



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

RAPHAEL GOMES PINHEIRO

**ENTROPIA DIFERENCIAL MULTIESCALA NA DESCRIÇÃO E ANÁLISE DE
FORMAS**

FORTALEZA

2021

RAPHAEL GOMES PINHEIRO

ENTROPIA DIFERENCIAL MULTIESCALA NA DESCRIÇÃO E ANÁLISE DE FORMAS

Dissertação apresentada ao 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 mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Orientadora: Prof^a. Dr^a. Fátima Nelsi-zeuma Sombra de Medeiros

FORTALEZA

2021

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)

P722e Pinheiro, Raphael Gomes.
Entropia Diferencial Multiescala na Descrição e Análise de Formas / Raphael Gomes Pinheiro. – 2021.
53 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Teleinformática, Fortaleza, 2021.

Orientação: Profa. Dra. Fátima Nelsizeuma Sombra de Medeiros.

1. Análise de formas. 2. Entropia diferencial. 3. Descritores multiescala. 4. Otimização meta-heurística.
5. Classificação de formas. I. Título.

CDD 621.38

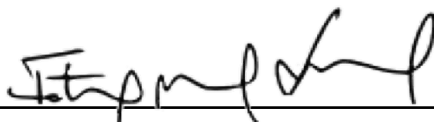
RAPHAEL GOMES PINHEIRO

ENTROPIA DIFERENCIAL MULTIESCALA NA DESCRIÇÃO E ANÁLISE DE FORMAS

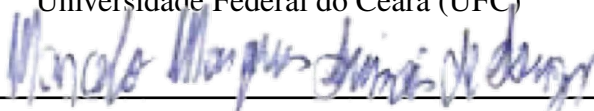
Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Teleinformática do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia de Teleinformática. Área de Concentração: Sinais e Sistemas

Aprovada em: 17 de Junho de 2021

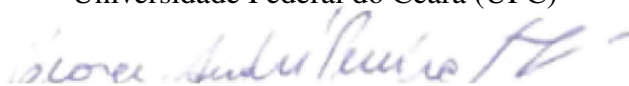
BANCA EXAMINADORA



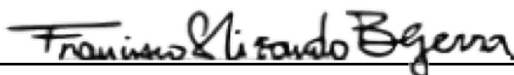
Prof^a. Dr^a. Fátima Nelsizeuma Sombra de Medeiros (Orientadora)
Universidade Federal do Ceará (UFC)



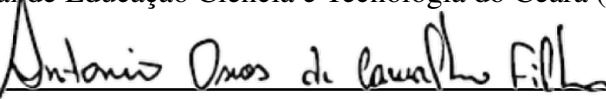
Prof. Dr. Marcelo Marques Simões de Souza
Universidade Federal do Ceará (UFC)



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



Prof. Dr. Francisco Nivando Bezerra
Instituto Federal de Educação Ciência e Tecnologia do Ceará (IFCE)



Prof. Dr. Antônio Oseas de Carvalho Filho
Universidade Federal do Piauí (UFPI)

“Como poderei retribuir ao Senhor Deus por tudo
aquilo que Ele fez em meu favor?”
(Salmos 116,12)

AGRADECIMENTOS

Primeiramente, venho agradecer a Deus, que com sua infinita bondade me permitiu chegar até aqui. Obrigado Senhor por ter me agraciado com muita saúde e força de vontade para evoluir com sabedoria ao longo de minha vida. Para mim és o maior mestre!

Aos meus pais, por todo o amor, carinho, dedicação, incentivo e apoio incondicional a todas as minhas decisões. Vocês são a minha base e o meu orgulho!

À minha irmã, por todo o companheirismo, afetividade e por estar sempre disponível para me ajudar juntamente com a minha tia Branca, que foi presente durante toda a minha caminhada acadêmica e afetivamente é minha segunda mãe.

À minha orientadora, Prof^ª. Dr^ª. Fátima Nelsizeuma Sombra de Medeiros, por todas as oportunidades que me concedeu, pela paciência, pelo incentivo constante e pela confiança.

Aos amigos do Grupo de Processamento de Imagens, por toda a convivência e acolhimento. Com vocês aprendi muito!

À todo o corpo docente do Departamento de Engenharia de Teleinformática, coordenação e secretaria, por serem profissionais competentes, responsáveis e muito prestativos.

E por fim, meu sincero agradecimento à CAPES, por todo o apoio financeiro.

"Não vos amoldeis às estruturas deste mundo, mas transformai-vos pela renovação da mente, a fim de distinguir qual é a vontade de Deus: o que é bom, o que Lhe é agradável, o que é perfeito."

(Bíblia Sagrada, Romanos 12, 2)

RESUMO

A análise e o reconhecimento de formas são fundamentais no projeto de sistemas baseados em visão computacional. O grande desafio consiste em desenvolver métodos robustos capazes de extrair características significativas dessas formas para representá-las. Nesse contexto, os descritores multiescala são uma alternativa na caracterização de formas, utilizando-se de uma ferramenta versátil e eficiente. Este trabalho propõe um descritor de formas com base na entropia diferencial da curvatura multiescala (*MEC*), que apresenta melhor poder de discriminação em formas quando comparado ao já consolidado descritor energia de dobramento multiescala normalizado (*NMBE*). Em uma primeira abordagem, esses dois descritores são comparados com descritores monoescala em experimentos de classificação com formas de três bases públicas: Kimia-99, MPEG7-CE e Flavia, sendo esta uma base de folhas de plantas. Os resultados mostraram que os descritores monoescala foram mais efetivos nas formas da base Kimia-99, enquanto o *MEC* apresentou melhor desempenho nas duas últimas bases, que são inclusive as mais desafiadoras. Na segunda abordagem, os parâmetros de ajuste dos descritores ou escalas do *MEC* e do *NMBE* são otimizadas por um algoritmo meta-heurístico denominado *Simulated Annealing*. Para tanto, empregamos uma função custo de validação de agrupamentos, a *Silhouette*, visando encontrar, através da otimização, o melhor conjunto de escalas que maximize a coesão intraclasse e a separação entre classes. Após a otimização, realizamos experimentos de classificação cujos resultados apresentaram um ganho considerável nas taxas de precisão, de revocação e de Acurácia, com destaque especial para o descritor proposto na base de folhas, que superou em 10% o ganho na precisão. Por fim, a terceira abordagem proposta considera a concatenação de ambos os descritores multiescala, sendo esta última a que refletiu em maior ganho de desempenho nos experimentos de classificação.

Palavras-chave: Análise de formas. Curvatura multiescala. Entropia diferencial. Descritores multiescala. Otimização meta-heurística. Classificação de formas.

ABSTRACT

Shape analysis and recognition are fundamental in the design of systems based on computer vision. The greatest challenge is to develop robust methods capable of extracting significant features from these shapes to represent them. In this context, multiscale descriptors are an alternative in the characterization of shapes, using a versatile and efficient tool. This work proposes a shape descriptor based on the differential entropy of the multiscale curvature (MEC), which presents better shape discrimination power when compared to the already consolidated descriptor normalized multiscale bending energy (NMBE). In a first approach, these two descriptors are compared with monoscale descriptors in shape classification experiments from three public bases: Kimia-99, MPEG7-CE and Flavia, which is a dataset of plant leaves. The results showed that the monoscale descriptors were more effective in Kimia-99 dataset, while the MEC had the best performance in the last two datasets, which are even the most challenging. In the second approach, the adjustment parameters or scales of the MEC and NMBE descriptors are optimized by a meta-heuristic algorithm called Simulated Annealing. For that, we employ a clustering validation cost function, the Silhouette, aiming to find, through optimization, the best set of scales that maximize the intraclass cohesion and the separation between classes. After optimization, we carried out classification experiments whose results showed a considerable gain in precision, recall and accuracy measures, with special emphasis on the descriptor proposed on the leaves dataset, which exceeded by 10% the gain in precision. Finally, the proposed third approach considers the concatenation of both multiscale descriptors, the latter being the one that reflected the greatest performance gain in the classification experiments.

Keywords: Shape analysis. Multiscale curvature. Differential entropy. Multiscale descriptors. Metaheuristic optimization. Shape classification.

LISTA DE FIGURAS

Figura 1 – Exemplos de formas com diferentes circularidades.	16
Figura 2 – Exemplos de formas com diferentes excentricidades.	17
Figura 3 – Exemplos de formas com diferentes razões de aspecto.	17
Figura 4 – Exemplo de aplicação do conceito de solidez em uma forma.	18
Figura 5 – Formas da classe de arraias da base Kimia-99 e seus respectivos contornos externos	18
Figura 6 – Processo de extração do contorno de uma forma. (a) Processamento da imagem, segmentação da forma e aquisição de seu contorno. (b) Amostragem do contorno. (c) Gráfico das abscissas e ordenadas do contorno.	19
Figura 7 – Curvatura estimada do contorno de uma folha.	20
Figura 8 – Contorno em múltiplas escalas de uma forma da classe de arraias da base Kimia-99.	22
Figura 9 – Curvatura multiescala do contorno de uma forma de folha da Figura 6. . . .	23
Figura 10 – Metodologia geral para a descrição multiescala de formas, com ênfase na otimização.	29
Figura 11 – Todas as formas da base Kimia-99.	31
Figura 12 – Uma amostra por classe da base MPEG7-CE.	31
Figura 13 – Uma amostra por classe da base Flavia.	31
Figura 14 – Efeito da rotação de uma forma sobre a curvatura.	33
Figura 15 – Efeito da mudança de escala de uma forma sobre a curvatura.	34
Figura 16 – Efeito da mudança de escala de uma forma sobre o descritor. (a) Sem o fator de normalização. (b) Com o fator de normalização.	35
Figura 17 – Efeito da rotação de uma forma sobre o descritor.	35
Figura 18 – Experimento de validação cruzada realizado para avaliar o desempenho dos descritores na classificação de formas.	42
Figura 19 – Exemplo de aplicação dos descritores multiescala em uma forma da base MPEG7-CE.	43
Figura 20 – Matrizes de distância euclidiana par-a-par nas classes da base Kimia-99 a partir do descritor <i>MEC</i>	45
Figura 21 – Matrizes de distâncias obtidas com <i>NMBE</i> , <i>MEC</i> e <i>MONO</i> para (a) Kimia-99; (b) MPEG7-CE e (c) Flavia.	47

LISTA DE TABELAS

Tabela 1 – Escalas utilizadas para o descritor <i>NMBE</i>	26
Tabela 2 – Resultados (sem otimização) de precisão (P) e de revocação (R) para os testes de validação cruzada com <i>NMBE</i> , <i>MEC</i> e MONO para as três bases de formas.	46
Tabela 3 – Resultados (sem otimização) de <i>F-score</i> e de Acurácia para os testes de validação cruzada com <i>NMBE</i> , <i>MEC</i> e MONO para as três bases de formas.	46
Tabela 4 – Resultados (com otimização) de precisão (P) e de revocação (R) para os testes de validação cruzada com <i>NMBE</i> , <i>MEC</i> e <i>NMBE + MEC</i> para as três bases de formas.	48
Tabela 5 – Resultados (com otimização) de <i>F-score</i> e de Acurácia para os testes de validação cruzada com <i>NMBE</i> , <i>MEC</i> e <i>NMBE + MEC</i> para as três bases de formas.	48

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Objetivos	13
1.2	Produção científica	14
1.3	Organização da dissertação	14
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Extração de características	15
2.1.1	<i>Características monoescala</i>	15
2.1.2	<i>Representação paramétrica do contorno de formas</i>	18
2.1.3	<i>Curvatura</i>	20
2.1.4	<i>Representações multiescala</i>	21
2.2	Energia de Dobramento Multiescala Normalizada (NMBE)	23
2.3	Conceitos de Teoria da Informação	27
2.3.1	<i>Entropia de Shannon</i>	27
2.3.2	<i>Entropia diferencial</i>	27
3	MATERIAIS E MÉTODOS	29
3.1	Bases de formas	30
3.2	Entropia como descritor multiescala	31
3.2.1	<i>Entropia da Curvatura Multiescala (MEC)</i>	32
3.3	Algoritmo de otimização	35
3.3.1	<i>Simulated Annealing (SA)</i>	36
3.4	Função custo para validação de agrupamentos	38
3.4.1	<i>Silhouette (SI)</i>	39
3.5	Algoritmo de classificação KNN	39
3.6	Métricas de avaliação da classificação	40
3.7	Experimentos	41
4	RESULTADOS	44
4.1	Experimentos de classificação sem otimização	44
4.2	Experimentos de classificação com otimização	48
5	CONCLUSÕES E TRABALHOS FUTUROS	50
	REFERÊNCIAS	51

1 INTRODUÇÃO

Visão computacional é um campo de estudo da inteligência artificial que busca compreender como os sistemas visuais funcionam e podem ser implementados, a fim de encontrar soluções computacionais eficientes para tarefas que têm imagens como sua principal entrada (SOUZA *et al.*, 2016). Nesse campo, a modelagem de formas desempenha um papel importante, uma vez que uma das mais relevantes informações visuais para a identificação de um objeto está em seu contorno. No entanto, encontrar descritores de forma que representem objetos a partir de sua geometria é, em geral, desafiador, embora seja fundamental em tarefas de reconhecimento (CARNEIRO *et al.*, 2019).

Nesse contexto, as redes neurais convolucionais (CNNs) ganharam destaque e se apresentam hoje bem sucedidas em diversos problemas de reconhecimento de padrões (CARNEIRO *et al.*, 2017). Além disso, a análise de formas a partir de descritores multiescala recebeu atenção crescente nas últimas décadas (YANG; YU, 2019), uma vez que estes são capazes de caracterizar formas a partir de uma arquitetura menos complexa, mais transparente e eficiente. Uma abordagem na descrição multiescala é a caracterização das formas em termos das curvaturas de seus contornos digitais, em diferentes escalas espaciais.

Este trabalho enfatiza a extração de características a partir da geometria do contorno das formas e também destaca a aplicação de técnicas de processamento de sinais para a obtenção de assinaturas, como a curvatura, que tem sido empregada recentemente e com sucesso em tarefas de descrição de formas (CARNEIRO *et al.*, 2019).

Apesar da curvatura ser um recurso importante na descrição de formas, sua estimativa é difícil, uma vez que ruídos e distorções estão constantemente presentes nos sinais amostrados. No entanto, a partir de conceitos de geometria diferencial e de técnicas de processamento de sinais, é possível estimá-la com precisão (CESAR JR; COSTA, 1997). A motivação é desenvolver um descritor de formas robusto o suficiente para ser comparável ao *NMBE* e até mesmo superá-lo em tarefas de classificação.

Nos últimos anos, os pesquisadores exploraram muitos conceitos da teoria da informação para resolver problemas de reconhecimento de padrões (RUIZ *et al.*, 2009; DAS *et al.*, 2019), uma vez que medidas como a entropia e a informação mútua se relacionam diretamente com as informações de conteúdo dos sinais (SINGH; PRINCIPE, 2011). Com base nesses estudos, esta dissertação propõe um descritor de formas a partir da entropia diferencial (*MEC*), utilizando os conceitos de curvatura multiescala que foram aplicados com sucesso para

representar formas em trabalhos recentes da literatura (SOUZA *et al.*, 2016; CARNEIRO *et al.*, 2017; CARNEIRO *et al.*, 2019). Porém, até então, a energia de dobramento multiescala normalizada (*NMBE*) foi o descritor de formas utilizado, por ser robusto a ruídos e invariante à escala, rotação e translação.

