
Caracterização de Texturas com o auxílio de Redes Complexas, Projeções Topológicas e Padrões Semânticos

Thiago Pirola Ribeiro



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2019

Thiago Pirola Ribeiro

**Caracterização de Texturas com o auxílio de
Redes Complexas, Projeções Topológicas e
Padrões Semânticos**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Celia A. Zorzo Barcelos

Uberlândia

2019

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

R484c Ribeiro, Thiago Pirola, 1978-
2019 Caracterização de texturas com o auxílio de redes complexas, projeções topológicas e padrões semânticos [recurso eletrônico] / Thiago Pirola Ribeiro. - 2019.

Orientadora: Celia Aparecida Zorzo Barcelos.
Tese (Doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Ciência da Computação.

Modo de acesso: Internet.

Disponível em: <http://dx.doi.org/10.14393/ufu.te.2019.4>

Inclui bibliografia.

Inclui ilustrações.

1. Computação. 2. Processamento de imagens. 3. Modelagem de processos. 4. Topologia. I. Barcelos, Celia Aparecida Zorzo, 1954-, (Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU: 681.3

UNIVERSIDADE FEDERAL DE UBERLÂNDIA – UFU
FACULDADE DE COMPUTAÇÃO – FACOM
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO – PPGCO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da Tese de doutorado intitulada “**Caracterização de Texturas com o auxílio de Redes Complexas, Projeções Topológicas e Padrões Semânticos**” por **Thiago Pirola Ribeiro** como parte dos requisitos exigidos para a obtenção do título de **Doutor em Ciência da Computação**.

Uberlândia, 28 de Março de 2019.

Orientadora:

Prof. Dr. Celia Aparecida Zorzo Barcelos
Universidade Federal de Uberlândia (FACOM/UFU)

Membros da Banca Examinadora:

Prof. Dr. Odemir Martinez Bruno
Universidade de São Paulo (IFSC/USP)

Prof. Dr. João do Espirito Santo Batista Neto
Universidade de São Paulo (ICMC/USP)

Prof. Dr. Bruno Augusto Nassif Travençolo
Universidade Federal de Uberlândia (FACOM/UFU)

Prof. Dr. Marcelo Zanchetta do Nascimento
Universidade Federal de Uberlândia (FACOM/UFU)

Este trabalho é dedicado a Ana Cláudia, minha esposa e companheira, aos meus filhos Gustavo e André, aos meus pais, Valter e Maria Helena, minha irmã Rachel com sua filha Luna, a todos os meus familiares e a Deus.

Agradecimentos

Agradeço à Deus por tudo.

À minha esposa Ana e meus filhos Gustavo e André tenho muito que agradecer pela ajuda e por entenderem os vários momentos em que tive que me dedicar mais ao trabalho, deixando de me dedicar a eles.

Agradeço aos meus familiares pelo apoio e incentivo ao meu trabalho.

Agradeço aos amigos, que mesmo sabendo que eu estava trabalhando na Tese, enviavam mensagens de apoio com as fotos dos churrascos e festas.

Agradeço aos colegas de trabalho.

Agradeço à minha orientadora profa. Dra. Celia Aparecida Zorzo Barcelos pela orientação e ajuda durante o trabalho.

Agradeço ao prof. Dr. André Luiz Naves de Oliveira da Universidade Federal de Uberlândia - Faculdade de Matemática pelo grande auxílio nas formulações da Garrafa de Klein e pela amizade cultivada nesses anos todos.

Agradeço ao prof. Dr. André Ricardo Backes da Universidade Federal de Uberlândia - Faculdade de Computação pela disponibilização do código de geração de Redes Complexas e ao prof. Dr. Eraldo Ribeiro do *Florida Institute of Technology* pela disponibilização da base de imagens de Pólen.

Agradeço à Faculdade de Computação (FACOM) e à Universidade Federal de Uberlândia (UFU) por ter concedido meu afastamento das funções para a realização do Doutorado.

*"Is this the end of the beginning?
Or the beginning of the end?"
(Black Sabbath, 2013)*

Resumo

A área de processamento de imagens está presente em diversos segmentos da sociedade. Esta área está em constante desenvolvimento e, por meio do aprimoramento de diversas técnicas, pode-se auxiliar na resolução de problemas diversificados que sejam baseados em imagens. Existem diversos métodos na literatura que tratam da modelagem e da caracterização de texturas que, neste trabalho, são utilizados em texturas naturais e artificiais.

Neste trabalho foram utilizadas as redes complexas para representar texturas, considerou-se os histogramas de energia do mapa de graus da rede complexa, selecionando os picos e vales desses histogramas para a definição de limiares a serem adotados para a geração do vetor de característica baseado nas imagens de textura a serem classificadas.

Foram utilizadas, também, formas alternativas para a representação de imagens: foi proposto um método que analisa a modelagem dos dados utilizando padrões semânticos, para análises de imagens de grãos de pólen, e foram exploradas características que influenciam a formação do vetor de características utilizando a topologia da garrafa de Klein para projetar *patches* da imagem em espaços topológicos. Foi explorado o cálculo dos Coeficientes Estimados de K-Fourier, a composição do vetor de características bem como novas combinações de tamanhos, que resultariam em uma redução da quantidade de descritores e melhoria na caracterização das bases de texturas.

Foram realizados diversos experimentos sendo validados com a utilização de bases de imagens de texturas públicas já classificadas: Brodatz, CURET, Outex_TC_00013, KTH-TIPS, UIUCTex, VisTex e uma base de imagens de grãos de pólen.

Este trabalho demonstra que os métodos propostos conseguiram atingir as expectativas, conseguindo se sobressair pela quantidade reduzida de descritores utilizados e pelas acurácias obtidas nos experimentos.

Palavras-chave: Caracterização de Texturas, Redes Complexas, Padrões Semânticos, Projeções Topológicas.

Abstract

Image processing is a computing area that is present in many segments of society. This area is in constant development and, through the improvement of several techniques, it can help resolving many different problems which are based on images. There are several methods in the literature that deal with the modeling and the characterization of textures. This work is focused on natural and artificial textures.

The proposed approach uses complex networks to represent textures. Energy histograms generated from maps of degrees of complex networks were used, selecting their peaks and valleys for defining thresholds to be adopted to generate the feature vector based on the target images to be classified.

We also used alternative forms for image representation: first, a method that analyzes the data modeling using semantic patterns for pollen grain image analysis was proposed, and, second, the analysis of influence of characteristics of the feature vector composition using the Klein bottle topology for designing image patches in topological spaces. The computation of the K-Fourier Estimated Coefficients was explored, the composition of the feature vector as well as new combinations of sizes, which would result in a reduction of the amount of descriptors and improvement in the characterization of texture datasets.

Several experiments were carried out on some public domain datasets already classified: Brodatz, CURET, Outex_TC_00013, KTH-TIPS, UIUCTex, VisTex, and a pollen grains image dataset.

This work demonstrates that the proposed methods were able to reach out the expectations, being able to stand out for the reduced amount of used descriptors and the accuracy obtained in the experiments.

Keywords: Characterization of Texture, Complex Networks, Semantics Patterns, Topological Projections.

Lista de ilustrações

Figura 1 – Exemplo da definição do limiar: em (b) exemplo do conjunto E relativo ao <i>pixel</i> central da imagem dada em (a) considerando $r = 3$ e em (c) os pesos ($w(e)$) que estão associados aos vértices ligados. Um limiar t é aplicado ao conjunto original de arestas (neste exemplo $t = 0,25$) para executar uma transformação na rede, resultando, assim, em um subconjunto de arestas (E_t) (<i>pixels</i> em cinza) (BACKES; CASANOVA; BRUNO, 2013).	35
Figura 2 – Etapas do Método Proposto: (a) Imagem de entrada; (b) Rede Complexa gerada para cada valor de limiar t e amostra da textura no conjunto de dados; (c) Grau da distribuição da amostra de textura para cada limiar selecionado; (d) Exemplo do histograma médio de uma medida estatística descritiva; (e) Limiares Selecionados por meio dos picos e vales, em destaque, gerando um conjunto S de limiares; (f) Vetor de atributos composto pela concatenação dos valores da energia, entropia e contraste dos S limiares ($F_{1,1}, F_{1,2}, F_{1,3}, \dots, F_{ S ,1}, F_{ S ,2}, F_{ S ,3}$).	38
Figura 3 – Gráfico dos valores médios das imagens da base Brodatz para os 190 limiares: (A) contraste; (B) energia; (C) entropia e (D) média.	39
Figura 4 – Gráfico dos valores médios das imagens da base Brodatz para os 190 limiares: (A) variância; (B) <i>skewness</i> e (C) <i>kurtosis</i>	40
Figura 5 – Gráfico dos valores médios das medidas estatísticas das imagens da base Vistex para os 190 limiares: (A) média; (B) contraste; (C) variância; (D) energia; (E) entropia; (F) <i>kurtosis</i> e (G) <i>skewness</i>	41
Figura 6 – Garrafa de Klein: (A) Modelo da Garrafa de Klein; (B) Exemplo de um resultado da projeção no espaço \mathcal{H} . Modificado de (PEREA; CARLSSON, 2014).	49
Figura 7 – Representação visual da garrafa de Klein.	50
Figura 8 – Fluxo das etapas do processamento	50

Figura 9 – Exemplo de recorte de <i>Patch</i> : (A) sem sobreposição e (B) com janela deslizante. As marcações são para exemplificação e não estão em escala com as matrizes.	62
Figura 10 – Imagens de texturas utilizadas nos experimentos de rotação.	63
Figura 11 – Exemplos de <i>Patches</i> criados. Em (a) o <i>patch</i> original, (b) rotacionado em 180°, (c) rotacionado em 90° e (d) rotacionado em -90°. Todas as rotações em sentido anti-horário.	66
Figura 12 – Exemplo de extração de <i>Patches</i> . Itens 1 e 4 estão na mesma posição nas imagens, assim como os itens 2 e 5 e 3 e 6, respectivamente. As marcações são para exemplificação e não estão em escala com as matrizes. A rotação está no sentido anti-horário.	67
Figura 13 – Imagens exemplificando 10 classes da base Pollen. As imagens não estão em escala.	83
Figura 14 – Diagrama Geral do Método Proposto para a Semântica.	83
Figura 15 – Exemplos de 5 conceitos em tons de cinza de imagens da base Pollen. As imagens não estão em escala.	84
Figura 16 – Exemplos de conceitos. À esquerda, a imagem original em tons de cinza (ao fundo) e, em destaque, o exemplo do conceito selecionado da imagem em preto e branco. À direita estão 3 exemplos, em preto e branco, de conceitos selecionados de outras imagens da base Pollen. . .	85
Figura 17 – Todos os L conceitos selecionados para a base Pollen. Para melhor visualização, os conceitos não estão em escala.	87
Figura 18 – Diagrama detalhado do Método Proposto para a Semântica.	88
Figura 19 – Exemplos de imagens das classes 2, 4, 7 e 8, que ficaram indistintas da classe 23 para o classificador.	88
Figura 20 – Diagrama Geral do procedimento alternativo.	93
Figura 21 – Exemplos de imagens das bases: (a) CURET; (b) KTH-TIPS; (c) UIUCTex; (d) Brodatz; (e) VisTex; (f) Outex_TC_00013; (g) Pollen. .	115

Lista de tabelas

Tabela 1 – Resultados da classificação para valores médios e medianos (em %) baseado na medida descritiva de energia.	42
Tabela 2 – Tabela com a quantidade de limiares obtidos automaticamente pelo método proposto para cada base de imagem	42
Tabela 3 – Resultados do método proposto e de métodos da literatura para as seis bases de imagens.	43
Tabela 4 – Resultados da Classificação com e sem a utilização do valor absoluto (em %).	44
Tabela 5 – Comparativo dos Resultados da Classificação com entre as propostas para a definição automática de limiares (em %).	45
Tabela 6 – Resultados dos Métodos Propostos e de Métodos da literatura para as seis bases de imagens.	45
Tabela 7 – Resultados da Classificação das imagens da base KTH-TIPS projetadas na garrafa de Klein (em %).	60
Tabela 8 – Resultados da Classificação das imagens da base Pollen projetadas na garrafa de Klein (em %).	60
Tabela 9 – Resultados da Classificação das imagens da base KTH-TIPS projetadas na garrafa de Klein sem invariância à rotação (em %).	61
Tabela 10 – Resultados da Classificação das imagens da base Pollen projetadas na garrafa de Klein sem invariância à rotação (em %).	61
Tabela 11 – Resultado da Classificação da base KTH-TIPS para os 2 tipos de janelas (acurácias em %).	62
Tabela 12 – Resultados das distâncias das imagens considerando caracterização usando os coeficientes estimados de K-Fourier e os coeficientes estimados de K-Fourier invariantes à rotação.	64
Tabela 13 – Resultados da Classificação da base USC-SIPI com o classificador LMNN (em %).	64

Tabela 14 – Resultados da Classificação da base KTH-TIPS com o classificador LMNN para <i>patches</i> sem e com rotação (em %).	65
Tabela 15 – Resultados da Classificação com LMNN para a base KTH-TIPS (em %).	69
Tabela 16 – Resultados da Classificação com LMNN para a base Pollen (em %).	69
Tabela 17 – Resultado da Classificação das 13 escalas individuais para a base KTH-TIPS.	71
Tabela 18 – Maiores Acurácias resultantes da Classificação aplicada à base KTH-TIPS (em %).	71
Tabela 19 – Maiores Acurácias resultantes da Classificação aplicada à base Brodatz (em %).	71
Tabela 20 – Maiores Acurácias resultantes da Classificação aplicada à base CURET (em %).	72
Tabela 21 – Maiores Acurácias resultantes da Classificação aplicada à base Outex (em %).	72
Tabela 22 – Maiores Acurácias resultantes da Classificação aplicada à base UIUC-Tex (em %).	72
Tabela 23 – Maiores Acurácias resultantes da Classificação aplicada à base VisTex (em %).	73
Tabela 24 – Maiores Acurácias resultantes da Classificação aplicada à base Pollen (em %).	73
Tabela 25 – Resultado da Classificação da base KHT-TIPS para novas escalas baseadas em Fibonacci (em %).	73
Tabela 26 – Resultado da Classificação da base KHT-TIPS para novas escalas (em %).	74
Tabela 27 – Resultado da Classificação das bases para as escalas $n = 3, 4, 5$ (em %).	75
Tabela 28 – Comparativo do método proposto com outros métodos da literatura para algumas das bases testadas.	76
Tabela 29 – Resultado da Classificação com LMNN da base KHT-TIPS para as escalas $n = 3, 7, 11, 15, 19$ com frequências de 2 à 12 (em %).	77
Tabela 30 – Resultado da Classificação com LMNN da base KHT-TIPS para as escalas $n = 3, 5$ com frequências de 2 à 12 (em %).	78
Tabela 31 – Resultado da Classificação com LMNN para as escalas $n = 3, 5$ com frequência 4 para as diversas bases (em %).	78
Tabela 32 – Resultado da Classificação com LMNN para as escalas $n = 3, 4, 5$ para a base KTH-TIPS (em %).	78
Tabela 33 – Comparativo do método proposto com outros métodos da literatura para algumas das bases testadas.	79
Tabela 34 – Resultado do Método Proposto para a base Pollen utilizando o classificador KNN (em %).	87

Tabela 35 – Resultados do Método Proposto para a base Pollen por classe.	89
Tabela 36 – Matriz de Confusão da classificação com KNN do vetor de características gerado pelo Método Proposto para a base Pollen. Em destaque umas das classes com alto índice de confusão.	90
Tabela 37 – Resultados da classificação do Método Proposto para a base Pollen por classe e conjunto de Conceito (%)	91
Tabela 38 – Resultados de classificação do Método Proposto e Métodos Seleccionados que utilizaram a mesma base Pollen.	91
Tabela 39 – Resultados de classificação do Método Proposto e Métodos Seleccionados com bases de pólen variadas.	92

Lista de siglas

Ac. Acurácia - *Accuracy*

CCN Correlação Cruzada Normalizada - *Normalized Cross-Correlation*

CLBP Padrão Binário Local Completo - *Completed Local Binary Pattern*

CoV Coeficiente de Variação - *Coefficient of Variation*

CRoV Coeficiente Relativo de Variação - *Relative Coefficient of Variation*

Dim. Dimensão - *Dimension*

DCT Transformada Discreta do Cosseno - *Discrete Cosine Transform*

DTM Momentos Discretos de Tchebichef - *Discrete Tchebichef Moments*

EKFC Coeficientes Estimados de K-Fourier Invariantes à Rotação - *Estimated Rotation-Invariant K-Fourier Coefficients*

GLCM Matriz de Coocorrência de Nível de Cinza - *Gray Level Co-occurrence Matrix*

GLDM Método de Diferenças de Nível de Cinza - *Gray Level Difference Method*

HOG Histograma de Gradientes Orientados - *Histogram of Oriented Gradients*

KNN K-Vizinhos mais Próximos - *K-Nearest Neighbor*

LBP Padrão Binário Local - *Local Binary Pattern*

LBPV Padrões Binários Locais Variantes - *Local Binary Pattern Variance*

LDA Análise de Discriminantes Lineares - *Linear Discriminant Analysis*

LGF Filtro Log de Gabor - *Log Gabor Filter*

LMNN Grande Margem de Vizinho mais Próximo - *Large Margin Nearest Neighbor*

LTP Padrão Ternário Local - *Local Ternary Pattern*

MMI Índice de Informação Mútua - *Mutual Information Index*

QDA Análise de Discriminantes Quadráticos - *Quadratic Discriminant Analysis*

RGB-D Vermelho Verde Azul - Profundidade - *Red Green Blue - Depth*

SVM Máquina de Vetores de Suporte - *Support Vector Machine*

Sumário

1	INTRODUÇÃO	25
1.1	Textura caracterizando imagem	25
1.2	Visão Geral da Tese	27
1.3	Objetivos	29
1.4	Organização da Tese	29
2	REDES COMPLEXAS	31
2.1	Fundamentação Teórica	33
2.1.1	Modelagem do Grafo	33
2.1.2	Evolução Dinâmica do Grafo	34
2.1.3	Medidas Estatísticas Descritivas	36
2.2	Proposta	36
2.3	Experimentos e Análise dos Resultados	37
2.3.1	Definição dos Limiares	37
2.3.2	Validação da Proposta	42
2.3.3	Alternativa para a Definição de Limiares	44
2.4	Considerações Finais	45
3	PROJEÇÕES TOPOLÓGICAS	47
3.1	Fundamentação Teórica	47
3.1.1	Projeção	50
3.1.2	Coefficientes Estimados de K-Fourier	53
3.1.3	Coefficientes Estimados de K-Fourier invariantes à rotação	56
3.2	Propostas	57
3.3	Experimentos e Análise dos Resultados	58
3.3.1	Obtenção dos <i>Patches</i>	61
3.3.2	Análise da Rotação	63
3.3.3	Vetor de Características	66

3.3.4	Análise das Escalas	70
3.3.5	Análise das Frequências	76
3.4	Considerações Finais	79
4	PADRÕES SEMÂNTICOS	81
4.1	Fundamentação Teórica	81
4.2	Propostas	82
4.2.1	Criação da Semântica Específica para a base Pollen	83
4.2.2	Definição do Espaço Semântico	84
4.2.3	Mapeamento do Espaço Semântico	85
4.3	Experimentos e Análise dos Resultados	86
4.3.1	Desenvolvimento da Proposta	86
4.3.2	Confusões	88
4.4	Considerações Finais	93
5	CONCLUSÃO	95
5.1	Contribuições em Produção Bibliográfica	97
5.2	Trabalhos Futuros	98
	REFERÊNCIAS	99
	APÊNDICE A FUNDAMENTAÇÃO TEÓRICA	107
A.1	Projeções Topológicas	107
A.1.1	Obtenção do <i>Theta</i>	107
A.1.2	Obtenção dos Coeficientes de K-Fourier	109
A.1.3	Detalhes de Implementação	111
A.2	Padrões Semânticos	112
A.2.1	Correlação Cruzada Normalizada	112
A.2.2	Coeficiente de Variação	112
A.2.3	Coeficiente Relativo de Variação	112
	APÊNDICE B BASES DE IMAGENS DE TEXTURAS	113

Introdução

A área de processamento de imagens tem assumido um papel de destaque em diversos segmentos da sociedade, desde um simples filtro aplicado a uma foto, até mesmo o controle de um veículo em movimento. Diversos algoritmos são desenvolvidos, estudados e utilizados para trabalhar com problemas específicos. Em muitos casos, a aplicação de filtros ou convoluções em imagens resolvem especificidades e, em outros casos, a identificação ou a classificação de um determinado objeto ou cena se faz necessária, um exemplo sendo o reconhecimento de uma face no meio de uma multidão. Aprimorar as técnicas para a classificação de objetos é um ponto necessário, porém para classificar um objeto, inicialmente, deve-se obter informações que possam caracterizá-lo e diferenciá-lo dos demais objetos.

A utilização de descritores auxilia nesse processo de caracterização. Esses descritores podem ser desde o formato de um objeto, a quantificação da área desse objeto, suas características relacionadas a cor ou mesmo localização espacial. Analisar a textura do objeto ou imagem auxilia na sua caracterização, pois consegue-se, em muitos casos, obter descritores que os diferenciam, por exemplo, por meio de características de frequência e direcionamentos de gradientes.

1.1 Textura caracterizando imagem

Textura é um dos atributos usados para representar uma imagem isoladamente ou em conjunto com outros descritores, uma vez que um bom descritor deve descrever uma imagem adequadamente.

Embora para alguns autores não haja uma definição exata para o termo textura, para Haralick, Shanmugam e Dinstein (1973) as texturas definem uniformidade, densidade, aspereza, regularidade, intensidade, entre outras características em uma imagem.

Segundo Backes, Casanova e Bruno (2013) o atributo de textura é facilmente compreendido por humanos, sendo uma rica fonte de informação visual (considerando objetos físicos tridimensionais). Ainda, segundo Raghuvanshi e Tyagi (2016), em geral,

imagens de textura são caracterizadas pelas propriedades da superfície e o aspecto de um objeto, tal como a forma, a densidade e arranjo. A análise de imagens de textura é diferente de imagens não-texturizadas e, para analisá-las, verificam-se propriedades de um objeto, como padrões repetidos, periodicidade local, periodicidade global, etc.

Na literatura a caracterização por meio de texturas é abordada por diferentes métodos (MATERKA; STRZELECKI, 1998), categorizados em: estruturais, estatísticos, baseados em filtros e baseados em modelos.

- Métodos estruturais consideram a textura como arranjos hierárquicos de elementos. Geralmente as ferramentas de morfologia matemática (fechamento e abertura) são utilizadas nas imagens para a caracterização dos elementos de textura (HARALICK; STERNBERG; ZHUANG, 1987).
- Métodos estatísticos representam a textura por meio de distribuições estatísticas em relação aos níveis de cinza da imagem. Normalmente utiliza-se do histograma da imagem para trabalhar esses métodos. Porém, essa abordagem limita-se a texturas irregulares ou estruturais. Métodos como a matriz de coocorrência e o Padrão Binário Local (LBP) sanam essa limitação do histograma. Matriz de Coocorrência de Nível de Cinza (GLCM) foi proposta em (HARALICK; SHANMUGAM; DINSTEN, 1973), e caracteriza-se como um dos mais conhecidos métodos que exploram repetições de ocorrências de padrões, sendo que esses padrões são caracterizados pela dependência entre os níveis de cinza dos *pixels* das imagens e por sua distribuição espacial. Trabalhos como o de Xie (2008) utilizam a GLCM para detectar defeitos em superfícies. Ojala, Pietikainen e Maenpaa (2002) propõem o uso do LBP para classificação de texturas invariantes a rotação, porém o método é suscetível a ruídos.
- Métodos baseados em filtros utilizam informações espectrais da imagem para representar a textura. Basicamente é por meio de uma estimativa da frequência espacial da intensidade dos pixels que se correlaciona texturas mais rugosas às baixas frequências e texturas mais finas com as altas frequências. Weldon, Higgins e Dunn (1996) utilizaram Gabor para segmentar imagem por meio da modelagem das texturas da imagem. Chen et al. (2015) utilizaram a transformada Wavelet para a filtragem do sinal devido à sua capacidade de destacar características de sinais locais e não homogêneos. Singh et al. (2015) utilizaram a média dos filtros de Gabor e Wavelet como técnica para extração de características para reconhecimento de expressões faciais.
- Métodos baseados em modelos procuram representar a textura por meio da construção de um modelo e a obtenção dos seus parâmetros. Cross e Jain (1983) descreveram um modelo de textura como sendo um procedimento matemático

capaz de produzir e descrever uma imagem de textura, considerando a textura de uma imagem bidimensional como estocástica e possivelmente periódica. Lu et al. (2015) propuseram propagar a semelhança no campo aleatório de Markov para a recuperação de objetos, pois as semelhanças entre a imagem de consulta e as imagens do conjunto de dados são modelados como probabilidades e relações entre as imagens do conjunto de dados também são modelados estatisticamente a priori. Para obterem um melhor desempenho de recuperação, as estatísticas a priori e a probabilidade são combinados. Backes, Casanova e Bruno (2009) utilizaram os descritores fractais e analisaram a complexidade da superfície gerada a partir de uma textura, obtendo a descrição e caracterização da mesma, produzindo uma assinatura. Uma outra ferramenta nesta classe aplicada a imagens são as redes complexas. A análise de redes complexas é feita por meio de grafos e se dá em duas etapas, sendo a primeira: representar o sistema como uma rede complexa e a segunda: analisar as características topológicas utilizando algumas medidas para identificar as diferentes categorias de estruturas presentes na imagem (COSTA et al., 2007). Newman (2003), Backes, Casanova e Bruno (2013) e Boccaletti et al. (2006) apresentaram diferentes descritores para a caracterização de texturas por meio de redes complexas. Existem na literatura, formas alternativas de modelagem de textura. A utilização de topologia (COURANT; ROBBINS, 2000) aplicada à análise de dados e imagens está presente na literatura em diversos trabalhos, por exemplo, em Lee, Pedersen e Mumford (2003), Carlsson et al. (2008), Lum et al. (2013). Perea e Carlsson (2014) utilizaram projeções de pequenos pedaços da imagem em um modelo projetado em um espaço topológico da garrafa de Klein. Outra forma de modelagem de textura é a utilização da semântica que está presente na literatura em diversos trabalhos, por exemplo, He et al. (2003), Cao e Wang (2015), Pandey, Khanna e Yokota (2015), Zhang et al. (2015a). Kwitt et al. (2012) propuseram a modelagem de textura utilizando uma codificação semântica para analisar imagens endoscópicas da mucosa do cólon por meio do estabelecimento de um espaço semântico baseado em conceitos.

