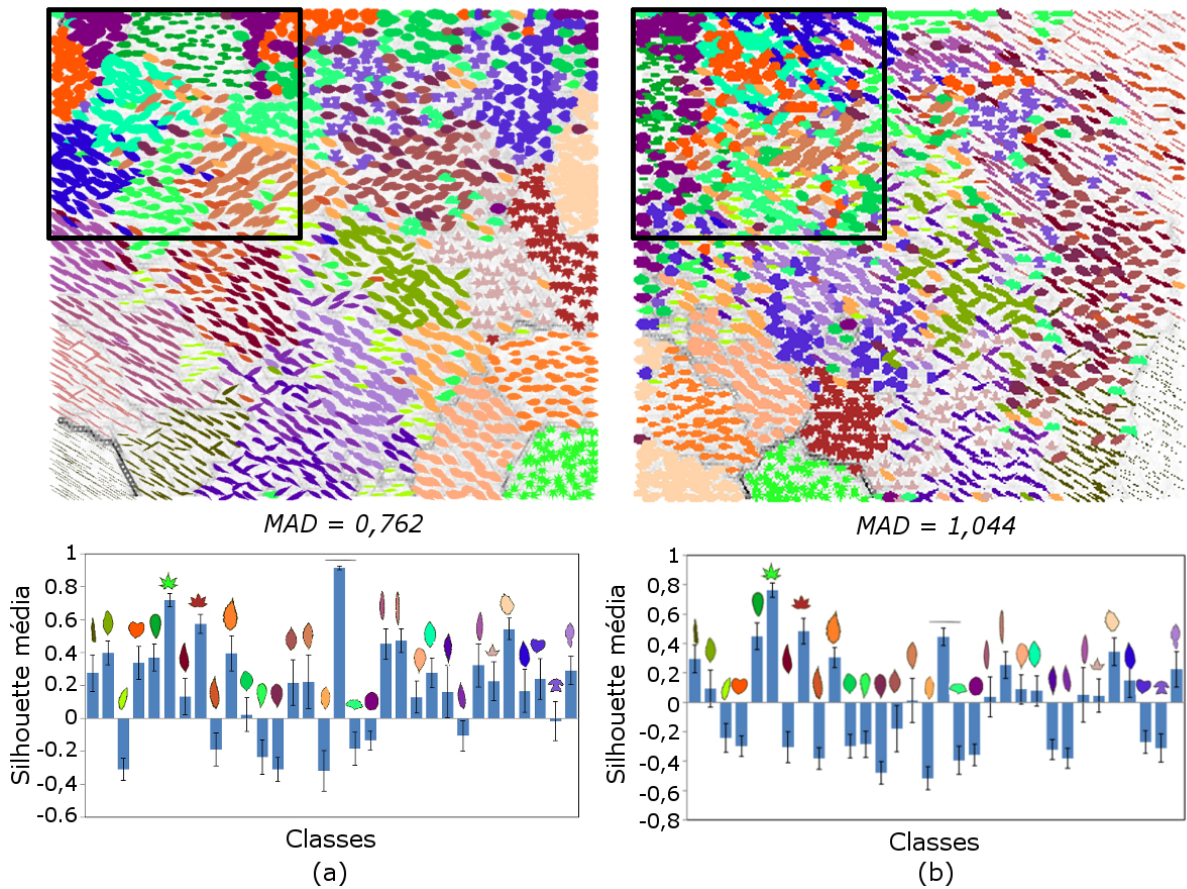
Figura 23 – Resultados dos descritores NMBE e DFM para a base Kimia-216. (a) Matriz-U obtida com o descritor NMBE e respectiva *silhouette* média por classe. (b) Matriz-U obtida com o descritor DFM e respectiva *silhouette* média por classe.



dois arranjos promovidos pelos referidos descritores. A cada posição ou elemento da matriz, os valores numéricos foram substituídos pela imagem das folhas correspondentes, estando as cores das folhas em correspondência com o rótulo das classes exibidas na Figura 19. Nessas matrizes-U, estão representadas as 1907 formas de folhas da base Flavia, sendo cada folha representada por sua descrição multiescala.

Os gráficos representativos da *silhouette* média nas Figuras 24a e 24b demonstram que as classes das folhas bem caracterizadas e, portanto, agrupadas adequadamente são as que apresentam valores de *silhouette* média positivos. Por outro lado, as classes de folhas cujo

Figura 24 – Matrizes-U e *silhouettes* média para as descrições das folhas da base Flavia com o descritor NMBE: (a) otimizado e (b) não otimizado.



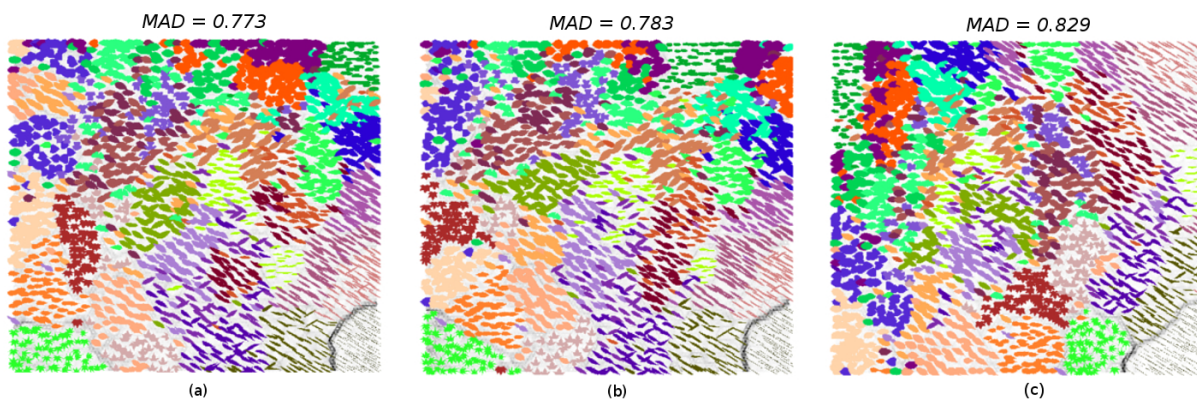
descritor não foi capaz de caracterizar e agrupar adequadamente são as que apresentaram valores negativos de *silhouette* média. Comparando estes resultados, observa-se que na Figura 24a há expressiva redução dos valores negativos de *silhouette* média quando se utiliza o descritor $NMBE_{opt}$.