Carneiro *et al.* (2019) realizaram o ajuste dos parâmetros de escalas do descritor de formas *NMBE* por otimização, através da meta-heurística *simulated annealing*, adotando como função custo o índice de validação de agrupamentos *Silhouette*. Inspirados por Carneiro *et al.* (2019), aplicamos esta mesma metodologia de otimização ao descritor *MEC*, buscando comparar o desempenho de ambos descritores em experimentos de classificação supervisionada sem otimização e com otimização. As bases de formas utilizadas nos testes foram duas sintéticas: Kimia-99 (SEBASTIAN *et al.*, 2004) e MPEG7-CE (LING *et al.*, 2010), e uma base de folhas de plantas reais denominada Flavia (WU *et al.*, 2007). Os experimentos de classificação supervisionada foram guiados pelo algoritmo k vizinhos mais próximos (KNN), utilizando as medidas de desempenho *F-score* e Acurácia. Por fim, experimentos foram realizados com vetores de características do descritor proposto (*MEC*) otimizado concatenado ao do outro descritor multiescala (*NMBE*) otimizado, a fim de verificar se a adição de informação por descritores diferentes promove ganho de desempenho.

1.1 Objetivos

Este trabalho tem como objetivo geral a proposição de um descritor de formas com base na entropia diferencial da curvatura multiescala (*MEC*).

Dentro desta abordagem e com a finalidade de alcançar o objetivo geral, alguns objetivos específicos são listados:

- Garantir robustez ao *MEC* a partir de invariância à rotação, à translação e à escala;
- Avaliar qualitativa e quantitativamente a capacidade discriminativa de ambos os descritores multiescala, frente a descritores monoescala;
- Aplicar ao descritor de formas proposto (*MEC*) a mesma metodologia de otimização de escalas utilizada no *NMBE* em (CARNEIRO *et al.*, 2019), comparando os ganhos de desempenho em ambos;
- Validar a eficiência do *MEC* em experimentos de classificação, a partir de formas de uso geral e de folhas de plantas reais;
- Concatenar as escalas ótimas de ambos os descritores multiescala (*NMBE+MEC*), visando

melhorar o ganho de desempenho na classificação;

- Otimizar conjuntamente as escalas de ambos os descritores multiescala.
- Comparar o descritor proposto à energia de dobramento multiescala normalizado (*NMBE*), avaliando o desempenho dos mesmos em experimentos de classificação.

1.2 Produção científica

1. PINHEIRO, R. G.; LOPES, J. G. F.; CARNEIRO, A. C.; SOUZA, M. M. S.; MEDEIROS, F. N. S. *A multiscale shape descriptor based on differential entropy*. **14º Simpósio Brasileiro de Automação Inteligente (SBAI)**, Ouro Preto, 2019.

1.3 Organização da dissertação

Esta dissertação segue a seguinte estrutura:

- **Capítulo 2:** apresenta os conceitos teóricos fundamentais para o entendimento das técnicas aplicadas na metodologia do trabalho, dentre os quais o processo de extração de características das formas, o descritor Energia de Dobramento Multiescala Normalizada (*NMBE*) e os conceitos de entropia discreta e diferencial.
- **Capítulo 3:** apresenta as bases de imagens que foram utilizadas para a realização dos experimentos de classificação; introduz a entropia como um novo descritor multiescala; descreve uma metodologia de otimização dos descritores, as funções custo empregadas, bem como o classificador escolhido e as métricas de avaliação dos resultados.
- **Capítulo 4:** exhibe e discute os resultados dos experimentos realizados com os descritores de formas, incluindo uma avaliação qualitativa e quantitativa dos mesmos.
- **Capítulo 5:** destaca as principais conclusões alcançadas bem como as limitações das abordagens adotadas, possíveis melhorias e trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados os conceitos básicos necessários para o entendimento das técnicas utilizadas na metodologia deste trabalho. Inicialmente descrevemos os métodos de extração de características com base na geometria das formas, destacando métricas de área, perímetro, comprimento e largura. Em seguida, tratamos dos métodos baseados em assinaturas do contorno, em particular a curvatura (KINDRATENKO, 2003), a curvatura multiescala (MOKHTARIAN; MACKWORTH, 1992) e, por fim, o descritor *NMBE* (COSTA; CESAR JR, 2009). Este último é o descritor multiescala que será comparado ao novo descritor, *MEC*, proposto na metodologia deste trabalho. Para isso, conceitos gerais de teoria da informação também serão discutidos ao fim deste capítulo.

2.1 Extração de características

Dentre os atributos dos quais se realizam a extração de características, a forma é considerada a mais relevante em diversas aplicações de visão computacional pela riqueza de informações que esta possui (SOUZA *et al.*, 2016).

De acordo com Zhang e Lu (2004), uma forma é obtida quando um objeto de interesse é identificado e segmentado em uma imagem. A representação dessas formas envolve a busca por informações que sejam perceptualmente relevantes. Porém, obter tal representação é uma tarefa complexa, pois em uma imagem os objetos tridimensionais são projetados em duas dimensões, gerando perda de informações visuais importantes para sua descrição. O problema torna-se ainda mais complexo se levarmos em conta que a forma é frequentemente corrompida por ruídos (como digitalização, borramento, etc), defeitos, distorções arbitrárias ou oclusões (SINGH; SINGH, 2019).

2.1.1 Características monoescala

Os descritores monoescala são em sua maioria relacionados a características geométricas das formas e podem ser usados para classificação e agrupamento (COSTA; CESAR JR, 2009). Entre as diversas características monoescala existentes, as exploradas nesse trabalho foram: circularidade, excentricidade, razão de aspecto e solidez.

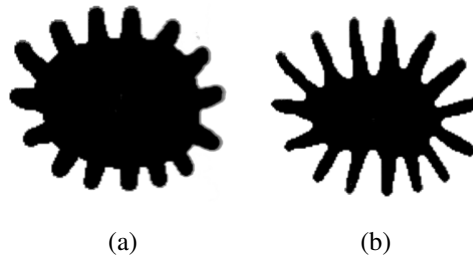
Circularidade

A circularidade ou compacidade de um objeto é definida como:

$$C = \frac{P^2}{A}, \quad (2.1)$$

onde P e A são o perímetro e a área da região, respectivamente, ambos medidos em unidades de pixels. Esta característica descreve o quanto a forma se aproxima de uma circunferência, o que significa que quanto maior for este valor, mais circular é a forma (PEDRINI; SCHWARTZ, 2007). Algumas das propriedades desse descritor são: a sua invariância à translação, à rotação e à escala. A Figura 1 ilustra formas que apresentam diferentes circularidades, $C_{(a)} > C_{(b)}$.

Figura 1 – Exemplos de formas com diferentes circularidades.



Excentricidade

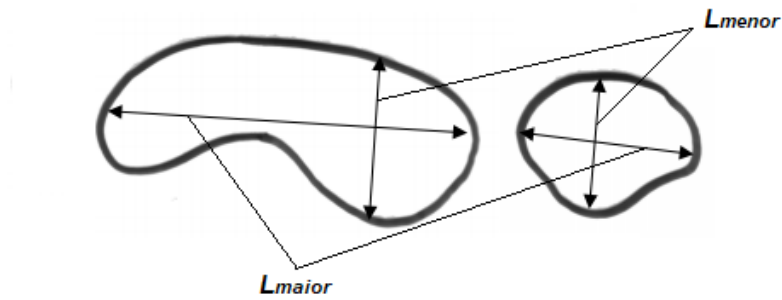
Esta característica, definida em termos da razão entre o comprimento do menor e do maior eixo horizontal ou vertical de uma forma, identifica como o objeto está distribuído espacialmente em relação aos seus eixos, em um intervalo que varia entre $[0,1]$ (GONZALEZ; WOODS, 2006).

Desse modo, se o valor da excentricidade (E) for próximo a 1, significa que o objeto é mais circular, e caso seja próximo a 0, menos circular, tendo assim uma baixa excentricidade. A sua Equação é dada por:

$$E = \frac{L_{menor}}{L_{maior}}, \quad (2.2)$$

em que L_{menor} corresponde ao comprimento do eixo menor e L_{maior} ao comprimento do eixo maior, ambos indicados em exemplos na Figura 2.

Figura 2 – Exemplos de formas com diferentes excentricidades.

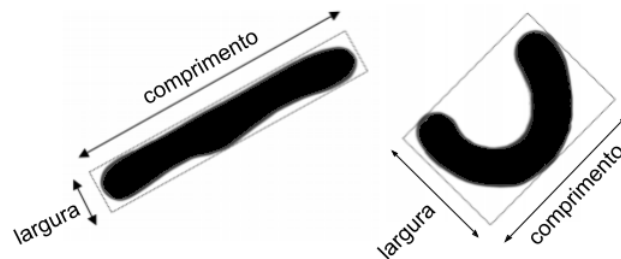


Razão de aspecto

A razão de aspecto, característica ilustrada na Figura 3, é mais conhecida pelo termo em inglês *aspect ratio* como a medida da razão entre o comprimento e a largura de uma forma (COSTA; CESAR JR, 2009), sendo definida por:

$$P = \frac{\text{comprimento}}{\text{largura}}. \quad (2.3)$$

Figura 3 – Exemplos de formas com diferentes razões de aspecto.



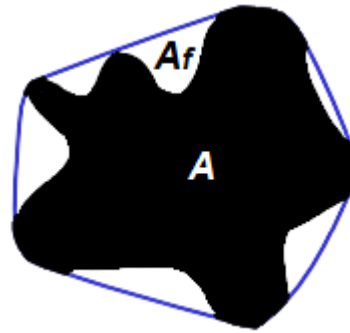
Solidez

A solidez é uma medida de densidade de um objeto e pode ser obtida pela relação entre a área de um objeto e a área do fecho convexo do objeto. Entende-se por fecho convexo de um conjunto de pontos o menor polígono convexo que contém esse pontos (BARBER *et al.*, 1996).

$$S = \frac{A}{A_f}, \quad (2.4)$$

em que A representa a área do objeto e A_f a área do fecho convexo (ver Figura 4). O valor máximo de solidez é 1 para um objeto sólido convexo e menor que 1 para objetos não convexos (PEDRINI; SCHWARTZ, 2007).

Figura 4 – Exemplo de aplicação do conceito de solidez em uma forma.

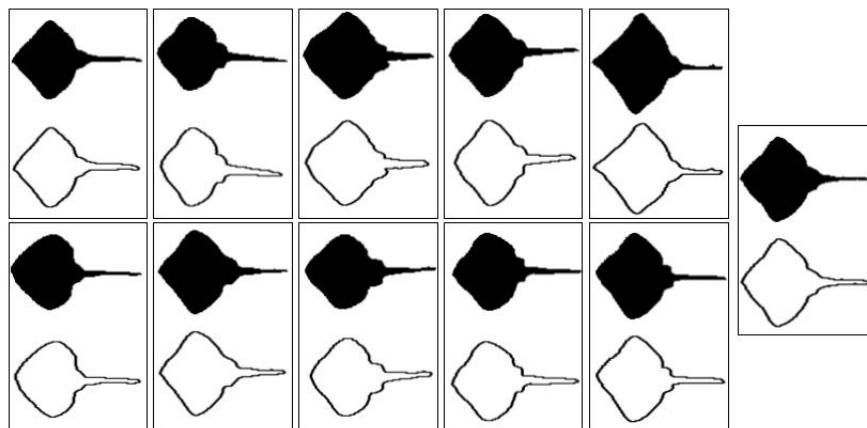


2.1.2 Representação paramétrica do contorno de formas

No processo de descrição de formas, busca-se por características que as especifiquem apropriadamente, determinando o que é necessário para que um sistema computacional possa distingui-las, agrupá-las ou classificá-las corretamente (ZHANG; LU, 2004).

Os descritores multiescala de formas explorados neste trabalho (*NMBE* e *MEC*) utilizam como base as relações entre os elementos do contorno, o qual pode ser definido como o conjunto de pontos que delimitam uma forma. Aqui vamos considerar apenas o contorno externo, ou seja, o contorno que separa um objeto bidimensional de seu exterior, como exemplificado na Figura 5.

Figura 5 – Formas da classe de arraias da base Kimia-99 e seus respectivos contornos externos .



Para extrair os pontos do contorno de um objeto, define-se um ponto inicial em sua borda e, a partir desse, percorre-se todo o contorno em sentido horário (ou anti-horário). Desse processo, resultam pares ordenados que podem ser representados como dois vetores (COSTA; CESAR JR, 2009):

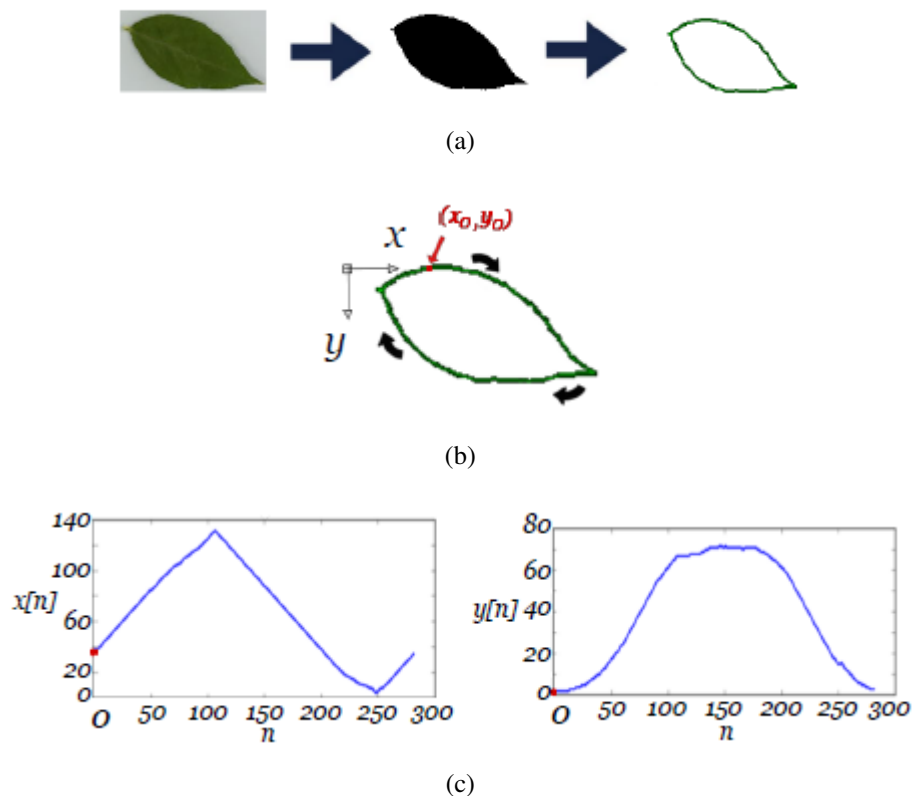
$$x = [x(1), x(2), \dots, x(N)], \quad y = [y(1), y(2), \dots, y(N)], \quad (2.5)$$

sendo N o número de pontos do contorno.