1.2 Visão Geral da Tese

Neste trabalho focou-se não simplesmente na modelagem da textura, mas também na topologia dos dados, que são fatores que podem ser analisados separadamente ou como complemento um do outro. Com isso, as Redes Complexas e Padrões Semânticos foram utilizados para analisar a modelagem dos dados e as projeções no Espaço \mathcal{K} (CARLSSON et al., 2008) aplicadas à topologia dos dados, especificamente projeções na topologia da Garrafa de Klein, para estimar e representar a distribuição dos dados em espaços de baixa dimensionalidade que são subvariedades do espaço dos *pixels*.

As redes complexas conseguem representar diversos sistemas do mundo real, tornando-as muito populares nas diversas áreas do conhecimento. A utilização das redes complexas tem sido abordada por diversos pesquisadores nas mais diversificadas áreas com resultados bastante significativos, como por exemplo, os trabalhos de Rubinov e Sporns (2010), Backes, Casanova e Bruno (2013), Hu et al. (2014) e Mari, Lee e Memon (2015).

Backes, Casanova e Bruno (2013) propõem representar as texturas por um vetor de características obtido utilizando valores da energia, entropia e contraste do histograma do mapa de graus de uma rede complexa. Para a obtenção desse vetor de características os autores utilizaram limiares para a criação das redes complexas e mapas de graus. A escolha desses limiares foi tomada de forma empírica para apenas algumas bases de imagens. Observando que a escolha desses limiares interferem no resultado obtido quando utilizados em outras bases de imagens, investigou-se neste trabalho não a escolha empírica, mas considerou-se os histogramas de energia do mapa de graus da rede complexa, selecionando os picos e vales desse histograma para a definição dos novos limiares a serem adotados para a geração do vetor de característica baseado nas imagens a serem classificadas. Conseguiu-se com esse procedimento, a obtenção de descritores competitivos e obtidos de forma automática criando assim, uma extração natural de limiares dependente da base a ser processada.

Utilizando a topologia dos dados, Perea e Carlsson (2014) propuseram a utilização da projeção de *patches* da imagem em espaços topológicos. Segundo esses autores, diversos trabalhos utilizando histogramas derivados de métodos de agrupamento adaptativos podem sofrer do mal da dimensionalidade com dados não lineares e, além disso, a criação de dicionários de texturas de tamanho fixo ou o uso de bancos de filtros para a classificação dos padrões não obtinham bons resultados ou tinham alta dimensionalidade. A proposta dos autores foi estimar e representar distribuições em torno de um espaço utilizando-se da topologia da Garrafa de Klein e que ao se projetar as amostras, estas se acumularão com uma alta densidade em torno da superfície $\mathcal{K} \subset \mathbb{R}^{n^2}$. Para a caracterização foi proposto um descritor multi-escala invariante à rotação utilizando-se os Coeficientes Estimados de K-Fourier.

Observando-se que não foram explorados todos os itens que poderiam influenciar no cálculo dos Coeficientes Estimados de K-Fourier, a composição do vetor de características e também, explorar novas combinações de tamanhos, que resultariam em uma redução da quantidade de descritores e melhoria na caracterização das bases de texturas, neste trabalho explorou-se o uso da projeção no espaço da Garrafa de Klein para a redução da dimensionalidade do vetor de características proposto por Perea e Carlsson (2014).

Propôs-se, neste trabalho, o uso da semântica para a analisar a modelagem dos dados. A proposta para este trabalho foi a utilização de semânticas para a definição de padrões e vetores de características para a caracterização de texturas.

Padrões semânticos dependem das imagens e, naturalmente, do Banco de Dados a ser

analisado. Existem características semânticas gerais e específicas para um dado conjunto de imagens. Características semânticas deste banco podem ser um subconjunto das características gerais de qualquer imagem. No intuito de obter um vetor de características reduzido, foram consideradas apenas características semânticas relacionadas com o banco a ser estudado.

Diversos trabalhos na literatura tentam autenticar da origem do pólen de abelhas visando detectar fraudes. Existem métodos precisos para identificar a origem do pólen que utilizam análises laboratoriais por meio de microscopia dos grãos de pólen para reconhecer seus diversos tipos (CARRIÓN et al., 2004). Trabalhos como os de Chica (2012), Chudyk et al. (2015), Marcos et al. (2015), Redondo et al. (2015), Daood, Ribeiro e Bush (2016b), Daood, Ribeiro e Bush (2016c) e Daood, Ribeiro e Bush (2016a) caracterizaram grãos de pólen por meio de diferentes técnicas. Com isso, neste trabalho propõem-se uma nova vertente para a caracterização dos grãos de pólen por meio de Padrões Semânticos.

Para validação dos métodos propostos foram utilizadas bases de imagens de texturas Públicas: Brodatz (BRODATZ, 1966), CURET (DANA et al., 1999), Outex_TC_00013 (OJALA et al., 2002), KTH-TIPS (HAYMAN et al., 2004), UIUCTex (LAZEBNIK; SCHMID; PONCE, 2005) e VisTex (VISTEX, 2009). Estas bases já foram validadas por outros métodos da literatura e, com isso, a pesquisa torna-se empírica. A base de imagens referente à pólen foi disponibilizada pelo prof. Dr. Eraldo Ribeiro da *Florida Institute of Technology* - Melbourne - USA.

1.3 Objetivos

O objetivo geral da tese proposta é analisar e propor técnicas para caracterizar texturas artificiais e naturais em imagens digitais.

Os objetivos específicos da tese são:

- ❑ Aprimorar os limiares para obter redes complexas com alto poder discriminativo a partir das imagens das texturas;
- ❑ Aplicar as projeções na topologia da Garrafa de Klein para a caracterização das texturas por meio de vetores de características reduzidos;
- ❑ Caracterizar imagens de grãos de pólen por meio de Padrões Semânticos.

1.4 Organização da Tese

No capítulo 2 abordam-se as Redes Complexas elencando a fundamentação teórica, a proposta e os experimentos e resultados obtidos. No Capítulo 3 são apresentadas as Projeções Topológicas elencando a fundamentação teórica, a proposta e os experimentos

e resultados obtidos. No Capítulo 4 são abordados os Padrões Semânticos elencando a fundamentação teórica, a proposta, os experimentos e os resultados obtidos. Finalizando com um capítulo com as conclusões elencando as principais contribuições obtidas com o trabalho, juntamente com as publicações geradas e trabalhos futuros.

Redes Complexas aplicadas a texturas de imagens

As redes complexas são estruturas que representam muitos sistemas do mundo real, devido às suas flexibilidades e generalidades para representar praticamente qualquer estrutura natural, incluindo as estruturas com mudanças dinâmicas de topologia. Com essa flexibilidade, as redes complexas tem sido abordadas por diversos pesquisadores nas mais diversificadas áreas com resultados bastante significativos, como por exemplo, os trabalhos de Cuadros et al. (2012), Mari, Lee e Memon (2015), Li et al. (2017), Shakibian e Charkari (2017).

Na literatura são encontrados diversos exemplos da utilização das redes complexas para auxiliar na resolução de problemas variados.

Segundo Mari, Lee e Memon (2015), a globalização dentro das cadeias de fornecimento resultou em cadeias de abastecimento altamente complexas, instáveis e frágeis. Devido à importância da teoria de redes complexas na cadeia de suprimentos, os autores abordaram as métricas de resiliência para a cadeia de suprimentos e as desenvolveram com base nessa teoria. Os resultados mostraram que a rede da cadeia de suprimentos pode ser projetada com base na teoria de redes complexas, especialmente como uma rede sem escala. Concluíram que os modelos sem escala possuem algumas limitações e não puderam representar com precisão uma cadeia de abastecimento eficiente e resiliente, sugerindo que fosse desenvolvido um modelo sem escala que pudesse representar redes de cadeia de suprimentos mais robustas e resilientes.

Li et al. (2017) analisaram a estrutura global de investimento em energia com base nas relações de participação entre as empresas de energia listadas e seus acionistas, tratando essas relações como uma rede complexa heterogênea a partir de um ângulo global. Construíram uma rede de investimento primitiva de empresas de energia listadas e seus acionistas. Basearam-se nas relações de afiliação e construíram uma rede de investimentos em derivativos de 112 países. Calcularam diferentes características topológicas e analisaram quantitativamente: a diversidade nacional do investimento em

energia externo e interno, a força da relação bilateral de investimento em energia entre países e o grupo mais poderoso de países que investem em energia com base nas relações de participação das empresas de energia global listadas. Verificaram que a grande maioria das relações de investimento externo e interno ainda estão em mãos de alguns países: países desenvolvidos tradicionais, países emergentes e alguns poucos países insulares particulares que são paraísos fiscais.

Outra área ativa de pesquisa em redes complexas é a de previsão de links. Shakibian e Charkari (2017) propuseram um modelo chamado de Índice de Informação Mútua (MMI) baseado em meta-caminho que introduz a entropia de ligação baseada em meta-caminho para estimar a probabilidade de ligação, sendo realizada em um conjunto de meta-caminhos disponíveis. Essa estimativa mede a quantidade de informações utilizando-se dos caminhos ao invés de medir a quantidade de conectividade entre os pares de nós. Segundo os autores, os resultados experimentais em uma rede de bibliografia mostraram que o MMI obteve alta precisão de previsão em comparação com outros índices populares como PathSim, HeteSim e caminhada aleatória, pois estes são dependentes dos graus de conectividade dos pares de nós, não consideraram as informações adicionais fornecidas pelo meta-caminho dado, além de serem obrigados a utilizar um meta-caminho único e geralmente simétrico definido com antecedência.

Na área de imagens, dentre as várias abordagens de segmentação estão as baseadas no particionamento de grafos, apresentando desvantagens como a alta taxa de processamento. No entanto, a segmentação de imagens baseada em redes complexas reduz essa taxa, porém apresenta a limitação do número excessivo de nós na rede. Cuadros et al. (2012) apresentaram uma abordagem de rede complexa para segmentação de imagem de grandes dimensões incorporando o conceito de super *pixels*, para reduzir o número de nós na rede, obtendo bons resultados.

Backes, Casanova e Bruno (2013) apresentaram um método que utilizou redes complexas para caracterizar texturas. A proposta dos autores foi uma abordagem para representar e caracterizar a relação entre os elementos estruturais da textura usando a Teoria de Redes Complexas. Com isso, a textura foi representada como uma rede complexa, incluindo as informações sobre os *pixels* da imagem e de seus vizinhos convertidos para vértices e arestas, seguido de uma análise das características topológicas da rede gerada.

Neste trabalho são analisados novos limiares para a proposta de Backes, Casanova e Bruno (2013), além de torná-los automatizados e flexíveis utilizando-se dos picos e vales baseado no gráfico da energia média da base de textura, reduzindo a quantidade de descritores e melhorando a acurácia.

A organização deste capítulo contempla a fundamentação teórica do trabalho de Backes, Casanova e Bruno (2013), a apresentação de uma metodologia para a definição automatizada dos limiares, validações, resultados e as conclusões obtidas.

2.1 Fundamentação Teórica

Para representar as redes complexas na caracterização de texturas, Backes, Casanova e Bruno (2013) utilizaram grafos, sendo os *pixels* representados pelos nós e, a similaridade entre os *pixels*, mapeadas como as ligações entre os nós. As conectividades da rede foram mensuradas para criar o vetor de características da imagem.

2.1.1 Modelagem do Grafo

Um grafo $G = (V, E)$ é construído considerando cada *pixel* (x, y) da imagem (I) , em escala de cinza, como um vértice $v_{xy} \in V$ do grafo G . Os vértices associados a dois *pixels* (x, y) e (x', y') são conectados por uma aresta não direcional $e = (v_{x,y}, v_{x',y'})$ quando a distância Euclidiana entre eles não for maior que um valor r . O conjunto de todas as arestas de um vértice $(v_{x,y})$ é dado pelo conjunto E , onde

$$E = \left\{ e = (v_{x,y}, v_{x',y'}) \in I \times I \mid \sqrt{(x - x')^2 + (y - y')^2} \leq r \right\}. \quad (1)$$

Com o intuito de incluir informações de contexto, Backes, Casanova e Bruno (2013) definiram uma ponderação para cada aresta de E . O peso $w(e)$ definido por:

$$w(e) = \frac{1}{2r^2} \left((x - x')^2 + (y - y')^2 + r^2 \frac{I(x, y) - I(x', y')}{L} \right) \quad \forall e \in E \quad (2)$$

onde $L = \max(I(x, y)) \quad \forall x, y$ pertencente ao domínio da imagem.

Esses pesos $w(e)$ são normalizados para o intervalo $[0, 1]$.

O grau do vértice v ($\deg(v)$) representa o número de arestas conectadas à v . Sendo $\partial v = \{v' \in V \mid \{v, v'\} \in E\}$ o conjunto de vizinhos de v , define-se o grau do vértice v como:

$$D = \deg(v) = |\{v' \in V \mid \{v, v'\} \in E\}| = |\partial v| \quad (3)$$

Com isso, o grau de um nó contém informações sobre como o *pixel* correspondente se relaciona com seus vizinhos.

O histograma para cada valor i dos valores referentes ao grau assumido por um vértice é obtido por:

$$h(i) = \sum_{v \in V} \delta(\deg(v), i); \quad (4)$$

onde $\delta(j, i)$ é o delta de Kronecker

$$\delta(j, i) = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases} \quad (5)$$

Muitos atributos podem ser calculados analisando-se o histograma dos graus. Basicamente são atributos estatísticos de primeira ordem, que são calculados baseados na função de densidade de probabilidade $p(i)$ do histograma $h(i)$:

$$p(i) = \frac{h(i)}{\sum_{i=0}^k h(i)} \quad i = 0, 1, \dots, k; \quad (6)$$

onde $k = \max_{v \in V} \text{deg}(v)$ é o grau máximo do grafo G e $i = 0$ para o caso de vértices isolados.

Os histogramas de graus gerados foram analisados utilizando-se as medidas estatísticas de contraste (C), energia (E) e entropia (H), dadas pelas equações 7, 8 e 9, respectivamente. Essas medidas concatenadas compuseram o vetor de características do método proposto por Backes, Casanova e Bruno (2013).

$$C = \sum_{i=0}^k p(i) i^2 \quad (7)$$

$$E = \sum_{i=0}^k [p(i)^2] \quad (8)$$

$$H = - \sum_{i=0}^k p(i) \log_2[p(i)] \quad (9)$$

2.1.2 Evolução Dinâmica do Grafo

O grafo obtido a partir de uma imagem de textura é um grafo regular. Isso ocorre porque cada vértice está conectado ao mesmo número de outros vértices do grafo e esse número de vértices é limitado pelo valor do raio r . Assim, o grau de seus vértices não contém nenhuma informação desejável para a caracterização do grafo/textura.

Os autores Backes, Casanova e Bruno (2013) investigaram as propriedades do grafo em função do tempo, ou seja, consideraram que o grafo muda ao longo do tempo (da criação a extinção), bem como suas propriedades. Esse processo permitiu estudar como as propriedades mudam e suas trajetórias, e usá-las como informações adicionais para analisar e classificar uma rede complexa.

O grafo obtido a partir de uma textura é regular, mas surgem dinâmicas topológicas aplicando um limiar t sobre o conjunto de arestas E para selecionar um subconjunto E_t , $E_t \subseteq E$. Backes, Casanova e Bruno (2013) propuseram selecionar arestas de $e \in E$ para compor E_t , onde o peso da aresta, $w(e)$, seja igual ou menor ao limiar t . E_t é representado por:

$$E_t = \delta_t(E) = \{e \in E \mid w(e) \leq t\}. \quad (10)$$

A Figura 1 (a) ilustra uma imagem de exemplo com um pixel marcado. Em (b) são ilustrados os vizinhos (em cinza) do *pixel* marcado em (a) $(x, y)/I(x, y) = 52$, considerando $r = 3$. Em (c), em cinza, o resultado da aplicação do limiar com $t = 0,25$.

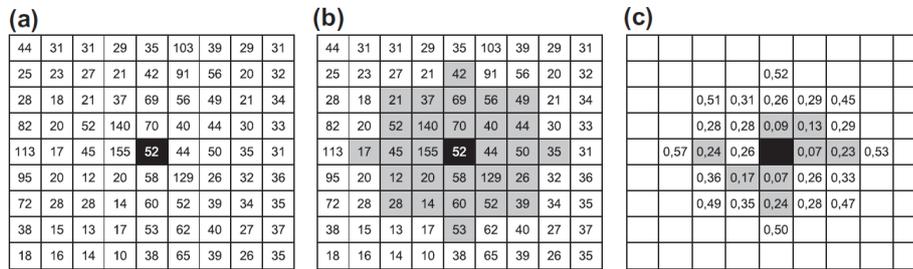


Figura 1 – Exemplo da definição do limiar: em (b) exemplo do conjunto E relativo ao *pixel* central da imagem dada em (a) considerando $r = 3$ e em (c) os pesos ($w(e)$) que estão associados aos vértices ligados. Um limiar t é aplicado ao conjunto original de arestas (neste exemplo $t = 0,25$) para executar uma transformação na rede, resultando, assim, em um subconjunto de arestas (E_t) (*pixels* em cinza) (BACKES; CASANOVA; BRUNO, 2013).

Ao considerar o conjunto original de vértices V , obtém-se um novo grafo $G_t = (V, E_t)$ representando uma etapa intermediária na evolução do grafo.

O grau de um nó contém informações sobre como o *pixel* correspondente se relaciona com seus vizinhos. Sendo D_t a matriz de graus dos nós $v_{x,y} \in V$ para um dado t . Cada $v_{x,y}$ tem o correspondente $D_t(x, y) \in D_t$, onde $D_t(x, y) = deg(v_t(x, y))$. Energia, entropia e contraste são extraídos dessas matrizes de grau para compor o vetor de característica usado para caracterizar a amostra de textura.

Para a obtenção do vetor de características, Backes, Casanova e Bruno (2013) utilizaram um conjunto T de limiares, sendo $t_{inicial} = 0,005$, com aumento gradual de $t_{incr} = 0,015$, até atingir o máximo $t_{final} = 0,530$. Segundo os autores, a escolha desse conjunto de limiares foi tomada de forma empírica. As transformações obtidas tornaram o grafo regular original um rede rede complexa.

As bases de texturas públicas Brodatz, Vistex e Outex foram utilizadas no trabalho para avaliação do vetor de atributos gerado pelo método proposto por Backes, Casanova e Bruno (2013), sendo analisados utilizando-se o classificador LDA (FUKUNAGA, 1990, p. 135) com validação cruzada *leave-one-out*.

Backes, Casanova e Bruno (2013) compararam o método proposto com: histogramas de primeira ordem, Fourier, Gabor, Matriz de coocorrência de nível de cinza GLCM, Transformada Discreta do Cosseno (DCT), Método de diferenças de nível de cinza GLDM, wavelets, Padrão Binário Local Completo (CLBP), Padrões Binários Locais Variantes (LBPV) e Padrão Ternário Local (LTP). Os autores, utilizando-se de 108 descritores, conseguiram obter taxas de acertos bem próximas aos melhores resultados, visto que o método CLBP obteve 95,32% para a Brodatz e 98,03% para a Vistex utilizando-se de 648 descritores.

2.1.3 Medidas Estatísticas Descritivas

Além das medidas estatísticas descritivas como contraste (Eq. 7), energia (Eq. 8) e entropia (Eq. 9), outras podem ser utilizadas para as análises: média (Eq. 11), variância (Eq. 12), *skewness* (Eq. 13) e *kurtosis* (Eq. 14).

$$\bar{p} = \sum_{i=0}^k p(i)i \quad (11)$$

$$\sigma^2 = \sum_{i=0}^k \frac{p(i)(i - \bar{p})^2}{k} \quad (12)$$

$$skw = \sqrt{\sigma^2}^{(-3)} \sum_{i=0}^k p(i)(i - \bar{p})^3 \quad (13)$$

$$kurt = \sqrt{\sigma^2}^{(-4)} \sum_{i=0}^k p(i)(i - \bar{p})^4 \quad (14)$$

2.2 Proposta

Backes, Casanova e Bruno (2013) representam as texturas presentes nas imagens utilizando-se de redes complexas. Utilizaram com sucesso medidas estatísticas do histograma de graus de uma rede complexa geradas a partir de uma imagem para caracterizar texturas, mostrando que o histograma pode ser uma ferramenta poderosa para tais tarefas. No entanto, o sucesso do método depende da escolha apropriada dos parâmetros para gerar o grafo a partir da textura, ou seja, o quão discriminante pode ser esse grafo depende da escolha do valor de limiar que determinam as conexões entre os nós da rede complexa. Diferentes limiares produzem grafos muito diferentes. Segundo os autores, foi definido um conjunto T de limiares para a caracterização das texturas de forma empírica.

Considerando que diferentes bases de imagens apresentam diferentes comportamentos, o conjunto de limiares definido pelo método de Backes, Casanova e Bruno (2013) para a caracterização das texturas nem sempre se adequam à essas diferenças comportamentais das bases. Com isso, propõe-se um método que consiga analisar as características presentes na base de imagens e selecionar automaticamente um conjunto de limiares que melhor consiga caracterizar as texturas presentes na base.

Baseado na proposta de Backes, Casanova e Bruno (2013) serão gerados, para cada imagem da base, os N mapas de graus obtidos por conjunto de limiares $T = \{t_1, t_2, \dots, t_N\}$ igualmente espaçados.

O método proposto para a seleção automática dos limiares é ilustrado na Figura 2 (setas em verde). Inicialmente são selecionadas, aleatoriamente, 10% das imagens da Base. Para cada imagem selecionada, esta é utilizada como imagem de entrada (a), constrói-se um grafo para cada limiar do conjunto T (G_t) (b). Calcula-se as matrizes de graus (D_t) obtidas dos grafos (G_t) para cada limiar do conjunto T . Para cada D_t são gerados histogramas e calculadas as medidas estatísticas descritivas (contraste (C), energia (E), entropia (H), média (\bar{p}), variância (σ^2), *skewness* (skw) ou *kurtosis* ($kurt$)) (c). Analisam-se as oscilações presentes nas curvas obtidas pela média das medidas estatísticas descritivas (d) (as análises das medidas estatísticas descritivas, média e mediana estão descritas na Seção 2.3.1). São selecionados todos os picos e vales locais destas curvas (e). O conjunto de picos e vales selecionados resultará em um novo conjunto de limiares S ($S \subset T$).

Após a obtenção do novo conjunto de limiares S , processa-se o restante das imagens da Base (Figura 2, setas em preto). Utilizando-se as imagens não selecionadas para a criação dos limiares, para cada imagem faz-se a entrada (a), constrói-se um grafo para cada limiar $t_i \in S$ (G_t) (b). Calcula-se as matrizes de graus (D_t) obtidas dos grafos (G_t) para cada limiar do conjunto S (c). Para cada D_t geram-se histogramas desses graus e são calculados a energia, a entropia e o contraste. Estas medidas serão concatenadas para compor o vetor de características do método proposto (f), mantendo o mesmo vetor de características proposto por Backes, Casanova e Bruno (2013).

2.3 Experimentos e Análise dos Resultados

Nesta seção serão apresentados os experimentos e análises referentes às redes complexas, apresentando as análises para a escolha da medida estatística descritiva a ser utilizada para a definição dos limiares automáticos e o desenvolvimento da proposta.

2.3.1 Definição dos Limiares

Definiu-se o conjunto $T = \{t_1, t_2, \dots, t_N\}$ de limiares igualmente espaçados, sendo $N = 190$ e $t_i = 0,005 + 0,005(i - 1)$; $i = 1, \dots, N$. Além disso, para modelar a

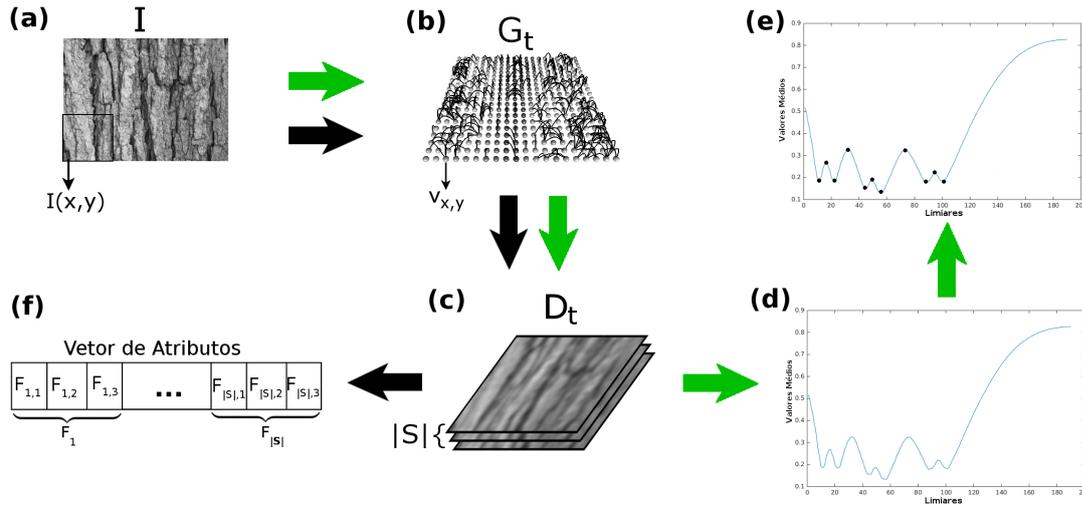


Figura 2 – Etapas do Método Proposto: (a) Imagem de entrada; (b) Rede Complexa gerada para cada valor de limiar t e amostra da textura no conjunto de dados; (c) Grau da distribuição da amostra de textura para cada limiar selecionado; (d) Exemplo do histograma médio de uma medida estatística descritiva; (e) Limiares Selecionados por meio dos picos e vales, em destaque, gerando um conjunto S de limiares; (f) Vetor de atributos composto pela concatenação dos valores da energia, entropia e contraste dos S limiares $(F_{1,1}, F_{1,2}, F_{1,3}, \dots, F_{|S|,1}, F_{|S|,2}, F_{|S|,3})$.

rede, foi utilizado um raio $r = 3$, o que resultou em um nó $v_{x,y}$ de grau máximo 28. Para que as análises não fossem influenciadas por outros parâmetros, com exceção do N e t_i , todos os parâmetros são os mesmos definidos por Backes, Casanova e Bruno (2013).

Para essas avaliações foram utilizadas seis bases de imagens descritas no Apêndice B: Brodatz, CURET, Outex, KTH-TIPS, UIUCTex e VisTex.