O quadrado preto inserido na Figura 24 destaca o desempenho do descritor otimizado em termos de melhoria no arranjo das formas na matriz-U. É possível ainda observar quão bem o descritor otimizado mapeou as folhas em grupos de acordo com as respectivas classes rotuladas das mesmas. Além do mais, o resultado com as escalas otimizadas melhorou significativamente a *silhouette* média por classe para quase todas as classes de folhas da base. Por outro lado, a região interna ao quadrado preto na Figura 24b destaca que o descritor não otimizado não foi capaz de mapear satisfatoriamente as formas de folhas em suas classes correspondentes. Deduzimos que essa melhoria se deve ao fato de que as escalas otimizadas dos descritores são mais sensíveis a variações intrínsecas das formas, ou seja, incorporam nuances presentes nos seus contornos, os quais os métodos tradicionais não fazem adequadamente. Ademais, o ajuste automático das

escalas é de fato uma tarefa alternativa para o ajuste arbitrário ou empírico, sendo que esta última não é uma tarefa simples e está sujeita à fadiga e subjetividade do realizador da mesma.

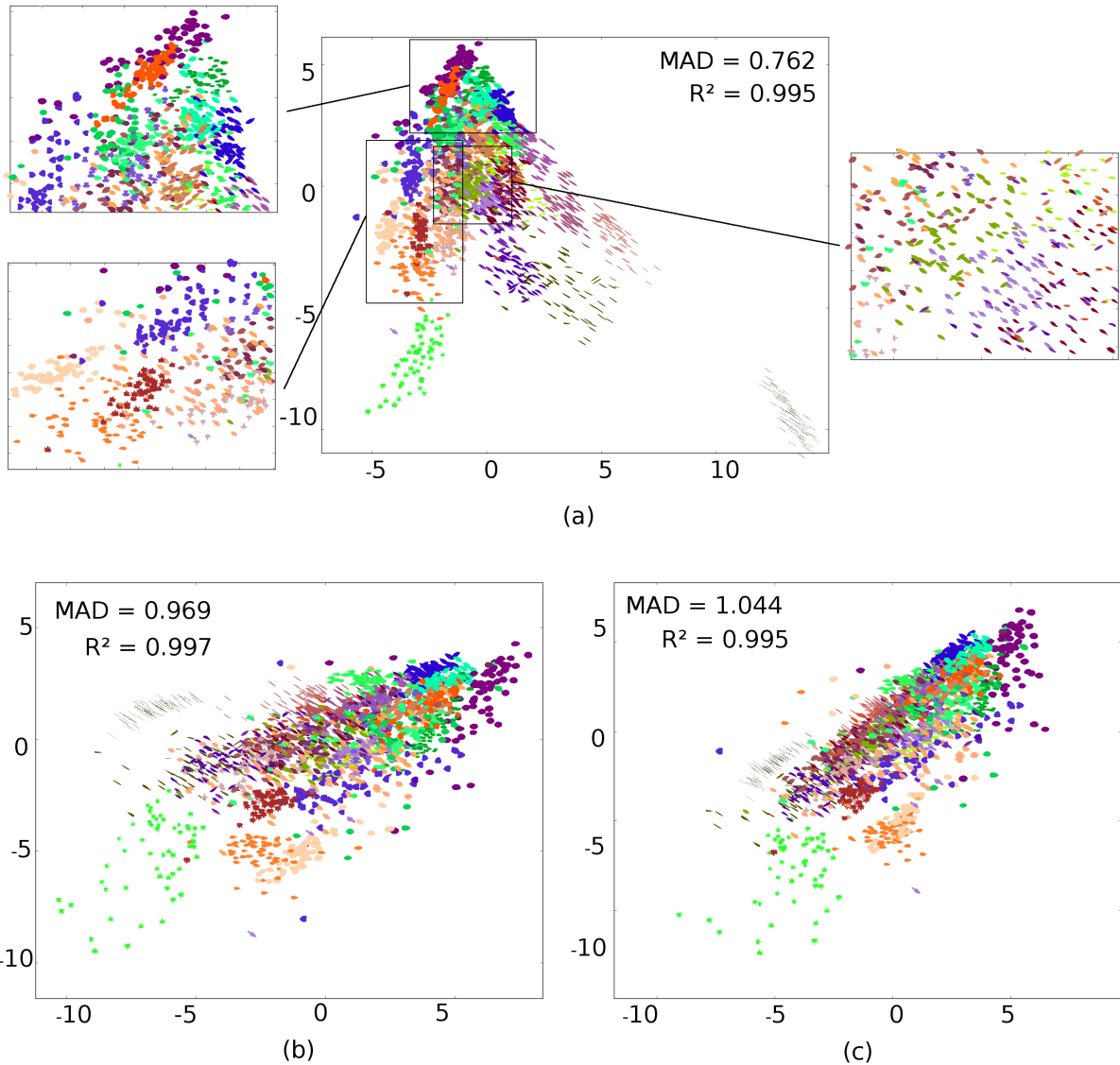
As Figuras 25a, 25b and 25c exibem as matrizes-U correspondentes aos valores MAD dos experimentos na base Flavia utilizando os algoritmos de otimização SA, DE e PSO, respectivamente. Um aspecto interessante a examinar é o fato dos três algoritmos apresentarem diferentes soluções, ou valores de MAD. Estas soluções resultaram em diferentes arranjos de folhas e conseqüentemente diferentes relações entre as estruturas de vizinhança dos grupos ou classes. Por exemplo, a dispersão geral dos elementos nas Figuras 25a e 25b tende a ser muito semelhante uma vez que seus respectivos valores da função objetivo minimizada são bem próximos em magnitude. Ao mesmo tempo, concluímos que o resultado que corresponde ao mais elevado valor de MAD (0,829) exhibe uma dispersão particularmente elevada no centro da matriz-U (Figura 25c). Este é um detalhe que reforça a observação anterior de que o PSO convergiu para um mínimo local. Estes achados apontam para a importante conclusão de que quanto menor o valor de MAD, melhor o arranjo dos grupos ou do agrupamento, e conseqüentemente a qualidade do agrupamento. Por isso, afirmamos que a qualidade dos grupos e o valor de MAD são negativamente correlacionados. Nossa análise ainda considera que devam existir outras soluções no espaço de busca, além da solução mínima, que sejam adequadas ao problema em estudo.

Figura 25 – Matrizes-U e valores MADs obtidos para a base de imagens Flavia por uso da NMBE otimizada pelos algoritmos: (a) SA, (b) DE e (c) PSO.