Cada par $(x(n), y(n))$ corresponde a localização cartesiana de um ponto do contorno e n é a sua ordem. Assim, obtém-se o contorno parametrizado pela posição em que seus elementos foram extraídos.

A Figura 6 exemplifica o processo de extração do contorno de uma forma a partir de uma folha da base de imagens Flavia. A Figura 6a destaca as etapas envolvidas para a representação de uma forma. Temos nesse caso a segmentação da imagem por limiar seguida da extração de seu contorno. A Figura 6b ilustra a amostragem do contorno, na qual está destacado em vermelho o ponto de origem de onde a varredura se inicia, em sentido horário. Na Figura 6c estão representados os sinais obtidos do contorno da folha, com amostragem $N = 280$ pontos. Os dois pontos observados aonde a evolução do sinal $x[n]$ 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 composto das ordenadas $y[n]$ corresponde a região da parte inferior da folha, em que quase não se observam variações do contorno ao longo do eixo y .

Figura 6 – Processo de extração do contorno de uma forma. (a) Processamento da imagem, segmentação da forma e aquisição de seu contorno. (b) Amostragem do contorno. (c) Gráfico das abscissas e ordenadas do contorno.



Fonte: (SOUZA, 2016)

2.1.3 Curvatura

A curvatura é uma assinatura do contorno de uma forma com importantes propriedades geométricas, o que motiva sua utilização para obtenção de descritores. Há até mesmo 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).

Para uma curva contínua fechada, a curvatura pode ser definida através das coordenadas paramétricas do contorno como (KINDRATENKO, 2003):

$$\kappa(n) = \frac{\dot{x}(n)\ddot{y}(n) - \ddot{x}(n)\dot{y}(n)}{\{[\dot{x}(n)]^2 + [\dot{y}(n)]^2\}^{\frac{3}{2}}}, \quad (2.6)$$

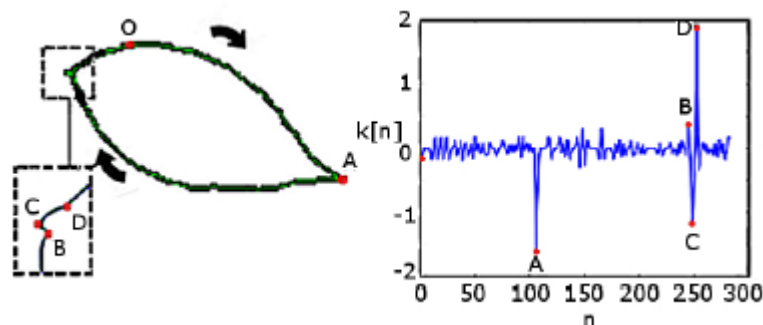
em que $[\dot{x}(n), \dot{y}(n)]$ representam as derivadas de primeira ordem e $[\ddot{x}(n), \ddot{y}(n)]$ são as derivadas de segunda ordem das coordenadas paramétricas da curva.

A Equação 2.6 resulta em valores positivos ou negativos dependendo da concavidade local e, além disso, a curvatura é muito influenciada pela geometria do contorno. Analiticamente, a curvatura de uma reta é zero, a de uma circunferência é constante e a de um objeto irregular repleto de cantos e junções apresenta grande variação.

Porém, sob o ponto de vista computacional, o cálculo da curvatura requer que o contorno seja espacialmente amostrado e discretizado. Tal processo torna o cálculo das derivadas da Equação 2.6 muito sensível ao ruído, o que limita a aplicação direta da curvatura para a obtenção de descritores.

Diversas estratégias foram propostas na literatura para contornar o problema da sensibilidade ao ruído no cálculo computacional da curvatura. Entre elas está um método proposto por Mokhtarian e Mackworth (1992), que suaviza o contorno através de uma convolução do mesmo com um filtro gaussiano, antes de calcular a curvatura pela Equação 2.6.

Figura 7 – Curvatura estimada do contorno de uma folha.



Fonte: (SOUZA, 2016)

A Figura 7 mostra a curvatura estimada para o contorno da folha da Figura 6, através do método proposto por Mokhtarian e Mackworth. Nela, o contorno foi previamente suavizado com um filtro gaussiano de desvio padrão $\sigma=20$ e os pontos vermelhos aonde a curvatura apresenta os picos correspondem aos pontos mais salientes da folha, ou seja, aos pontos de mais altas curvaturas.

2.1.4 Representações multiescala

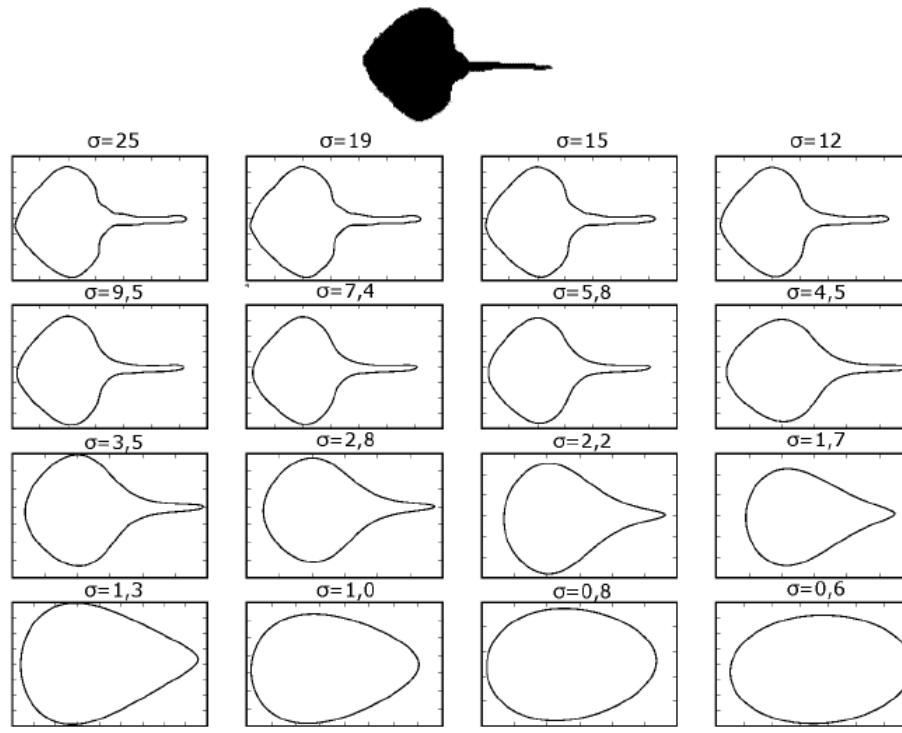
O contorno carrega muita informação a respeito das formas, porém, a sensibilidade dele a ruídos deve ser tratado com a devida atenção. Além disso, em alguns casos, o contorno não se apresenta completamente disponível, com regiões disjuntas, descontínuas ou oclusas, comprometendo a confiabilidade e a precisão de suas assinaturas.

Para superar tais problemas, a descrição multiescala do contorno das formas surgiu como uma alternativa. Koenderink (1984) e um pouco mais tarde Witkin (1987) propuseram técnicas para análise de sinais em vários níveis de resolução. Eles introduziram o conceito de fator de escala, que permite representar os atributos das formas em vários níveis de detalhes. Esse processo é baseado em uma função de transformação $F(n, \sigma)$, geralmente com características de filtragem passa-baixa 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.

A abordagem multiescala no estudo de formas torna a representação mais discriminativa que os métodos que empregam características monoescala globais, uma vez que a análise multiescala permite a descrição em vários níveis de detalhes e de abstração.

Inspirados nas técnicas de Koenderink (1984) e Witkin (1987), 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-baixa gaussiano 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.

Figura 8 – Contorno em múltiplas escalas de uma forma da classe de arraias da base Kimia-99.



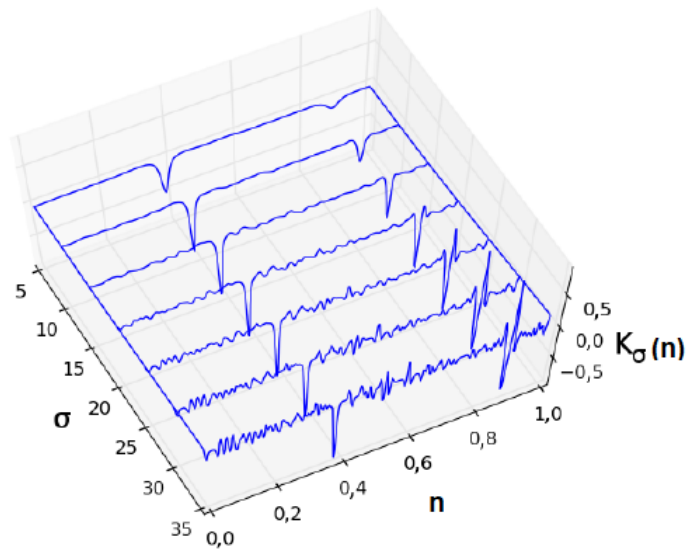
Fonte: (SOUZA, 2016)

A Figura 8 ilustra o efeito da suavização do contorno para diferentes valores de escala em uma das formas da Figura 5. Observa-se nesta figura que, no domínio da frequência, à medida que σ diminui, o grau de suavização 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.

O cálculo da curvatura, considerando agora a suavização do contorno, passa pela inclusão do fator de escala (σ) na Equação 2.6, ou seja:

$$\kappa_{\sigma}(n) = \frac{\dot{x}_{\sigma}(n)\ddot{y}_{\sigma}(n) - \ddot{x}_{\sigma}(n)\dot{y}_{\sigma}(n)}{\{[\dot{x}_{\sigma}(n)]^2 + [\dot{y}_{\sigma}(n)]^2\}^{\frac{3}{2}}}. \quad (2.7)$$

Figura 9 – Curvatura multiescala do contorno de uma forma de folha da Figura 6.



Fonte: (SOUZA, 2016)

A Figura 9 ilustra a evolução do sinal de curvatura do contorno da folha ilustrada na Figura 6 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 multiescala também é conhecida na literatura como curvograma (COSTA; CESAR JR, 2009).

2.2 Energia de Dobramento Multiescala Normalizada (NMBE)

A energia de dobramento, do inglês *bending energy*, expressa a quantidade de energia necessária para deformar um determinado contorno fechado a um contorno de menor energia equivalente, ou seja, transformá-lo em um círculo de perímetro igual ao do contorno original (COSTA; CESAR JR, 2009). Essa energia é calculada a partir da curvatura do objeto $k(n)$ do contorno discreto representado na forma complexa $u(n) = x(n) + jy(n)$, em que n representa a quantidade de pontos do contorno.

A curvatura discreta $k(n)$ pode então ser obtida por meio da expressão:

$$k(n) = \frac{-Im\{\dot{u}(n) \ddot{u}^*(n)\}}{|\dot{u}(n)|^3}, \quad (2.8)$$

em que Im é a parte imaginária do número complexo, $\dot{u}(n)$ e $\ddot{u}(n)$ representam, respectivamente,

a primeira e a segunda derivada do contorno da forma, $u^*(n)$ é o complexo conjugado de $u(n)$ e o operador $|\cdot|$ representa o módulo.

Essa dependência direta de derivar as coordenadas paramétricas do contorno, faz com que o sinal de curvatura seja contaminado, de forma amplificada, pelos ruídos inerentes aos processos de aquisição, amostragem e quantização dos contornos dos objetos das imagens, o que restringe o seu uso direto como descritor. Uma solução para isso é a suavização prévia do contorno $u(n)$, por meio de um filtro passa-baixas gaussiano (COSTA; CESAR JR, 2009):

$$g_{\sigma}(n) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-n^2}{2\sigma^2}\right). \quad (2.9)$$

Portanto, pode-se definir a versão filtrada de $u(n)$ por $g_{\sigma}(n)$ como:

$$\hat{u}(n, \sigma) = u(n) * g_{\sigma}(n), \quad (2.10)$$

em que o símbolo $*$ representa o operador de convolução.

O parâmetro σ permite controlar o nível de detalhamento do contorno na descrição da forma. Dessa forma, à medida em que ocorre a suavização, há uma diminuição do ruído, mas também temos a perda de informações em relação à forma inicial.

A primeira e segunda derivadas de $\hat{u}(n, \sigma)$ em relação a n são definidas como:

$$\hat{\dot{u}}(n, \sigma) = \dot{u}(n) * g_{\sigma}(n), \quad (2.11)$$

$$\hat{\ddot{u}}(n, \sigma) = \ddot{u}(n) * g_{\sigma}(n). \quad (2.12)$$

Em (COSTA; CESAR JR, 2009), foi proposta uma abordagem para calcular a curvatura multiescala operando com a representação do contorno no domínio da frequência, conforme adotado nesta dissertação. A partir da transformada de Fourier das coordenadas do contorno suavizado, os referidos autores calculam as derivadas utilizando a propriedade da derivada da transformada de Fourier, ou seja:

$$\hat{u}(n, \sigma) = \mathcal{F}^{-1}\{U(f)G_{\frac{1}{\sigma}}(f)\}, \quad (2.13)$$

$$\hat{\dot{u}}(n, \sigma) = \mathcal{F}^{-1}\{j2\pi fU(f)G_{\frac{1}{\sigma}}(f)\}, \quad (2.14)$$

$$\hat{\ddot{u}}(n, \sigma) = \mathcal{F}^{-1}\{-(2\pi f)^2U(f)G_{\frac{1}{\sigma}}(f)\}, \quad (2.15)$$

onde $U(f)$ e $G_{\frac{1}{\sigma}}(f)$ são definidas por:

$$U(f) = \mathcal{F}\{u(n)\}, \quad (2.16)$$

$$G_{\frac{1}{\sigma}}(f) = \mathcal{F}\{g_{\sigma}(n)\} = e^{-2\pi^2 f^2 \sigma^2}, \quad (2.17)$$

em que $\mathcal{F}\{\cdot\}$ e $\mathcal{F}^{-1}\{\cdot\}$ denotam as transformadas de Fourier e sua inversa, respectivamente.

O processo de filtragem passa-baixas reduz a energia espectral da representação complexa do contorno, causando o encolhimento do seu perímetro. Uma estratégia para compensar tal efeito é normalizar o contorno suavizado com a razão entre o perímetro do contorno original (L) e o seu perímetro suavizado (L_σ) (COSTA; CESAR JR, 2009), de acordo com a seguinte equação:

$$P_\sigma = \frac{L}{L_\sigma}, \quad (2.18)$$

sendo que L e L_σ são estimados por:

$$L = \frac{2\pi}{N} \sum_{n=0}^{N-1} |\dot{u}(n)|, \quad (2.19)$$

$$L_\sigma = \frac{2\pi}{N} \sum_{n=0}^{N-1} |\hat{u}(n, \sigma)|. \quad (2.20)$$

O contorno e suas derivadas normalizadas são dados, respectivamente, por:

$$u(n, \sigma) = \hat{u}(n, \sigma)P_\sigma, \quad (2.21)$$

$$\dot{u}(n, \sigma) = \hat{\dot{u}}(n, \sigma)P_\sigma, \quad (2.22)$$

$$\ddot{u}(n, \sigma) = \hat{\ddot{u}}(n, \sigma)P_\sigma. \quad (2.23)$$

Com as Equações 2.21, 2.22 e 2.23, a curvatura multiescala pode ser definida como:

$$k(n, \sigma) = \frac{-Im\{\dot{u}(n, \sigma) \ddot{u}^*(n, \sigma)\}}{|\dot{u}(n, \sigma)|^3}, \quad (2.24)$$

em que cada valor σ corresponde a uma escala, que permite a extração de diferentes detalhes da forma.