O objetivo inicial é avaliar a variação das medidas estatísticas em relação aos limiares. Para esta análise utilizou-se os passos do método proposto na Seção 2.2 ilustrados na Figura 2. A partir da imagem de entrada, construiu-se um grafo para cada limiar do conjunto T (G_t). Calculou-se as matrizes de graus (D_t) obtidas dos grafos (G_t) para cada limiar do conjunto T . Para cada G_t foram calculadas as sete medidas estatísticas descritivas. Realizou-se uma análise das oscilações presentes nas curvas obtidas pela média das medidas estatísticas, selecionando todos os picos e vales locais destas curvas. Cada conjunto de picos e vales, resultou em um novo conjunto de limiares S ($S \subset T$), que foram utilizados para modelar o padrão de textura para a rede complexa. Para cada rede complexa gerada utilizando-se um limiar $t_i \in S$, calculou-se as medidas descritivas de energia, entropia e contraste dos histogramas de graus das imagens. Estas medidas descritivas obtidas foram concatenadas para compor o vetor de características da base analisada.

As Figuras 3 e 4 ilustram as curvas referentes aos valores estatísticos médios obtidos para os 190 limiares para a base Brodatz.

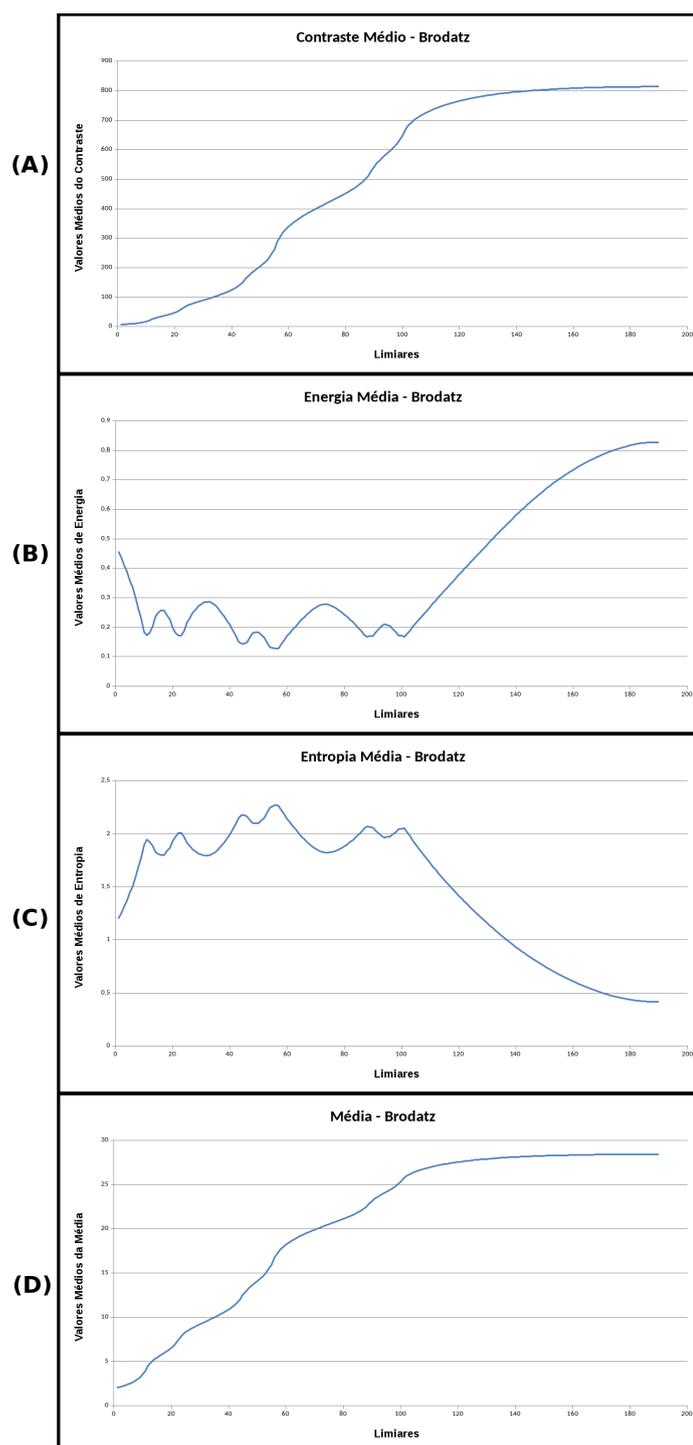


Figura 3 – Gráfico dos valores médios das imagens da base Brodatz para os 190 limiares: (A) contraste; (B) energia; (C) entropia e (D) média.

Observou-se que para as medidas estatísticas descritivas média, contraste, *kurtosis* e *skewness*, as curvas obtidas para os valores médios não apresentaram picos e vales locais, impossibilitando sua utilização para a determinação dos limiares. Observou-se um comportamento semelhante nas outras bases analisadas para essas medidas.

A Figura 5 ilustra as curvas obtidas dos valores médios para a base Vistex. Quando

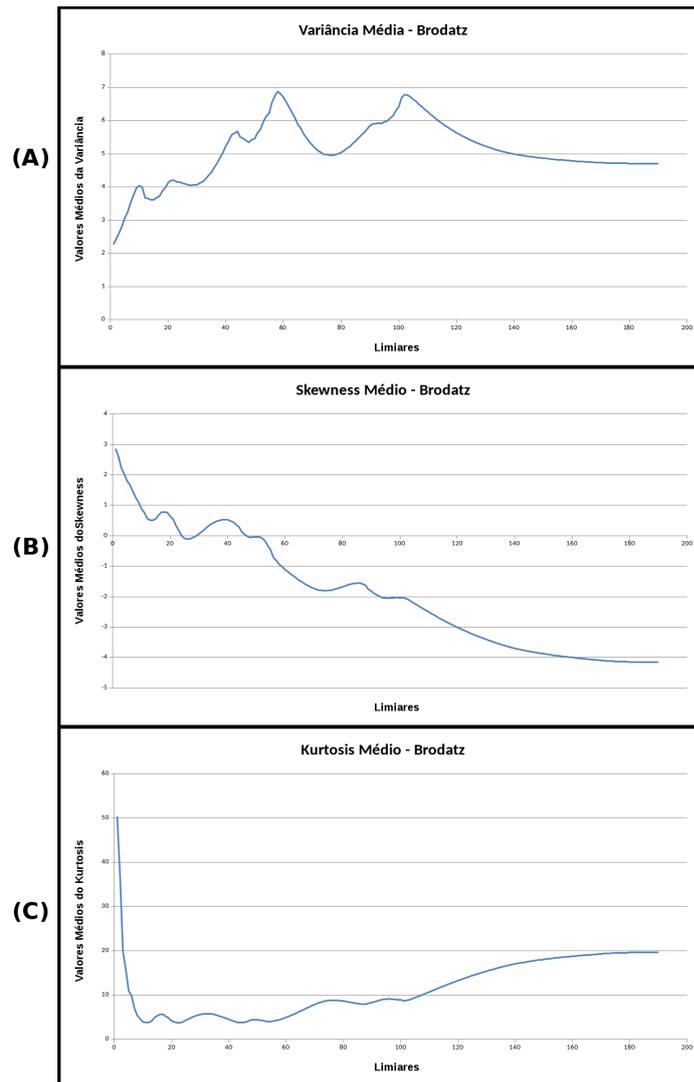


Figura 4 – Gráfico dos valores médios das imagens da base Brodatz para os 190 limiares: (A) variância; (B) *skewness* e (C) *kurtosis*.

comparadas as curvas da base Brodatz (Figuras 3 e 4) com as curvas da base Vistex (Figura 5) pode-se observar o comportamento semelhantes entre as curvas das medidas estatísticas descritivas analisadas. O mesmo comportamento se repetiu para as outras bases analisadas.

Analisando-se as sete curvas das medidas estatísticas obtidas para as bases analisadas, observa-se que apenas nas medidas descritivas de energia, entropia e variância ocorrem grandes oscilações.

A entropia é a medida da energia não disponível para a realização do trabalho e, se a entropia aumenta, uma quantidade de energia torna-se indisponível para a realização do trabalho (TIPLER; MOSCA, 2009). Analisando-se estas definições de entropia e as curvas obtidas para a entropia e para a energia, observa-se que a energia é inversamente proporcional à entropia. Ao se analisar empiricamente as curvas obtidas, observou-se que a variância contém menos oscilações que a energia e entropia e, não conseguiu representar

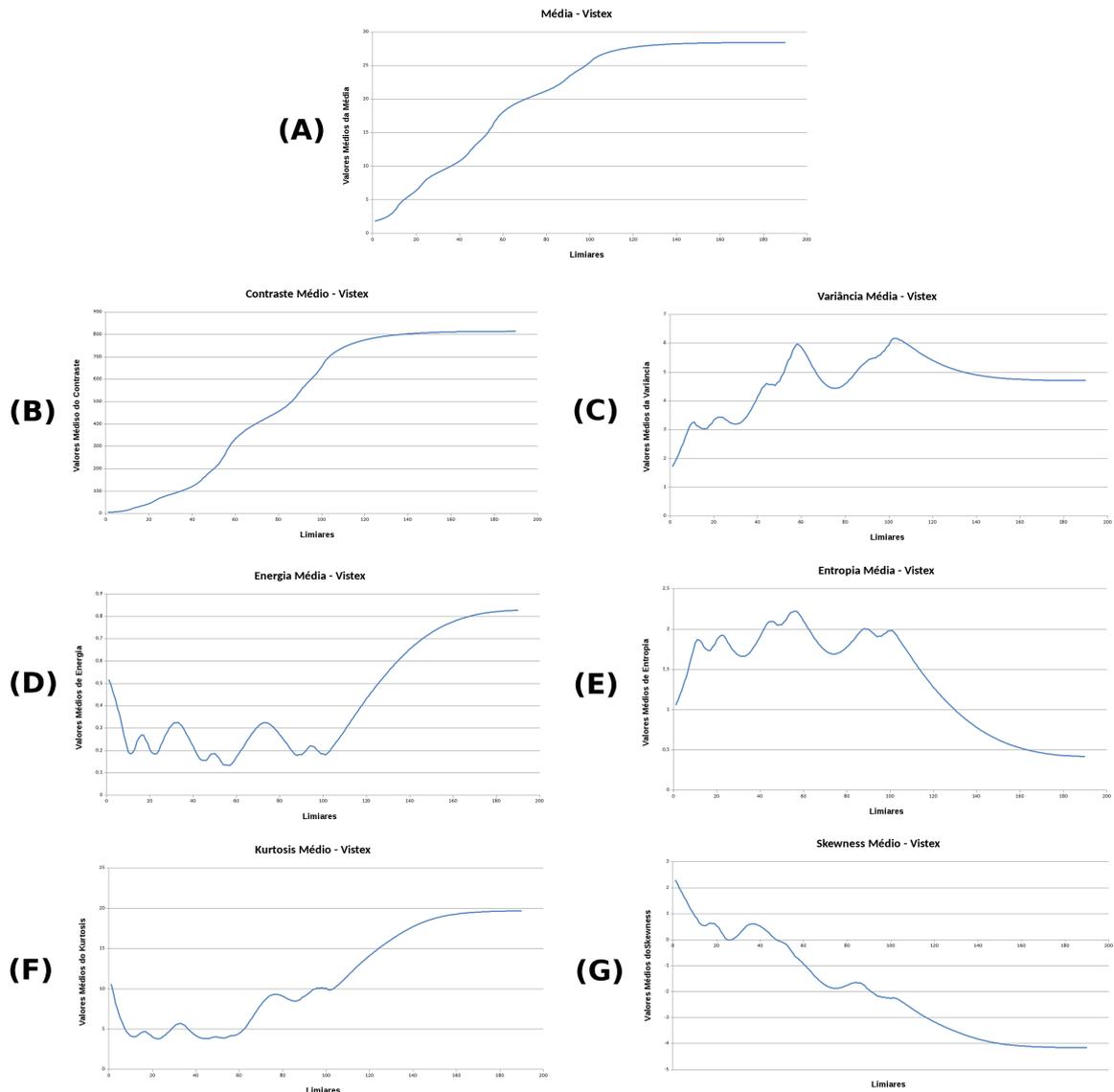


Figura 5 – Gráfico dos valores médios das medidas estatísticas das imagens da base Vistex para os 190 limiares: (A) média; (B) contraste; (C) variância; (D) energia; (E) entropia; (F) *kurtosis* e (G) *skewness*.

as bases de imagens testadas. Com isso, optou-se pela energia como medida estatística seletora para os limiares.

Após a definição do seletor de limiares, analisou-se outra abordagem para a redução do grande volume de imagens e limiares resultantes para cada base, ao invés de utilizar os valores médios seriam utilizados os valores medianos. Avaliaram-se as duas abordagens (média e mediana) nas seis diferentes bases de imagens com as sete medidas estatísticas descritivas. Manteve-se o mesmo procedimento descrito anteriormente para a escolha da energia, porém foi substituído o cálculo dos valores médios pelos valores medianos. Os vetores de características obtidos neste trabalho foram submetidos ao classificador LDA com validação cruzada *leave-one-out*.

A Tabela 1 ilustra os resultados obtidos para as duas abordagens (média e mediana) para as seis bases de imagens. Nota-se que cada abordagem atinge um número diferente de limiares (Dim.) para cada conjunto de imagens. Na maioria dos resultados alcançou-se a melhor acurácia utilizando-se os valores médios ao invés dos medianos. Observou-se que a base Brodatz não obteve o mesmo procedimento das demais bases, mas utilizou um fator maior de limiares.

Tabela 1 – Resultados da classificação para valores médios e medianos (em %) baseado na medida descritiva de energia.

Métodos	Brodatz		Vistex		Outex		KTH-TIPS		UIUCTex		CURET	
	Dim.	Ac.										
Média	36	93,30	48	96,88	42	83,60	48	84,20	54	82,20	48	78,90
Mediana	54	94,14	54	96,64	36	81,91	42	82,84	54	81,80	42	77,08

A Tabela 2 ilustra a quantidade de limiares, utilizando-se os valores médios da energia, encontrados automaticamente para cada base de imagem analisada. Observa-se que a quantidade de limiares obtida pelo método proposto é, no máximo, a metade da quantidade de 36 limiares fixos proposta por Backes, Casanova e Bruno (2013).

Tabela 2 – Tabela com a quantidade de limiares obtidos automaticamente pelo método proposto para cada base de imagem

Bases	Limiares
Brodatz	12
Outex	14
Vistex	16
CURET	16
KTH-TIPs	16
UICteX	18

Analisando-se os resultados e as curvas obtidas nos testes para as seis bases, verificou-se que ocorrem variações entre as bases de imagens analisadas para diversas medidas estatísticas descritivas e na quantidade de limiares encontrados automaticamente. Isso se deve as características peculiares de cada base. Após as análises e testes com as sete medidas estatísticas descritivas e as seis bases de imagens, optou-se para a escolha dos limiares a utilização dos picos e vales dos valores médios da energia para cada base.

2.3.2 Validação da Proposta

Para validar o método proposto de automatizar a definição do limiar t para a geração das redes complexas (Seção 2.2) foi realizada uma análise estatística utilizando LDA com validação cruzada *leave-one-out* para avaliar a precisão da abordagem proposta. Utilizou-se em todos os experimentos, LDA com validação cruzada *leave-one-out* para

que os resultados fossem comparados com o método originalmente proposto por Backes, Casanova e Bruno (2013).

Para realizar os testes, foram criados conjuntos para treinamento e teste, sendo que 10% das imagens de cada classe foram utilizadas para o treinamento. Aplicou-se 10 vezes o esquema de treinamento e testes e, os resultados apresentados foram a média das 10 acurácias obtidas.

Também foram analisados alguns métodos do estado da arte para análise de textura. A Tabela 3 ilustra os resultados do método proposto e de métodos selecionados da literatura para as seis bases de imagens. Os resultados apresentados para os métodos de Varma e Zisserman (2004), Crosier e Griffin (2010), Perea e Carlsson (2014) e Zhang et al. (2015b) foram obtidos diretamente dos respectivos artigos. Dos resultados tabelados, foram implementados os métodos de primeira ordem, wavelets, LBP, Gabor, Backes, Casanova e Bruno (2013) e o método proposto e, para esses métodos foi utilizado o LDA com validação cruzada *leave-one-out*. Observa-se na Tabela 3 que o número de descritores para todos os métodos é fixa, com exceção do método proposto que, para cada base de imagem, calculou um conjunto próprio de limiares. Com isso, o método proposto utilizou 12 limiares, o que resultou em 36 descritores para a base Brodatz, utilizou 14 limiares, o que resultou em 42 descritores para a base Outex, utilizou 16 limiares, o que resultou em 48 descritores para as bases Vistex, CURET e KTH-TIPS e utilizou-se 18 limiares, o que resultou em 54 descritores para a base UIUCTex.

Tabela 3 – Resultados do método proposto e de métodos da literatura para as seis bases de imagens.

Métodos	Nº de Descritores	Acurácia (%)					
		Brodatz	Vistex	CURET	Outex	UIUCTex	KTH-TIPS
Primeira Ordem	5	40,20	52,66	19,73	54,63	58,30	57,78
Wavelets	36	70,27	72,11	51,00	-	57,70	58,52
LBP	25	89,58	88,89	67,37	77,87	77,00	74,07
Gabor	64	82,49	91,67	82,61	78,60	59,50	80,12
(BACKES; CASANOVA; BRUNO, 2013)	108	95,27	98,03	84,32	86,76	86,50	90,37
(PEREA; CARLSSON, 2014)	215	-	-	95,66 ± 0,45	-	91,23 ± 1,13	94,77 ± 1,3
(ZHANG et al., 2015b)	648	-	-	94,44 ± 1,13	-	-	98,86 ± 1,12
(CROSIER; GRIFFIN, 2010)	1.296	-	-	98,6 ± 0,2	-	98,8 ± 0,5	98,5 ± 0,7
(VARMA; ZISSERMAN, 2004)	2.440	-	-	97,43	-	-	-
Método Proposto	36, 48, 48, 42, 54, 48	93,30	96,88	78,90	83,60	82,20	84,20

Analisando-se os resultados para a base Brodatz, comparando o resultado de Backes, Casanova e Bruno (2013) com o método proposto, observa-se que o método proposto obteve uma acurácia inferior de 1,97%, porém quando comparada a quantidade de limiares utilizada por cada método, o método proposto utilizou apenas 1/3 (36) da quantidade de descritores utilizado por Backes, Casanova e Bruno (2013).

Além dos experimentos explicitados anteriormente, foram realizados testes, para analisar a Equação 2, considerando o valor absoluto $\frac{|I(x, y) - I(x', y')|}{L}$ ao invés de

$\frac{I(x, y) - I(x', y')}{L}$. Tal modificação, apesar de parecer mais adequada, não consegue expressar a direção da mudança de intensidades entre $I(x, y)$ e $I(x', y')$, o que influenciou nos resultados conforme pode ser observado na Tabela 4. Para fins comparativos, foram copiados os valores obtidos para o método proposto da Tabela 3.

Tabela 4 – Resultados da Classificação com e sem a utilização do valor absoluto (em %).

	Com Absoluto		Sem Absoluto	
	Descritores	Acurácia	Descritores	Acurácia
Brodatz	33	90,82	36	93,30
Outex	33	76,15	42	83,60
Vistex	33	92,79	48	96,88
CUReT	33	67,63	48	78,90
KTH-TIPS	33	80,54	48	84,20
UIUCTex	33	78,16	54	82,20

Observa-se que para todas as bases testadas, a não utilização do absoluto obteve-se melhores resultados sem gerar um grande aumento na quantidade de descritores, conforme já havia sido demonstrado anteriormente.

2.3.3 Alternativa para a Definição de Limiares

A proposta de utilização de picos e vales para a definição automática de limiares obteve bons resultados, conforme apresentado anteriormente. Conceitualmente, quanto mais conexões um vértice de um grafo tem, melhor representada por este grafo a imagem estará. Com isso, quanto maior o número de conexões, mais próximo de zero o gráfico da energia estará.

Partindo desse princípio, utilizaram-se todos os vales (mínimos), retirando-se os picos da proposta inicial. A Tabela 5 ilustra os resultados obtidos com o método publicado em Ribeiro et al. (2015), utilizando picos e vales, e os resultados da utilização dos vales (mínimos) para a definição de limiares automáticos. Observa-se que com a utilização dos vales ocorreu uma redução significativa na quantidade de descritores e uma redução menos expressiva na acurácia, por exemplo, apenas 3,01% de redução na acurácia da Base Brodatz, por exemplo. Estas reduções

A Tabela 6 ilustra os resultados do método proposto com vales, do método com picos e vales e de métodos selecionados da literatura para as seis bases de imagens. Os resultados apresentados para os métodos de Varma e Zisserman (2004), Crosier e Griffin (2010), Perea e Carlsson (2014) e Zhang et al. (2015b) foram obtidos diretamente dos respectivos artigos. Dos resultados tabelados, foram implementados os métodos de primeira ordem, wavelets, LBP, Gabor, Backes, Casanova e Bruno (2013), os métodos propostos de vales e picos e vales. Observa-se na Tabela 6 que o número de descritores para todos os métodos

Tabela 5 – Comparativo dos Resultados da Classificação com entre as propostas para a definição automática de limiares (em %).

	Picos e Vales		Vales	
	Descritores	Acurácia	Descritores	Acurácia
Brodatz	36	93,30	19	90,29
Vistex	48	96,88	23	93,60
CUReT	48	78,90	25	69,12
Outex	42	83,60	21	81,15
UIUCTex	54	82,20	28	73,95
KTH-TIPS	48	84,20	24	75,52

é fixa, com exceção dos métodos propostos que, para cada base de imagem, calculou um conjunto próprio de limiares.

Observa-se que os métodos propostos obtiveram valores próximos de acurácia, porém a quantidade de descritores variou. Além disso, quando comparado o método proposto de vales aos demais métodos, em sua maioria, a acurácia do método não ficou muito abaixo dos outros métodos com uma grande redução na quantidade de descritores.

Tabela 6 – Resultados dos Métodos Propostos e de Métodos da literatura para as seis bases de imagens.

Métodos	Nº de Descritores	Acurácia (%)					
		Brodatz	Vistex	CUReT	Outex	UIUCTex	KTH-TIPS
Primeira Ordem	5	40,20	52,66	19,73	54,63	58,30	57,78
Wavelets	36	70,27	72,11	51,00	-	57,70	58,52
LBP	25	89,58	88,89	67,37	77,87	77,00	74,07
Gabor	64	82,49	91,67	82,61	78,60	59,50	80,12
(BACKES; CASANOVA; BRUNO, 2013)	108	95,27	98,03	84,32	86,76	86,50	90,37
(PEREA; CARLSSON, 2014)	215	-	-	95,66 ± 0,45	-	91,23 ± 1,13	94,77 ± 1,3
(ZHANG et al., 2015b)	648	-	-	94,44 ± 1,13	-	-	98,86 ± 1,12
(CROSIER; GRIFFIN, 2010)	1.296	-	-	98,6 ± 0,2	-	98,8 ± 0,5	98,5 ± 0,7
(VARMA; ZISSERMAN, 2004)	2.440	-	-	97,43	-	-	-
Método de Picos e Vales	36, 48, 48, 42, 54, 48	93,30	96,88	78,90	83,60	82,20	84,20
Método de Vales	19, 23, 25, 21, 28, 24	90,29	93,60	81,15	75,52	73,95	69,12

Ao se analisar os resultados, observa-se que ainda necessita ser expandida a proposta do uso de vales para conseguir obter acurácias ainda melhores, além de se testar a proposta em outras bases de texturas.

2.4 Considerações Finais

Baseado no trabalho de Backes, Casanova e Bruno (2013), foi proposta a modificação na escolha dos limiares, tornando-os automatizados e flexíveis conforme a base de texturas a ser analisada.

Foram avaliadas diversas medidas estatísticas descritivas sendo a energia (Eq. 8) a medida utilizada nos experimentos apresentados para auxiliar na definição dos limiares. Com isso, foram selecionados os picos e vales das curvas obtidas com a energia para as bases. Através de conjuntos de treinamento e teste, obteve-se os limiares automáticos e realizou-se os testes.

Os resultados indicam que os métodos propostos (vales ou picos e vales) apresentam um alto desempenho consistente em todas as bases de textura testadas. O Método baseado em picos e vales, dentre os seis resultados obtidos, obteve-se a maior taxa de acertos para a base Vistex e o segundo melhor resultado na base Brodatz comparando com os outros métodos. Quando comparados com os melhores resultados obtidos na Tabela 6, os métodos propostos utilizaram menos do que a metade do número de descritores para a formação dos vetores característicos e o valor da acurácia não reduziu significativamente.

Ao se analisar o método proposto baseado apenas em vales (mínimos) obteve-se melhores resultados quando comparado ao método proposto utilizando picos e vales (Tabela 6). Com isso, a proposta baseada apenas em vales está sendo preparada para publicação.

O tamanho do vetor descritor é uma característica importante a ser considerada em um método de modelagem matemática, especialmente nos últimos anos em que os sistemas embarcados de baixo custo e os problemas de *Big Data* tornaram-se relevantes. Se dois descritores oferecem taxas de classificação semelhantes, um método que apresenta um vetor de característica de dimensões muito menores é preferível, pois ocupa menos espaço de armazenamento e memória e geralmente é mais rápido para processar e classificar. Os métodos propostos neste trabalho produziram vetores de características menores do que outros métodos do estado da arte, enquanto manteve a acurácia das taxas de classificação em várias bases de imagens, conforme demonstrado pelos experimentos realizados. Este trabalho também apresentou como as medidas de rede complexas podem ser uma ferramenta poderosa para descrever adequadamente texturas. Neste caso, a análise estatística do grau de histograma produziram informações não apenas sobre a natureza geral dos dados, mas também sobre os limiares mais significativos. Através da seleção de limiar automatizado, os métodos propostos podem se adaptar às particularidades de cada conjunto específico de dados gerando conjuntos próprios de limiares. A redução da dimensionalidade é, portanto, baseada na complexidade dos dados.

Projeções Topológicas

Diversos trabalhos da literatura propõem a caracterização de texturas por diversos meios e, dentre eles, Perea e Carlsson (PEREA; CARLSSON, 2014) propuseram a utilização da projeção de pequenos pedaços da textura de tamanho $n \times n$ no espaço topológico da garrafa de Klein, criando um descritor multi-escala e invariante à rotação.

Neste capítulo são apresentados os resultados da análise de diferentes combinações de tamanhos $n \times n$ que chamamos de escala (n) na composição do vetor de característica. Além dessa análise, outros parâmetros e passos que são necessários para criação do descritor foram investigados e propôs-se modificações que melhoraram a caracterização por meio do método.

A organização deste capítulo contempla a fundamentação teórica com uma revisão da bibliografia e trabalhos do estado-da-arte sobre Projeções Topológicas, além de uma análise dos trabalhos relacionados da literatura, a apresentação das modificações nas escalas objetivando melhor caracterizar as bases de imagens. São apresentadas as análises para a obtenção dos pedaços das imagens, a invariância à rotação, o vetor de características, a variação das escalas, validações, resultados e as conclusões obtidas.