A Figura 26 ilustra as projeções MDS para três valores distintos de MAD. Na Figura 26a, temos o arranjo dos agrupamentos após a convergência do algoritmo de otimização SA, ou seja, para o descritor cujos parâmetros otimizados correspondem ao mínimo valor da função MAD encontrado nas otimizações. Analisando essas projeções, observa-se que a minimização

Figura 26 – Projeções MDS dos descritores NMBE para formas da base de folhas Flavia. (a) Descritor NMBE otimizado pelo algoritmo SA com valor de função $MAD = 0,762$ (b) Descritor NMBE não otimizado com valor de função $MAD = 0,969$ e (c) Descritor NMBE não otimizado com valor de função $MAD = 1,044$.



da função objetivo resultou em agrupamentos de folhas com maior coesão intra classe e maior separação entre classes. A análise visual das três regiões detalhadas nessa imagem indica mais claramente que o descritor NMBE otimizado melhorou a representação das formas de folhas em comparação aos resultados das Figuras 26b e 26c. Ademais, os valores do coeficiente R^2 próximos de 1,0 indicam que a representação no espaço bi-dimensional preservou a relação de distâncias entre as amostras no espaço de alta dimensão.

4.3 Recuperação de imagens baseada em conteúdo

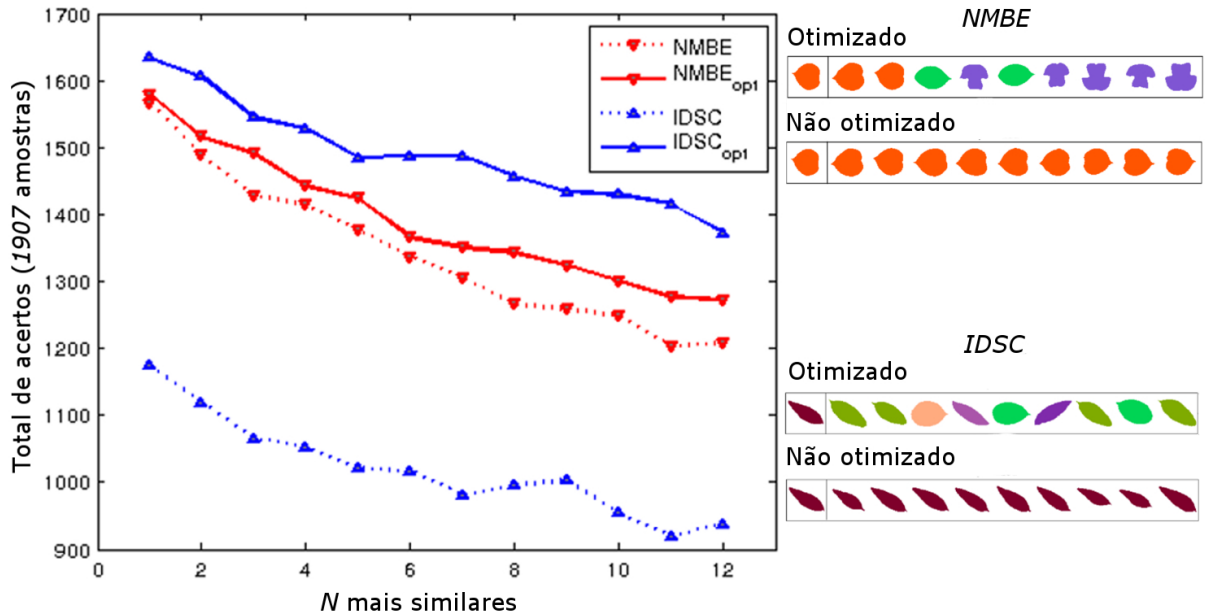
O descritor *inner distance shape context* (IDSC) (LING; JACOBS, 2007) foi otimizado para fins de comparação com o descritor multiescala. Desenvolvido para aplicações de recuperação de formas, o IDSC utiliza algoritmos de programação dinâmica, tal como o *dynamic time warping* (DTW) (PALAZÓN-GONZÁLEZ; MARZAL, 2012), para melhorar a precisão da avaliação de similaridade entre formas. Assim sendo, substituiu-se a norma L_2 pelo algoritmo DTW no cálculo da medida *silhouette*, uma vez que esta última é requerida para a avaliação da função objetivo MAD na Equação 3.1.

A Figura 27a exibe gráficos que permitem a comparação dos resultados dos experimentos realizados para as versões otimizadas ($NMBE_{opt}$, $IDSC_{opt}$) e não otimizadas (NMBE, IDSC) de ambos descritores. Os gráficos mostram que o ajuste dos parâmetros dos descritores pelo método de otimização melhorou o resultado (taxa) de recuperação de folhas consideravelmente. Observa-se ainda que o descritor $IDSC_{opt}$ superou ambos os descritores $NMBE_{opt}$ e NMBE não otimizado com relação à taxa de recuperação. Logo, pode-se afirmar que a otimização de parâmetros melhorou a capacidade de discriminação do $IDSC_{opt}$, incorporando ao descritor um mecanismo de extração de informações relevantes de detalhes intrínsecos das formas das folhas. Por outro lado, observa-se que o desempenho do IDSC não otimizado foi inferior ao do NMBE não otimizado.

As Figuras 27b e 27c ilustram resultados obtidos de recuperação de dois exemplares da base de imagens de folhas para experimentos realizados com os descritores NMBE não otimizado e $NMBE_{opt}$, IDSC não otimizado com parâmetros ajustados conforme recomendado em (WANG *et al.*, 2015) e $IDSC_{opt}$. Observa-se na Figura 27b que o descritor NMBE não otimizado falhou parcialmente na recuperação das formas, enquanto a versão otimizada ($NMBE_{opt}$) com sucesso recuperou todas as formas. O mesmo comentário se aplica ao descritor $IDSC_{opt}$, cujos resultados mostram o sucesso da recuperação. Ainda na Figura 27c, o descritor IDSC não otimizado não foi capaz de recuperar as amostras corretamente.

A Tabela 5 apresenta o resultado da avaliação de desempenho dos descritores $NMBE_{opt}$ e NMBE não otimizado com escalas ajustadas conforme recomendado por Cesar Jr. e Costa (1997), $IDSC_{opt}$ e o IDSC não otimizado. Esses resultados mostram que a taxa de recuperação para ambos os descritores otimizados superou os resultados dos pares correspondentes não otimizados. A melhoria significativa observada nas medidas Bulls-eye reforça a hipótese de que a metodologia de otimização proposta é adequada para ajuste dos descritores em aplicações

Figura 27 – Experimentos realizados com formas da base de imagens de folhas Flavia. (a) taxa de recuperação obtida com os descritores NMBE e IDSC originais e suas versões otimizadas, (b) e (c) recuperação de duas amostras de folhas utilizando os descritores NMBE e IDSC otimizadas e não otimizadas, respectivamente.



de recuperação e análise de formas de folhas.