A partir de todas essas definições, a energia de dobramento (E) pode então ser representada. No caso discreto, temos:

$$E = \frac{1}{N} \sum_{n=0}^{N-1} k(n)^2, \quad (2.25)$$

em que N é o número de pontos do contorno.

Para que a energia de dobramento seja invariante à escala, faz-se necessário a sua normalização acrescentando um termo L^2 à Equação 2.25. A energia de dobramento normalizada (E_n) fica então definida por:

$$E_n = \frac{L^2}{N} \sum_{n=0}^{N-1} k(n)^2. \quad (2.26)$$

Por fim, utilizando a definição de curvatura multiescala da Equação 2.24, pode-se definir a energia de dobramento multiescala normalizada, do inglês *Normalized Multiscale Bending Energy (NMBE)* (COSTA; CESAR JR, 2009):

$$E_n(\sigma) = \frac{L^2}{N} \sum_{n=0}^{N-1} k(n, \sigma)^2, \quad (2.27)$$

onde $E_n(\sigma)$ é a energia de dobramento normalizada na escala σ .

$E_n(\sigma)$ apresenta valores com alta amplitude e elevada ordem de grandeza. A critério de ajuste, Souza *et al.* (2016) aplicaram a operação $\log(\cdot)$ nos diferentes valores de energia, redefinindo assim o descritor *NMBE* segundo a expressão:

$$NMBE = [\log E_n(\sigma_1), \log E_n(\sigma_2), \dots, \log E_n(\sigma_m)], \quad (2.28)$$

em que m indica o número de escalas e σ_i é o i -ésimo parâmetro de escala.

Uma questão importante é o critério para a escolha das escalas do descritor. Cesar Jr e Costa (1997) propuseram um método de ajuste para o descritor *NMBE* de acordo com as seguintes equações:

$$\sigma_l = \left[\left(\frac{\sigma_{min}^{-1} - \sigma_{max}^{-1}}{oct_{max} - \sqrt{2}} \right) (oct_l - \sqrt{2}) + \sigma_{max}^{-1} \right]^{-1}, \quad (2.29)$$

sendo que

$$oct_l = (\sqrt{2})^l; l = [1, 2, \dots, N_s]. \quad (2.30)$$

Nas Equações 2.29 e 2.30, o índice l indica a escala, e o número de escalas é definido como N_s . O termo *oct* indica o incremento em oitavas, enquanto σ_{min} e σ_{max} correspondem aos valores mínimos e máximos que um dado parâmetro de escala σ pode assumir. Por fim, o termo oct_{max} é o valor máximo que *oct* pode assumir.

As escalas obtidas por Cesar Jr e Costa (1997) adotam os seguintes parâmetros: $m = 32$, $\sigma_{min} = \frac{1}{64}$ e $\sigma_{max} = 2$, $oct_{min} = \sqrt{2}$ e $oct_{max} = (\sqrt{2})^{32}$. As cinco escalas utilizadas nesse trabalho são exibidas na Tabela 1:

Tabela 1 – Escalas utilizadas para o descritor *NMBE*.

Parâmetros	σ_1	σ_2	σ_3	σ_4	σ_5
Valor	0,4	0,654	1,1905	2,826	100

2.3 Conceitos de Teoria da Informação

A teoria matemática da informação estuda a codificação, armazenamento e comunicação da informação. Seus conceitos têm sido aplicados, recentemente e com sucesso, em variados problemas de visão computacional e reconhecimento de padrões. Isso porque algumas grandezas têm se mostrado promissoras no treinamento de sistemas adaptativos e de aprendizagem de máquina pelo fato de terem relação direta com a informação contidas nos objetos de estudo (SINGH; PRINCIPE, 2011). Dentre essas grandezas destaca-se uma medida chave em teoria da informação: a entropia.

2.3.1 Entropia de Shannon

A entropia de Shannon é uma medida do grau de incerteza que se tem, em média, dos possíveis estados que uma variável aleatória discreta pode assumir.

Seja X uma variável aleatória discreta, que pode assumir N diferentes valores $\{x_1, \dots, x_k, \dots, x_N\}$, na qual a probabilidade de ocorrência de cada valor seja $p_k = P(X = x_k)$, com $0 \leq p_k \leq 1$ e $\sum_{k=1}^N p_k = 1$. A entropia de Shannon de X pode ser definida então pela seguinte Equação:

$$H(X) = - \sum_{k=1}^N p_k \log p_k, \quad (2.31)$$

apresentando as seguintes propriedades (COVER; THOMAS, 2006):

- $0 \leq H(X) \leq \log N$;
- $H(X) = 0$, se $p_k = 1$ para um único valor de k e $p_k = 0$ para os demais valores de k ;
- $H(X) = \log N$, se $p_k = \frac{1}{N}$ para todos os valores de k .

A entropia de Shannon indica, em média, quanto de informação uma variável aleatória carrega. Quanto mais informativa uma variável aleatória é, maior o grau de novidade que esta apresenta, ou seja, menos previsível ela é.

2.3.2 Entropia diferencial

Para uma variável aleatória contínua X , com função densidade de probabilidade $p_X(x)$, a entropia diferencial é definida como (COVER; THOMAS, 2006):

$$h(X) = - \int_{-\infty}^{\infty} p_X(x) \log p_X(x) dx. \quad (2.32)$$

Essa definição é uma extensão do conceito da entropia discreta de Shannon para os casos de variáveis aleatórias contínuas. No entanto, diferente daquela, a entropia diferencial não mede o grau de incerteza da variável aleatória contínua X , pois tal incerteza é infinita, já que a quantidade de estados que um sistema contínuo pode abranger é infinito.

É possível demonstrar que $H(X)$ e $h(X)$ guardam uma relação entre si, pois considerando que uma variável aleatória contínua é o caso limite da variável aleatória discreta X , que assume valores $x_k = k\delta x$, $k = 0, \pm 1, \pm 2, \dots$, quando $\delta x \mapsto 0$, tem-se (COVER; THOMAS, 2006):

$$H(X) = h(X) - \lim_{\delta x \rightarrow 0} \log \delta x. \quad (2.33)$$

A Equação 2.33 evidencia que, à medida que $\delta x \mapsto 0$, a entropia de Shannon, $H(X)$ tende ao infinito.

A entropia diferencial apresenta a propriedade de ser invariante à translação da variável aleatória, ou seja:

$$h(X + a) = h(X). \quad (2.34)$$

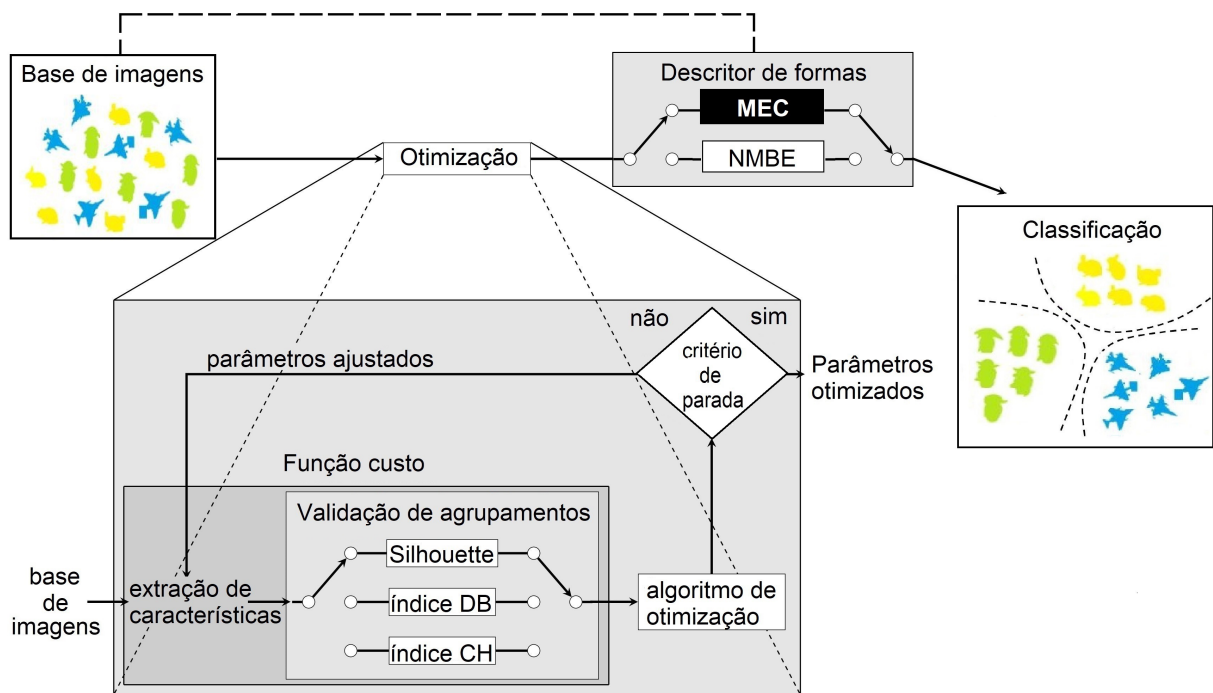
No caso de mudança da escala da variável aleatória por um fator α , temos a seguinte propriedade:

$$h(\alpha X) = h(X) + \log \alpha. \quad (2.35)$$

3 MATERIAIS E MÉTODOS

Este capítulo apresenta as ferramentas e definições necessárias para o desenvolvimento e a análise do descritor de formas proposto nesse trabalho, o *MEC*, destacado no retângulo preto da Figura 10, que ilustra as etapas da metodologia adotada.

Figura 10 – Metodologia geral para a descrição multiescala de formas, com ênfase na otimização.



O caminho tracejado identifica a descrição direta da forma, ou seja, sem otimização. Por outro lado, o caminho contínuo indica o procedimento de otimização, no qual o algoritmo seleciona aleatoriamente um conjunto de escalas para descrever as formas. Em seguida, extraem-se as características das formas da base e calcula-se a função custo. Esse processo é repetido até que se atinja um critério de parada, neste caso, o número de rodadas. A função custo guia o algoritmo na busca de um conjunto de parâmetros ótimos para o descritor de formas que maximize a separação interclasse e minimize a separação intraclasse. Quando o critério de parada é atingido, os parâmetros (escalas do descritor) obtidos são utilizados para uma nova etapa de descrição da base de imagens, gerando uma base de assinaturas, a qual é submetida a experimentos de classificação.

Nas próximas seções, as etapas mostradas na Figura 10 serão apresentadas e co-

mentadas, com exceção do descritor de formas *NMBE*, o qual foi detalhado no Capítulo 2. A condução dos experimentos será discutida na última seção deste capítulo.

3.1 Bases de formas

No estudo de análise de formas, Dryden e Mardia (2016) definem que se efeitos de rotação, de translação e/ou de escala forem aplicados a uma imagem ou objeto e todas as informações geométricas se mantiverem, podemos conceituar essa estrutura como uma forma. Dessa forma, diferentes transformações aplicadas a um mesmo objeto produziriam imagens diferentes, mas preservando a forma. Desse conceito surge a importância das propriedades de invariância do descritor multiescala proposto, que serão discutidas na Subseção 3.2.1.

Nesse trabalho foram utilizadas três bases públicas de formas rotuladas, sendo duas binárias sintéticas e outra de formas de folhas de plantas. Esta última tem sido utilizada em diversos trabalhos da literatura que desenvolvem aplicações de visão computacional no reconhecimento automático de espécies de plantas (WANG *et al.*, 2015; HALL *et al.*, 2015; QUADRI; SIRSHAR, 2015; CHAKI *et al.*, 2015).

A base Kimia-99 (Figura 11) é composta por 99 formas distribuídas uniformemente em 9 classes. A base MPEG7-CE (Figura 12) contém 1400 formas distribuídas uniformemente em 70 classes. Por fim, a base Flavia (Figura 13) é composta por 1907 formas de folhas de 32 espécies distintas, sendo esta base bastante desafiadora pois, além de ser desbalanceada (apresentar classes com quantidades distintas de folhas), essa apresenta espécies com grande variabilidade intra classe e pequena variabilidade entre classes.

É importante também salientar que a base Flavia, por se tratar de uma base de folhas reais, não apresenta suas formas binarizadas por padrão. Portanto, é necessária uma etapa de pré-processamento antes da extração dos contornos, pois aspectos de textura (como as nervuras evidentes em algumas folhas), poderiam atrapalhar na detecção das bordas. Assim, as imagens das folhas precisam ser previamente segmentadas por limiarização antes do processo de extração dos contornos (rever Figura 6).

Figura 11 – Todas as formas da base Kimia-99.

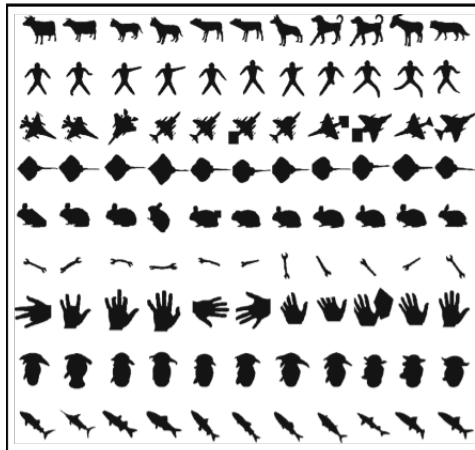


Figura 12 – Uma amostra por classe da base MPEG7-CE.

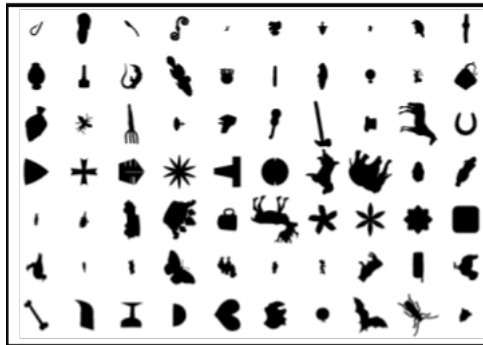


Figura 13 – Uma amostra por classe da base Flavia.



3.2 Entropia como descritor multiescala

Na Subseção 2.1.3 foi mostrada a utilização da curvatura multiescala como uma assinatura do contorno de uma forma. Em seguida, na Seção 2.2, apresentamos um descritor multiescala construído a partir da versão discreta dessa assinatura, mais especificamente a

Energia de Dobramento Multiescala.

Nesta seção propomos um novo descritor multiescala a partir da mesma assinatura, porém explorando os conceitos de entropia diferencial, discutidos anteriormente na Subseção 2.3.2.