3.1 Fundamentação Teórica

Diversos trabalhos presentes na literatura abordam o uso de topologia para representar problemas com imagens e alta dimensionalidade de dados.

Lum et al. (2013) utilizaram métodos topológicos para analisar conjuntos complexos de dados com alta dimensionalidade, extraíndo e obtendo informações por meio de padrões. O método combina análise da componente principal e análises de cluster de conjuntos de dados complexos para fornecer uma representação geométrica desses dados. O método que os autores propuseram, além de analisar conjuntos de dados individualmente, analisa as relações entre esses conjuntos de dados. Os autores aplicaram o método em três tipos de dados de áreas diferentes: dois conjuntos de expressão de genes diferentes em tumores mamários; basearam-se em dados de 22 anos (1990 - 2011) de comportamento de votação

dos membros da Câmara dos Deputados dos Estados Unidos e, um conjunto de dados das estatísticas de desempenho de 425 jogadores de basquete da NBA (*National Basketball Association*) da temporada de 2010 - 2011. Segundo os autores, foi possível identificar subgrupos sutis de câncer de mama por meio das formas das redes de expressão gênica, conseguindo comparar visualmente formas em várias redes para câncer de mama. No caso do comportamento da votação, o formato das redes geradas ao longo dos anos indicam o quão coesivo ou fragmentado é o comportamento eleitoral para a Câmara dos Deputados dos EUA. No caso da NBA, conseguiu-se identificar 13 estilos diferentes de jogos utilizados pelos jogadores naquela temporada.

Segundo Carlsson et al. (2008) estatísticas de imagens naturais é um campo extremamente importante para diversas áreas como a visão computacional, a neurociência e a fisiologia, por exemplo. Os autores propuseram o estudo do comportamento local de imagens por meio da análise de um espaço de pequenas regiões de alto contraste extraídas (*patches*) a partir das imagens.

O estudo das propriedades geométricas e estatísticas do espaço de pequenos *patches* de imagens naturais foi realizado por Lee, Pedersen e Mumford (2003).

Um dos objetivos era entender como *patches* de tamanho 3×3 , de alto contraste era distribuídos no espaço de *pixel* \mathbb{R}^9 .

Para modelar a distribuição foi criado um conjunto \mathcal{M} de dados e este estudo forneceu evidências da existência de uma variedade de dimensão 2 em torno da qual os pontos de \mathcal{M} se acumulam com alta densidade. Este é denominado modelo de Lee.

Posteriormente, em (CARLSSON et al., 2008), usando todo o formalismo da homologia foi detectado a existência de um núcleo coletor de alta densidade e dimensão 2, subjacente ao conjunto \mathcal{M} . Foi provado que esse coletor possuía a topologia de uma garrafa de Klein e que consistia dos *patches* do modelo de Lee, bem como barras de iluminação claras e escuras em todas as orientações. Foi também fornecido por Carlsson et al. (2008) uma parametrização para este espaço em termos de grau 2. O resultado principal foi que 50% dos pontos na superfície \mathcal{M} se acumulavam com alta densidade em uma outra superfície \mathcal{K} com a topologia da garrafa de Klein.

Existem vários modelos para espaços com essa topologia. Perea e Carlsson (2014) apresentaram uma técnica para caracterização de texturas de imagens por meio da projeção de pequenos pedaços da imagem no espaço topológico da garrafa de Klein.

A proposta foi criar um descritor multi-escala e invariante à rotação conseguindo estimar e representar distribuições em torno do espaço da garrafa de Klein, pois utilizando-se amostras (*patches*) de alto contraste, com dimensão $n \times n$, de uma imagem em escala de cinza e, utilizando-se da topologia da garrafa de Klein ao projetar as amostras, estas se acumulam com uma alta densidade em torno da superfície da garrafa.

Perea e Carlsson (2014) explicitam em seu artigo que a maioria dos *patches* podem ser continuamente projetadas no modelo do espaço da garrafa de Klein. O modelo

apresentado dos dados considerou o espaço K definido pelo retângulo $R = \left[\frac{\pi}{4}, \frac{5\pi}{4} \right] \times \left[\frac{-\pi}{2}, \frac{3\pi}{2} \right]$ onde cada ponto $\left(\alpha, \frac{-\pi}{2} \right)$ é identificado com $\left(\alpha, \frac{3\pi}{2} \right)$ e cada $\left(\frac{\pi}{4}, \theta \right)$ é identificado com $\left(\frac{5\pi}{4}, \pi - \theta \right)$ para todo $(\alpha, \theta) \in R$, conforme ilustrado na Figura 6(A). Com isso conseguiu-se criar o espaço \mathcal{K} com a topologia da garrafa de Klein. \mathcal{K} é composto de *patches* representando arestas e barras em todas as direções. Esses *patches* podem ser parametrizados por meio de sua direção $\alpha \in \left[\frac{\pi}{4}, \frac{5\pi}{4} \right]$ e pela sua estrutura de borda definida pelo ângulo θ de transição. Diferentes pares de (α, θ) podem descrever uma mesma amostra, por exemplo $\left(\frac{\pi}{4}, 0 \right)$ e $\left(\frac{5\pi}{4}, \pi \right)$ descrevem os degraus de bordas com direção do gradiente em noroeste.

A Figura 6(B) exemplifica os *patches* projetados no espaço \mathcal{K} . A representação visual da garrafa de Klein obtida por meio do MathWorks MATLAB[®] pelo comando `xpklein` está ilustrada pela Figura 7.

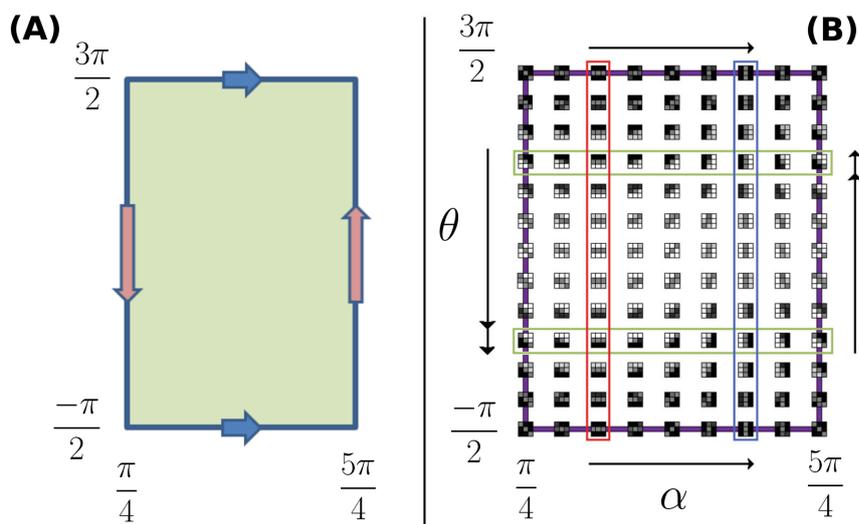


Figura 6 – Garrafa de Klein: (A) Modelo da Garrafa de Klein; (B) Exemplo de um resultado da projeção no espaço \mathcal{K} . Modificado de (PEREA; CARLSSON, 2014).

Utilizando-se dessa projeção, o conjunto de *patches* produz uma amostra $S \subset K$. Consegue-se construir bases ortonormais para $L^2(K, \mathbb{R})$. Sendo $f : K \rightarrow \mathbb{R}$ a função de densidade de probabilidade da amostra S , os seus coeficientes em relação a qualquer base ortonormal para $L^2(K, \mathbb{R})$ podem ser estimados a partir de S (ver detalhes em (REED; SIMON, 1980; SILVERMAN, 1986)).

Os passos simplificados realizados por Perea e Carlsson (2014) para conseguir gerar o descritor multi-escala e invariante à rotação estão ilustrados na Figura 8. Inicialmente seleciona-se a imagem e extrai-se os *patches*. Estes são projetados no espaço da garrafa de Klein. Utilizando-se dos *patches* projetados, cria-se um vetor para descrever a imagem por meio de um descritor baseado em Fourier. Este processo é realizado para todas

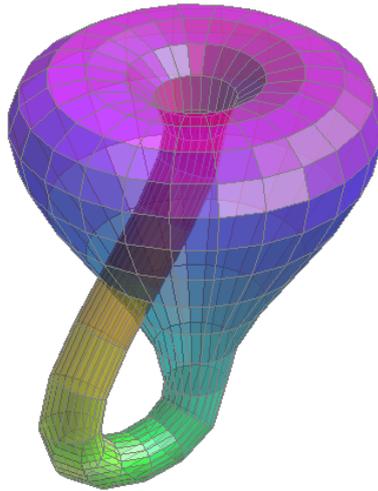


Figura 7 – Representação visual da garrafa de Klein.

as imagens da base de texturas. Tendo gerado, para todas as imagens, os vetores que descrevem as imagens, este conjunto de vetores é submetido a um classificador. Como resultado, pode-se analisar da acurácia obtida pelo método.

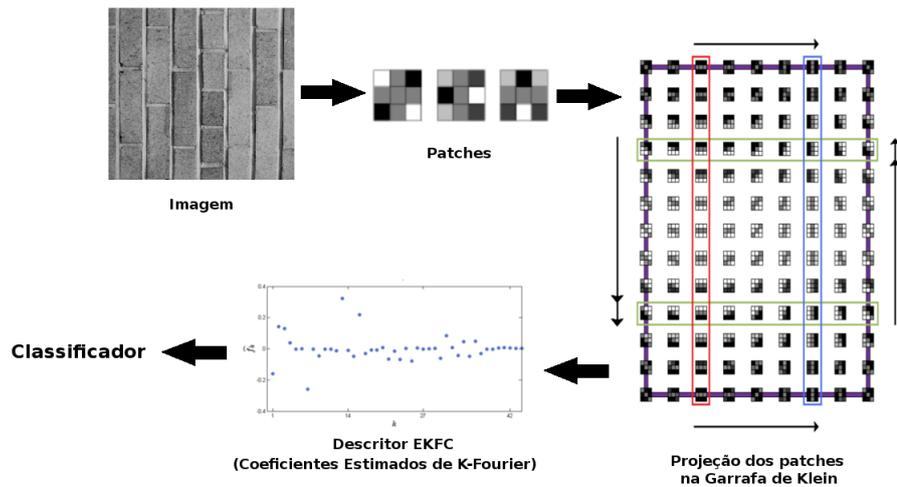


Figura 8 – Fluxo das etapas do processamento

3.1.1 Projeção

Para realizar a projeção na garrafa de Klein é necessário calcular o par (α, θ) para cada *patch* e projetá-lo na superfície \mathcal{K} .

Para obter $(\alpha, \theta) \in K$, que determina um único *patch*, tomamos $a + ib = e^{i\alpha}$ e $c + id = e^{i\theta}$. a, b, c e d serão obtidos via aproximação da função intensidade por um polinômio ortogonal de grau 2, dado por

$$p(x, y) = c \frac{(ax + by)}{2} + d \frac{\sqrt{3}(ax + by)^2}{4}, \quad a^2 + b^2 = c^2 + d^2 = 1. \quad (15)$$

Inicialmente extrai-se *patches* (a_{ij}) da imagem para a projeção, onde

$$a_{ij} = \int_{1-\frac{2i}{n}}^{1-\frac{2i-2}{n}} \int_{-1+\frac{2j-2}{n}}^{-1+\frac{2j}{n}} p(x, y) dx dy, \quad i, j = 1, \dots, n. \quad (16)$$

A extração desses *patches* baseia-se no trabalho de Lee, Pedersen e Mumford (2003), no qual retira-se *patches* com dimensões $n \times n$ *pixels*. Após a extração são selecionados uma quantidade aleatoriamente de *patches*.

Para cada um dos *patches* selecionados, $A = [a_{ij}]$, aplica-se o log e calcula-se a D-norma utilizando a Equação 17.

$$\|A\|_D^2 = \sum_{ij \sim kl} (a_{ij} - a_{kl})^2, \quad (17)$$

com $a_{ij} \sim a_{kl}$ se e somente se $|i - k| + |j - l| \leq 1$.

Considera-se todos os *patches* com D-norma maior ou igual a um certo limiar (os autores utilizaram 0.01). Seleciona-se então, um subconjunto de *patches* relevantes mantendo apenas os 1.000 *patches* com maior valor de D-norma ou todos os que restaram, caso tenha menos que 1.000 elementos.

Em seguida subtrai-se a média e normaliza-se cada *patch*, obtendo-se, como *patch* $P = [p_{ij}]$ tal que

$$\mathbf{P} = \frac{A - \frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^n a_{ij} B}{\left\| A - \frac{1}{n^2} \sum_{j=1}^n \sum_{i=1}^n a_{ij} B \right\|_D} \quad (18)$$

onde B é uma matriz $n \times n$ com todos os elementos iguais a 1.

Seja I_P a função intensidade relacionada com P por meio da média local

$$p_{ij} = \int_{1-\frac{2i}{n}}^{1-\frac{2i-2}{n}} \int_{-1+\frac{2j-2}{n}}^{-1+\frac{2j}{n}} I_P(x, y) dx dy, \quad i, j = 1, \dots, n. \quad (19)$$

Tem-se que ∇I_P (gradiente da função intensidade) é seccionalmente constante e igual ao gradiente discreto de P , com isso, $\nabla I_P \cong \nabla P$. Para a discretização utiliza-se a discretização centrada, que será exposta a seguir.

Se $2 \leq i, j \leq n - 1$ então

$$\nabla \mathbf{P}(i, j) = \frac{1}{2} \begin{bmatrix} p_{i+1, j} - p_{i-1, j} \\ p_{i, j+1} - p_{i, j-1} \end{bmatrix} \quad (20)$$

$$HP(i, j) = \begin{bmatrix} p_{i+1, j} - 2p_{i, j} + p_{i-1, j} & H_{xy}P(i, j) \\ H_{xy}P(i, j) & p_{i, j+1} - 2p_{i, j} + p_{i, j-1} \end{bmatrix} \quad (21)$$

onde

$$H_{xy}P(i, j) = \frac{p_{i+1, j+1} - p_{i-1, j+1} - p_{i+1, j-1} + p_{i-1, j-1}}{4}. \quad (22)$$

Se $i \in \{1, n\}$ ou $t \in \{1, n\}$, então existe um único $(i, j) \in \{2, \dots, n-1\}^2$ que minimiza $|r - i| + |t - j|$, tendo

$$\nabla \mathbf{P}(r, t) = \nabla \mathbf{P}(i, j) + HP(i, j) \begin{bmatrix} r - i \\ t - j \end{bmatrix} \quad (23)$$

Utilizando a expansão de Taylor de primeira ordem, aproximou-se o gradiente de P em (r, t) próximo a (i, j) . Para o gradiente da expansão I_P , fazemos $\nabla \mathbf{I}_P(x, y) = \nabla \mathbf{P}(i, j)$ se

$$\left| x - \left(-1 + \frac{2j-1}{n} \right) \right| + \left| y - \left(1 - \frac{2i-1}{n} \right) \right| < \frac{1}{n} \quad (24)$$

para algum $(i, j) \in \{1, \dots, n\}^2$, e $\mathbf{0}$ caso contrário.

Vide Seção A.1.3 do Apêndice A para detalhes de implementação.

Se os autovalores da matriz associada A_P (Equação 25) são distintos e reais, então define-se $\alpha_P \in \left[\frac{\pi}{4}, \frac{5\pi}{4} \right)$ como a direção do autoespaço correspondente ao maior autovalor e, caso contrário, o *patch* é descartado.

$$A_P(i, j) = \iint_{[-1, 1]^2} \frac{\partial I_P}{\partial x_i} \frac{\partial I_P}{\partial x_j} dx dy \quad i, j = 1, 2, \quad (25)$$

onde

$$\iint_{[-1, 1]^2} \frac{\partial I_P}{\partial x_i} \frac{\partial I_P}{\partial x_j} dx dy \approx \sum_{j=1}^n \sum_{i=1}^n \frac{\partial P}{\partial x_i}(i, j) \frac{\partial P}{\partial x_j}(i, j) \quad (26)$$

e $x_1 = x$ e $x_2 = y$.

Tendo encontrado α_P e sendo $a + ib = \cos \alpha_P + i \sin \alpha_P = e^{i\alpha_P}$, encontra-se a e b .

Seja

$$\langle f, g \rangle_D = \iint_{[-1, 1]^2} \langle \nabla f(x, y), \nabla g(x, y) \rangle dx dy \quad (27)$$

o produto interno que induz a D-norma $\|\cdot\|_D$.

Se $u = \frac{(ax + by)}{2}$, então o vetor $\begin{bmatrix} c^* \\ d^* \end{bmatrix} \in S^1$ que minimiza:

$$\Phi(c, d) = \left\| I_P - (cu + d\sqrt{3}u^2) \right\|_D, \quad c^2 + d^2 = 1 \quad (28)$$

é dado por

$$c^* = \frac{\langle I_P, u \rangle_D}{\sqrt{\langle I_P, u \rangle_D^2 + 3 \langle I_P, u^2 \rangle_D^2}} \quad d^* = \frac{\sqrt{3} \langle I_P, u^2 \rangle_D}{\sqrt{\langle I_P, u \rangle_D^2 + 3 \langle I_P, u^2 \rangle_D^2}} \quad (29)$$

sempre que

$$\varphi(I_P, \alpha) = \langle I_P, u \rangle_D^2 + 3 \langle I_P, u^2 \rangle_D^2 \neq 0 \quad (30)$$

e tal cálculo determina um único $\theta_P \in \left[\frac{-\pi}{2}, \frac{3\pi}{2} \right)$ tal que $c^* + id^* = \cos(\theta_P) + i \sin \theta_P = e^{i\theta_P}$.

$$\begin{aligned} \Phi(c^*, d^*) &= \left\| I_P - (c^*u + d^*\sqrt{3}u^2) \right\|_D \\ &= \sqrt{2 \left(1 - \sqrt{\langle I_P, \nabla u \rangle^2 + 3 \langle I_P, \nabla u^2 \rangle^2} \right)} \\ &= \sqrt{2 \left(1 - \sqrt{\varphi(I_P, \alpha_P)} \right)} \end{aligned} \quad (31)$$

$\Phi(c^*, d^*)$ pode ser pensado como a distância de P a K e $\Phi(c^*, d^*) \leq \sqrt{2}$.

Tendo encontrado a, b, c e d para um *patch* P , consegue-se obter α_P e θ_P considerando $a + ib = e^{i\alpha_P}$ e $c + id = e^{i\theta_P}$.

Após encontrar (α_P, θ_P) para cada um dos *patches* selecionados, consegue-se projetar estes no espaço \mathcal{K} .

Na Seção A.1.1 do Apêndice A está a demonstração para a obtenção da Equação 31.

3.1.2 Coeficientes Estimados de K-Fourier

Com os *patches* já projetados, calculam-se os coeficientes estimados de K-Fourier, sendo a estimativa correspondente a \hat{f} para a função de densidade de probabilidade $f : K \rightarrow \mathbb{R}$ em que

$$\hat{f}(\alpha, \theta) = \sum_{k \in N_\omega} \hat{f}_k \phi_k(\alpha, \theta) \quad (32)$$

onde $\{\phi_k\}_{k \in \mathbb{N}}$ é a base trigonométrica para $L^2(K, \mathbb{R})$ e $N_\omega \subset \mathbb{N}$ é tal que ϕ_k tem frequência menor ou igual a ω se $k \in N_\omega$.

Sejam M um espaço métrico geral e η uma métrica sobre M . Encontrar boas bases ortonormais para $L^2(M)$ quando (M, η) é um espaço métrico geral, é algo difícil. Mesmo quando há garantias teóricas para a sua existência, não é trivial realizar essa escolha. Neste caso, ter a garrafa de Klein como espaço subjacente torna-se crucial para a análise.

Um dos resultados mais fundamentais na teoria de funções periódicas de quadrado integráveis é o Teorema de Fourier, em que o conjunto $\{e^{in\alpha} : n \in \mathbb{Z}\}$ de exponenciais complexas é uma base ortonormal para $L^2(S^1)$, na qual $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Seja $\{e^{in\alpha + im\theta} : n, m \in \mathbb{Z}\}$ uma base ortonormal para $L^2(T)$, em que $T = S^1 \times S^1$ denota o toro bidimensional. Com isso, é possível escolher bases convenientes para $L^2(S^1)$, incluindo, mas não limitando-se a exponenciais trigonométricas e, assim, fornecer bases para $L^2(T)$. Como a garrafa de Klein pode ser coberta por T via identificações simples baseadas em

simetrias, encontrar bases para $L^2(K)$ será encontrar bases para $L^2(T)$ tendo em mente as identificações.

Essas simetrias surgem naturalmente do modelo da garrafa de Klein baseado em *patches* do tipo borda e barra. No Apêndice A.1.2 está o detalhamento da obtenção dos Coeficientes de K-Fourier de $K\mathcal{F}(f)$. Seja $S = \{(\alpha_1, \theta_1), \dots, (\alpha_N, \theta_N)\} \subset K$ a amostra obtida da projeção dos $n \times n$ *patches* sobre a Garrafa de Klein. A parametrização $\widehat{K\mathcal{F}}_\omega(f, S)$ consiste nos coeficientes estimados de K-Fourier da função densidade de probabilidade $f(\alpha, \theta)$ que possuem frequência menor ou igual a ω .

Perea e Carlsson (2014) selecionaram uma frequência de corte $\omega = 6$ (WIT; FLORIANI, 1998) de modo que $\widehat{K\mathcal{F}}_\omega(f, S)$ exclua o primeiro regime onde \hat{f} tenda a zero e inclua os coeficientes com alta variância. As altas frequências dos Coeficientes de K-Fourier ajustam os detalhes finos da aproximação da série e, com isso, as informações mais relevantes de $f(\alpha, \theta)$ são codificadas nos termos de baixa frequência.

Se N é a quantidade de *patches* projetados, conforme Corolário 3, o conjunto $\{\phi_k\}_{k \in \mathbb{N}}$ é formado pelos elementos

$$\begin{aligned} & 1, \\ & \sqrt{2} \cos \left(m\theta_P - \frac{(1 - (-1)^m)\pi}{4} \right), \\ & \sqrt{2} \cos(2n\alpha_P), \\ & \sqrt{2} \sin(2n\alpha_P), \\ & 2 \cos(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{m+n})\pi}{4} \right), \\ & 2 \sin(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{m+n})\pi}{4} \right). \end{aligned}$$

e será a base trigonométrica para $L^2(K, \mathbb{R})$.

Os coeficientes estimados de K-Fourier serão:

$$\hat{a}_m = \frac{1}{N} \sum_{k=1}^N \sqrt{2} \cos \left(m\theta_k - \frac{(1 - (-1)^m)\pi}{4} \right), \quad (33)$$

$$\hat{b}_n = \frac{1}{N} \sum_{k=1}^N \sqrt{2} \cos(2n\alpha_k), \quad (34)$$

$$\hat{c}_n = \frac{1}{N} \sum_{k=1}^N \sqrt{2} \sin(2n\alpha_k), \quad (35)$$

$$\hat{d}_{n,m} = \frac{1}{N} \sum_{k=1}^N 2 \cos(n\alpha_k) \cos \left(m\theta_k - \frac{(1 - (-1)^{m+n})\pi}{4} \right), \quad (36)$$

$$\hat{e}_{n,m} = \frac{1}{N} \sum_{k=1}^N 2 \sin(n\alpha_k) \cos \left(m\theta_k - \frac{(1 - (-1)^{m+n})\pi}{4} \right). \quad (37)$$

onde a somatória é sobre todos os $N(\alpha_{\mathbf{P}}, \theta_{\mathbf{P}})$ patches selecionados.

A frequência (grau) de cada um dos coeficientes estimados de K-Fourier será dada por:

$$\begin{aligned} \text{grau}(\hat{a}_m) &= m \\ \text{grau}(\hat{b}_n) &= n \\ \text{grau}(\hat{c}_n) &= n \\ \text{grau}(\hat{d}_{mn}) &= m + n \\ \text{grau}(\hat{e}_{mn}) &= m + n \end{aligned}$$

Seja $C(I) = \{\hat{a}_m, \hat{b}_n, \hat{c}_n, \hat{d}_{nm}, \hat{e}_{nm}\}$ o conjunto dos coeficientes estimados de K-Fourier da imagem I . Dota-se este conjunto de uma ordem parcial \preceq . Sejam $u, v \in C(I)$ distintos (não necessariamente como valores, mas como elementos do conjunto),

- Se $\text{grau}(u) < \text{grau}(v)$ então $u \prec v$.
- Se $\text{grau}(u) = \text{grau}(v)$ e u é menor do que v com relação à ordem gramatical, então $u \prec v$. A ordem gramatical em $C(I)$ é dada pelas letras sem seus sub-índices (e.g. $\hat{a}_m \prec \hat{b}_n$, para todos m, n).
- Se $\text{grau}(u) = \text{grau}(v)$ e $u \approx v$ com respeito à ordem gramatical, (e.g. \hat{d}_{12} e \hat{d}_{21}) então $u \prec v$ se u é menor do que v com respeito à ordem natural $(n, m) \mapsto (n + m - 2)(n + m - 1)/2 + n$.

Se ordenar os coeficientes baseado nas frequências (totais) e no posicionamento alfabético tem-se

$$\hat{a}_1, \underbrace{\hat{a}_2, \hat{b}_1, \hat{c}_1, \hat{d}_{1,1}, \hat{e}_{1,1}}_{\text{frequência}=2}, \underbrace{\hat{a}_3, \hat{d}_{1,2}, \hat{d}_{2,1}, \hat{e}_{1,2}, \hat{e}_{2,1}}_{\text{frequência}=3}, \hat{a}_4, \hat{b}_2, \dots$$

então obtém-se a sequência ordenada $\widehat{K\mathcal{F}}(f)$ e a parametrização $\widehat{K\mathcal{F}}_{\omega}(f, S)$ consiste nos coeficientes estimados de K-Fourier da função de densidade de probabilidade estimada $\hat{f}(\alpha, \theta)$ com uma frequência menor ou igual a ω .

A função estimadora \hat{f} para a função densidade de probabilidade $f : K \rightarrow \mathbb{R}$ será:

$$\begin{aligned}
\widehat{f}(\alpha, \theta) &= \sum_{k \in N_w} \widehat{f}_k \phi_k(\alpha, \theta) \\
&= 1 + \\
&+ \sum_{m=1}^w \widehat{a}_m \sqrt{2} \cos \left(m\theta_P - \frac{(1 - (-1)^m)}{4} \pi \right) \\
&+ \sum_{n=1}^{w_0} \left[\widehat{b}_n \sqrt{2} \cos(2n\alpha_P) + \widehat{c}_n \sqrt{2} \sin(2n\alpha_P) \right] \\
&+ \sum_{n+m=1}^w \left[\widehat{d}_{n,m} 2 \cos(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{n+m})}{4} \pi \right) \right] \\
&+ \sum_{n+m=1}^w \left[\widehat{e}_{n,m} 2 \sin(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{n+m})}{4} \pi \right) \right], \tag{38}
\end{aligned}$$

onde $w_0 = \lfloor \frac{w}{2} \rfloor$.