Tabela 5 – Taxa Bulls-eye para a base de imagens Flavia.

NMBE	NMBE _{opt}	IDSC	IDSC _{opt}
63,86 %	71,16 %	53,38%	77,50%

4.4 Custo computacional da otimização

Avaliou-se a complexidade computacional dos três algoritmos de otimização como exige a Tabela 6. Os algoritmos SA e PSO apresentam complexidades similares e estas dependem do número de iterações para convergir (N_{iter}), do tamanho da população (N_{pop}) e do parâmetro P . Por outro lado, o algoritmo DE demanda maior complexidade, pois depende da dimensão (D) do problema a ser otimizado, do tamanho da população (N_{pop}) e do número de interações (N_{iter}).

Tabela 6 – Complexidade computacional dos métodos de otimização.

Método	Complexidade	Custo
SA	$O(P \cdot N_{iter} \cdot \log N_{iter})$	7.431
DE	$O(N_{pop} \cdot N_{iter} \cdot D)$	19.500
PSO	$O(N_{pop} \cdot N_{iter} \cdot \log N_{iter})$	1.107

Para fins de comparação dos métodos de otimização, sob o aspecto de custo computacional, assumimos como custo dos algoritmos o número de vezes que a função objetivo é demandada ao longo do processo de otimização. Considerando para cada método de otimização o ajuste de parâmetros adotado, bem como a complexidade computacional correspondente, os custos computacionais calculados para os algoritmos SA, DE e PSO resultaram em 7.431, 19.500 e 1.107, respectivamente.

Vale destacar que existe um compromisso entre o custo computacional e a qualidade da solução obtida. Assim sendo a redução do número de amostras de formas, da resolução das imagens e do tamanho da população dos métodos (N_{pop}) ameniza o custo computacional, porém estas reduções degradam a qualidade do resultado da otimização. Ademais, o uso de processamento paralelo pode contribuir para a redução do custo computacional, sem degradar a qualidade da solução ótima, entretanto o paralelismo adiciona complexidade aos algoritmos de otimização.

5 CONCLUSÕES E TRABALHOS FUTUROS

Esta tese abordou a caracterização e análise de formas por meio de descritores multiescala baseados em conceitos de curvatura como a energia de dobramento. Estes descritores foram calculados a partir do contorno de formas binárias de bases de uso geral e uma base pública de folhas. Considerando que estes descritores podem ser utilizados em diferentes aplicações de visão computacional e, portanto, distintas bases de imagens, introduzimos uma metodologia versátil de ajuste automático de parâmetros por otimização evolucionária.

Vale destacar a importância e ao mesmo tempo a dificuldade inerente ao processo de ajuste manual ou empírico de parâmetros multiescala a um determinado problema ou aplicação. Assim sendo, neste trabalho apresentamos uma alternativa automática para o delineamento do descritor e seus parâmetros às particularidades do problema e da base em estudo. Além disso, a alternativa do ajuste automático se destaca por não ser fatigante como o ajuste manual arbitrário ou empírico.

Na metodologia proposta para análise de formas de folhas de plantas, a função objetivo, a ser minimizada no processo de busca dos parâmetros, é fundamentada na medida de qualidade de agrupamento *silhouette*. Como resultado da minimização desta função objetivo, encontramos um conjunto otimizado de parâmetros de escala dos descritores de forma que são utilizados na análise de formas, em particular, de folhas de plantas.

A versatilidade desta metodologia se deve, portanto, ao fato de que a mesma pode ser ajustada à função objetivo ou função custo e moldada assim para as bases de imagens do problema abordado. Ademais, seu desempenho se mostrou satisfatório e promissor, uma vez que os parâmetros otimizados incorporaram características e detalhes sutis das formas, o que foi confirmado pelas técnicas de avaliação qualitativa e quantitativa.

De fato, o conjunto otimizado de parâmetros, moldado pela minimização da função objetivo, embute informações de nuances da forma. A comprovação deste achado se deu pela considerável melhoria alcançada na organização dos agrupamentos, quantificada pela medida da qualidade de agrupamento e pela elevação na taxa de acerto da classificação das formas ao utilizarmos os descritores otimizados em bases com elevada similaridade de formas entre classes.

As versões otimizadas dos descritores estudados discriminaram diferenças de formas dentro de uma mesma classe e entre classes. Observamos ainda que determinadas classes de formas apresentaram-se mais desafiadoras que as demais para a representação das mesmas, e isso foi comprovado pela medida *silhouette* e o arranjo espacial exibido pelas técnicas de visualização

exploratória de agrupamentos.

Experimentos com uma base pública de imagem de folhas indicaram a adequação das metodologias propostas em problemas de taxonomia de folhas de plantas. Vale ressaltar que a base de imagens de folhas de plantas Flavia é bastante desafiadora pois a mesma apresenta uma elevada similaridade entre formas de classes distintas. Isso significa que formas de classes distintas não apresentam significativas variações nos contornos das mesmas. Entretanto, essas particularidades podem ser captadas pelos descritores multiescala, os quais se mostraram bastante efetivos no agrupamento e classificação de formas de folhas. Esta importante característica da base Flavia é portanto um desafio para a metodologia e para os descritores, de modo geral.

Logo, a metodologia proposta constitui uma ferramenta adicional e fonte de informação para taxonomistas discriminarem e agruparem espécies de plantas.

Trabalhos futuros

Deste trabalho se desdobram outras ações futuras relacionadas a seguir:

- melhoria do processo de otimização dos parâmetros, buscando outras alternativas para a função objetivo, de modo que se reduza o custo computacional da metodologia e melhore as taxas de classificação e recuperação;
- análise de desempenho da metodologia proposta em outros problemas de visão computacional como reconhecimento automático de células saudáveis e com carcinoma a partir de descritores multiescala de formas;
- aplicação da metodologia de otimização em classificação taxonômica de outras espécies vegetais e animais;
- investigação e ajuste dos parâmetros dos algoritmos de otimização de modo que resultem em maior estabilidade na convergência, assim como em maior velocidade de processamento;
- seleção e testes de outros algoritmos de otimização evolucionária em análise de formas.