3.2.1 Entropia da Curvatura Multiescala (MEC)

O cálculo da entropia diferencial da curvatura multiescala, através da Equação 2.32, requer a estimativa das funções de densidade de probabilidade do sinal de curvatura em diferentes escalas. Tal estimativa pode ser realizada através de uma técnica não paramétrica conhecida como janela de Parzen (WEBB, 2002), que estima uma distribuição de probabilidade diretamente das amostras de dados.

Para uma dada escala σ , temos a curvatura amostrada em N valores como sendo $\kappa[\sigma, n] = (\kappa[\sigma, 1], \kappa[\sigma, 2], \dots, \kappa[\sigma, N])$. A função de densidade de probabilidade estimada da curvatura multiescala, $p_\sigma(\kappa)$, na escala σ e para um dado valor de curvatura κ , através da janela de Parzen, é dada por (WEBB, 2002):

$$p_\sigma(\kappa) = \frac{1}{bN} \sum_{i=1}^N \Psi \left(\frac{\kappa - \kappa[\sigma, i]}{b} \right), \quad (3.1)$$

onde

$$\Psi \left(\frac{\kappa - \kappa[\sigma, i]}{b} \right) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{(\kappa - \kappa[\sigma, i])^2}{2b^2} \right), \quad (3.2)$$

é uma janela gaussiana e b é a sua largura de banda.

No processo de estimação pela janela de Parzen, a escolha do parâmetro largura de banda é crítica, pois esse altera o grau de suavização da curva de densidade obtida. Tendo como referência o desvio padrão entre as amostras de curvatura, valores elevados desse parâmetro tendem a suavizar em demasia a curva de densidade estimada, acarretando perda de detalhes importantes. Por outro lado, valores pequenos resultam em densidades estimadas com mais detalhes, porém mais ruidosa. Assim, é importante estabelecer um critério para escolha do parâmetro de largura de banda de modo que se tenha uma boa estimativa da função de densidade de probabilidade. Adotamos neste trabalho o método proposto por Silverman e Green (1986). Este método determina uma escolha ótima de largura de banda, assume que a distribuição de probabilidade dos dados é normal e que a janela utilizada é gaussiana.

A largura de banda por este método é dada por (SILVERMAN; GREEN, 1986):

$$b = \left(\frac{4\hat{\sigma}^5}{3N} \right)^{\frac{1}{5}}, \quad (3.3)$$

sendo $\hat{\sigma}$ o desvio padrão das amostras de curvatura e N o número de amostras envolvidas na estimação.

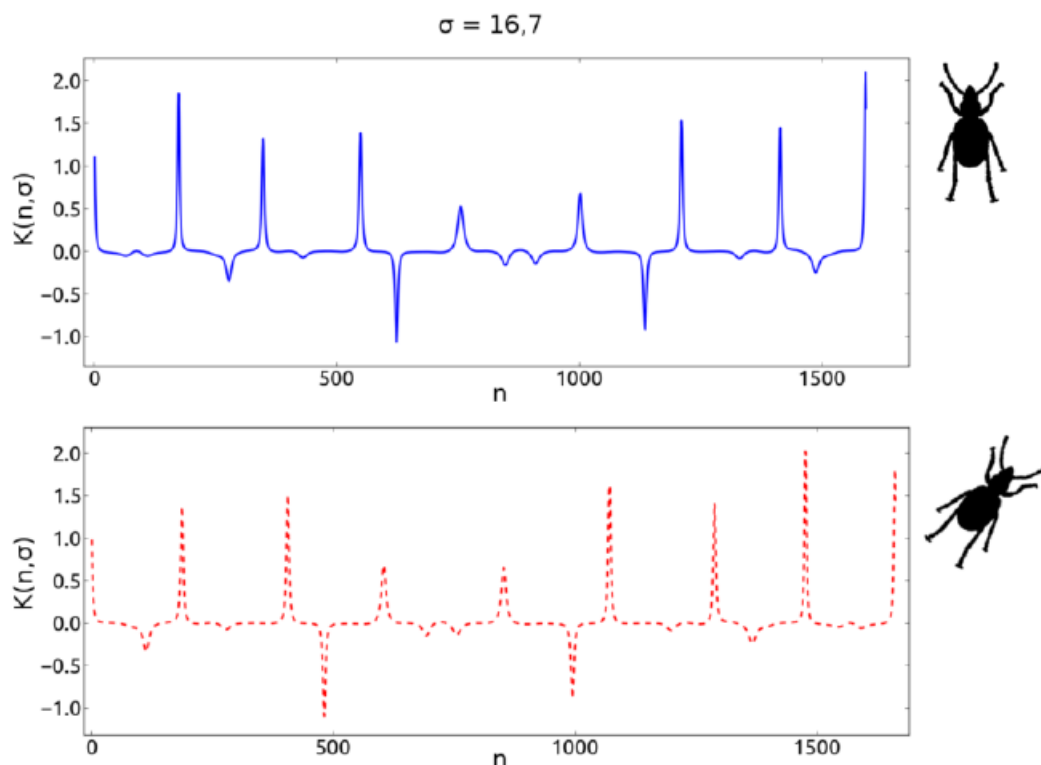
Então, a entropia diferencial da curvatura κ multiescala na escala σ , aplicando o resultado obtido na Equação 3.1 na Equação 2.32, pode ser aproximada por:

$$\hat{h}_{\sigma}(\kappa) = - \int_{\kappa_{min}}^{\kappa_{max}} p_{\sigma}(\kappa) \log p_{\sigma}(\kappa) d\kappa, \quad (3.4)$$

sendo a integral calculada através do método integração de Simpson (SPERANDIO *et al.*, 2003) e os limites de integração $\kappa_{min} = E\{\kappa[\sigma, n]\} - 5\hat{\sigma}$ e $\kappa_{max} = E\{\kappa[\sigma, n]\} + 5\hat{\sigma}$, onde $E\{\cdot\}$ representam o valor esperado e $\pm 5\hat{\sigma}$ seria para garantir a abrangência de mais de 99% da área da curva da distribuição gaussiana.

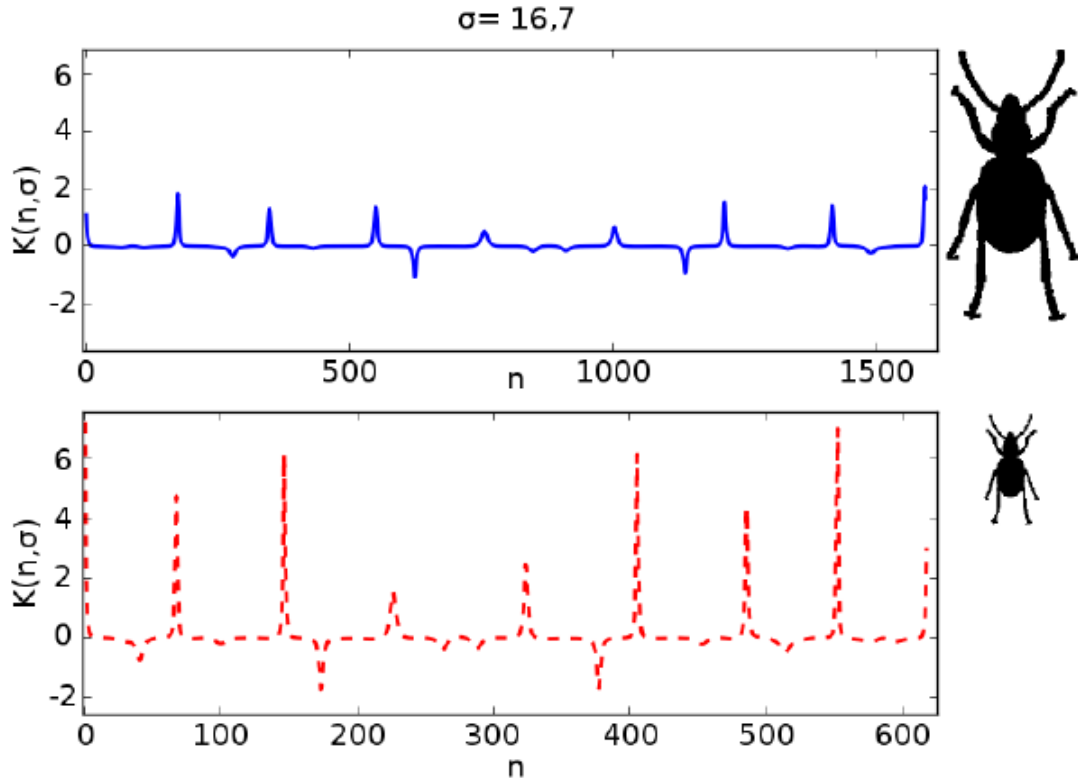
O descritor obtido a partir da Equação 3.4 apresenta propriedades de invariância à translação e à rotação. Isso acontece porque a translação da forma não afeta o sinal da curvatura. Embora a rotação desloque em fase o sinal da curvatura, como ilustra a Figura 14, esse deslocamento não interfere na função densidade de probabilidade estimada, e portanto não interfere no cálculo do descritor.

Figura 14 – Efeito da rotação de uma forma sobre a curvatura.



Porém, o descritor não apresenta invariância à mudança de escala da forma. Isso porque tal mudança reflete efetivamente no sinal de curvatura (Figura 15), o que interfere na função de densidade de probabilidade estimada, e por conseguinte no cálculo do descritor.

Figura 15 – Efeito da mudança de escala de uma forma sobre a curvatura.



No entanto, é possível conseguir a invariância do descritor à mudança de escala utilizando o perímetro da forma (L_σ) como fator de normalização. Assim, aplicando a propriedade da Equação 2.35, temos a versão do descritor invariante à mudança de escala:

$$\hat{h}_\sigma(\kappa) = - \int_{\kappa_{min}}^{\kappa_{max}} p_\sigma(\kappa) \log p_\sigma(\kappa) d\kappa + \log L_\sigma. \quad (3.5)$$

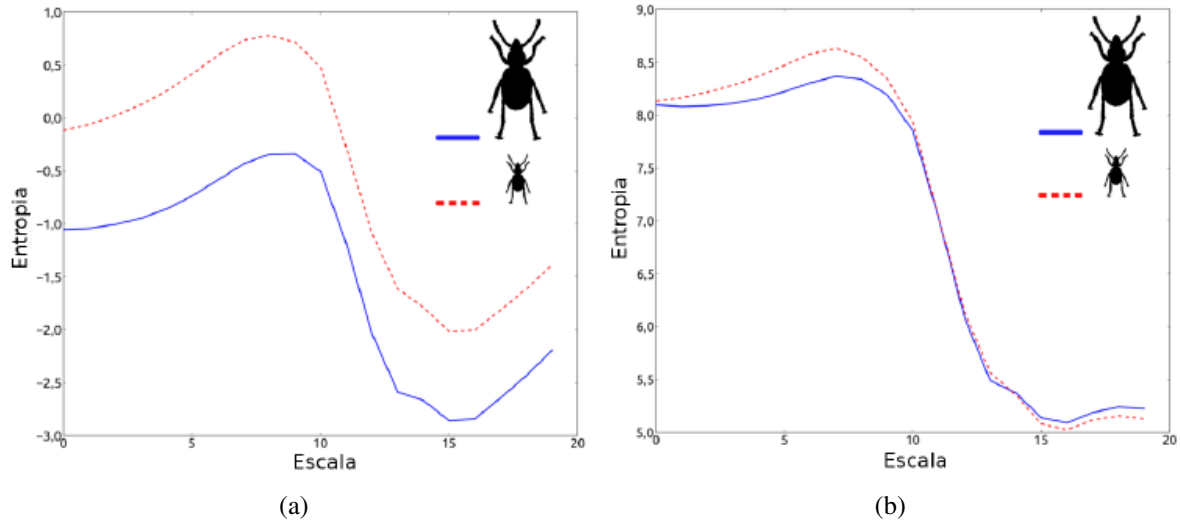
O descritor entropia da curvatura multiescala, do inglês *multiscale entropy of curvature* (MEC), para um dado vetor discreto de escalas ($\sigma_1 \ \sigma_2 \ \dots \ \sigma_N$) é então definido, a partir da Equação 3.5, como:

$$MEC = \left(\hat{h}_{\sigma_1}(\kappa) \ \hat{h}_{\sigma_2}(\kappa) \ \dots \ \hat{h}_{\sigma_N}(\kappa) \right). \quad (3.6)$$

Será ilustrado a seguir o descritor entropia diferencial da curvatura multiescala, calculado para uma mesma forma, mas em escalas diferentes. Na Figura 16a, o cálculo foi realizado a partir da Equação 3.4 e o descritor não apresentou invariância à escala da forma. Já na

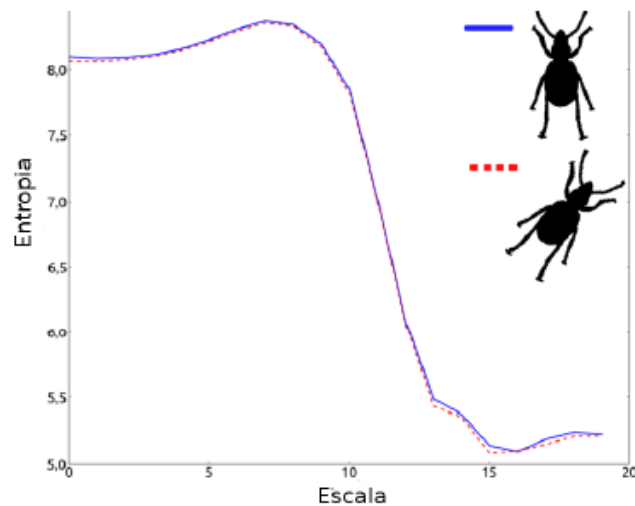
Figura 16b, o cálculo foi realizado a partir da Equação 3.5 e o descritor tornou-se parcialmente invariante, com erros provenientes da amostragem do contorno.

Figura 16 – Efeito da mudança de escala de uma forma sobre o descritor. (a) Sem o fator de normalização. (b) Com o fator de normalização.



A Figura 17 ilustra a invariância do descritor à rotação da forma. Nesse caso, o cálculo já foi realizado a partir da Equação 3.5.

Figura 17 – Efeito da rotação de uma forma sobre o descritor.



3.3 Algoritmo de otimização

Souza *et al.* (2016) foram pioneiros no processo de otimização do descritor *NMBE* em descrição de formas. Nesse processo, os referidos autores avaliaram três algoritmos: o *Simulated Annealing (SA)*, a *Evolução Diferencial (DE)* (STORN; PRICE, 1997) e a otimização

por enxame de partículas (*PSO*) (KENNEDY, 1995), mostrando que o *SA* e o *DE* eram os mais promissores para tal tarefa. Além disso, os autores concluíram que o *SA* tem um custo computacional significativamente menor que o *DE*.

Seguindo essa abordagem, a otimização do descritor *MEC* será realizada pelo *SA*, estabelecendo uma análise comparativa justa do descritor proposto com o *NMBE*, otimizados ambos pelo mesmo algoritmo.