Quando aplicado o corte $\omega = 6$, a função estimadora \widehat{f} torna-se:

$$\begin{aligned}
\widehat{f}(\alpha, \theta) &= \sum_{k \in N_6} \widehat{f}_k \phi_k(\alpha, \theta) \\
&= 1 + \\
&+ \sum_{m=1}^6 \widehat{a}_m \sqrt{2} \cos \left(m\theta_P - \frac{(1 - (-1)^m)}{4} \pi \right) \\
&+ \sum_{n=1}^3 \left[\widehat{b}_n \sqrt{2} \cos(2n\alpha_P) + \widehat{c}_n \sqrt{2} \sin(2n\alpha_P) \right] \\
&+ \sum_{n+m=1}^6 \left[\widehat{d}_{n,m} 2 \cos(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{n+m})}{4} \pi \right) \right] \\
&+ \sum_{n+m=1}^6 \left[\widehat{e}_{n,m} 2 \sin(n\alpha_P) \cos \left(m\theta_P - \frac{(1 - (-1)^{n+m})}{4} \pi \right) \right] \tag{39}
\end{aligned}$$

Defini-se, então, um descritor multi-escala obtido pela concatenação de $KFchap_n$ para os valores de $n = 3, 7, 11, 15, 19$, onde

$$KFchap_n = \widehat{K\mathcal{F}}_\omega(f, S)|_n \tag{40}$$

onde

$$KFchap = [KFchap_3, KFchap_7, KFchap_{11}, KFchap_{15}, KFchap_{19}] \tag{41}$$

3.1.3 Coeficientes Estimados de K-Fourier invariantes à rotação

Define-se que os coeficientes estimados de K-Fourier ($\widehat{K\mathcal{F}}_\omega(f, S)$) de uma imagem são invariantes por rotação se estiverem na forma canônica ($\widehat{d}_{1,1} > 0$ e $\widehat{e}_{1,1} = 0$). Se

$\sigma(f) \in [0, 2\pi)$ é tal que os coeficientes estimados de K-Fourier de uma imagem estão na forma canônica após rotacionados de $\sigma(f)$, denota-se o EKFC desta imagem por $T_{\sigma(f)}(\widehat{K\mathcal{F}}_{\omega}(f, S))$, em que

$$\begin{aligned} T_{\sigma} \hat{a}_m &= \hat{a}_m \\ T_{\sigma} \begin{bmatrix} \hat{b}_n \\ \hat{c}_n \end{bmatrix} &= \begin{bmatrix} \cos 2n\sigma & -\sin 2n\sigma \\ \sin 2n\sigma & -\cos 2n\sigma \end{bmatrix} \begin{bmatrix} \hat{b}_n \\ \hat{c}_n \end{bmatrix} \\ T_{\sigma} \begin{bmatrix} \hat{d}_{n,m} \\ \hat{e}_{n,m} \end{bmatrix} &= \begin{bmatrix} \cos n\sigma & -\sin n\sigma \\ \sin n\sigma & \cos n\sigma \end{bmatrix} \begin{bmatrix} \hat{d}_{n,m} \\ \hat{e}_{n,m} \end{bmatrix} \end{aligned}$$

Neste ponto, já foram obtidos (α, θ) para cada *patch* e estes foram projetados na superfície \mathcal{K} , os coeficientes estimados de K-Fourier será $\widehat{K\mathcal{F}}_{\omega}(f, S)$ e os coeficientes estimados de K-Fourier invariantes à rotação será $T_{\sigma}(\widehat{K\mathcal{F}}_{\omega}(f, S))$.

Se $M + N$ é o número total de $n \times n$ *patches* em uma determinada imagem, M sendo o número de *patches* que estão muito longe da garrafa de Klein para serem projetados, e seja S a K -amostra resultante de seus $n \times n$ *patches*.

O descritor multi-escala e invariante à rotação é obtido pela concatenação de $EKFC_n$ para os valores de $n = 3, 7, 11, 15, 19$, onde

$$EKFC_n = \left[1 - \frac{1}{M}, \frac{T_{\sigma(f)}(\widehat{K\mathcal{F}}_{\omega}(f, S))}{M} \right]_n \quad (42)$$

onde

$$\begin{aligned} M &= \text{quantidade de patches não projetados} \\ T_{\sigma(f)}(\widehat{K\mathcal{F}}_{\omega}(f, S)) &= \text{coeficientes estimados de K-Fourier invariantes à rotação} \\ \omega &= \text{frequência de corte (WIT; FLORIANI, 1998)} \end{aligned}$$

ou seja:

$$EKFC = [EKFC_3, EKFC_7, EKFC_{11}, EKFC_{15}, EKFC_{19}] \quad (43)$$

3.2 Propostas

Analisando-se a proposta de Perea e Carlsson (2014), foi observado que as escalas (tamanhos) dos *patches* é composta de 5 valores ímpares. Verificando as equações que compõem o algoritmo de projeção, percebeu-se que não há impedimentos para a utilização de outros valores inteiros desde que $n \geq 3$, sendo essa restrição imposta pelas Equações 20 e 21.

Além disso, conforme explicitado na mesma proposta de Perea e Carlsson (2014), foi utilizado o mesmo corte das frequências ($\omega = 6$) proposto por Wit e Floriani (1998) para o cálculo da função estimadora \hat{f} . Esse valor influencia diretamente no cálculo dos coeficientes estimados de K-Fourier (Eq. 41) e não foram analisadas outras frequências de corte.

Baseado nesses fatos, propõe-se investigar a criação do vetor de características, o corte das frequências e a obtenção dos *patches*.

Um pesquisa mais aprofundada para a utilização de novos métodos para a seleção dos *patches* deve ser feita, para que estes deixem de ser escolhidos aleatoriamente e sejam selecionados os *patches* que melhor representem a imagem.

Para criar o vetor de características, Perea e Carlsson (2014) utilizaram a concatenação de 5 diferentes escalas ($n = 3, 7, 11, 15, 19$) no descritor *EKFC* (Eq. 43). Deve ser feito um estudo para a obtenção de uma nova combinação de escalas, inicialmente aplicado às bases Brodatz, CURET, Outex, KTH-TIPS, UIUCTex e VisTex (algumas já utilizadas pelos autores), que pode resultar em uma redução na quantidade de descritores e melhorar a caracterização das bases. Apesar da aplicação ser realizada para as bases anteriormente especificadas, será proposta a generalização de escalas para quaisquer bases de texturas.

Neste trabalho serão analisados outros valores para o corte das frequências (ω), e analisados os resultados para a representação de uma imagem.

3.3 Experimentos e Análise dos Resultados

Nesta seção serão apresentados os experimentos e análises referentes às Projeções no Espaço \mathcal{X} , apresentando o desenvolvimento da proposta para o uso da projeção de *patches* das imagens no Espaço da garrafa de Klein. Também serão apresentados nesta seção: uma análise da invariância à rotação, uma análise do vetor de características obtido, uma análise das escalas e uma análise das frequências de corte.

Desenvolvimento da Proposta

Para a caracterização de texturas por meio da garrafa de Klein, utilizou-se a formulação apresentada na Seção 3.1. Foram implementados os algoritmos para: a projeção dos *patches*, para a obtenção dos coeficientes estimados de K-Fourier, para aplicar a invariância à rotação. Para os experimentos iniciais foram utilizadas como entrada do algoritmo as imagens da base Pollen (Apêndice B) e as imagens da base KTH-TIPS. A base KTH-TIPS será utilizada nos experimentos para que os resultados possam ser comparados aos resultados de Perea e Carlsson (2014).

Para a criação do vetor de características para cada imagem, conforme já explicitado na Figura 8, após a seleção dos *patches*, faz-se: a projeção dos *patches*, obtém-se os coeficientes estimados de K-Fourier e aplica-se a invariância à rotação. Tendo criado

os vetores de características invariantes à rotação para todas as imagens, envia-se este conjunto de vetores ao classificador. Além do classificador LMNN (WEINBERGER; SAUL, 2009), o conjunto de vetores de características obtidos foram submetido aos classificadores K-Vizinhos mais Próximos (KNN) (FUKUNAGA, 1990, p. 306) e Análise de Discriminantes Quadráticos (QDA) (FRIEDMAN, 1989). O classificador LMNN foi utilizado para fins de comparação de resultados com o trabalho de Perea e Carlsson (2014). Os outros dois classificadores foram aplicados para uma posterior análise da eficiência dos classificadores perante os descritores gerados pelo algoritmo baseado na projeção da garrafa de Klein.

No classificador LMNN foram utilizados, como parâmetros, a vizinhança 3 para aprendizado da métrica global e 20% dos dados para o conjunto de treinamento da validação cruzada. Foram utilizadas, de cada classe, 50% para o conjunto de treinamento e 50% para o conjunto de testes, conforme já utilizado por Perea e Carlsson (2014). Manteve-se o mesmo princípio utilizado por Perea e Carlsson (2014) para classificar os dados, sendo executado 100 vezes o classificador LMNN. Para cada uma dessas execuções, os conjuntos de treinamento e teste foram gerados aleatoriamente (mantendo as proporções de 50% para cada um). Após a execução das 100 vezes, calculou-se a média e variância das acurácias obtidas e, também, o *Precision*, o *Recall* e o *F-Score* (RIJSBERGEN, 1979).

Foram realizados 2 experimentos aplicados às bases KTH-TIPS e Pollen para ilustrar os resultados da implementação do algoritmo. Em ambos os experimentos, para aplicar a Equação 17 (Seção 3.1.1), foram utilizados todos os *patches* extraídos da imagem, sem a realizar nenhuma escolha aleatória ou restringir a utilização de 1.000 *patches* com maior D-Norma. Isto foi feito para que todos os *patches* das imagens fossem utilizados nos dois experimentos e não houvesse nenhuma diferença nas entradas dos experimentos.

O primeiro experimento utilizou o descritor EKFC para a criação do vetor de características das bases. No segundo experimento, foi utilizado o descritor *KFchap* para a criação do vetor de características das bases. Em ambos os experimentos utilizou-se 3 classificadores distintos: KNN, QDA e LMNN.

No primeiro experimento utilizou-se a implementação da formulação da Seção 3.1, tendo como descritor o EKFC para a criação do vetor de características. A Tabela 7 ilustra os resultados dos classificadores aplicados aos vetores de características das imagens da base KTH-TIPS. Verifica-se que o classificador LMNN obteve uma taxa de acurácia média de 23,39% e, 20,73% e 48,63% para os classificadores KNN e QDA, respectivamente.

Observou-se aqui que essa base KTH-TIPS foi utilizada no trabalho de Perea e Carlsson (2014), sendo classificada com o LMNN, e os autores apresentaram um resultado de $94,77\% \pm 1.13$, seguindo o procedimento descrito. No processamento realizado por nós seguindo a metodologia descrita (PEREA; CARLSSON, 2014), obteve-se apenas 28,29% em seu melhor resultado.

Tabela 7 – Resultados da Classificação das imagens da base KTH-TIPS projetadas na garrafa de Klein (em %).

Classificadores	KNN	Discriminante Quadrático	LMNN
Descritores	215	215	215
Maior Acurácia	21,73	49,01	28,29
Acurácia Média	20,73	48,63	23,39
<i>Precision</i> Médio	28,68	57,30	-
<i>Recall</i> Médio	21,73	49,01	-
<i>F-Score</i> Médio	21,84	49,47	-

A Tabela 8 ilustra os resultados dos classificadores aplicados aos vetores de características das imagens da base Pollen. Verifica-se que o classificador LMNN obteve uma taxa de acurácia média de 32,10% e, 17,26% e 32,43% para os classificadores KNN e QDA, respectivamente.

Tabela 8 – Resultados da Classificação das imagens da base Pollen projetadas na garrafa de Klein (em %).

Classificadores	KNN	Discriminante Quadrático	LMNN
Descritores	215	215	215
Maior Acurácia	17,87	32,83	36,20
Acurácia Média	17,26	32,43	32,10
<i>Precision</i> Médio	22,62	36,08	-
<i>Recall</i> Médio	19,32	35,49	-
<i>F-Score</i> Médio	18,94	33,02	-

No segundo experimento utilizou-se a implementação da formulação da Seção 3.1, agora tendo como descritor os valores dos coeficientes estimados de K-Fourier (*KFchap*) para a criação do vetor de características. A Tabela 9 apresenta os resultados obtidos para a base KTH-TIPS, novamente utilizando-se os valores dos coeficientes estimados de K-Fourier (*KFchap*) e mantendo a utilização dos 5 escalas de *patches*. Observa-se que o resultado do classificador LMNN (93,17%) está muito próximo ao resultado obtido por Perea e Carlsson (2014).

A Tabela 10 apresenta os resultados obtidos para a base Pollen. Pode-se observar que o resultado do classificador LMNN aumentou de 32,10% para 64,84% em sua acurácia média.

Analisando-se os resultados obtidos sem a utilização da invariância à rotação, percebeu-se que a transformação nos coeficientes estimados de K-Fourier para torná-los invariantes à rotação ($T_\sigma(\widehat{K\mathcal{F}}_\omega(f, S))$) modifica as informações das imagens. Uma análise da invariância à rotação que está sendo aplicada aos *patches* para verificar a influência em imagens rotacionadas e não rotacionadas é apresentada na seção 3.3.2.

Tabela 9 – Resultados da Classificação das imagens da base KTH-TIPS projetadas na garrafa de Klein sem invariância à rotação (em %).

Classificadores	KNN	Discriminante Quadrático	LMNN
Descritores	210	210	210
Maior Acurácia	55,80	71,98	93,17
Acurácia Média	54,32	70,94	87,62
<i>Precision</i> Médio	63,76	73,27	-
<i>Recall</i> Médio	55,80	71,98	-
<i>F-Score</i> Médio	58,46	71,80	-

Tabela 10 – Resultados da Classificação das imagens da base Pollen projetadas na garrafa de Klein sem invariância à rotação (em %).

Classificadores	KNN	Discriminante Quadrático	LMNN
Descritores	210	210	210
Maior Acurácia	41,96	52,78	69,63
Acurácia Média	40,95	51,81	64,84
<i>Precision</i> Médio	48,54	57,26	-
<i>Recall</i> Médio	43,72	54,80	-
<i>F-Score</i> Médio	44,59	53,19	-

Passou-se, então, a uma análise mais criteriosa do processo. Nas próximas sub-seções serão analisados a obtenção dos *patches*, a invariância à rotação, o vetor de características e a definição das escalas e das frequências.

3.3.1 Obtenção dos *Patches*

Algo que pode influenciar nos resultados é o modo de obtenção dos *patches*. Realizou-se 2 experimentos para a obtenção dos *patches* de maneiras diferentes: uma janela sem sobreposição e uma janela deslizante. Além disso, realizou-se uma análise sobre o uso da aleatoriedade na obtenção dos *patches*.

A Figura 9(A) ilustra a obtenção de 2 *patches* na imagem. Em 1 e 2 são ilustrados os valores correspondentes à intensidade dos *pixels* presentes nas janelas, neste caso $n = 3$ (janela 3×3). A Figura 9(B) ilustra a obtenção de outros 2 *patches* da mesma imagem. Em 3 e 4 são ilustrados os valores correspondentes à intensidade dos *pixels* presentes nas janelas que, novamente são 3×3 . As marcações presentes nas imagens são apenas para exemplificar o posicionamento e, não estão em escala com as matrizes.

Realizou-se os experimentos com a base KTH-TIPS utilizando-se dos métodos para a obtenção dos *patches* (sem sobreposição e deslizante) e, como descritor, foram utilizados os coeficientes estimados de K-Fourier (*KFchap*). Os resultados da classificação utilizando LMNN estão ilustrados na Tabela 11.

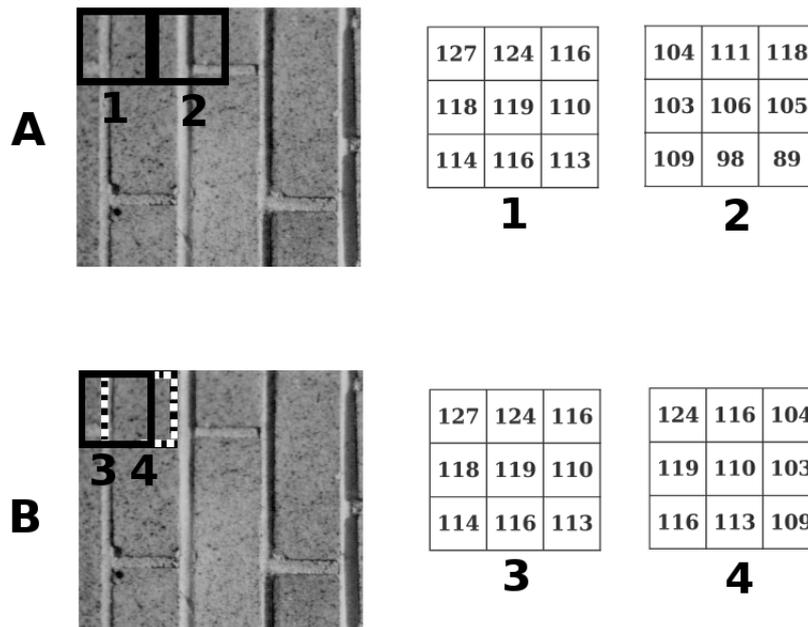


Figura 9 – Exemplo de recorte de *Patch*: (A) sem sobreposição e (B) com janela deslizante. As marcações são para exemplificação e não estão em escala com as matrizes.

Tabela 11 – Resultado da Classificação da base KTH-TIPS para os 2 tipos de janelas (acurácias em %).

	Sem Sobreposição	Deslizante
Maior Acurácia	92,20	66,83
Acurácia Média	87,26	60,00
Variância	0,0004	0,0004
Desvio Padrão	0,0195	0,0210

Pensou-se na utilização da janela deslizante por conseguir percorrer todos os *pixels* da imagem e, conseguiria-se obter todos os *patches* que poderiam representar a imagem.

Observou-se com os resultados que a utilização das janelas deslizantes, apesar de obter todos os possíveis *patches* de uma imagem, para a utilização no algoritmo da garrafa de Klein, não produziu bom resultado. Isso se deve ao fato da obtenção de *patches* com janelas deslizantes gerarem, pelo menos, 3 vezes mais *patches* do que a obtenção sem sobreposição. Neste sentido, ao se aplicar o algoritmo para projetar os *patches*, são selecionados aleatoriamente 5.000 *patches* e destes, apenas os 1.000 com maior D-Norma serão supostamente projetados. Com isso, para todos os testes utilizados neste trabalho os *patches* foram obtidos sem sobreposição.

Além do modo como são obtidos os *patches*, a utilização da aleatoriedade poderia influenciar os resultados finais. O algoritmo proposto por Perea e Carlsson (2014) utiliza seleção aleatória de *patches* para o processamento. Formas alternativas de escolha foram

testadas, mas não apresentaram resultados significantes.

3.3.2 Análise da Rotação

Para analisar a invariância à rotação ($T_\sigma(\widehat{K\mathcal{F}}_\omega(f, S))$) foram selecionadas imagens de diferentes bases de texturas para se realizar uma análise do algoritmo referente à rotação.

A Figura 10, composta por 10 imagens, foi criada a partir de 3 bases de texturas: USC-SIPI (WEBER, 1997), Outex (OJALA; PIETIKAINEN; MAENPAA, 2002) e UIUCTex (LAZEBNIK; SCHMID; PONCE, 2005).

As imagens 1 e 7 foram retiradas da base de texturas USC-SIPI (WEBER, 1997), sendo que as imagens 2, 3 e 4 foram rotações da imagem 1 em -90° , 90° e 180° , respectivamente, no sentido anti-horário. As imagens 5 e 6 são recortes da imagem 1. As imagens 8 e 9 foram retiradas da base de texturas Outex e a imagem 10 foi retirada da base de texturas UIUCTex.

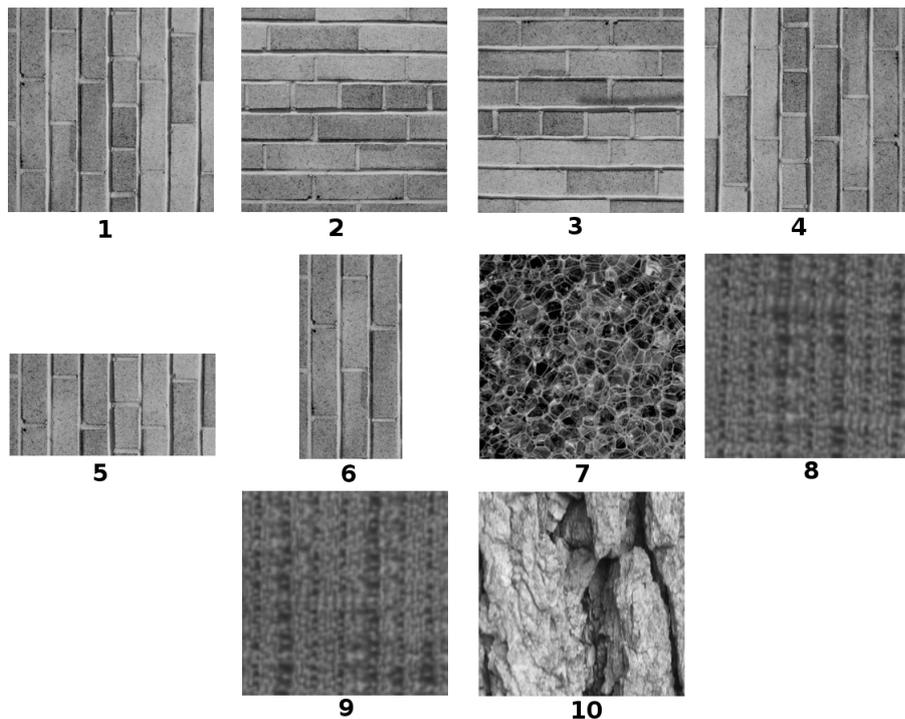


Figura 10 – Imagens de texturas utilizadas nos experimentos de rotação.

Aplicou-se, às imagens da Figura 10, o algoritmo de caracterização proposto por Perea e Carlsson (2014), utilizando-se da formulação apresentada na Seção 3.1, e foram calculadas a distância Euclidiana entre os vetores de características dessas imagens. A Tabela 12 ilustra os resultados obtidos no cálculo das distâncias entre as imagens.

Observa-se que os valores das distâncias da imagem 1 para as suas rotações estão bem próximos, quando aplicado o coeficientes estimados de K-Fourier invariante à rotação ($EKFC$). Quando comparado os valores das distâncias obtidos dos coeficientes estimados

Tabela 12 – Resultados das distâncias das imagens considerando caracterização usando os coeficientes estimados de K-Fourier e os coeficientes estimados de K-Fourier invariantes à rotação.

Imagens		<i>KFchap</i>	<i>EKFC</i>
1	2	2,9317	0,5394
1	3	2,6674	0,5513
2	3	2,0013	0,5541
1	4	1,9349	1,0389
1	5	1,1122	0,5851
1	6	1,3666	0,4530
1	7	2,0936	0,5322
1	10	2,6948	2,9747
2	7	1,8709	0,6871
5	7	1,8291	0,5633
8	9	2,1099	2,4603
10	8	2,8806	3,5410
10	9	2,9937	3,2508

de K-Fourier (*KFchap*), percebe-se que esses valores não mantêm a proximidade. Comparações entre imagens de bases diferentes devem ter valores de distâncias relativamente maiores do que imagens de mesma base/classe. Isto pode ser observado, por exemplo, na comparação entre as imagens 10 e 8 que são de bases diferentes. Porém quando analisadas imagens da mesma base (imagens 8 e 9) sem rotação, as distâncias estão maiores que 2.

Com os resultados obtidos, considerando-se os coeficientes estimados de K-Fourier invariante à rotação (*EKFC*) das imagens, percebe-se que a invariância à rotação aplicada modificou as informações das imagens, não conseguindo obter vetores de características muito semelhantes. Fato explicitado pelas distâncias entre rotações da mesma imagem não estarem próximas à 0 e/ou não estarem com distâncias muito próximas entre elas, conforme pode ser observado na comparação entre as imagens 1 e 4 que são a mesma imagem, porém a imagem 4 está rotacionada e, a distância nos dois casos foi maior que 1.

Para avaliar o efeito da rotação, aplicou-se o algoritmo proposto por Perea e Carlsson (2014) baseado na formulação apresentada na Seção 3.1 para a base USC-SIPI, que contém imagens rotacionadas. A Tabela 13 ilustra os resultados obtidos pelo classificador LMNN para a base USC_SIPI.

Tabela 13 – Resultados da Classificação da base USC-SIPI com o classificador LMNN (em %).

Descritores	<i>KFchap</i>	<i>EKFC</i>
Quantidade	210	215
Maior Acurácia	84,62	76,92
Acurácia Média	73,23	60,13

Observa-se na Tabela 13 que mesmo para uma base contendo imagens rotacionadas, os coeficientes estimados de K-Fourier invariante à rotação (*EKFC*) das imagens não produzem boa separação das classes, pois obteve-se maior acurácia utilizando apenas os coeficientes estimados de K-Fourier (*KFchap*).

Realizou-se outros 2 experimentos. No primeiro, denominado Original, foram recortados os *patches* e, aplicou-se o processamento padrão descrito anteriormente para a obtenção dos coeficientes estimados de K-Fourier. Após isso, classificou-se o conjunto de vetores característicos gerados por meio do classificador LMNN.

O segundo experimento, denominado Rotacionado, utilizou-se dos mesmos *patches* recortados para o primeiro experimento, porém estes foram rotacionados em 90° (sentido anti-horário) e, novamente, aplicou-se o processamento padrão para a obtenção dos coeficientes estimados de K-Fourier. Apenas o conjunto de vetores característicos gerados com os *patches* rotacionados foi submetido ao classificador LMNN. Os resultados dos dois experimentos estão ilustrados na Tabela 14.

Tabela 14 – Resultados da Classificação da base KTH-TIPS com o classificador LMNN para *patches* sem e com rotação (em %).

	<i>KFchap</i>		<i>EKFC</i>	
	Original	Rotacionado	Original	Rotacionado
Acurácia	93,90	90,00	29,76	42,68
Acurácia Média	87,57	83,66	24,48	37,78
Desvio Padrão	0,0173	0,0191	0,0201	0,0211

Pode ser observado na Tabela 14 que, ao se utilizar como descritor os coeficientes estimados de K-Fourier invariante à rotação (*EKFC*) para os *patches* rotacionados, obteve-se maior acurácia quando comparada a acurácia obtida nos *patches* não rotacionados. Porém, quando utilizados os coeficientes estimados de K-Fourier (*KFchap*) como descritor obteve-se melhores resultados quando comparados aos resultados obtidos com o uso do descritor *EKFC*, em ambos os casos (com e sem a rotação dos *patches*). Observou-se que apesar da acurácia e acurácia média terem variações na casa da unidade, o desvio padrão apresentado é resultado da execução do classificador por 100 vezes, conforme descrito anteriormente.