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**, s.l., v. 23, n. 1, p. 7–14, jan. 2011. Elsevier BV. <http://dx.doi.org/10.1016/j.jksuci.2010.01.001>.
- ANUAR, F. M.; SETCHI, R.; LAI, Y. Trademark image retrieval using an integrated shape descriptor. **Expert Systems with Applications**, [s.l.], v. 40, n. 1, p. 105–121, jan. 2013. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2012.07.031>.
- 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**, [s.l.], v. 32, n. 5, p. 861–874, maio 2010. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tpami.2009.85>.
- BELONGIE, S.; MALIK, J.; PUZICHA, J. Shape matching and object recognition using shape contexts. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [s.l.], v. 24, n. 4, p. 509–522, abr. 2002. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/34.993558>.
- BERNIER, T.; LANDRY, J. A new method for representing and matching shapes of natural objects. **Pattern Recognition**, [s.l.], v. 36, n. 8, p. 1711–1723, ago. 2003. Elsevier BV. [http://dx.doi.org/10.1016/s0031-3203\(02\)00352-7](http://dx.doi.org/10.1016/s0031-3203(02)00352-7).
- BEUS, H. L.; TIU, S. S. An improved corner detection algorithm based on chain-coded plane curves. **Pattern Recognition**, [s.l.], v. 20, n. 3, p. 291–296, jan. 1987. Elsevier BV. [http://dx.doi.org/10.1016/0031-3203\(87\)90004-5](http://dx.doi.org/10.1016/0031-3203(87)90004-5).
- BLUM, H. A transformation for extracting new descriptors of shape. In: **Models for the Perception of Speech and Visual Form**. Cambridge: [s.n.], 1967. p. 362–381. MIT Press.
- CAZALS, F.; POUGET, M. Estimating differential quantities using polynomial fitting of osculating jets. **Computer Aided Geometric Design**, [s.l.], v. 22, n. 2, p. 121–146, fev. 2005. Elsevier BV. <http://dx.doi.org/10.1016/j.cagd.2004.09.004>.
- CESAR JR., R. M.; COSTA, L. F. Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. **Pattern Recognition**, [s.l.], v. 29, n. 9, p. 1559–1569, set. 1996. Elsevier BV. [http://dx.doi.org/10.1016/0031-3203\(96\)00007-6](http://dx.doi.org/10.1016/0031-3203(96)00007-6).
- CESAR JR., R. M.; COSTA, L. F. Application and assessment of multiscale bending energy for morphometric characterization of neural cells. **Review Of Scientific Instruments**, [s.l.], v. 68, n. 5, p. 2177–2186, maio 1997. AIP Publishing. <http://dx.doi.org/10.1063/1.1148112>.
- CHAKI, J.; PAREKH, R.; BHATTACHARYA, S. Plant leaf recognition using texture and shape features with neural classifiers. **Pattern Recognition Letters**, [s.l.], v. 58, p. 61–68, jun. 2015. Elsevier BV. <https://doi.org/10.1016/j.patrec.2015.02.010>.
- COPE, J. S.; CORNEY, D.; CLARK, J. Y.; REMAGNINO, P.; WILKIN, P. Plant species identification using digital morphometrics: A review. **Expert Systems with Applications**, [s.l.], v. 39, n. 8, p. 7562–7573, jun. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2012.01.073>.

COSTA, L. F.; CESAR JR., R. M. **Shape Classification and Analysis: Theory and Practice**. 2. ed. Boca Raton, FL, USA: CRC Press Taylor & Francis Group, 2009.

COX, T. F.; COX, M. A. A. **Multidimensional Scaling**. 2. ed. Boca Raton, FL, USA: Chapman and Hall/CRC, 2000.

CUI, M.; FEMIANI, J.; HU, J.; WONKA, P.; RAZDAN, A. Curve matching for open 2D curves. **Pattern Recognition Letters**, [s.l.], v. 30, n. 1, p. 1–10, jan. 2009. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2008.08.013>.

DIREKOGLU, C.; NIXON, M. S. Image-based multiscale shape description using gaussian filter. In: **Sixth Indian Conference on Computer Vision, Graphics Image Processing**. Bhudaneswar, India: [s.n.], 2008. p. 673–678. Institute of Electrical and Electronics Engineers (IEEE).

DIREKOGLU, C.; NIXON, M. S. Shape classification via image-based multiscale description. **Pattern Recognition**, [s.l.], v. 44, p. 2134–2146, set. 2011. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2011.02.016>.

EIBEN, A. E.; SMITH, J. From evolutionary computation to the evolution of things. **Nature**, [s.l.], v. 521, n. 7553, p. 476–482, maio 2015. Springer Nature. <http://dx.doi.org/10.1038/nature14544>.

ENGELBRECHT, A. P. **Computational intelligence**. 2. ed. [S.l.]: John Wiley & Sons Ltd, 2007.

ESCOLANO, F.; SUAUI, P.; BONEV, B. **Information Theory in Computer Vision**. 1. ed. London: Springer, 2009.

FENG, D.; SIU, W. C.; ZHANG, H. J. **Multimedia Information Retrieval and Management: Technological Fundamentals and Applications**. [S.l.]: Springer, 2003.

FLORINDO, J. B.; BACKES, A. R.; CASTRO, M.; BRUNO, O. M. A comparative study on multiscale fractal dimension descriptors. **Pattern Recognition Letters**, [s.l.], v. 33, n. 6, p. 798–806, abr. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2011.12.016>.

FLORINDO, J. B.; BRUNO, O. M. Multiscale Fractal Descriptors Applied to Texture Classification. **Journal of Physics: Conference Series**, Budapest, Hungary, 410, n. 1, p. 012022, set. 2013. IOP Publishing Ltd. <https://doi.org/10.1088/1742-6596/410/1/012022>.

FOTOPOULOU, F.; LASKARIS, N.; ECONOMOU, G.; FOTOPOULOS, S. Advanced leaf image retrieval via Multidimensional Embedding Sequence Similarity (MESS) method. **Pattern Analysis and Applications**, [s.l.], v. 16, n. 3, p. 381–392, dez. 2011. Springer Nature. <http://dx.doi.org/10.1007>.

FREITAS, A. M.; TORRES, R. da S.; MIRANDA, P. A. V. TSS & TSB: Tensor scale descriptors within circular sectors for fast shape retrieval. **Pattern Recognition Letters**, [s.l.], v. 83, p. 303–311, nov. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2016.06.005>.

FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**. 2. ed. [S.l.]: Morgan Kaufmann (Academic Press), 1990.