3.3.1 *Simulated Annealing (SA)*

A expressão *annealing* remonta ao processo utilizado para fundição de um metal, em que este é aquecido a uma temperatura elevada e em seguida é resfriado lentamente, de modo que o produto final seja uma massa homogênea. O algoritmo denominado *Simulated Annealing* surgiu no contexto da mecânica estatística, desenvolvido por Kirkpatrick *et al.* (1983) e independentemente por Černý (1985), utilizando o algoritmo de Metropolis (METROPOLIS *et al.*, 1953).

Considerado um problema de minimização, o mecanismo desse algoritmo para exploração do espaço de busca consiste numa variável de temperatura que determina a probabilidade de aceitação de uma solução pior do que aquela encontrada até o momento. Soluções melhores são sempre aceitas, enquanto que a probabilidade de aceitar soluções piores decresce exponencialmente à medida que ocorrem as iterações.

Definindo a solução atual da t -ésima época do *SA* como $s_a(t)$, a probabilidade de aceitação de $s_c(t)$ ao invés de $s_a(t)$ é dada por:

$$P(t) = \begin{cases} 1 & \text{se } f(s_c(t)) < f(s_a(t)) \\ e^{-\frac{f(s_c(t)) - f(s_a(t))}{\theta T(t)}} & \text{caso contrário} \end{cases} \quad (3.7)$$

em que $f(\cdot)$ corresponde à função custo, θ é uma constante e $T(t)$ é a temperatura do sistema na t -ésima época. Abaixo, o Algoritmo 1 descreve um pseudocódigo da implementação do método de otimização *SA*.

Algoritmo 1: Simulated Annealing (SA)

```

1 Criar uma solução inicial,  $s_a(0)$ ;
2 Calcular a função custo da solução inicial,  $f(s_a(0))$ ;
3 Definir  $t = 1$ ;
4 Definir o valor inicial da temperatura,  $T(1)$ ;
5 repeat
6    $contador = 1$ ;
7   while  $contador \leq L_e$  do
8     Gera solução candidata,  $s_c(t)$ ;
9     Calcula a função custo da solução candidata,  $f(s_c(t))$ ;
10    Calcula a probabilidade de aceitação usando a Equação 3.7;
11    if  $rand\{[0,1]\} \leq \textit{probabilidade de aceitação}$  then
12      | Solução candidata passa a ser a atual,  $s_a(t) = s_c(t)$ ;
13    else
14      |  $s_a(t) = s_a(t - 1)$ ;
15    end
16     $contador = contador + 1$ ;
17  end
18   $t = t + 1$ ;
19  Atualiza a temperatura  $T(t)$ 
20 until critério de parada ser atingido;
21 return  $s_a(t)$ 

```

Fonte: adaptado de (ENGELBRECHT, 2007).

No algoritmo 1, temos que a função $rand\{[0,1]\}$ retorna uma amostra de uma distribuição uniforme no intervalo $[0,1)$. O parâmetro L_e consiste no número de vezes que a solução atual é testada por diferentes soluções candidatas em uma determinada época. Como critério de parada, pode-se estabelecer o número de épocas (N_e), onde uma época é o passo do algoritmo SA na qual a temperatura é atualizada.

Segundo Engelbrecht (2007), o *Simulated Annealing* requer também a escolha de mais dois critérios para a execução do algoritmo:

- Um critério de perturbação da solução atual para geração das soluções candidatas. Uma

possível estratégia é adicionar um vetor aleatório da seguinte forma:

$$s_c(t) = s_a(t) + r(t)D(t), \quad (3.8)$$

em que s_c é o vetor linha que indica a solução candidata e s_a é o vetor linha que corresponde à solução atual. O índice t denota a t -ésima época, $D(t)$ é uma matriz diagonal que define o máximo de mudança permitida para cada parâmetro que se deseja otimizar e $r(t)$ é um vetor aleatório de mesma dimensão da solução com elementos uniformemente distribuídos no intervalo de $[-1,1]$;

- Um critério de resfriamento para diminuição da temperatura ao longo das épocas, sendo a temperatura inicial definida previamente. Uma estratégia de resfriamento é dada por:

$$T(t+1) = \alpha T(t), \quad (3.9)$$

sendo α a taxa de resfriamento ($\alpha \in [0, 1]$) e t a época atual.

Para os experimentos de otimização realizados nesse trabalho foi utilizado o Algoritmo 1, adotando o critério de resfriamento conforme a Equação 3.9 e o número de épocas (N_e) como critério de parada. Os parâmetros foram escolhidos empiricamente: $\theta = 1$ como constante para a Equação 3.7, a temperatura inicial ($T(1) = 100$), a taxa de resfriamento ($\alpha = 0,9$), o número de testes por época ($L_e = 10$) e o número de épocas ($N_e = 200$). Se a solução candidata ultrapassar os limites do espaço de busca, essa receberá o valor correspondente ao limite que ultrapassou. Como critério de perturbação da solução atual para geração de soluções candidatas, utilizou-se a Equação 3.8.

3.4 Função custo para validação de agrupamentos

Em (CARNEIRO *et al.*, 2019) três funções custo foram avaliadas na metodologia de otimização proposta para medir a qualidade de agrupamentos nos experimentos: a *Silhouette* (*SI*), o índice *Calinski-Harabasz* (*CH*) e o índice *Davies–Bouldin* (*DB*). Embora a complexidade computacional da *SI* seja da ordem de $O(\eta^2)$, em comparação com os demais índices que são $O(\eta)$, sendo η o número de formas da base (SAITTA *et al.*, 2008), os resultados de classificação se mostraram superiores quando a *Silhouette* foi utilizada como função custo na otimização de parâmetros.

Com base nessas evidências, *SI* foi a escolhida como função custo para validar os melhores agrupamentos nos experimentos de classificação deste trabalho, sendo descrita a seguir.

3.4.1 *Silhouette (SI)*

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 atribuído a uma dada classe. Ela é definida como

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \in [-1, 1], \quad (3.10)$$

sendo a_i a dissimilaridade média entre o objeto i e aqueles que pertencem à sua classe e b_i é a dissimilaridade média entre o objeto i e aqueles da classe vizinha mais próxima, excluindo a classe a que i pertence.

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. Valores positivos indicam alta afinidade de um objeto à classe. Um valor de *Silhouette* próximo a zero indica que o objeto está na fronteira entre duas classes e que há, portanto, um grau de incerteza a respeito de qual classe este pertence.

Uma vez definida a função custo, o processo de otimização dos parâmetros do descritor visa ajustá-los ao problema em estudo (CARNEIRO *et al.*, 2017). A ideia é encontrar o melhor vetor de escalas $\sigma_{otimo} = (\sigma_1 \ \sigma_2 \ \sigma_3 \ \sigma_4 \ \sigma_5)$ para os descritores *MEC* e *NMBE* que maximize a Equação 3.10. Utilizando-se as escalas ótimas encontradas é possível então realizar, com o descritor multiescala, a extração de características para toda uma base de formas e, em seguida, realizar o experimento de classificação.

3.5 Algoritmo de classificação KNN

O algoritmo k vizinhos mais próximos, do inglês *K-Nearest Neighbors* (KNN) é um classificador não paramétrico, comumente utilizado em situações onde a superfície de decisão entre classes é irregular (FUKUNAGA, 2013). Por não ser paramétrico, este classificador não constrói um modelo interno geral utilizando o conjunto de treinamento, mas simplesmente compara o padrão de teste com as amostras de treinamento (RAJAGURU; PRABHAKAR, 2017). 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, com tal k definido previamente.

Geralmente, valores ímpares são utilizados para k a fim de evitar empates na votação (RUSSELL; NORVIG, 2009). Se k for muito pequeno, a classificação fica sensível a pontos de ruído, mas se k for muito grande, a vizinhança pode incluir elementos de outras classes. Então o classificador KNN foi utilizado nos experimentos deste trabalho com $k = 3$ e a distância euclidiana como métrica, que é a mais comumente utilizada.

3.6 Métricas de avaliação da classificação

Nesta seção descrevemos as duas métricas que são a base para a avaliação quantitativa de desempenho do descritor nos experimentos de classificação realizados. É importante ressaltar que além dessas, uma matriz visual de distâncias para análise qualitativa também estará presente, a fim de comparar resultados da distância euclidiana entre as formas na avaliação da capacidade discriminativa de cada descritor.

F-score

Na avaliação de desempenho de classificadores binários, a métrica *F-score* ou *F-measure* é definida pela média harmônica entre a precisão (do inglês *precision*) e a revocação (do inglês *recall*). A precisão (Equação 3.11) é uma métrica relacionada aos objetos classificados como relevantes, quantos são realmente relevantes. Já a revocação (Equação 3.12) mede quantos objetos relevantes foram selecionados. Matematicamente, temos as expressões destas dadas como:

$$\text{precisão} = \frac{\text{verdadeiros positivos}}{\text{verdadeiros positivos} + \text{falsos positivos}}. \quad (3.11)$$

$$\text{revocação} = \frac{\text{verdadeiros positivos}}{\text{verdadeiros positivos} + \text{falsos negativos}}. \quad (3.12)$$

Para uma determinada classe, os verdadeiros positivos representam o número de formas corretamente rotuladas pelo classificador como pertencentes a uma classe em particular. Os falsos positivos dizem respeito ao número de formas rotuladas como pertencentes a uma classe a qual não pertencem, e os falsos negativos correspondem ao número de formas que pertencem a uma classe específica, mas foram rotuladas como pertencente a outra.

A partir desses conceitos, a métrica *F-score* pode ser definida na Equação 3.13 por:

$$F\text{-score} = \frac{2 \cdot \text{precisão} \cdot \text{revocação}}{\text{precisão} + \text{revocação}}. \quad (3.13)$$

O valor mais alto possível da *F-score* é 1, indicando 100% de precisão e de revocação, e o valor mais baixo é 0, se a precisão ou a revocação for zero.

Acurácia

A Acurácia de um experimento de classificação é definida pela razão entre a quantidade de elementos de uma determinada base de dados que tiveram a classe corretamente predita e o número total dos elementos. Assim a Acurácia é definida pela equação:

$$\text{Acurácia} = \frac{\text{verdadeiros positivos} + \text{verdadeiros negativos}}{\text{total de amostras}}. \quad (3.14)$$

É importante ressaltar que o valor da Acurácia está no intervalo [0,1], sendo o valor 1 a indicação da taxa de acerto de 100%.

3.7 Experimentos

Todos os experimentos apresentados nesta dissertação foram implementados em *Python* e executados em uma máquina Intel Core i7-6800K, com processador de 3,40 GHz e 64GB de memória.

Com o propósito de avaliar e comparar a metodologia proposta para extração de características, este trabalho conduz os experimentos de classificação a partir de formas sintéticas e reais, ilustradas na Seção 3.1. Os experimentos foram divididos em duas etapas: a primeira, com escalas pré-definidas e a segunda, com as escalas obtidas por otimização.

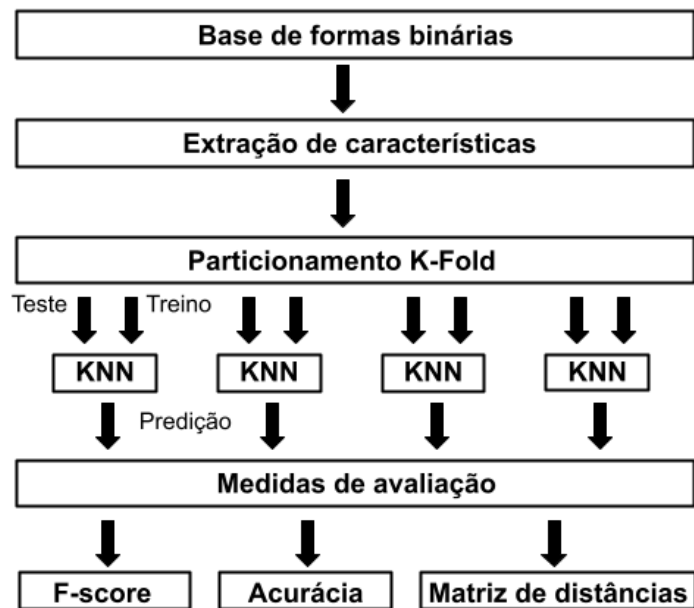
Na primeira etapa, as formas foram caracterizadas pelos descritores multiescala (*MEC* e *NMBE*) ajustando as escalas conforme a recomendação proposta em (CESAR JR; COSTA, 1997). Desta forma, foram adotadas as cinco escalas de suavização do contorno da Tabela 1 para ambos descritores. Foram então comparadas as habilidade de discriminação do *MEC* frente ao *NMBE*, e de ambos para os descritores monoescala descritos na Subseção 2.1.1.

Na segunda etapa, a proposta é comparar os descritores multiescala em tarefas de classificação, após ambos terem suas escalas otimizadas pela função custo *Silhouette*. Uma vez empregada no algoritmo de otimização descrito na Seção 3.3, essa função permite encontrar um conjunto de cinco escalas ótimas para os descritores, a partir de um certo espaço de busca (variando de 0,4 a 100) e de um determinado número de execuções (200 vezes).

Esse processo de otimização é repetido por 15 vezes para a base Kimia-99 e por 5 vezes para as demais bases, uma vez que estas são numericamente maiores e, por consequência,

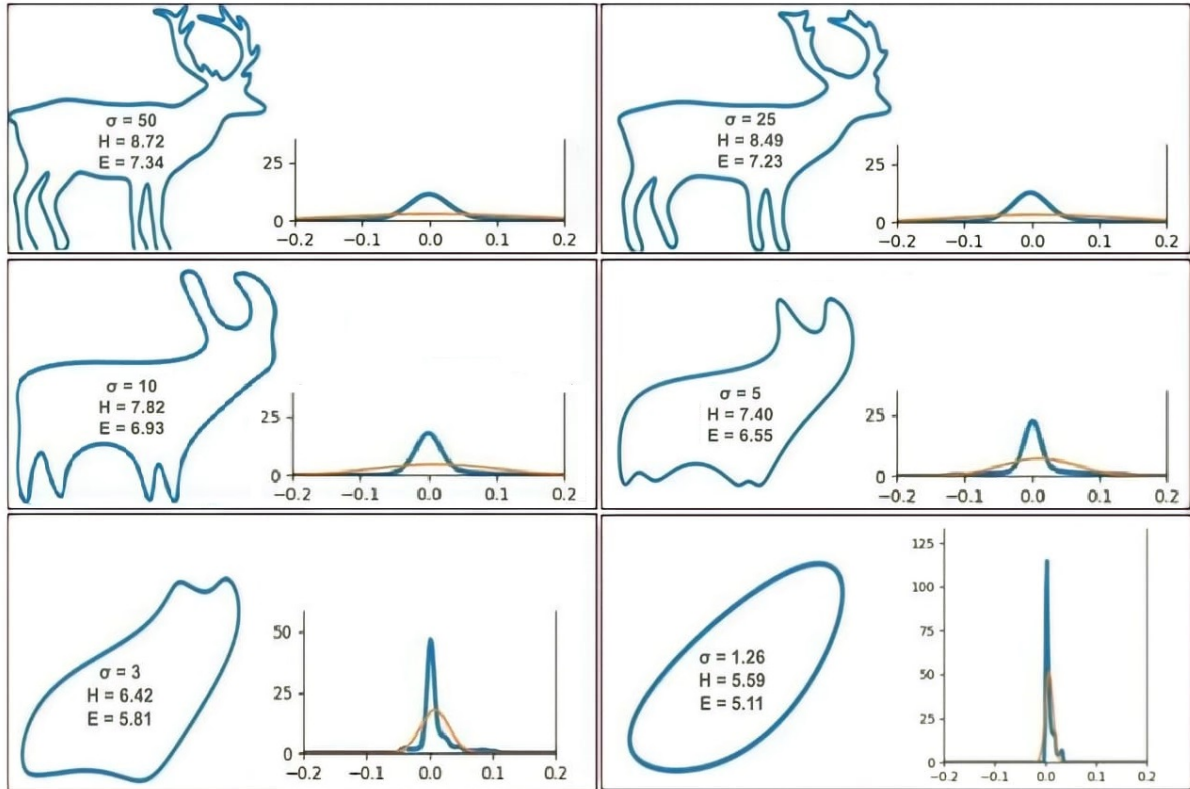
mais custosas na otimização. De posse do conjunto de escalas ótimas, as formas são descritas escolhendo-se o descritor multiescala e classificadas pelo KNN em um experimento de validação cruzada com a técnica de particionamento *K-Fold*. Na classificação, os descritores obtidos no procedimento anterior são divididos em 4 grupos de amostras treino/teste. A Figura 18 mostra o *pipeline* empregado nesses experimentos:

Figura 18 – Experimento de validação cruzada realizado para avaliar o desempenho dos descritores na classificação de formas.