Analisando-se os resultados obtidos com esses experimentos de rotação, definiu-se um novo experimento com apenas 1 *patch* e suas rotações em sentido anti-horário (Figura 11). Ao se aplicar o algoritmo de projeção na garrafa de Klein nos *patches* definidos na Figura 11, obteve-se os vetores de características para cada um deles por meio dos coeficientes estimados de K-Fourier invariante à rotação (*EKFC*).

Para todas as 4 matrizes (Figura 11), os valores de todos os 215 itens do descritor *EKFC* obtidos foram praticamente iguais, ocorrendo variações nas casas decimais acima da 18 posição.

26	23	50	52	33	38
3	3	45	45	3	3
38	33	52	50	23	26
a) Original			b) 180°		
50	45	52	38	3	26
23	3	33	33	3	23
26	3	38	52	45	50
c) 90°			d) -90°		

Figura 11 – Exemplos de *Patches* criados. Em (a) o *patch* original, (b) rotacionado em 180° , (c) rotacionado em 90° e (d) rotacionado em -90° . Todas as rotações em sentido anti-horário.

Baseado nos resultados dos experimentos realizados para analisar a rotação, pôde-se observar que o cálculo da invariância à rotação está correto se aplicado a um *patch* rotacionado. Quando rotaciona-se a imagem, não existe a garantia de que um *patch* na imagem original é o mesmo *patch* rotacionado da imagem rotacionada.

A Figura 12 ilustra a obtenção de *patches* 3×3 ($n = 3$) em uma imagem de 256×256 *pixels*. As matrizes 1, 2 e 3 foram obtidas da imagem original iniciando (canto superior esquerdo) nas posições: (1, 1), (1, 254), (124, 124), respectivamente. As matrizes 4, 5 e 6 foram obtidas da imagem original rotacionada em 90° (sentido anti-horário) iniciando nas mesmas posições das outras 3 matrizes anteriores. As marcações nas imagens são para exemplificação e não estão em escala com as matrizes.

Observa-se na Figura 12 que as matrizes 2 e 4 são semelhantes por uma ser a rotação da outra. Porém, as matrizes 3 e 6 foram obtidas nas mesmas posições (centro da imagem) e estão com valores diferentes. Apesar de uma imagem ser rotação da outra, não se pode garantir que todos os valores estarão em suas posições "relativas". Com isso, diferentemente da rotação de um *patch*, a invariância à rotação aplicada à imagem não consegue descrever corretamente as imagens rotacionadas para um classificador por se tratar de um novo *patch* e não de sua rotação.

Com isso, deve-se analisar a Equação 42 ($EKFC_n$) que é a base para compor o vetor de características, podendo influenciar no resultado final.

3.3.3 Vetor de Características

Realizou-se uma análise para a formação do vetor de características obtido e percebeu-se a influência negativa do denominador M na Equação 42 para composição do descritor invariante à rotação. Na formulação proposta em (PEREA; CARLSSON, 2014), M é a quantidade de *patches* que não foram projetados por ficarem com distância superior a $\sqrt{2}$ (Eq. 31). Os experimentos utilizaram o classificador LMNN e foram divididos em:

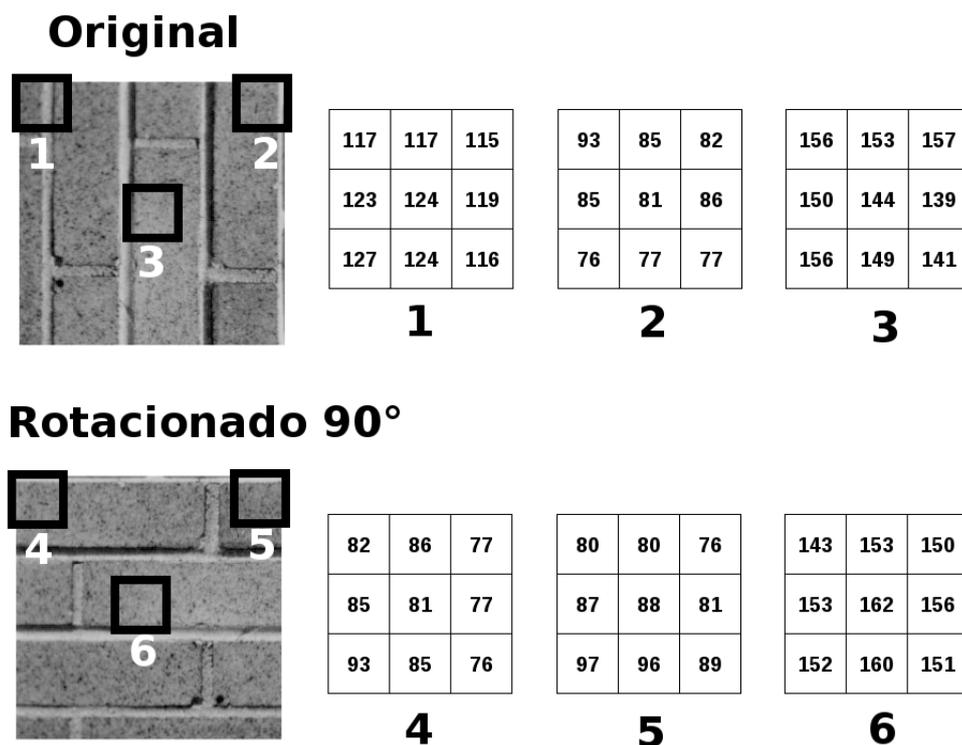


Figura 12 – Exemplo de extração de *Patches*. Itens 1 e 4 estão na mesma posição nas imagens, assim como os itens 2 e 5 e 3 e 6, respectivamente. As marcações são para exemplificação e não estão em escala com as matrizes. A rotação está no sentido anti-horário.

1. *EKFC*: utiliza a caracterização via Equação 42;

2. *Ind*: utiliza a caracterização via Equação 44;

3. *Sdiv*: utiliza a caracterização via Equação 45;

4. *Kdiv*: utiliza a caracterização via Equação 46;

5. *KInd*: utiliza a caracterização via Equação 47;

6. *KFchap*: utiliza a caracterização via Equação 40.

$$\begin{aligned} Ind_n &= \left[1 - \frac{1}{M}, T_{\sigma(f)} \left(\widehat{K\mathcal{F}}_{\omega}(f, S) \right) \right]_n \\ Ind &= [Ind_3, Ind_7, Ind_{11}, Ind_{15}, Ind_{19}] \end{aligned} \quad (44)$$

$$\begin{aligned} SDiv_n &= T_{\sigma(f)} \left(\widehat{K\mathcal{F}}_{\omega}(f, S) \right) \Big|_n \\ SDiv &= [SDiv_3, SDiv_7, SDiv_{11}, SDiv_{15}, SDiv_{19}] \end{aligned} \quad (45)$$

$$\begin{aligned} KDiv_n &= \left[1 - \frac{1}{M}, \frac{\widehat{K\mathcal{F}}_{\omega}(f, S)}{M} \right]_n \\ KDiv &= [KDiv_3, KDiv_7, KDiv_{11}, KDiv_{15}, KDiv_{19}] \end{aligned} \quad (46)$$

$$\begin{aligned} KInd_n &= \left[1 - \frac{1}{M}, \widehat{K\mathcal{F}}_{\omega}(f, S) \right]_n \\ KInd &= [KInd_3, KInd_7, KInd_{11}, KInd_{15}, KInd_{19}] \end{aligned} \quad (47)$$

Para todos os 6 experimentos realizados para a análise do vetor de características, seguiu-se as etapas descritas anteriormente pelo método, porém para a criação do vetor de características foram utilizadas as formulações propostas para os experimentos em questão. Além disso, todos os 6 experimentos utilizaram-se das bases KTH-TIPS e Pollen.

Os 3 primeiros experimentos basearam-se na Equação 42 que aplica a invariância à rotação nos coeficientes estimados de K-Fourier. O primeiro experimento foi realizado utilizando-se a Equação 42. O segundo experimento utilizou-se da Equação 44 para criar o vetor de característica, porém com a retirada do denominador M no segundo termo da equação. Com essa retirada, pretende-se verificar se o denominador M estava modificando o segundo termo. No terceiro experimento, apenas os coeficientes estimados de K-Fourier invariantes à rotação foram mantidos e, com isso, pretende-se verificar se a retirada total de M influenciará nos resultados.

Os 3 últimos experimentos basearam-se nos coeficientes estimados de K-Fourier (Eq. 40). O quarto experimento foi realizado utilizando-se a mesma estrutura da Equação 42, porém o termo $T_{\sigma(f)} \left(\widehat{K\mathcal{F}}_{\omega}(f, S) \right)$ foi substituído por $\widehat{K\mathcal{F}}_{\omega}(f, S)$. O quinto experimento utilizou-se da mesma estrutura da Equação 44 e, novamente, o termo $T_{\sigma(f)} \left(\widehat{K\mathcal{F}}_{\omega}(f, S) \right)$ foi substituído por $\widehat{K\mathcal{F}}_{\omega}(f, S)$. No último experimento utilizou-se para criar o vetor característico apenas os coeficientes estimados de K-Fourier (Eq. 40).

As Tabelas 15 e 16 ilustram os valores obtidos para os 6 experimentos aplicados às bases KTH-TIPS e Pollen.

Os resultados dos experimentos para a base KTH-TIPS (Tabela 15) ilustram que quando utilizado o *EKFC* obteve-se a menor acurácia. Ao manter $1 - \frac{1}{M}$ em *Ind*, observa-se um aumento aproximado de 54% em relação ao *EKFC*. Comparando-se *Ind* com a retirada de toda divisão por M (*Sdiv*), a melhora não é expressiva. A modificação

aplicada no quarto experimento consegui melhorar a acurácia comparada ao *EKFC*. Ao se utilizar apenas o *KFchap* foi obtida a maior acurácia com 210 descritores.

Tabela 15 – Resultados da Classificação com LMNN para a base KTH-TIPS (em %).

Descritores	<i>EKFC</i>	<i>Ind</i>	<i>Sdiv</i>	<i>KDiv</i>	<i>KInd</i>	<i>KFchap</i>
Quantidade	215	215	210	215	215	210
Maior Acurácia	50,24	81,95	82,44	72,20	90,24	91,46
Acurácia Média	44,62	77,38	76,98	68,58	86,49	87,72
Desvio Padrão	0,02522	0,02045	0,01960	0,01672	0,01614	0,01826

Os resultados dos experimentos para a base Pollen (Tabela 16) ilustram que quando utilizado o *Sdiv* obteve-se a menor acurácia. Utilizando-se do *EKFC* não foi obtida uma boa acurácia. No segundo experimento (*Ind*) que ainda se baseia na invariância à rotação, obteve-se a melhor acurácia entre os três primeiros experimentos. A modificação aplicada aos três últimos experimentos conseguiram melhorar a acurácia comparada aos experimentos baseados na invariância à rotação. Ao se utilizar apenas o *KFchap* foi obtida a maior acurácia com 210 descritores.

Tabela 16 – Resultados da Classificação com LMNN para a base Pollen (em %).

Descritores	<i>EKFC</i>	<i>Ind</i>	<i>Sdiv</i>	<i>KDiv</i>	<i>KInd</i>	<i>KFchap</i>
Quantidade	215	215	210	215	215	210
Maior Acurácia	32,82	43,56	26,99	56,13	58,90	74,23
Acurácia Média	27,77	37,51	23,07	50,17	51,26	69,26
Desvio Padrão	0,01841	0,02018	0,01890	0,02267	0,02420	0,02071

As imagens da Base Pollen tem um maior contraste em pequenas regiões do pólen ("bordas" ou estruturas da textura) e, durante a obtenção dos *patches*, muitos deles não tem alto contraste para poderem ajudar na caracterização da imagem. Com isso, durante o processo de projeção muitos *patches* são descartados, conseqüentemente aumentando o valor de M (*patches* não projetados). Os vetores de características gerados pelas equações 42, 44, 46 e 47 sofrem influência do valor de M .

Com esses experimentos, observa-se que ainda deve ser investigado com mais profundidade a influência do denominador M para outras bases de texturas. Mesmo assim, conseguiu-se verificar que a utilização dos coeficientes estimados de K-Fourier (Eq. 41) obteve melhores resultados do que a utilização coeficientes estimados de K-Fourier invariantes à rotação (Eq. 43). Com isso, para os próximos experimentos deste capítulo serão utilizados os coeficientes estimados de K-Fourier (*KFchap* - Eq. 41) para a criação do vetor de características.

3.3.4 Análise das Escalas

Um fato que mereceu nossa atenção foi a escolha das escalas (tamanhos). Ao se analisar as equações que compõem o algoritmo de projeção, verificou-se que não há impedimentos para a utilização de valores inteiros desde que $n \geq 3$, sendo essa restrição imposta pelas Equações 20 e 21.

Para criar o vetor de características em (PEREA; CARLSSON, 2014) foi utilizada a concatenação de 5 diferentes escalas ($n = 3, 7, 11, 15, 19$) no descritor *EKFC*. Os autores não explicitaram como definiram as 5 escalas utilizadas.

O propósito desta análise é verificar e validar as escalas utilizadas no vetor de características que melhorem os resultados de caracterização. Para analisar modificações nas escalas utilizadas em (PEREA; CARLSSON, 2014), foram realizados 11 experimentos divididos em:

- Combinação de Escalas;
- Análise de combinações utilizando Sequência de Fibonacci.
- Análise de combinações utilizando Sequência de Números Inteiros;

3.3.4.1 Combinação de Escalas

Para este primeiro experimento aplicou-se o algoritmo de projeção, conforme descrito anteriormente, porém para a definição de *KFchap_n* (Eq. 40) foram tomadas 13 escalas de números naturais ímpares ($n = 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27$), mantendo o corte das frequências ($\omega = 6$). A escolha dos números ímpares foi para seguir o procedimento adotado por Perea e Carlsson (2014).

Para avaliar as 13 escalas propostas, após a criação dos descritores para cada escala, aplicou-se o classificador LMNN. A Tabela 17 ilustra o resultado do classificador para cada uma das escalas aplicado à base KTH-TIPS. Neste caso o descritor final foi composto apenas por 1 escala e com isso, o vetor de características é composto de apenas 42 posições.

Após essa etapa inicial, combinou-se as 13 escalas em 2 a 2, 3 a 3, 4 a 4 e 5 a 5 para formar os vetores de características totalizando 2.379 combinações. Todos os resultados dessas combinações foram tabulados e, para ilustrar esses resultados, na Tabela 18 estão apenas as maiores acurácias para cada combinação de tamanho (1 a 1, 2 a 2, 3 a 3, 4 a 4 e 5 a 5), incluindo na última coluna a combinação proposta em (PEREA; CARLSSON, 2014) ($n = 3, 7, 11, 15, 19$), aplicados à base KTH-TIPS. Além desta base, também foi realizado o mesmo experimento de combinação de escalas e, apresentadas apenas as maiores acurácias para cada uma das combinações, para as bases Brodatz (Tabela 19), CURET (Tabela 20), Outex (Tabela 21), UIUCTex (Tabela 22), VisTex (Tabela 23) e Pollen (Tabela 24).

Conforme ilustrado pelos resultados das Tabelas 18, 19, 20, 21, 22, 23 e 24, observa-se que ao modificar as escalas, reduzindo a quantidade de descritores, não ocorreu redução

Tabela 17 – Resultado da Classificação das 13 escalas individuais para a base KTH-TIPS.

Escala (n)	Maior Acurácia (%)	Acurácia Média (%)	Desvio Padrão
3	88,05	84,35	0,01938
5	79,76	76,36	0,02074
7	71,22	66,03	0,02276
9	57,32	51,51	0,02155
11	51,71	47,35	0,01814
13	46,59	41,25	0,02111
15	42,20	37,53	0,01871
17	36,34	31,55	0,01739
19	35,12	30,99	0,01726
21	34,88	29,84	0,02017
23	30,24	26,50	0,01698
25	34,39	30,64	0,01949
27	27,32	22,70	0,01756

Tabela 18 – Maiores Acurácias resultantes da Classificação aplicada à base KTH-TIPS (em %).

Escalas (n)	3	[3 5]	[3 5 9]	[3 5 7 13]	[3 5 15 17 19]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	88,05	92,93	92,93	94,39	95,61	92,20
Acurácia Média	84,35	88,65	88,70	89,05	90,80	86,34
Desvio Padrão	0,019384	0,017045	0,015229	0,017030	0,016734	0,016814

Tabela 19 – Maiores Acurácias resultantes da Classificação aplicada à base Brodatz (em %).

Escalas (n)	3	[3 5]	[3 5 7]	[3 5 7 17]	[3 5 13 17 21]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	85,25	90,65	93,02	93,36	94,93	91,10
Acurácia Média	82,03	88,15	90,87	90,78	92,13	88,90
Desvio Padrão	0,010010	0,008902	0,008181	0,008333	0,007879	0,009507

significativa na acurácia. Um bom exemplo disso é a Tabela 18 na qual as escalas [3, 5] obtiveram 92,93% de acurácia com 84 descritores enquanto utilizando-se dos 210 descritores propostos em (PEREA; CARLSSON, 2014) obteve-se 92,20% de acurácia. Para as outras bases, como por exemplo a base Vistex, modificando as escalas e mantendo a quantidade de 5 escalas, obteve-se maior acurácia comparada ao resultado obtido pelas escalas definidas em (PEREA; CARLSSON, 2014) [3, 7, 11, 15, 19].

Observou-se na tabulação de todas as combinações para as bases analisadas que, em quase todos os experimentos, as escalas [3, 5] obtiveram os melhores resultados proporcionalmente à quantidade de descritores utilizados na caracterização.

Tabela 20 – Maiores Acurácias resultantes da Classificação aplicada à base CURET (em %).

Escalas (n)	3	[3 5]	[3 5 7]	[3 5 7 9]	[3 5 7 9 11]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	77,15	83,04	85,68	85,78	87,32	79,86
Acurácia Média	75,73	81,28	84,33	84,21	85,17	78,58
Desvio Padrão	0,006800	0,006126	0,006157	0,005723	0,005577	0,005974

Tabela 21 – Maiores Acurácias resultantes da Classificação aplicada à base Outex (em %).

Escalas (n)	3	[3 5]	[3 5 7]	[3 5 19 21]	[3 5 13 21 23]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	72,06	78,09	80,88	82,79	88,09	83,97
Acurácia Média	69,27	74,98	77,71	80,00	84,41	80,66
Desvio Padrão	0,012503	0,012962	0,012877	0,012037	0,012141	0,012696

Tabela 22 – Maiores Acurácias resultantes da Classificação aplicada à base UIUCTex (em %).

Escalas (n)	5	[3 5]	[3 5 17]	[3 5 13 25]	[3 5 7 13 25]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	53,80	63,60	71,00	76,40	78,00	71,40
Acurácia Média	47,81	59,38	65,28	69,64	72,05	66,41
Desvio Padrão	0,020784	0,020308	0,021328	0,020413	0,022383	0,019967

3.3.4.2 Outras Combinações de Escalas: Sequência de Fibonacci

Após realizar a análise das combinações (seção 3.3.4.1), resolveu-se realizar experimentos com uma sequência de números conhecida como Sequência de Fibonacci (CHANDRA; WEISSTEIN, 2018) $(0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \dots)$, mantendo o corte das frequências ($\omega = 6$).

Para os experimentos foram utilizados todos os números da sequência de Fibonacci maiores que 3 e menores que 55, bem como combinações desses números. Foram considerados também as caracterizações tomando os primeiros 5 e 4 elementos da sequência de Fibonacci. Considerou-se, também, apenas os números ímpares deste sequência ($n = 3, 5, 11, 21, 55$), bem como os 4 primeiros elementos ímpares da sequência de Fibonacci.

A Tabela 25 apresenta o resultado da classificação utilizando a base KTH-TIPS com as escalas baseadas na sequência de Fibonacci. Para a sequência completa (coluna A), definida por 7 escalas ($n = 3, 5, 8, 13, 21, 34, 55$), obteve-se 294 descritores e uma acurácia de 94,88%. Apesar do bom resultado, a quantidade de descritores foi elevada.

Reduziu-se a quantidade de escalas e conseqüentemente a quantidade de descritores

Tabela 23 – Maiores Acurácias resultantes da Classificação aplicada à base VisTex (em %).

Escalas (n)	3	[3 5]	[3 5 9]	[3 5 7 11]	[3 5 15 19 23]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	86,57	91,67	90,97	92,82	95,83	88,89
Acurácia Média	82,82	87,96	87,16	89,51	90,65	84,64
Desvio Padrão	0,015547	0,013053	0,014057	0,013884	0,014398	0,015320

Tabela 24 – Maiores Acurácias resultantes da Classificação aplicada à base Pollen (em %).

Escalas (n)	11	[7 15]	[9 11 17]	[5 11 17 19]	[9 11 13 21 25]	[3 7 11 15 19]
Descritores	42	84	126	168	210	210
Maior Acurácia	51,84	62,58	68,71	73,01	76,69	69,02
Acurácia Média	46,42	57,31	61,47	64,98	68,80	63,37
Desvio Padrão	0,020408	0,024331	0,024517	0,022018	0,022374	0,021567

para 5 escalas ($n = 3, 5, 8, 13, 21$) obtendo-se 210 descritores e acurácia de 93,66% (Tabela 25 coluna B). Observa-se que a redução para 84 descritores não prejudicou a acurácia.

Do mesmo modo explicitado anteriormente, utilizou-se apenas os valores ímpares da sequência ($n = 3, 5, 13, 21, 55$) resultando em 5 escalas e 210 descritores. Nesse experimento, observou-se que as escalas maiores interferiram e a acurácia ficou em 45,12% (Tabela 25 coluna C). Mantendo-se a mesma ideia dos valores ímpares, retirou-se a maior escala utilizada (55), tornando a sequência ($n = 3, 5, 13, 21$) com apenas 168 descritores (Tabela 25 coluna D) e, com isso, a acurácia ficou em 92,44%, comprovando o indicativo de que as escalas maiores não auxiliam na caracterização das imagens.

Outro experimento, utilizando-se os 4 primeiros elementos da sequência de Fibonacci ($n = 3, 5, 8, 13$), manteve-se 168 descritores, a mesma quantidade utilizada na coluna D, porém obteve-se acurácia de 93,90% (Tabela 25 coluna E), superando a acurácia obtida pelas escalas propostas em (PEREA; CARLSSON, 2014) (Tabela 25 coluna F) com menos descritores.

Tabela 25 – Resultado da Classificação da base KHT-TIPS para novas escalas baseadas em Fibonacci (em %).

	A	B	C	D	E	F
Escalas	[3 5 8 13 21 34 55]	[3 5 8 13 21]	[3 5 13 21 55]	[3 5 13 21]	[3 5 8 13]	[3 7 11 15 19]
Descritores	294	210	210	168	168	210
Maior Acurácia	94,88	93,66	45,12	92,44	93,90	92,68
Acurácia Média	91,37	89,97	37,75	89,04	89,63	87,52
Desvio Padrão	0,01281	0,01624	0,02269	0,01643	0,01754	0,01823

Observa-se que algumas acurácias obtidas com a sequência de Fibonacci (Tabela 26 colunas A, B, D e E) ficaram bem próximas à obtida na coluna F, porém com

uma quantidade de descritores reduzida. Novamente, consegue afirmar que a sequência proposta em (PEREA; CARLSSON, 2014) pode ser modificada, obtendo melhores acurácias.

3.3.4.3 Outras Combinações de Escalas: Sequência de Números Inteiros

Além da Sequência de Fibonacci, foram realizados experimentos com uma sequência de números inteiros maiores que 2 ($n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$). Novamente, não foram inseridos nos experimentos os valores iniciais da sequência (0, 1, 2) ou negativos por serem inferiores a 3 (restrição imposta pelas Equações 20 e 21). Optou-se por restringir a sequência numérica de inteiros a 12, pois conforme já demonstrado em experimentos anteriores, escalas maiores acabam prejudicando a caracterização das imagens.

Para o primeiro experimento foi utilizada uma sequência de números inteiros ($n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$) como escalas, mantendo o corte das frequências ($\omega = 6$). A coluna A da Tabela 26 ilustra o resultado dessa proposta na qual se aplicou o classificador LMNN nos descritores obtidos em 10 escalas definidas aplicadas à base KTH-TIPS, obtendo-se uma acurácia de 98,05%, porém com 420 descritores.

Tabela 26 – Resultado da Classificação da base KHT-TIPS para novas escalas (em %).

	A	B	C	D	E	F	G
Escalas	[3 4 5 6 7 8 9 10 11 12]	[3 4 5 6 7]	[4 6 8 10 12]	[3 5 7 9 11]	[3 4 5 8 10]	[3 4 5]	[3 7 11 15 19]
Descritores	420	210	210	210	210	126	210
Maior Acurácia	98,05	92,20	94,15	94,15	95,12	95,12	92,68
Acurácia Média	94,33	87,09	90,05	89,75	92,07	90,99	87,52
Desvio Padrão	0,01559	0,01811	0,01911	0,01772	0,01605	0,01525	0,01823

No segundo experimento reduziu-se a quantidade de escalas escolhendo os 5 primeiros n 's ($n = 3, 4, 5, 6, 7$), a mesma quantidade utilizada em (PEREA; CARLSSON, 2014), resultando em 210 descritores. Os resultados são apresentados na coluna B. Nota-se que o valor de 92,68% foi obtido como maior acurácia, próximo ao obtido pelas escalas da coluna G.

O terceiro experimento utilizou apenas os valores pares maiores que 2 e menores que 12, o que mantém a mesma quantidade de descritores (Tabela 26 coluna C). Verifica-se que com essa mudança obteve-se 94,15% de acurácia, mantendo a mesma quantidade de descritores da coluna G, porém com um aumento na acurácia.

No quarto experimento utilizou-se apenas os valores ímpares para as escalas ($n = 3, 5, 7, 9, 11$), o que mantém a mesma quantidade de descritores (Tabela 26 coluna D). Verifica-se que com essa mudança a acurácia ficou semelhante à obtida para os valores pares.

Ao se comparar os resultados das colunas B, C e D percebeu-se que introduzindo escalas de tamanho maiores junto a escalas de tamanhos menores obtém-se melhores

resultados. Isto motivou um quinto experimento realizado com a sequência $n = 3, 4, 5, 8, 10$, iniciando sequencialmente e depois inserindo valores maiores. Os valores 8 e 10 foram escolhidos sem um critério, poderiam ter sido outros dois valores fora da sequência inicial. Observa-se que a acurácia desse experimento (Tabela 26 coluna E) foi superior a todos os outros experimentos desta seção, ficando em 95,12% com 210 descritores.