GOMES, A. C. S.; ANDRADE, A.; BARRETO-SILVA, J. S.; BRENES-ARGUEDAS, T.; LÓPEZ, D. C.; FREITAS, C. C. de; LANG, C.; OLIVEIRA, A. A. de; PÉREZ, A. J.; PEREZ, R.; SILVA, J. B. da; SILVEIRA, A. M.; VAZ, M. C.; VENDRAMI, J.; VICENTINI, A. Local plant species delimitation in a highly diverse amazonian forest: do we all see the same species? **Journal of Vegetation Science**, [s.l.], v. 24, n. 1, p. 70–79, jul. 2012. Wiley-Blackwell. <http://dx.doi.org/10.1111/j.1654-1103.2012.01441.x>.

GONZALEZ, R. C.; WOODS, R. E. **Digital Imaging Processing**. 3. ed. Upper Saddle River, NJ, USA: Prentice Hall, Inc, 2006.

HALL, D.; MCCOOL, C.; DAYOUB, F.; SUNDERHAUF, N.; UPCROFT, B. Evaluation of features for leaf classification in challenging conditions. In: **IEEE Winter Conference on Applications of Computer Vision (WACV 2015)**. Big Island, Hawaii, USA: [s.n.], 2015. p. 797–804. Institute of Electrical and Electronics Engineers (IEEE).

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man and Cybernetics**, [s.l.], v. 3, n. 6, p. 610–621, nov. 1973. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tsmc.1973.4309314>.

HU, R.; JIA, W.; ZHANG, D.; GUI, J.; SONG, L. Hand shape recognition based on coherent distance shape contexts. **Pattern Recognition**, [s.l.], v. 45, n. 9, p. 3348–3359, set. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2012.02.018>.

JIANG, M.; LUO, Y. P.; YANG, S. Y. Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. **Information Processing Letters**, [s.l.], v. 102, n. 1, p. 8–16, abr. 2007. Elsevier BV. <http://dx.doi.org/10.1016/j.ipl.2006.10.005>.

KIM, W.; KIM, Y. A region-based shape descriptor using Zernike moments. **Signal Processing: Image Communication**, [s.l.], v. 16, n. 1–2, p. 95–102, set. 2000. Elsevier BV. [http://dx.doi.org/10.1016/s0923-5965\(00\)00019-9](http://dx.doi.org/10.1016/s0923-5965(00)00019-9).

KINDRATENKO, V. V. On using functions to describe the shape. **Journal of Mathematical Imaging and Vision**, [s.l.], v. 18, n. 3, p. 225–245, mai. 2003. Springer Nature. <http://dx.doi.org/10.1023/a:1022843426320>.

KOENDERINK, J. J. The structure of images. **Biological Cybernetics**, [s.l.], v. 50, n. 5, p. 363–370, ago. 1984. Springer Nature. <http://dx.doi.org/10.1007/bf00336961>.

KOHONEN, T. Self-organized formation of topologically correct feature maps. **Biological Cybernetics**, [s.l.], v. 43, n. 1, p. 59–69, jan. 1982. Springer Nature. <http://dx.doi.org/10.1007/bf00337288>.

KOHONEN, T.; SCHROEDER, M. R.; HUANG, T. S. **Self-Organizing Maps**. 3. ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2001.

KPALMA, K.; RONSIN, J. Turning angle based representation for planar objects. **Electronics Letters**, [s.l.], v. 43, n. 10, p. 561–562, mai. 2007. Institution of Engineering and Technology (IET). <http://dx.doi.org/10.1049/el:20073895>.

KUROIWA, J.; INAWASHIRO, S.; MIYAKE, S.; ASO, H. Self-organization of orientation maps in a formal neuron model using a cluster learning rule. **Neural Networks**, [s.l.], v. 13, n. 1, p. 31–40, jan. 2000. Elsevier BV. [http://dx.doi.org/10.1016/s0893-6080\(99\)00081-7](http://dx.doi.org/10.1016/s0893-6080(99)00081-7).

LATECKI, L. J.; LAKAMPER, R.; ECKHARDT, T. Shape descriptors for non-rigid shapes with a single closed contour. In: **Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000**. Hilton Head Island, South Carolina, USA: [s.n.], 2000. v. 1, p. 424–429. Institute of Electrical and Electronics Engineers (IEEE).

LATECKI, L. J.; MEGALOOIKONOMOU, V.; WANG, Q.; YU, D. An elastic partial shape matching technique. **Pattern Recognition**, [s.l.], v. 40, n. 11, p. 3069–3080, nov. 2007. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2007.03.004>.

LEE, D. J.; ANTANI, S.; CHANG, Y.; GLEDHILL, K.; LONG, L. R.; CHRISTENSEN, P. CBIR of spine X-ray images on inter-vertebral disc space and shape profiles using feature ranking and voting consensus. **Data & Knowledge Engineering**, [s.l.], v. 68, n. 12, p. 1359–1369, dec. 2009. Elsevier BV. <https://doi.org/10.1016/j.datak.2009.07.008>.

LI, Z.; KUANG, Z.; LIU, Y.; WANG, J. Multiscale shape context and re-ranking for deformable shape retrieval. **Computers & Graphics**, [s.l.], v. 54, p. 8–17, fev. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.cag.2015.07.002>.

LING, H.; JACOBS, D. W. Using the inner-distance for classification of articulated shapes. In: **Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2005**. San Diego, California, USA: [s.n.], 2005. v. 2, p. 719–726. Institute of Electrical and Electronics Engineers (IEEE).

LING, H.; JACOBS, D. W. Shape classification using the inner-distance. **IEEE Transactions on Pattern Analysis and Machine Intelligence.**, [s.l.], v. 29, n. 2, p. 286–299, fev. 2007. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tpami.2007.41>.

LIU, X.; YOU, X.; CHEUNG, Y. Texture image retrieval using non-separable wavelets and local binary patterns. In: **2009 International Conference on Computational Intelligence and Security. CIS 2009**. Beijing. China.: [s.n.], 2009. v. 1, p. 287–291. Institute of Electrical and Electronics Engineers (IEEE).

LIU, Y.; ZHANG, D.; LU, G.; MA, W. A survey of content-based image retrieval with high-level semantics. **Pattern Recognition**, [s.l.], v. 40, n. 1, p. 262–282, jan. 2007. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2006.04.045>.

MANDELBROT, B. B. **The Fractal Geometry of Nature**. 19. ed. New York, NY, USA: W. H. Freeman & Company, 2000.

MANJUNATH, B. S.; MA, W. Texture features for browsing and retrieval of image data. **IEEE Transactions on Pattern Analysis and Machine Intelligence.**, [s.l.], v. 18, n. 8, p. 837–842, 1996. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/34.531803>.