A Figura 19 ilustra o processo de descrição pelo *MEC* e pelo *NMBE* a partir de uma determinada forma da base MPEG7-CE.

Figura 19 – Exemplo de aplicação dos descritores multiescala em uma forma da base MPEG7-CE.



Cada quadrante ilustra uma dada escala de suavização do contorno, representada pelo desvio-padrão σ . Vale lembrar que toda a formulação está no domínio da frequência, por isso o princípio de deformação do contorno se inverte: menor desvio significa mais deformação.

A partir de um dado σ , seguem os respectivos valores de H (entropia) e de E (energia), com o respectivo escopo da função de densidade de probabilidade estimada para o MEC.

4 RESULTADOS

Neste capítulo, são apresentados os resultados obtidos com o descritor multiescala *MEC*, em experimentos de classificação.

Na Seção 4.1 temos os resultados sem as escalas otimizadas, sendo o desempenho obtido comparado ao do descritor *NMBE*, e também ao descritor monoescala.

Na Seção 4.2 estão os resultados com otimização de escalas, aonde as escalas do *MEC* são otimizadas e o desempenho em classificação do mesmo comparado ao do descritor *NMBE* otimizado. Será também avaliado nesta seção o ganho de desempenho com a concatenação dos atributos dos dois descritores multiescala.

4.1 Experimentos de classificação sem otimização

Nesta seção, os resultados são exibidos e avaliados tomando como premissa a não otimização dos descritores multiescala, ou seja, os vetores de características das formas em estudo serão calculados a partir das mesmas cinco escalas pré-definidas na Tabela 1, tanto para o *MEC* (vetores de entropias) quanto para o *NMBE* (vetores de energias). A título de comparação de desempenho entre multiescala e monoescala, utilizamos também um conjunto de quatro características monoescala para construir o descritor *MONO*, a saber: circularidade, excentricidade, razão de aspecto e solidez, todas elas devidamente descritas anteriormente na Subseção 2.1.1.

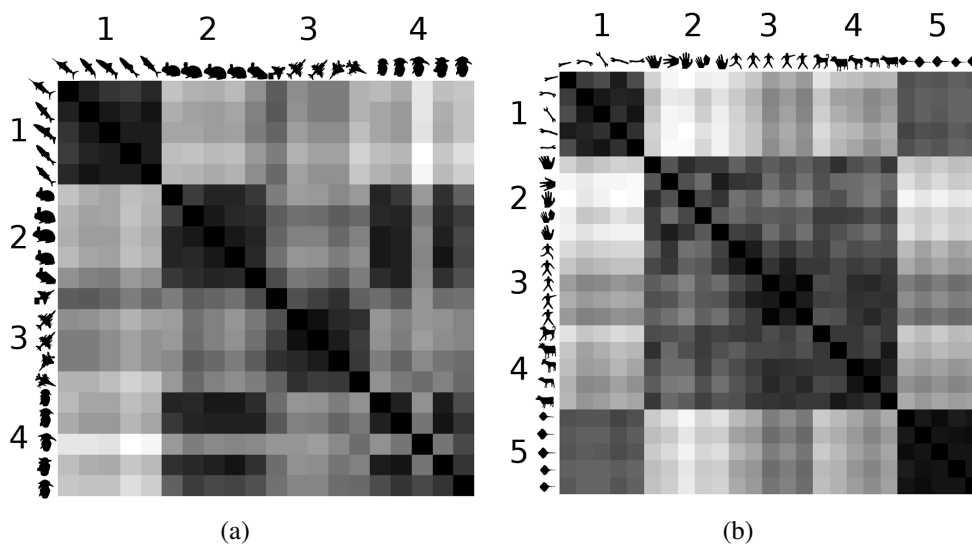
As matrizes visuais em escala de cinza na Figura 20 representam as matrizes de distância euclidiana entre pares a partir do descritor *MEC* em diferentes classes de formas da base *Kimia-99*. As áreas mais escuras nessas matrizes referem-se às distâncias menores, enquanto que as áreas mais claras correspondem às distâncias maiores. Essas ilustrações permitem avaliar a capacidade de discriminação do descritor multiescala qualitativamente.

Observamos nesta figura que as distâncias entre descritores de forma pertencentes à mesma classe são menores do que aquelas entre descritores de forma de classes distintas. O *MEC*, portanto, fornece baixa variabilidade intraclasse (presença de quadrados escuros na diagonal principal) e alta separabilidade interclasse para as formas em análise. No entanto, também identifica-se que o *MEC* não foi capaz de discriminar as formas que representam as classes de coelhos e de alienígenas, ambas mostradas na Figura 20a como classes 2 e 4, respectivamente. Os quadrados que constituem a intersecção dessas classes (linha 2, coluna 4 e linha 4, coluna

2) são predominantemente escuras, indicando que as distâncias entre as formas nessas classes distintas são pequenas. Assim, as formas em ambas as classes são realmente semelhantes e o descritor não foi capaz de discriminar corretamente as mesmas.

Na Figura 20b, identifica-se que as diferenças entre as formas nas classes de mãos, humanos e animais quadrúpedes (classes 2, 3 e 4, respectivamente) não são evidentes. Logo, o descritor não foi capaz de lidar com a variabilidade dentro dessas classes de formas, especialmente na de quadrúpedes (classe 4). No entanto, o *MEC* efetivamente discriminou objetos das classes de ferramentas e de arraias (classes 1 e 5, respectivamente).

Figura 20 – Matrizes de distância euclidiana par-a-par nas classes da base Kimia-99 a partir do descritor *MEC*.



As Tabelas 2 e 3 resumem os resultados quantitativos de classificação para as três bases de formas. O desempenho dos descritores multiescala foi semelhante para as bases Kimia-99 e Flavia. Entretanto, o descritor monoescala apresentou um ganho de aproximadamente 4% em relação aos multiescala, em termos de precisão e de revocação para a base Kimia-99. Entendemos que, por considerar aspectos geométricos locais das formas, o descritor monoescala é mais apto para discriminar formas da base Kimia-99, pois esta última apresenta alta variabilidade de formas entre classes e baixa variabilidade de formas intra classes. Em negrito nas tabelas temos o destaque dos resultados em que o *MEC* teve desempenho superior aos demais descritores.

Tabela 2 – Resultados (sem otimização) de precisão (P) e de revocação (R) para os testes de validação cruzada com *NMBE*, *MEC* e *MONO* para as três bases de formas.

Base de formas	P			R		
	<i>NMBE</i>	<i>MEC</i>	<i>MONO</i>	<i>NMBE</i>	<i>MEC</i>	<i>MONO</i>
Kimia-99	78,33 ± 9,46	78,37 ± 8,31	82,26 ± 8,76	75,47 ± 6,75	72,67 ± 6,62	76,99 ± 6,20
MPEG7-CE	62,67 ± 2,48	66,01 ± 2,60	63,08 ± 2,56	60,38 ± 2,10	63,05 ± 2,33	59,87 ± 1,92
Flavia	73,48 ± 1,65	76,17 ± 1,50	55,83 ± 1,89	72,60 ± 1,47	74,82 ± 1,55	52,81 ± 1,53

Tabela 3 – Resultados (sem otimização) de *F-score* e de Acurácia para os testes de validação cruzada com *NMBE*, *MEC* e *MONO* para as três bases de formas.

Base de formas	<i>F-score</i>			Acurácia		
	<i>NMBE</i>	<i>MEC</i>	<i>MONO</i>	<i>NMBE</i>	<i>MEC</i>	<i>MONO</i>
Kimia-99	74,78 ± 8,53	72,45 ± 8,26	77,57 ± 6,89	75.18 ± 8.32	73.03 ± 8.30	77.39 ± 6.81
MPEG7-CE	58,55 ± 2,25	61,61 ± 2,56	57,83 ± 2,18	59.09 ± 2.23	61.75 ± 2.54	57.73 ± 2.19
Flavia	72,09 ± 1,56	74,63 ± 1,44	52,56 ± 1,79	72.18 ± 1.58	75.07 ± 1.32	53.02 ± 1.78

Por outro lado, o descritor monoescala não foi capaz de discriminar classes de formas da base Flavia (ver Figura 21c). Isso se deve novamente à geometria local, uma vez que essa base apresenta características de forma que são semelhantes para folhas de diferentes espécies de plantas (baixa variabilidade entre classes). Além disso, tais atributos geométricos podem variar significativamente para as folhas da mesma espécie de planta (alta variabilidade intraclasse).

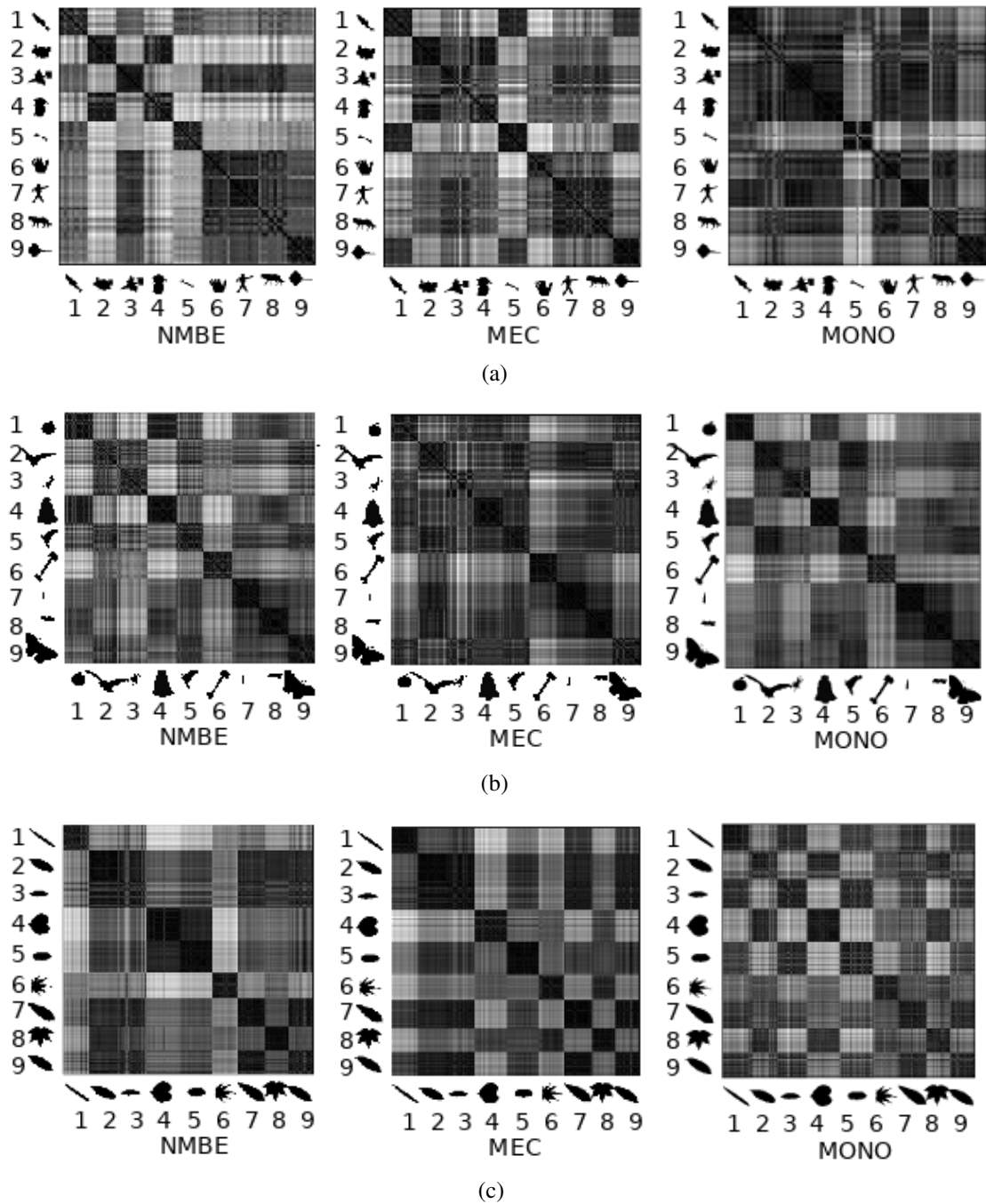
Mesmo sendo a Flavia uma base de formas desafiadora, o descritor *MEC* trouxe um ganho de aproximadamente 3% na métrica *F-score* em comparação ao descritor *NMBE*. Quanto à base MPEG7-CE, ambos os descritores multiescala não apresentaram bom desempenho, uma vez que essa base é complexa, apresentando formas entre classes com assinaturas de curvatura muito semelhantes. Ainda assim, o *MEC* obteve vantagem ao *NMBE*.

A Figura 21a mostra que ambos os descritores multiescala tiveram desempenho semelhante na descrição de formas da base de formas Kimia-99, e isso é confirmado pelos resultados de classificação nas Tabelas 2 e 3. Na Figura 21a também é possível observar que o *MEC* não diferenciou coelhos (classe 2) de alienígenas (classe 4), bem como peixes (classe 1) de ferramentas (classe 5) pela grande semelhança entre as formas dessas classes.

A Figura 21b mostra que o *NMBE* teve um desempenho marginalmente pior do que o *MEC* em formas da base MPEG7-CE, e a Figura 21c indica que várias formas da Flavia representam um desafio para ambos os descritores multiescala, conforme confirmado nas Tabelas

2 e 3. É importante notar que os padrões das matrizes de distância reforçam a conclusão de que a descrição de formas da MPEG7-CE e de parte da Flavia é uma tarefa desafiadora para todos os descritores considerados, uma vez que a quantidade de áreas quadradas mais escuras fora da diagonal principal foi maior do que o esperado.

Figura 21 – Matrizes de distâncias obtidas com *NMBE*, *MEC* e *MONO* para (a) Kimia-99; (b) MPEG7-CE e (c) Flavia.