Analisando-se os resultados obtidos, foi realizado outro experimento no qual optou-se por manter somente escalas menores e reduzir a quantidade de escalas resultando em $n = 3, 4, 5$ e o corte das frequências em $\omega = 6$, obtendo 126 descritores. A acurácia obtida para este experimento ficou em 95,12% e a acurácia média 90,99% com um desvio de 0,01525.

Observa-se que a maior acurácia obtida, com a quantidade de descritores inferiores à 215, para esses experimentos da Tabela 26 foi de 95,12% sendo atingida nas colunas E e F. Analisando-se os resultados dessas colunas, percebe-se que no item Maior Acurácia os dois experimentos tiveram o mesmo valor, porém a quantidade de descritores da coluna F é menor. Mesmo com a redução de valor no item Acurácia Média em pouco menos de 2%, ocorreu uma redução na quantidade de descritores no experimento da coluna F e, com isso, testou-se para as outras bases a mesma configuração utilizada nessa coluna: escalas $n = 3, 4, 5$.

A Tabela 27 ilustra o resultado da utilização das escalas $n = 3, 4, 5$ que obteve 126 descritores para as 6 bases testadas. Percebe-se que para algumas bases não se obteve altos valores de acurácia devido às características da própria base, por exemplo a UIUCTex que contém texturas rotacionadas. Mesmo assim, a quantidade de descritores foi reduzida e os valores da acurácia não reduziram tanto.

Tabela 27 – Resultado da Classificação das bases para as escalas $n = 3, 4, 5$ (em %).

	Brodatz	Vistex	CUReT	KTH-TIPS	Outex	UIUCTex	Pollen
Maior Acurácia	94,71	95,14	88,39	95,12	81,47	58,80	66,56
Acurácia Média	93,04	92,66	87,19	90,99	78,60	53,08	60,60
Desvio Padrão	0,00692	0,01140	0,00597	0,01525	0,01294	0,01881	0,02111

A Tabela 28 ilustra os resultados do método proposto, utilizando o descritor *KFchap* com 3 escalas ($n = 3, 4, 5$), e de métodos selecionados da literatura para as seis bases de imagens. Estes métodos utilizam outros classificadores como o K-NN ou Redes Neurais. Enfatiza-se este fato, pois os resultados apresentados para os métodos de Varma e Zisserman (2004), Crosier e Griffin (2010), Perea e Carlsson (2014) e Zhang et al. (2015b) foram obtidos diretamente dos respectivos artigos devido a complexidade do método ou falta de informações nos respectivos artigos para implementação dos mesmos. Com isso, para alguns métodos faltam resultados para uma ou mais bases. Dos resultados tabelados,

foram implementados os métodos de primeira ordem, wavelets, LBP, Gabor, Backes, Casanova e Bruno (2013) e o método proposto.

Tabela 28 – Comparativo do método proposto com outros métodos da literatura para algumas das bases testadas.

Método	Número de Descritores	Acurácia (%)			
		KTH-TIPS	CUReT	Brodatz	Vistex
Primeira Ordem	5	57,78	19,73	40,20	52,66
Wavelets	36	58,52	50,99	70,27	72,11
LBP	25	74,07	67,37	89,58	88,89
Gabor	64	80,12	82,60	82,49	91,67
(BACKES; CASANOVA; BRUNO, 2013)	108	90,37	84,32	95,27	86,76
(ZHANG et al., 2006)	100	94,90 ± 1,6	-	94,90 ± 0,7	-
(CROSIER; GRIFFIN, 2010)	1.296	-	98,6 ± 0,2	-	-
(PEREA; CARLSSON, 2014)	215	94,77 ± 1,3	95,66 ± 0,45	-	-
(VARMA; ZISSERMAN, 2004)	2.440	-	97,43	-	-
(Jun Zhang; Jimin Liang; Heng Zhao, 2013)	6.561	-	92,44 ± 1,04	-	-
(ZHANG et al., 2015b)	648	98,86 ± 1,12	94,44 ± 1,13	-	-
Método Proposto	126	95,12	88,39	93,04	95,14

Para esta tabela, a variância do método proposto ficou abaixo de 0,00023%.

Com a redução da quantidade de descritores por meio da redução da quantidade de escalas utilizadas, o corte das frequências (ω) torna-se um ponto a ser avaliado, pois dependendo da frequência utilizada pode-se reduzir ainda mais a quantidade de descritores ou até mesmo abre a possibilidade de aumentar a quantidade de escalas mantendo uma quantidade reduzida de descritores.

3.3.5 Análise das Frequências

Perea e Carlsson (2014) utilizaram o mesmo corte das frequências ($\omega = 6$) proposto por Wit e Floriani (1998) para o cálculo da função estimadora \hat{f} , influenciando no cálculo dos coeficientes estimados de K-Fourier (Eq. 41). Não foram analisadas outras frequências de corte e, aqui, propõe-se analisar o corte da frequência com ω variando de 2 a 12. Optou-se por não utilizar frequências maiores que 12, pois conforme já explicitado anteriormente, as altas frequências ajustam os detalhes finos e, com isso, as informações mais relevantes de $f(\alpha, \theta)$ são codificadas nos termos de baixa frequência.

3.3.5.1 Frequências de 2 a 12

Para iniciar as análises das frequências, realizou-se experimentos com a base KTH-TIPS mantendo-se as escalas definidas em (PEREA; CARLSSON, 2014) ($n = 3, 7, 11, 15, 19$), porém variando-se as frequências de 2 à 12. Os experimentos foram limitados a quantidade de descritores gerados por Perea e Carlsson (2014).

A Tabela 29 ilustra os resultados obtidos na classificação da base KTH-TIPS para as escalas $n = 3, 7, 11, 15, 19$ com com frequências de 2 à 12. Observa-se que a quantidade

de descritores está diretamente relacionada às escalas e à frequência de corte, conforme explicitado na Seção 3.1.2.

Tabela 29 – Resultado da Classificação com LMNN da base KHT-TIPS para as escalas $n = 3, 7, 11, 15, 19$ com frequências de 2 à 12 (em %).

Frequências	2	3	4	5	6	7
Descritores	30	55	100	145	210	275
Maior Acurácia	67,32	77,07	85,12	88,54	90,98	93,17
Acurácia Média	61,85	73,53	81,35	84,07	87,37	89,91
Desvio Padrão	0,01997	0,01947	0,01883	0,01913	0,01443	0,01792

Frequências	8	9	10	11	12
Descritores	360	445	550	655	780
Maior Acurácia	96,10	96,83	71,95	97,80	99,27
Acurácia Média	92,25	93,39	66,86	95,66	96,57
Desvio Padrão	0,01410	0,01660	0,02219	0,01225	0,01229

Conforme ilustram os resultados da Tabela 29, ao se aumentar o valor do corte das frequências (ω), conseqüentemente aumenta-se o tamanho do descritor e, na maioria dos casos, a acurácia também aumenta. Comparando-se os resultados das frequências 5 e 6, observa-se que para se ter um aumento de aproximadamente 2% na Acurácia Média ou na Maior Acurácia, a quantidade de descritores aumentou em 65 itens. O melhor resultado obtido utiliza a frequência 12 obtendo-se 99,27% de acurácia.

Utilizando-se das análises realizadas na Seção 3.3.4.1 (na qual para a grande maioria das bases analisadas, as escalas [3, 5] obtiveram os melhores resultados proporcionalmente à quantidade de descritores utilizados na caracterização) realizou-se um teste com as escalas $n = 3, 5$ e as frequências de 2 a 12. A Tabela 30 ilustra os resultados obtidos para a base KTH-TIPS.

Analisando-se os resultados da Tabela 30, observa-se que proporcionalmente a quantidade de descritores, a frequência 4 obteve um bom resultado. Com isso, realizou-se o mesmo teste para as outras bases de textura, mantendo as escalas $n = 3, 5$ com a frequência 4, obtendo 40 descritores para cada base. A Tabela 31 ilustra os resultados obtidos.

Os resultados utilizando o corte da frequência $\omega = 4$ e escalas $n = 3, 5$ ilustrados na Tabela 31 tiveram acurácias inferiores às obtidas para as mesmas escalas com $\omega = 6$, por exemplo, como pode ser visto na Tabela 18. Porém a quantidade de descritores com $\omega = 6$ é de 84 e com $\omega = 4$ ficou em 40 descritores, menos da metade.

Conforme apresentado na Seção 3.3.4 (Tabela 27) as escalas $n = 3, 4, 5$ obtiveram bons resultados e, com isso, realizou-se o experimento da variação das frequências de 2 a 12 para a base KTH-TIPS. A Tabela 32 ilustra os resultados obtidos.

Tabela 30 – Resultado da Classificação com LMNN da base KHT-TIPS para as escalas $n = 3, 5$ com frequências de 2 à 12 (em %).

Frequência	2	3	4	5	6	7
Descritores	12	22	40	58	84	110
Maior Acurácia	73,66	87,32	90,00	91,95	91,46	93,90
Acurácia Média	69,70	81,96	86,68	87,89	87,88	89,23
Desvio Padrão	0,01952	0,01892	0,01786	0,01897	0,01823	0,01730

Frequência	8	9	10	11	12
Descritores	144	178	220	262	312
Maior Acurácia	93,90	92,68	93,90	94,39	95,61
Acurácia Média	88,60	89,73	90,42	90,71	91,27
Desvio Padrão	0,01948	0,01417	0,01645	0,01612	0,01658

Tabela 31 – Resultado da Classificação com LMNN para as escalas $n = 3, 5$ com frequência 4 para as diversas bases (em %).

Bases	KTH-TIPS	Brodatz	CUReT	Outex	UIUCTex	Vistex	SIPI	Pollen
Maior Acurácia	90,00	89,08	79,83	75,59	58,40	89,58	89,74	55,52
Acurácia Média	86,68	87,25	78,46	72,57	53,15	86,06	76,18	50,56
Desvio Padrão	0,01786	0,00856	0,00655	0,01121	0,01894	0,01358	0,06440	0,02274

Tabela 32 – Resultado da Classificação com LMNN para as escalas $n = 3, 4, 5$ para a base KHT-TIPS (em %).

Frequência	2	3	4	5	6	7
Descritores	18	33	60	87	126	165
Maior Acurácia	81,95	90,24	93,41	94,63	91,71	92,68
Acurácia Média	75,17	85,30	88,97	90,33	87,16	88,53
Desvio Padrão	0,02391	0,02108	0,01692	0,01542	0,02027	0,01777

Frequência	8	9	10	11	12
Descritores	216	267	330	393	468
Maior Acurácia	91,95	96,10	96,10	95,61	92,20
Acurácia Média	87,53	92,22	92,40	92,96	87,10
Desvio Padrão	0,01641	0,01587	0,01553	0,01475	0,01894

Observando os resultados da Tabela 32, o corte da frequência $\omega = 5$ teve a melhor acurácia com a quantidade relativamente pequena de descritores obtendo 94,63%. Este corte de frequência ($\omega = 5$) foi utilizado para outras bases de texturas. Os resultados do método proposto para outras bases estão ilustrados na Tabela 33.

A Tabela 33 ilustra os resultados do método proposto, utilizando o descritor *KFchap* com 3 escalas ($n = 3, 4, 5$), e de métodos selecionados da literatura para quatro bases de

imagens. Os métodos da literatura utilizam outros classificadores como o K-NN ou Redes Neurais. Enfatiza-se este fato, pois não foi possível a implementação de todos os métodos comparados devido a complexidade do método ou falta de informações nos respectivos artigos. Com isso, para métodos que faltam resultados para uma ou mais bases, estes correspondem aos apresentados no respectivo artigo.

Os resultados apresentados na Tabela 33 para os métodos de Zhang et al. (2006), Perea e Carlsson (2014), Quan et al. (2014), Zhang et al. (2015b), Ji et al. (2018) e Yang et al. (2018) foram obtidos diretamente dos respectivos artigos devido a complexidade do método ou falta de informações nos respectivos artigos para implementação dos mesmos. Com isso, para alguns métodos faltam resultados para uma ou mais bases. Dos resultados tabelados, foram implementados os métodos owavelets, LBP, Gabor, Backes, Casanova e Bruno (2013) e o método proposto.

Tabela 33 – Comparativo do método proposto com outros métodos da literatura para algumas das bases testadas.

Método	Número de Descritores	Acurácia (%)			
		KTH-TIPS	CUReT	Brodatz	Vistex
Wavelets	36	58,52	50,99	70,27	72,11
LBP	25	74,07	67,37	89,58	88,89
Gabor	64	80,12	82,60	82,49	91,67
(BACKES; CASANOVA; BRUNO, 2013)	108	90,37	84,32	95,27	86,76
(JI et al., 2018)	3.888	-	-	98,55 ± 0,53	-
(PEREA; CARLSSON, 2014)	215	94,77 ± 1,3	95,66 ± 0,45	-	-
(QUAN et al., 2014)	140	98,4	-	-	-
(YANG et al., 2018)	20	-	-	85,02	71,23
(ZHANG et al., 2006)	100	94,90 ± 1,6	-	94,90 ± 0,7	-
(ZHANG et al., 2015b)	648	98,86 ± 1,12	94,44 ± 1,13	-	-
Método Proposto	87	94,63	86,78	94,14	95,14

Para esta tabela, o desvio padrão do método proposto ficou abaixo de 0,018.

3.4 Considerações Finais

Baseado no trabalho de Perea e Carlsson (2014), propôs-se investigar o uso da projeção de *patches* das imagens para a base de imagens no Espaço da garrafa de Klein para a criação de um vetor de características.

Analisou-se a escolha dos *patches* tanto na obtenção quanto na seleção dos mesmos e, pôde-se confirmar que o uso da aleatoriedade não influenciou nos resultados.

Analisou-se, também, a configuração do vetor de característica em função das escalas definidas por Perea e Carlsson (2014) e, percebeu-se que a utilização da combinação de novas escalas obteve melhores resultados quando comparados aos resultados obtidos na proposta inicial, além de uma redução na quantidade de descritores. Essa análise resultou na publicação do trabalho Ribeiro, Oliveira e Barcelos (2018).

Outro fator que influenciou na acurácia e na quantidade de descritores é o corte das frequências (ω) utilizado por Perea e Carlsson (2014) que manteve o mesmo valor ($\omega = 6$) proposto por Wit e Floriani (1998) para o cálculo da função estimadora \hat{f} . Utilizando-se dos experimentos com a modificação no corte das frequências (ω) combinado à modificação nas escalas obteve-se melhores resultados quando comparados à literatura e aos resultados de Perea e Carlsson (2014). Essas análises resultaram na publicação do trabalho (RIBEIRO; OLIVEIRA; BARCELOS, 2019).

Analisando-se todos os resultados obtidos neste capítulo, verifica-se que ainda deve ser investigado com mais profundidade a influência do denominador M em outras bases de texturas, pois com os testes realizados para as bases KTH-TIPS e Pollen, não foram obtidos os mesmos comportamentos em todos os experimentos realizados.

Outro fator que deverá ser investigado é a utilização de polinômios com termos de mais alto grau para a aproximação do $p(x, y)$ e p_{ij} (Equações 15 e 16, respectivamente) que podem resultar em uma melhora na caracterização de texturas por meio da garrafa de Klein.

Padrões Semânticos

Atributos visuais de uma imagem comumente são utilizados para sua caracterização. Esses atributos visuais devem ser escolhidos de forma que caracterizem a imagem de forma única, e que essas características encontradas estejam presentes em uma mesma classe de imagens, tornando-as diferenciáveis de outras classes. Em alguns casos, apenas atributos visuais não são suficientes para caracterizar uma imagem. Uma forma alternativa para a representação de imagens é a utilização de semântica.

As características semânticas podem ser gerais ou específicas para um dado objeto em estudo, logo dependem do problema a ser estudado e, portanto, do banco de dados a ser analisado.

Neste trabalho serão utilizados padrões semânticos para analisar imagens de grãos de pólen. Características semânticas desta base podem ser um subconjunto das características gerais de qualquer imagem da base. Para a obtenção de um vetor de características reduzido para a base Pollen (Apêndice B), propõe-se a caracterização semântica de grãos de pólen, não apenas utilizando-se da semântica global das imagens, mas criando uma semântica específica para a base de imagens.

A organização deste capítulo contempla a fundamentação teórica com uma revisão da bibliografia e trabalhos do estado-da-arte sobre Padrões Semânticos, além de uma análise de trabalhos relacionados da literatura, a apresentação de uma semântica específica para a base Pollen objetivando caracterizar semanticamente os grãos de pólen. São apresentados a definição do espaço semântico, o mapeamento dos dados nesse espaço semântico, validações, resultados e as conclusões obtidas.

4.1 Fundamentação Teórica

Os conceitos de semântica são utilizados nos mais diversos casos, como por exemplo, na recuperação de imagens em grandes volumes de dados, na mineração de dados através da rotulação de objetos em cenas ou mesmo na identificação textual em imagens.

Diversos exemplos da utilização da semântica para obter e recuperar informação estão descritos na literatura. Cao e Wang (2015) exploram a semântica para recuperação de imagens em mídias sociais, propondo um novo algoritmo para combinar conteúdo multimídia de forma robusta e *tags* associadas pela mineração a semântica latente que leva em conta as *tags* semanticamente semelhantes.

He et al. (2003) utilizaram-se de métodos espectrais para inferir um espaço semântico de *feedback* de relevância do usuário para a recuperação de imagens, fazendo com que por meio da interação dos usuários o sistema melhore o desempenho.

Foi utilizado por Pandey, Khanna e Yokota (2015) uma categorização semântica em uma base de dados de imagem hierárquica semanticamente categorizada, propondo um sistema de atribuição de semântica automática de imagens utilizando a adaptação e combinação de vários recursos visuais da imagem.

A proposta de Zhang et al. (2015a) é a combinação de representação de imagem e classificação em espaços semânticos aleatórios, codificando características locais para a representação de imagens. Selecionaram aleatoriamente as representações gerando espaços semânticos aleatórios para o mapeamento das imagens. A representação semântica mapeada é usada como a representação da imagem final.

Foi trabalhado por Banica e Sminchisescu (2015) o problema de segmentação semântica baseada em dados RGB-D (Vermelho Verde Azul - Profundidade), com ênfase em análise de cenas internas desordenadas que contêm muitas categorias visuais e instâncias destas. Utilizaram a base *NYU Depth Dataset V2* (SILBERMAN et al., 2012) que contém 1449 imagens RGB-D e, segundo os autores, obtiveram 82,42% de acurácia.

Sjöberg et al. () propuseram o uso de alguns métodos para recuperação de objetos multimídias por meio da combinação de informações semânticas de descritores visuais e textuais. Para isso, foram utilizados os mapas auto-organizáveis para o treinamento com descritores de baixo nível. Foram propostos nesse artigo, o uso dos descritores: textuais, cor, texturas, histograma de palavras, frequência das palavras, palavras-chave binárias.

4.2 Propostas

Características semânticas podem ser gerais ou específicas para um dado objeto em estudo, logo dependem do Banco de Dados a ser analisado. Neste trabalho serão utilizados padrões semânticos para analisar imagens de grãos de pólen. Características semânticas desta base podem ser um subconjunto das características gerais de qualquer imagem da base. Propõe-se a criação de uma semântica específica para a base Pollen objetivando um vetor de características que possa representar a base.

4.2.1 Criação da Semântica Específica para a base Pollen

A base Pollen utilizada no trabalho contém imagens de 30 diferentes espécies de pólen. A Figura 13 ilustra 10 dessas espécies e, pode ser observado que algumas espécies presentes nessa base são semelhantes no formato, porém a estrutura interna as diferenciam.

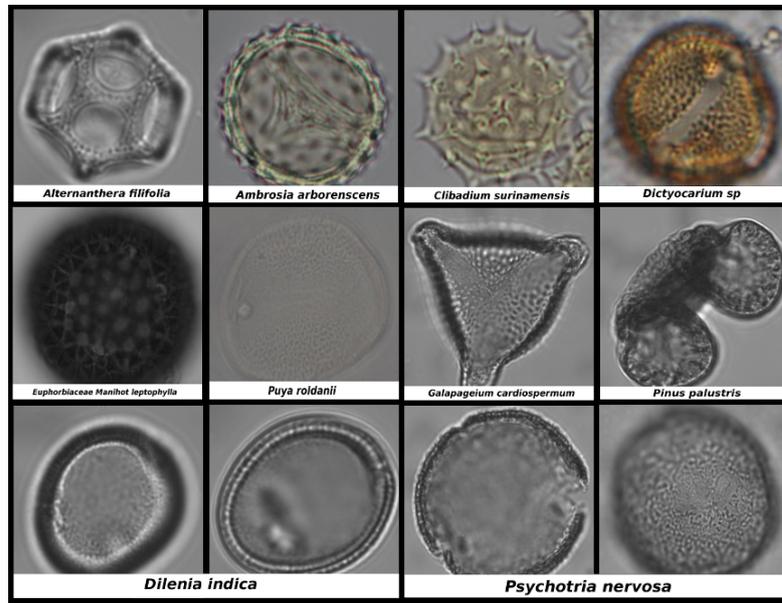


Figura 13 – Imagens exemplificando 10 classes da base Pollen. As imagens não estão em escala.

Para a criação da semântica específica para a base Pollen propõe-se um método que consiste, basicamente, de duas partes: a definição do espaço semântico, com a representação no nível semântico, e o mapeamento deste espaço semântico, conforme ilustrado pelo diagrama da Figura 14.

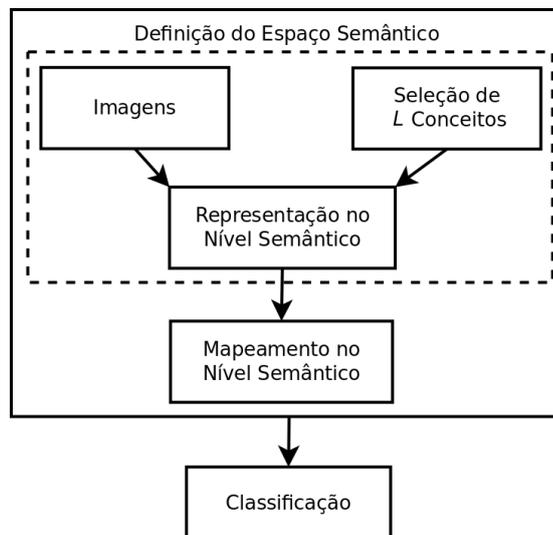


Figura 14 – Diagrama Geral do Método Proposto para a Semântica.

4.2.2 Definição do Espaço Semântico

O método proposto para a criação da semântica específica para a base Pollen utiliza partes das imagens dos grãos de pólen para a criação de padrões (conceitos) que representem as classes definidas na base.

Do mesmo modo que Kwitt et al. (2012), o conceito é criado selecionando-se partes de imagens que representem a base de dados, sendo selecionados por análise visual considerando o conceito de melhor representante de classe. Os conceitos selecionados podem ser em tons de cinza (Figura 15) ou em preto e branco. A Figura 16, lado esquerdo, ilustra uma imagem da base Pollen em tons de cinza (ao fundo) e, em destaque, o conceito selecionado em preto e branco desta imagem. À direita estão ilustrados conceitos criados a partir de outras imagens da base. Neste trabalho optou-se pelo uso de conceitos em preto e branco, porém devem ser investigados o uso de conceitos sem essa modificação, mantendo o padrão de cores da imagem original.

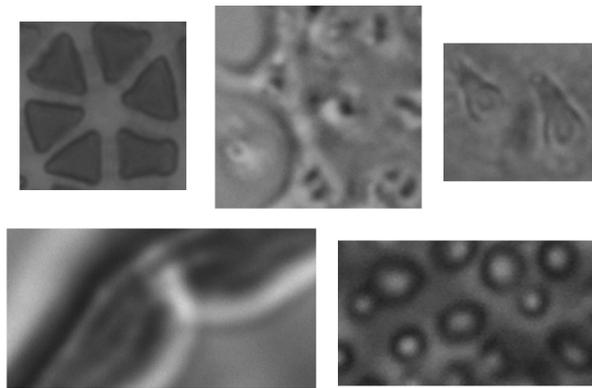


Figura 15 – Exemplos de 5 conceitos em tons de cinza de imagens da base Pollen. As imagens não estão em escala.

Para a definição do espaço semântico seleciona-se L Conceitos Semânticos retirados das imagens da base Pollen. A representação no nível semântico é realizada através de uma caracterização qualquer. Neste trabalho, cada um dos L conceitos é representado por um vetor $X_j = (X_{LBP}, X_{HOG})$ para a composição da Semântica. O vetor $X_j = (x_1, \dots, x_{j_n}, x_{j_n+1}, \dots, x_{j_m})$ onde $j = 1, 2, \dots, L$ é obtido utilizando-se o Padrão Binário Local (LBP) para a definição das j_n primeiras posições do vetor e o Histograma de Gradientes Orientados (HOG) para as demais posições do vetor. O mesmo procedimento utilizado para os conceitos foi adotado para todas as imagens da base Pollen, gerando para cada uma das N imagens da base um vetor $Z_i = (Z_{LBP}, Z_{HOG})$ onde $i = 1, 2, \dots, N$ para representar as imagens. A dimensão de cada vetor Z_i varia de acordo com o tamanho da imagem i e a dimensão do vetor X_j varia de acordo com o tamanho do conceito j .

A junção do HOG e do LBP tem sido utilizada em diversos trabalhos da literatura como Song et al. (2012), Nair et al. (2014), Qiao, Cappelle e Ruichek (2015), Li et al.

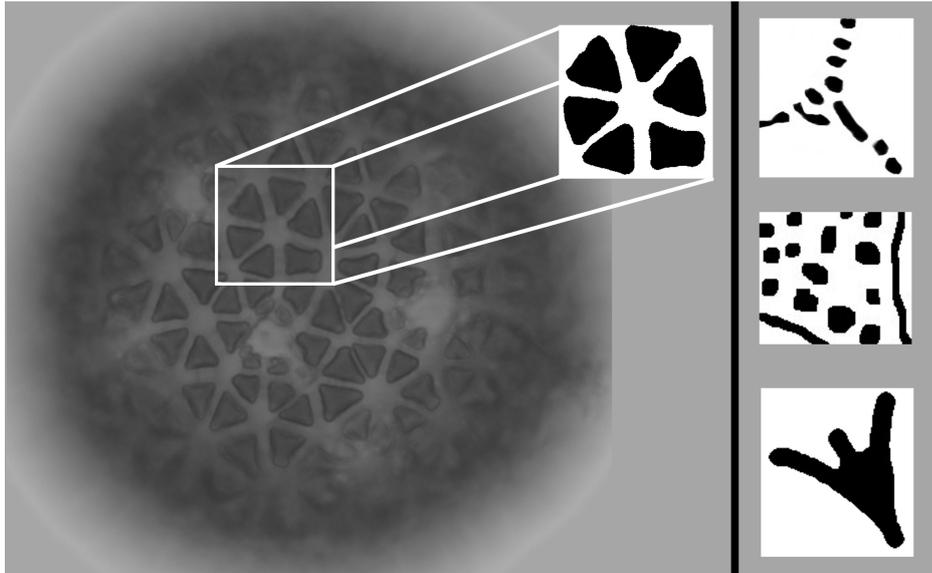


Figura 16 – Exemplos de conceitos. À esquerda, a imagem original em tons de cinza (ao fundo) e, em destaque, o exemplo do conceito selecionado da imagem em preto e branco. À direita estão 3 exemplos, em preto e branco, de conceitos selecionados de outras imagens da base Pollen.