MARANA, A. N.; COSTA, L. F.; VELASTIN, S. A.; LOTUFO, R. A. Oriented texture classification based on self-organizing neural network and Hough transform. In: **1997 IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP 1997**. Munich, Germany: [s.n.], 1997. v. 4, p. 2773–2775. Institute of Electrical and Electronics Engineers (IEEE).

MILIOS, E.; PETRAKIS, E. G. M. Shape retrieval based on dynamic programming. **IEEE Transactions on Image Processing**, [s.l.], v. 9, n. 1, p. 141–147, 2000. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/83.817606>.

MOKHTARIAN, F.; MACKWORTH, A. K. Scale-based description and recognition of planar curves and two-dimensional shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [s.l.], v. 8, n. 1, p. 34–43, jan. 1986. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tpami.1986.4767750>.

MOKHTARIAN, F.; MACKWORTH, A. K. A theory of multiscale, curvature-based shape representation for planar curves. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, [s.l.], v. 14, n. 8, p. 789–805, 1992. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/34.149591>.

MOKHTARIAN, F.; SUOMELA, R. Robust image corner detection through curvature scale space. **IEEE Transactions On Pattern Analysis And Machine Intelligence**, [s.l.], v. 20, n. 12, p. 1376–1381, 1998. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/34.735812>.

NAM, Y.; HWANG, E.; KIM, D. A similarity-based leaf image retrieval scheme: Joining shape and venation features. **Computer Vision and Image Understanding**, [s.l.], v. 110, n. 2, p. 245–259, maio 2008. Elsevier BV. <http://dx.doi.org/10.1016/j.cviu.2007.08.002>.

NANNI, L.; BRAHNAM, S.; LUMINI, A. Local phase quantization descriptor for improving shape retrieval/classification. **Pattern Recognition Letters**, [s.l.], v. 33, n. 16, p. 2254–2260, dez. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2012.07.007>.

PAGLIERONI, D. W.; JAIN, A. A control point theory for boundary representation and matching. In: **IEEE International Conference on Acoustics, Speech, and Signal Processing. ICASSP 1985**. Hyatt Regency Hotel, Tampa, Florida, USA: [s.n.], 1985. v. 10, p. 1851–1854. Institute of Electrical and Electronics Engineers (IEEE).

PALA, P.; BIMBO, A. D.; BERRETTI, S. Retrieval by shape similarity with perceptual distance and effective indexing. **IEEE Transactions on Multimedia**, [s.l.], v. 2, n. 4, p. 225–239, 2000. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/6046.890058>.

PALAZÓN-GONZÁLEZ, V.; MARZAL, A. On the dynamic time warping of cyclic sequences for shape retrieval. **Image and Vision Computing**, [s.l.], v. 30, n. 12, p. 978–990, dez. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.imavis.2012.08.012>.

PAULA JR., I. C.; MEDEIROS, F. N.; BEZERRA, F. N.; USHIZIMA, D. M. Multiscale corner detection in planar shapes. **Journal Of Mathematical Imaging And Vision**, [s.l.], v. 45, n. 3, p. 251–263, jul. 2012. Springer Nature. <http://dx.doi.org/10.1007/s10851-012-0365-8>.

QI, H.; LI, K.; SHEN, Y.; QU, W. An effective solution for trademark image retrieval by combining shape description and feature matching. **Pattern Recognition**, [s.l.], v. 43, n. 6, p. 2017 – 2027, jun. 2010. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2010.01.007>.

QUADRI, A. T.; SIRSHAR, M. Leaf recognition system using multi-class kernel support vector machine. **International Journal of Computer and Communication System Engineering**, [s.l.], v. 2, n. 2, p. 260–263, abr. 2015. IJCCS.

ROSENFELD, A.; JOHNSTON, E. Angle detection on digital curves. **IEEE Transactions on Computers**, [s.l.], C-22, n. 9, p. 875–878, set. 1973. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tc.1973.5009188>.

ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, [s.l.], v. 20, n. 1, p. 53–65, nov. 1987. Elsevier BV. [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7).

ROUSSEEUW, P. J.; LEROY, A. M. **Robust Regression and Outlier Detection**. New York, NY, USA: John Wiley & Sons, Inc., 1987.

RUBERTO, C. D.; CINQUE, L. Decomposition of two-dimensional shapes for efficient retrieval. **Image and Vision Computing**, [s.l.], v. 27, n. 8, p. 1097–1107, jul. 2009. Elsevier BV. <http://dx.doi.org/10.1016/j.imavis.2008.10.009>.

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**, [s.l.], v. 26, n. 5, p. 550–571, maio 2004. Institute of Electrical and Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tpami.2004.1273924>.

SHI, M.; FUJISAWA, Y.; WAKABAYASHI, T.; KIMURA, F. Handwritten numeral recognition using gradient and curvature of gray scale image. **Pattern Recognition**, [s.l.], v. 35, n. 10, p. 2051–2059, out. 2002. Elsevier BV. [http://dx.doi.org/10.1016/s0031-3203\(01\)00203-5](http://dx.doi.org/10.1016/s0031-3203(01)00203-5).

SHI, Y.; EBERHART, R. A modified particle swarm optimizer. In: **1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence**. Anchorage, Alaska, USA: [s.n.], 1998. p. 69–73. Institute of Electrical and Electronics Engineers (IEEE).

SHOJI, K. Generalized skeleton representation and adaptive rectangular decomposition of binary images. In: Gader, P. D.; Dougherty, E. R.; Serra, J. C. (Ed.). **Proceedings of SPIE, Image Algebra and Morphological Image Processing III**. San Diego, California, USA: [s.n.], 1992. v. 1769, p. 404–415.

SHU, X.; PAN, L.; WU, X. Multi-scale contour flexibility shape signature for Fourier descriptor. **Journal of Visual Communication and Image Representation**, [s.l.], v. 26, p. 161–167, jan. 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.jvcir.2014.11.007>.

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**, [s.l.], v. 22, n. 12, p. 1349–1380, dez. 2000. Institute of Electrical and Electronics Engineers (IEEE). <https://doi.org/10.1109/34.895972>.

SOUZA, M. M. de; MEDEIROS, F. N.; RAMALHO, G. L.; PAULA JR., I. C. de; OLIVEIRA, I. N. Evolutionary optimization of a multiscale descriptor for leaf shape analysis. **Expert Systems with Applications**, [s.l.], v. 63, p. 375–385, nov. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.eswa.2016.07.016>.