4.2 Experimentos de classificação com otimização

Nesta seção, os resultados são apresentados a partir de escalas ótimas, obtidas pelo procedimento de otimização dos parâmetros dos descritores multiescala, aonde foram buscadas representações que propiciem maior separação entre classes de formas e maior coesão intraclasse. Para as mesmas bases da seção anterior, os valores das Tabelas 4 e 5 confirmam um ganho de desempenho expressivo para os descritores multiescala otimizados em comparação aos resultados obtidos com esses descritores cujas escalas foram ajustadas pelo método de Cesar Jr e Costa (1997). Um detalhe importante a se destacar é que nas bases mais desafiadoras, a saber, MPEG7-CE (com alta variabilidade intraclasse) e Flavia (com baixa variabilidade entre classes), o descritor proposto neste trabalho (*MEC*) foi ainda mais efetivo na descrição frente ao *NMBE*, passando de aproximadamente 3% de ganho de Acurácia no experimento sem otimização para ganho de quase 15% no experimento otimizado, quando observado a base de folhas Flavia.

Tabela 4 – Resultados (com otimização) de precisão (P) e de revocação (R) para os testes de validação cruzada com *NMBE*, *MEC* e *NMBE + MEC* para as três bases de formas.

Base de formas	P			R		
	<i>NMBE</i>	<i>MEC</i>	<i>NMBE + MEC</i>	<i>NMBE</i>	<i>MEC</i>	<i>NMBE + MEC</i>
Kimia-99	88,17 ± 2,57	86,07 ± 0,54	90,34 ± 0,87	86,40 ± 2,20	84,21 ± 0,61	88,86 ± 0,93
MPEG7-CE	63,76 ± 2,08	67,71 ± 1,78	79,43 ± 0,40	62,33 ± 2,00	66,43 ± 1,06	77,41 ± 0,32
Flavia	73,79 ± 0,28	87,47 ± 0,18	90,84 ± 0,15	73,30 ± 0,25	87,30 ± 0,15	90,58 ± 0,18

Tabela 5 – Resultados (com otimização) de *F-score* e de Acurácia para os testes de validação cruzada com *NMBE*, *MEC* e *NMBE + MEC* para as três bases de formas.

Base de formas	<i>F-score</i>			Acurácia		
	<i>NMBE</i>	<i>MEC</i>	<i>NMBE + MEC</i>	<i>NMBE</i>	<i>MEC</i>	<i>NMBE + MEC</i>
Kimia-99	85,42 ± 2,37	83,39 ± 0,56	88,02 ± 0,87	86,54 ± 2,12	84,39 ± 0,48	88,66 ± 0,95
MPEG7-CE	60,63 ± 2,06	64,26 ± 1,28	76,20 ± 0,38	62,35 ± 2,10	66,43 ± 0,99	77,43 ± 0,29
Flavia	72,79 ± 0,30	86,96 ± 0,12	90,38 ± 0,15	73,33 ± 0,25	87,24 ± 0,10	90,59 ± 0,15

Com base nas Tabelas 4 e 5, a última coluna de cada métrica de avaliação (*NMBE + MEC*) diz respeito aos resultados obtidos em uma terceira abordagem, na qual o vetor de características otimizado dos descritores multiescala são concatenados para formar um único descritor. Nessa abordagem, cada forma de uma base passa a ser caracterizada por ambos os descritores, cada um impondo suas habilidades discriminativas. Como resposta, é possível

observar melhoria de desempenho que se apresentou superior a 10%, na base MPEG7-CE por exemplo, quando comparado à otimização individual de cada descritor. Vale ressaltar que também foram executados testes para otimizar conjuntamente ambos os descritores multiescala, porém não houve ganho, uma vez que os resultados foram similares ao do melhor descritor otimizado individualmente, ou seja, o aumento de custo na otimização não se justificaria.

5 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, a principal contribuição científica foi a proposta de um descritor de formas multiescala que depende da entropia diferencial e da curvatura: Entropia da Curvatura Multiescala (*MEC*). O desafio inicial foi garantir a robustez do *MEC*, pois para uma comparação justa com o descritor Energia de Dobramento Multiescala Normalizada (*NMBE*), ele também deveria ser invariante à rotação, à translação e à mudança de escala. Este objetivo foi alcançado com sucesso.

A metodologia empregada nos experimentos foi com base nos mesmos princípios, uma vez que ambos os descritores multiescala são fundamentados a partir da curvatura multiescala do contorno. Os resultados mostraram que o *MEC* é um descritor competitivo que pode revelar as características locais de formas muito variadas. A dificuldade inerente ao processo de ajuste manual ou empírico desses parâmetros multiescala a um determinado problema ou aplicação, bem como o ganho de desempenho evidenciado pelo método de ajuste por otimização justifica a utilização da proposta apresentada neste trabalho.

De fato, os resultados obtidos após a otimização são significativos e comprovam a superioridade do *MEC* em comparação ao *NMBE* em experimentos de classificação. O conjunto otimizado de parâmetros, moldado pela maximização da função custo *Silhouette*, embute informações de nuances da forma. A comprovação desde achado se deu pela considerável melhoria na taxa de acerto de classificação das formas ao serem utilizados os descritores otimizados em bases com elevada similaridade de formas entre classes. O descritor proposto foi capaz de melhorar o desempenho alcançado pelo *NMBE* para uma base real importante: a base de folhas Flavia; se consagrando portanto como uma ferramenta adicional e fonte de informação para taxonomistas discriminarem e classificarem espécies de plantas.

Como perspectivas de trabalhos para a continuidade desse projeto, é possível identificar as seguintes possibilidades:

- Utilização de outras funções custo no otimizador;
- Incluir os hiperparâmetros do otimizador no processo de otimização, uma vez que esses foram selecionados empiricamente;
- Buscar novas estratégias para reduzir o custo computacional do processo, visando o melhor *trade-off* para não prejudicar o desempenho.

REFERÊNCIAS

- BARBER, C. B.; DOBKIN, D. P.; HUHDANPAA, H. The quickhull algorithm for convex hulls. **ACM Transactions on Mathematical Software**, Association for Computing Machinery, New York, NY, USA, v. 22, n. 4, p. 469–483, 1996. ISSN 0098-3500.
- CARNEIRO, A. C.; LOPES, J. G.; ROCHA NETO, J. F.; SOUZA, M. M.; MEDEIROS, F. N. S.; BEZERRA, F. N. On the evaluation of cost functions for parameter optimization of a multiscale shape descriptor. In: IEEE. **2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)**. [S.l.], 2017. p. 045–050.
- CARNEIRO, A. C.; LOPES, J. G.; SOUZA, M. M.; ROCHA NETO, J. F.; ARAÚJO, F. H. D.; SILVA, R. R. V.; MEDEIROS, F. N. S.; BEZERRA, F. N. Parameter optimization of a multiscale descriptor for shape analysis on healthcare image datasets. **Pattern Recognition Letters**, Elsevier, v. 125, p. 694–700, 2019.
- CARNEIRO, A. C.; LOPES, J. G. F.; ARAÚJO, F. H. D.; SILVA, R. R. V.; PASSARINHO, C. J. P.; ROCHA NETO, J. F.; MEDEIROS, F. N. S. Análise de fotografias de pílulas por redes neurais convolucionais. In: **VIII Simpósio de Instrumentação e Imagens Médicas e VII Simpósio de Processamento de Sinais**. [S.l.: s.n.], 2017. p. 14.
- ČERNÝ, V. Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. **Journal of Optimization Theory and Applications**, Springer, v. 45, n. 1, p. 41–51, 1985.
- CESAR JR, R. M.; COSTA, L. F. Application and assessment of multiscale bending energy for morphometric characterization of neural cells. **Review of Scientific Instruments**, AIP Publishing, v. 68, n. 5, p. 2177–2186, 1997. ISSN 1089-7623.
- CHAKI, J.; PAREKH, R.; BHATTACHARYA, S. Plant leaf recognition using texture and shape features with neural classifiers. **Pattern Recognition Letters**, Elsevier, v. 58, p. 61–68, 2015.
- COSTA, L. F.; CESAR JR, R. M. **Shape Classification and Analysis: Theory and Practice**. 2. ed. Boca Raton, Florida, FL, USA: CRC Press Taylor & Francis Group, 2009. ISBN 978-0-8493-7929-1.
- COVER, T. M.; THOMAS, J. A. **Elements of Information Theory**. second. Hoboken, New Jersey: John Wiley & Sons, Inc., Publication, 2006. ISBN 978-0-471-24195-9.
- DAS, V.; DANDAPAT, S.; BORA, P. K. A novel diagnostic information based framework for super-resolution of retinal fundus images. **Computerized Medical Imaging and Graphics**, Elsevier, v. 72, p. 22–33, 2019.
- DRYDEN, I. L.; MARDIA, K. V. **Statistical Shape Analysis: with Applications in R**. [S.l.]: John Wiley & Sons, 2016. v. 995.
- ENGELBRECHT, A. P. **Computational Intelligence: An Introduction**. 2nd. ed. [S.l.]: Wiley Publishing, 2007. ISBN 0470035617.
- FUKUNAGA, K. **Introduction to Statistical Pattern Recognition**. [S.l.]: Elsevier, 2013.
- GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. USA: Prentice-Hall, Inc., 2006. ISBN 013168728X.

- HALL, D.; MCCOOL, C.; DAYOUB, F.; SUNDERHAUF, N.; UPCROFT, B. Evaluation of features for leaf classification in challenging conditions. In: **IEEE. 2015 IEEE Winter Conference on Applications of Computer Vision**. [S.l.], 2015. p. 797–804.
- KENNEDY, J. Particle swarm optimization/kennedy j., eberhart rc. In: **Proceeding of the IEEE International Conference on Neural Networks.–Perth: IEEE Service Center**. [S.l.: s.n.], 1995. p. 12–13.
- KINDRATENKO, V. V. On using functions to describe the shape. **Journal of Mathematical Imaging and Vision**, Springer, v. 18, n. 3, p. 225–245, 2003.
- KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. **Science**, American Association for the Advancement of Science, v. 220, n. 4598, p. 671–680, 1983. ISSN 0036-8075.
- KOENDERINK, J. J. The structure of images. **Biological Cybernetics**, Springer, v. 50, n. 5, p. 363–370, 1984.
- LING, H.; YANG, X.; LATECKI, L. J. Balancing deformability and discriminability for shape matching. In: DANIILIDIS, K.; MARAGOS, P.; PARAGIOS, N. (Ed.). **Computer Vision – ECCV 2010**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 411–424. ISBN 978-3-642-15558-1.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The Journal of Chemical Physics**, American Institute of Physics, v. 21, n. 6, p. 1087–1092, 1953.
- MOKHTARIAN, F.; MACKWORTH, A. K. A theory of multiscale, curvature-based shape representation for planar curves. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 14, n. 8, p. 789–805, 1992.
- PEDRINI, H.; SCHWARTZ, W. R. **Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações**. 1. ed. São Paulo, São Paulo, SP, BRA: Editora Thomson Learning Edições Ltda, 2007. ISBN 978-85-221-0595-3.
- QUADRI, A. T.; SIRSHAR, M. Leaf recognition system using multi-class kernel support vector machine. **International Journal of Computer and Communication System Engineering**, Citeseer, v. 2, n. 2, p. 260–263, 2015.
- RAJAGURU, H.; PRABHAKAR, S. K. **KNN Classifier and K-means Clustering for Robust Classification of Epilepsy from EEG Signals. A Detailed Analysis**. [S.l.]: diplom. de, 2017.
- ROUSSEEUW, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. **Journal of Computational and Applied Mathematics**, Elsevier, v. 20, p. 53–65, 1987.
- RUIZ, F. E.; PÉREZ, P. S.; BONEV, B. I. **Information Theory in Computer Vision and Pattern Recognition**. [S.l.]: Springer Science & Business Media, 2009.
- RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3. ed. Upper Saddle River, NJ, USA: Pearson, 2009. (Prentice Hall Series in Artificial Intelligence). ISBN 978-0-13-604259-4.

- SAITTA, S.; RAPHAEL, B.; SMITH, I. F. A comprehensive validity index for clustering. **Intelligent Data Analysis**, IOS Press, v. 12, n. 6, p. 529–548, 2008.
- 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**, IEEE, v. 26, n. 5, p. 550–571, 2004.
- SILVERMAN, B. W.; GREEN, P. J. **Density Estimation for Statistics and Data Analysis**. London: Chapman and Hall, 1986. ISBN 978-0-412-24620-3.
- SINGH, A.; PRINCIPE, J. C. Information theoretic learning with adaptive kernels. **Signal Processing**, Elsevier, v. 91, n. 2, p. 203–213, 2011.
- SINGH, C.; SINGH, J. Geometrically invariant color, shape and texture features for object recognition using multiple kernel learning classification approach. **Information Sciences**, Elsevier, v. 484, p. 135–152, 2019.
- SOUZA, M. M. de; MEDEIROS, F. N.; RAMALHO, G. L.; JR, I. C. de P.; OLIVEIRA, I. N. Evolutionary optimization of a multiscale descriptor for leaf shape analysis. **Expert Systems with Applications**, Elsevier, v. 63, p. 375–385, 2016.
- SOUZA, M. M. S. **Um método para otimização de descritores multiescala em análise de formas**. 2016. 77 f. Tese (Doutorado em Engenharia de Teleinformática) — Universidade Federal do Ceará, Fortaleza, 2016. Disponível em: <<http://www.repositoriobib.ufc.br/000030/00003057.pdf>>. Acesso em: 10 out. 2017.
- SPERANDIO, D.; MENDES, J. T.; SILVA, L. H. M. e. **Cálculo Numérico: Características Matemáticas e Computacionais dos Métodos Numéricos**. [S.l.]: Prentice Hall, 2003.
- STORN, R.; PRICE, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. **Journal of Global Optimization**, Springer, v. 11, n. 4, p. 341–359, 1997.
- WANG, B.; BROWN, D.; GAO, Y.; SALLE, J. L. March: Multiscale-arch-height description for mobile retrieval of leaf images. **Information Sciences**, Elsevier, v. 302, p. 132–148, 2015.
- WEBB, A. R. **Statistical Pattern Recognition**. 2. ed. [S.l.]: John Wiley & Sons Ltd, 2002. ISBN 0-470-84514-7.
- WITKIN, A. P. Scale-space filtering. In: **Readings in Computer Vision**. [S.l.]: Elsevier, 1987. p. 329–332.
- WU, S. G.; BAO, F. S.; XU, E. Y.; WANG, Y.-X.; CHANG, Y.-F.; XIANG, Q.-L. A leaf recognition algorithm for plant classification using probabilistic neural network. In: IEEE. **2007 IEEE International Symposium on Signal Processing and Information Technology**. [S.l.], 2007. p. 11–16.
- YANG, C.; YU, Q. Multiscale fourier descriptor based on triangular features for shape retrieval. **Signal Processing: Image Communication**, Elsevier, v. 71, p. 110–119, 2019.
- ZHANG, D.; LU, G. Review of shape representation and description techniques. **Pattern Recognition**, Elsevier, v. 37, n. 1, p. 1–19, 2004.