STORN, R. On the usage of differential evolution for function optimization. In: **1996 Biennial Conference of the North American Fuzzy Information Processing Society - NAFIPS**. Berkeley, California, USA: [s.n.], 1996. p. 519–523. Institute of Electrical and Electronics Engineers (IEEE).

STORN, R.; PRICE, K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. **Journal of Global Optimization**, [s.l.], v. 11, n. 4, p. 341–359, dez. 1997. Springer Nature. <http://dx.doi.org/10.1023/A:1008202821328>.

- STRONG, G.; GONG, M. Similarity-based image organization and browsing using multi-resolution self-organizing map. **Image and Vision Computing**, [s.l.], v. 29, n. 11, p. 774–786, out. 2011. Elsevier BV. <http://dx.doi.org/10.1016/j.imavis.2011.08.007>.
- SUZUKI, S.; BE, K. Topological structural analysis of digitized binary images by border following. **Computer Vision, Graphics, and Image Processing**, [s.l.], v. 30, n. 1, p. 32 – 46, abr. 1985. Elsevier BV. [http://dx.doi.org/10.1016/0734-189x\(85\)90016-7](http://dx.doi.org/10.1016/0734-189x(85)90016-7).
- TORRES, R. S.; FALCÃO, A. X. Content-based image retrieval: Theory and applications. **Revista de Informática Teórica e Aplicada**, [s.l.], v. 13, n. 2, p. 161–185, jan. 2006. Universidade Federal do Rio Grande do Sul.
- ULLAH, A. Entropy, divergence and distance measures with econometric applications. **Journal of Statistical Planning and Inference**, [s.l.], v. 49, n. 1, p. 137–162, jan. 1996. Elsevier BV. [http://dx.doi.org/10.1016/0378-3758\(95\)00034-8](http://dx.doi.org/10.1016/0378-3758(95)00034-8).
- ULLMAN, S. **High-level vision**. Cambridge, Massachusetts, USA: MIT Press, 1995.
- ULTSCH, A.; SIEMON, H. P. Kohonen's Self Organizing Feature Maps for Exploratory Data Analysis. In: **Proceedings of International Neural Network Conference (INNC 1990)**. Paris, France: [s.n.], 1990. p. 305–308. Kluwer Academic Publishers.
- USHIZIMA, D.; CARNEIRO, A.; SOUZA, M.; MEDEIROS, F. Investigating pill recognition methods for a new national library of medicine image dataset. In: **Advances in Visual Computing**. Las Vegas, NV, USA: [s.n.], 2015, (Lecture Notes in Computer Science, v. 9475). p. 410–419. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-27863-6_38.
- WANG, B.; BROWN, D.; GAO, Y.; SALLE, J. L. March: Multiscale-arch-height description for mobile retrieval of leaf images. **Information Sciences**, [s.l.], v. 302, p. 132–148, maio 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.ins.2014.07.028>.
- WANG, J.; BAI, X.; YOU, X.; LIU, W.; LATECKI, L. J. Shape matching and classification using height functions. **Pattern Recognition Letters**, [s.l.], v. 33, n. 2, p. 134–143, jan. 2012. Elsevier BV. <http://dx.doi.org/10.1016/j.patrec.2011.09.042>.
- WANG, Z.; CHI, Z.; FENG, D.; WANG, Q. Leaf image retrieval with shape features. In: **Advances in Visual Information Systems**. Lyon, France: [s.n.], 2000, (Lecture Notes in Computer Science, v. 1929). p. 477–487. Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-40053-2_42.
- WEBB, A. R. **Statistical Pattern Recognition**. 2. ed. [S.l.]: John Wiley & Sons Ltd, 2002.
- WITKIN, A. P. Scale-space filtering. In: **Readings in Computer Vision**. San Francisco, CA, USA: [s.n.], 1987. p. 329–332. Morgan Kaufmann Publishers Inc.
- WU, S. G.; BAO, F. S.; XU, E. Y.; WANG, Y.; CHANG, Y.; XIANG, Q. A leaf recognition algorithm for plant classification using probabilistic neural network. In: **2007 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2007)**. Cairo, Egypt: [s.n.], 2007. p. 11–16. Institute of Electrical and Electronics Engineers (IEEE).
- YANG, G. Y.; SHU, H. Z.; TOUMOULIN, C.; HAN, G. N.; LUO, L. M. Efficient Legendre moment computation for grey level images. **Pattern Recognition**, [s.l.], v. 39, n. 1, p. 74–80, jan. 2006. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2005.08.008>.

- YANG, J.; WANG, H.; YUAN, J.; LI, Y.; LIU, J. Invariant multi-scale descriptor for shape representation, matching and retrieval. **Computer Vision and Image Understanding**, [s.l.], v. 145, p. 43–58, abr. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.cviu.2016.01.005>.
- YANG, W.; ZHANG, S.; CHEN, Y.; LI, W.; CHEN, Y. Shape symmetry analysis of breast tumors on ultrasound images. **Computers in Biology and Medicine**, [s.l.], v. 39, n. 3, p. 231–238, mar. 2009. Elsevier BV. <http://dx.doi.org/10.1016/j.compbiomed.2008.12.007>.
- YOUNG, I. T.; WALKER, J. E.; BOWIE, J. E. An analysis technique for biological shape. **Information and Control**, [s.l.], v. 25, n. 4, p. 357–370, ago. 1974. Elsevier BV. [http://dx.doi.org/10.1016/s0019-9958\(74\)91038-9](http://dx.doi.org/10.1016/s0019-9958(74)91038-9).
- ZHANG, D.; LU, G. Shape-based image retrieval using generic Fourier descriptor. **Signal Processing: Image Communication**, [s.l.], v. 17, n. 10, p. 825–848, nov. 2002. Elsevier BV. [http://dx.doi.org/10.1016/s0923-5965\(02\)00084-x](http://dx.doi.org/10.1016/s0923-5965(02)00084-x).
- ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern Recognition**, [s.l.], v. 37, n. 1, p. 1–19, jan. 2004. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2003.07.008>.
- ZHANG, Y.; LIU, J. J.; ZHANG, L.; ANDA, J. C. D.; WANG, X. Z. Particle shape characterisation and classification using automated microscopy and shape descriptors in batch manufacture of particulate solids. **Particuology**, [s.l.], v. 24, p. 61–68, fev. 2016. Elsevier BV. <http://dx.doi.org/10.1016>.
- ZHAO, C.; CHAN, S. S. F.; CHAM, W.; CHU, L. M. Plant identification using leaf shapes—A pattern counting approach. **Pattern Recognition**, [s.l.], v. 48, n. 10, p. 3203–3215, out. 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.patcog.2015.04.004>.