(2016), Konstantinidis et al. (2017), com o objetivo de se complementarem.

4.2.3 Mapeamento do Espaço Semântico

Após a Definição do Espaço Semântico, faz-se o Mapeamento do Espaço Semântico criado aplicando-se a correlação cruzada normalizada (Eq. 65) entre cada vetor Z_i (Imagem) com o conceito X_j , obtendo para cada correlação um vetor CCN_{ij} onde $i = 1, 2, \dots, N$ e $j = 1, \dots, L$ com $\{(\text{dimensão de } Z_{LBP_i} + \text{dimensão de } Z_{HOG_i}) + (\text{dimensão de } X_{LBP_j} + \text{dimensão de } X_{HOG_j}) - 2\}$ atributos.

Para exemplificar os valores das dimensões, utilizando-se uma imagem de dimensões 256×256 pixels e um conceito com dimensões 50×50 pixels, ao se criar o espaço semântico desta imagem com este conceito, obtém-se um Z_i com 23.204 atributos e X_j com 8.243 atributos. Aplicando-se a correlação cruzada normalizada, apenas para esta imagem com este conceito, obtém-se um vetor com 31.445 atributos. Ao aplicar a CCN nas imagens de uma base, por exemplo a Pollen, que tem maiores dimensões, obtém-se vetores CCN_{ij} com dimensões superiores à 2.500.000 atributos.

Utilizou-se a correlação cruzada normalizada devido a esta ser invariante a mudanças de intensidade da imagem (LEWIS, 1995), como por exemplo, mudanças causadas pelas condições de aquisição das imagens impedindo que a correlação entre uma característica e uma região correspondendo exatamente à imagem possa ter valor inferior à correlação entre a mesma característica e a mesma região porém com valores mais altos de intensidade.

Devido a alta dimensionalidade dos vetores obtidos na correlação cruzada normalizada, para a realização da análise dos atributos, aplicou-se em cada vetor o Coeficiente de Variação (CoV) (Eq. 66) e o Coeficiente Relativo de Variação (CRoV) (Eq. 67), mapeando as informações no espaço semântico.

Desta forma, cada imagem da base será representada por um vetor de características de dimensão $4L$, sendo que, nas posições ímpares desse vetor estão os valores CoV e as posições pares estão os valores do CRoV.

4.3 Experimentos e Análise dos Resultados

Nesta seção serão apresentados os experimentos e análises referentes aos padrões semânticos, apresentando o desenvolvimento da proposta, análises da caracterização e um comparativo de resultados.

4.3.1 Desenvolvimento da Proposta

Conforme a proposta feita na Seção 4.2, inicialmente selecionou-se os conceitos para a definição do espaço semântico. Considerando as 30 classes da base Pollen ($N = 1063$ amostras) foram selecionados um total L de 12 conceitos. Estes conceitos foram selecionados por análise visual considerando o conceito de melhor representante de classe. Após a seleção de todos os conceitos, aplicou-se a limiarização através do método de Otsu (OTSU, 1979), deixando-os em preto e branco. A Figura 17 ilustra todos os conceitos selecionados em preto e branco.

Para a definição do espaço semântico, neste trabalho, optou-se por LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996) e HOG (DALAL; TRIGGS, 2005).

O HOG foi utilizado por conseguir obter características de bordas e informações locais de formatos dos objetos de uma imagem. Um ponto negativo é que o HOG é sensível a mudanças de orientações de gradiente do *pixel* e com isso, imagens com ruídos provenientes de sua obtenção acabam afetando o resultado do algoritmo.

O LBP foi utilizado por superar o problema dos ruídos, pois extrai a informação da textura local através da criação de um limiar para um número determinado de vizinhos baseado no valor do *pixel* central de uma vizinhança local.

Para realizar o mapeamento do espaço semântico criado serão utilizados coeficientes obtidos da Correlação Cruzada Normalizada (CCN) (LEWIS, 1995), conforme ilustra o diagrama da Figura 18. O classificador escolhido foi o KNN com $k = 1$ e $fold s = 10$ para que os resultados validem o método proposto, sem que a implementação do classificador interfira na acurácia.

Criou-se um conjunto de 1063 vetores com $4L$ características cada um para o mapeamento dos dados no espaço semântico para a base Pollen. O conjunto de vetores foi submetido ao classificador KNN. A Tabela 34 ilustra os resultados obtidos.

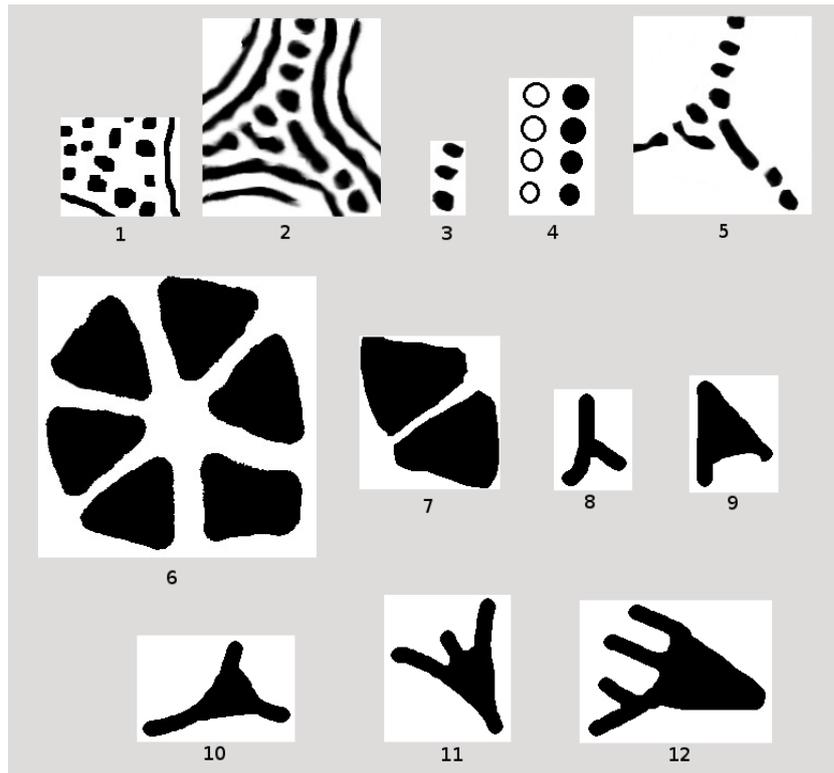


Figura 17 – Todos os L conceitos selecionados para a base Pollen. Para melhor visualização, os conceitos não estão em escala.

Tabela 34 – Resultado do Método Proposto para a base Pollen utilizando o classificador KNN (em %).

Atributos	48
Maior Acurácia	90,12
Acurácia Média	89,17
<i>Precision</i> Médio	89,98
<i>Recall</i> Médio	88,84
<i>F Score</i> Médio	89,13
Desvio Padrão	0,43098

Para uma melhor análise dos resultados da Tabela 34, especificou-se por classe o resultado da classificação, ilustrados na Tabela 35.

Nos resultados obtidos (Tabela 35) observa-se que algumas classes obtiveram um baixo valor de acurácia. Para todos os métodos aplicados, a classe 23 (*Sabal palmetto*) geralmente ficou com o menor valor de acurácia. Isso se deve principalmente pela quantidade reduzida de imagens na classe e o fato das texturas presentes nas imagens da classe terem propriedades que se assemelham a outras classes.

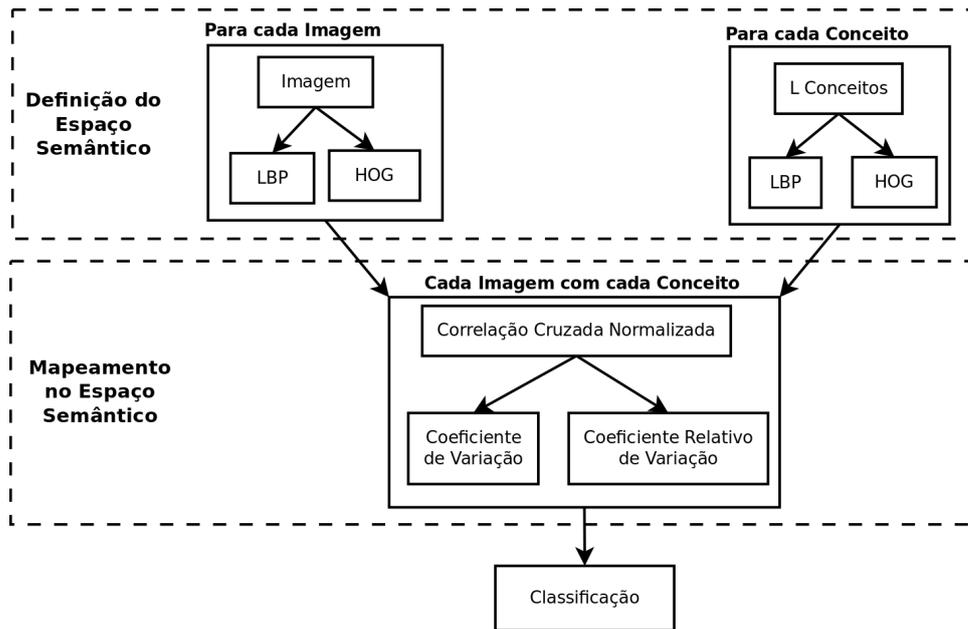


Figura 18 – Diagrama detalhado do Método Proposto para a Semântica.

4.3.2 Confusões

Algumas imagens de certas classes da base Pollen foram classificadas como sendo de outra classe. Isso se deve ao fato do vetor de características não conseguir identificar unicamente cada imagem da base. Observa-se na Tabela 36 a Matriz de Confusão obtida da classificação que gerou o dados da Tabela 34.

Analisando a Matriz de Confusão (Tabela 36), dentre as classes confundidas, para exemplificar, a classe 23 (em destaque) foi confundida com as classes 02, 04, 07 e 08. A Figura 19 ilustra imagens das classes confundidas. As semelhanças podem ser notadas entre as imagens em termos de forma, direção e bordas. As outras classes confundidas também foram analisadas.

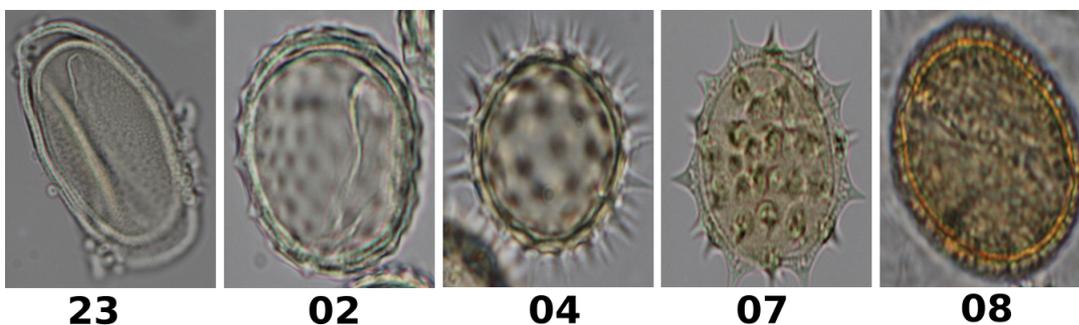


Figura 19 – Exemplos de imagens das classes 2, 4, 7 e 8, que ficaram indistintas da classe 23 para o classificador.

Para identificar quais conceitos estariam gerando a confusão no classificador, realizou-se experimentos utilizando combinações com os L conceitos. No primeiro experimento

Tabela 35 – Resultados do Método Proposto para a base Pollen por classe.

Classes	Erros	Qtd. Imagens	Acertos (%)
1	2	40	95,00
2	7	39	82,05
3	3	40	92,50
4	0	33	100,00
5	5	40	87,50
6	11	22	50,00
7	5	45	88,89
8	3	29	89,66
9	4	40	90,00
10	0	40	100,00
11	7	29	75,86
12	3	40	92,50
13	4	40	90,00
14	1	29	96,55
15	4	40	90,00
16	1	38	97,37
17	0	28	100,00
18	4	32	87,50
19	5	40	87,50
20	2	40	95,00
21	1	17	94,12
22	2	42	95,24
23	6	13	53,85
24	2	40	95,00
25	9	35	74,29
26	5	40	87,50
27	2	38	94,74
28	0	33	100,00
29	6	40	85,00
30	1	41	97,56

foram utilizados apenas os conceitos de 1 a 5. Para o segundo experimento foram utilizados apenas os conceitos 6 e 7. Para o terceiro experimento foram utilizados apenas os conceitos 8 a 12. Para o quarto experimento foram utilizados apenas os conceitos 1 a 7. Para o quinto experimento foram utilizados apenas os conceitos 6 a 12 e para o sexto experimento foram utilizados todos os 12 conceitos. A identificação dos números com os conceitos está presente na Figura 17.

A Tabela 37 apresenta os resultados da acurácia por classe para cada experimento descrito anteriormente. Cada coluna da Tabela ilustra os resultados para utilização de diferentes conceitos, conforme o experimento. Nas duas últimas linhas da Tabela, estão os valores Totais de acurácia e a quantidade de atributos obtidas para cada experimento.

Observa-se que a utilização dos 12 conceitos obteve a maior acurácia (90,12%). Ao se analisar os demais resultados, observa-se que utilizando os conceitos 6 a 12 obteve-se 86,92% com apenas 28 atributos. Nestes resultados, a classe 23, continua com baixa acurácia quando utilizados certos conceitos.

Baseado nos resultados e análises, pretende-se estudar outras combinações entre os 12 conceitos podendo manter ou elevar a acurácia do método. Pretende-se, também, estudar formas para a obtenção automática de conceitos utilizando de técnicas da literatura.

Além do uso de padrões semânticos para a caracterização de grãos de pólen, alguns

Tabela 37 – Resultados da classificação do Método Proposto para a base Pollen por classe e conjunto de Conceito (%)

Classes	Conceitos					
	1-5	6-7	8-12	1-7	6-12	1-12
1	87,50	82,50	85,00	92,50	90,00	92,50
2	82,05	71,79	82,05	79,49	84,62	89,74
3	80,00	85,00	75,00	92,50	95,00	95,00
4	100,00	100,00	100,00	100,00	100,00	100,00
5	82,50	72,50	77,50	77,50	85,00	92,50
6	68,18	45,45	68,18	54,55	45,45	50,00
7	88,89	73,33	82,22	88,89	86,67	88,89
8	96,55	89,66	100,00	100,00	100,00	100,00
9	55,00	60,00	82,50	80,00	92,50	92,50
10	100,00	87,50	97,50	100,00	100,00	100,00
11	58,62	44,83	58,62	62,07	68,97	68,97
12	75,00	32,50	80,00	85,00	80,00	90,00
13	82,50	55,00	92,50	92,50	95,00	95,00
14	89,66	93,10	96,55	93,10	100,00	96,55
15	92,50	70,00	92,50	92,50	95,00	97,50
16	84,21	81,58	86,84	100,00	92,11	97,37
17	100,00	85,71	92,86	100,00	100,00	100,00
18	87,50	62,50	84,38	87,50	78,13	87,50
19	75,00	50,00	75,00	72,50	80,00	85,00
20	82,50	67,50	75,00	87,50	85,00	85,00
21	94,12	94,12	94,12	94,12	94,12	94,12
22	78,57	61,90	85,71	85,71	92,86	92,86
23	53,85	38,46	38,46	30,77	38,46	53,85
24	92,50	75,00	90,00	95,00	87,50	95,00
25	45,71	45,71	62,86	54,29	60,00	57,14
26	85,00	52,50	82,50	85,00	77,50	92,50
27	94,74	86,84	89,47	92,11	92,11	92,11
28	96,97	100,00	96,97	100,00	100,00	100,00
29	80,00	65,00	70,00	85,00	70,00	82,50
30	92,68	0,00	97,56	100,00	100,00	97,56
Acurácia Total	83,25	70,84	83,91	86,83	86,92	90,12
Atributos	20	8	20	28	28	48

Estatísticas, Gabor, LBP, Backes, Casanova e Bruno (2013), Ribeiro et al. (2015) e o método proposto foram implementados e classificados com KNN. O trabalho de Daood, Ribeiro e Bush (2016c) utilizou a mesma base Pollen, porém com outro classificador.

Tabela 38 – Resultados de classificação do Método Proposto e Métodos Seleccionados que utilizaram a mesma base Pollen.

Métodos	Atributos	Acurácia (%)
Wavelet	36	19,90
Estatísticas	9	22,20
Gabor	36	23,10
LBP	27	24,10
(BACKES; CASANOVA; BRUNO, 2013)	108	76,70
(RIBEIRO et al., 2015)	36	79,90
Método Proposto	48	90,12
(DAOOD; RIBEIRO; BUSH, 2016c)	**	94,58

A Tabela 39 ilustra uma comparação do método proposto com outros métodos que trabalharam com bases de pólen e classificadores diversificadas, selecionados da literatura.

Os resultados apresentados para os métodos de Chica (2012), Chudyk et al. (2015), Marcos et al. (2015), Redondo et al. (2015), Daood, Ribeiro e Bush (2016b) e Daood, Ribeiro e Bush (2016a) foram obtidos diretamente dos respectivos artigos.

Tabela 39 – Resultados de classificação do Método Proposto e Métodos Seleccionados com bases de pólen variadas.

Métodos	Atributos	Acurácia (%)
(CHUDYK et al., 2015)	252	87,00
Método Proposto	48	90,12
(CHICA, 2012)	28	92,30
(DAOOD; RIBEIRO; BUSH, 2016b)	**	93,99
(MARCOS et al., 2015)	>470	94,83
(DAOOD; RIBEIRO; BUSH, 2016a)	**	98,66
(REDONDO et al., 2015)	6.320	99,40

O resultado obtido pelo método proposto utilizando a Representação Semântica para caracterização dos grãos de pólen obteve 90,12% de acurácia com apenas 48 descritores, mesmo tendo algumas classes da base com poucas amostras. Quando comparados os métodos clássicos (Estatísticas, Wavelet, LBP e Gabor) com o método proposto, observa-se que obtiveram acurácias inferiores a 25%. Os métodos de Backes, Casanova e Bruno (2013) e Ribeiro et al. (2015) foram inferiores a 80% de acurácia.

Nos resultados obtidos por Daood, Ribeiro e Bush (2016a), Daood, Ribeiro e Bush (2016b), Daood, Ribeiro e Bush (2016c) foram utilizados para compor o vetor de características: atributos de histograma, estatísticas de níveis de cinza, atributos geométricos, dimensão fractal, matriz de co-ocorrência de níveis de cinza, momentos invariantes, atributos de Gabor, HOG e LBP e não explicitaram nos trabalhos a quantidade de descritores gerada pelos métodos. Para a classificação do vetor de características gerado, utilizou-se Máquina de Vetores de Suporte (SVM).

Embora os resultados de outros métodos sejam superiores ao método proposto, o conjunto de imagens utilizado é composto por poucas classes e um número elevado de imagens por classe. Além disso, a quantidade de atributos utilizadas pelos métodos é elevada, como é o caso de Redondo et al. (2015), Daood, Ribeiro e Bush (2016a), Daood, Ribeiro e Bush (2016b) e Daood, Ribeiro e Bush (2016c).

Os resultados indicam que o método proposto apresenta um alto desempenho com um número reduzido de descritores para a base Pollen (Apêndice B). Quando comparados com outros métodos da literatura, pode-se observar que a grande maioria utiliza bases com um número reduzido de classes obtidas com alta precisão. Além disso, o método proposto apresenta uma nova vertente para a caracterização de grãos de pólen, sendo um método que não necessita de alto poder computacional para sua execução.

Um trabalho que está em desenvolvimento é uma caracterização alternativa ao uso do LBP e HOG para a definição do espaço semântico. Propõem-se uma caracterização

utilizando as Redes Complexas (Seção 2.2) e Projeções Topológicas (Seção 3.3). Com isso, a Figura 20 ilustra o procedimento proposto.

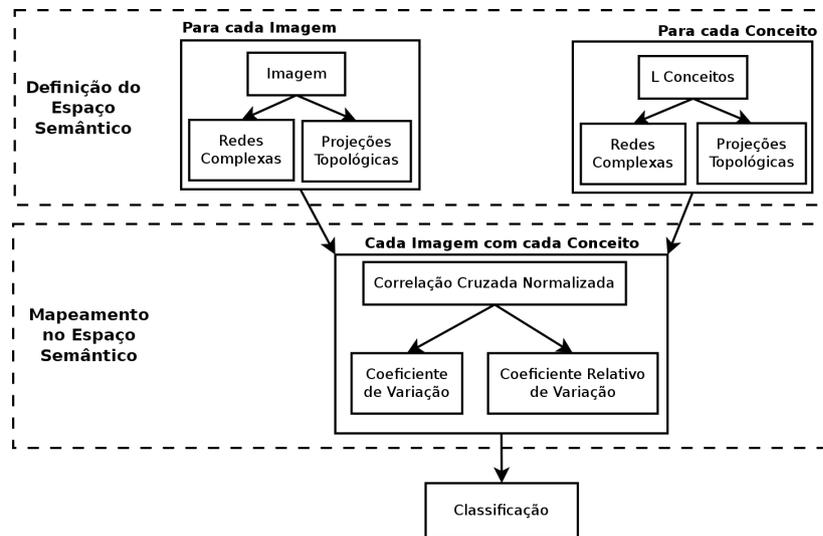


Figura 20 – Diagrama Geral do procedimento alternativo.

4.4 Considerações Finais

Utilizando-se de imagens de grãos de pólen da base Pollen propôs-se a caracterização semântica específica para classes da base de imagens, além da utilização da semântica global das imagens.

Propôs-se a representação no nível semântico através da caracterização por Padrão Binário Local (LBP) e Histograma de Gradientes Orientados (HOG) e, para o mapeamento do espaço semântico foram utilizados coeficientes obtidos da Correlação Cruzada Normalizada (CCN).

Após definidos os conceitos, para representar a base Pollen, realizou-se experimentos e obteve-se uma alta acurácia com uma quantidade reduzida de descritores, conforme comparações explicitadas na Tabela 38.

Analisou-se, também, o que estaria ocasionando erros na caracterização e por consequência, erros de classificação. Através da matriz de confusão conseguiu-se verificar que algumas classes estavam sendo confundidas. Contrapondo com alguns trabalhos da literatura que retiram algumas classes problemáticas para reduzir o problema de confusão, neste trabalho optou-se por não retirar nenhuma classe.

Quando comparados o método proposto com os métodos de Daood, Ribeiro e Bush (2016b), Daood, Ribeiro e Bush (2016c), Daood, Ribeiro e Bush (2016a) que utilizam a mesma base de imagens de pólen, observa-se que apesar desses métodos obterem resultados superiores, a quantidade de descritores é consideravelmente maior.

Em desenvolvimento está a modificação do espaço semântico substituindo o Padrão Binário Local (LBP) e o Histograma de Gradientes Orientados (HOG) pela caracterização obtida pelas redes complexas (Seção 2.2) e Projeções Topológicas (Seção 3.3).

Conclusão

Diversos métodos na literatura propõem a caracterização de texturas por diferentes meios, como explicitado por Materka e Strzelecki (1998). Neste trabalho foram analisadas duas frentes para conseguir a caracterização de texturas: a modelagem dos dados e a topologia dos dados, sendo que podem ser analisados separadamente ou como complemento um do outro.

Foram analisados, modificados e propostos métodos que trabalharam com redes complexas, projeções topológicas e padrões semânticos. As Redes Complexas e Padrões Semânticos foram utilizados para analisar a modelagem dos dados e a Projeção no Espaço \mathcal{K} aplicado a topologia dos dados para estimar e representar a distribuição dos dados em espaços de baixa dimensionalidade que são subvariedades do espaço dos "pixels".

O método proposto para seleção automática de limiares para a geração de redes complexas apresentou vetores de características de tamanho menores do que outros métodos do estado da arte, ao mesmo tempo que manteve a paridade com os resultados desses métodos em termos de acurácia, conforme mostrado pelos experimentos realizados (Seção 2.2). Este trabalho apresentou como as medidas descritivas aplicadas às redes complexas podem ser uma ferramenta poderosa para descrever adequadamente texturas, analisando as estatísticas dos histogramas dos mapas de graus, conseguiu-se informações não apenas sobre a natureza geral dos dados, mas também sobre os limiares mais significativos, adaptando-se às particularidades de cada conjunto específico de dados gerando conjuntos próprios de limiares.

As Projeções Topológicas no Espaço \mathcal{K} do trabalho de Perea e Carlsson (2014) obtiveram resultados significativos, porém observou-se que diversas partes do método poderiam ser modificados e, com isso, obteria-se melhores descritores para as texturas. Foram analisados os cálculos utilizados para tornar o descritor invariante à rotação e, através de experimentos realizados, percebeu-se que mesmo para uma base rotacionada, a invariância à rotação aplicada não estava ajudando na separação das classes. Com isso, analisou-se a influência do denominador M para composição do descritor invariante à rotação e percebeu-se que o cálculo (Eq. 42) imposto pelo descritor está reduzindo

os valores utilizados para descrever as texturas, tornando-os todos muito próximos e, com isso, indistinguíveis perante um classificador. Modificou-se o vetor de características, utilizando apenas os Coeficientes Estimados de K-Fourier (*KFchap*) para sua composição.

Além da modificação no vetor de características, foram realizados experimentos demonstrando que modificações na combinação de escalas (n) conseguiram reduzir a quantidade de descritores e ainda, melhoraram a separabilidade dos dados perante o classificador. Analisou-se também, através de experimentos realizados, que o valor do corte das frequências (ω) influencia diretamente nos Coeficientes Estimados de K-Fourier. Nessa análise, pôde-se perceber que não só a modificação da frequência, mas combinado com a modificação das escalas, conseguiu-se obter resultados com uma quantidade reduzida de descritores, melhorando a acurácia.

Ainda deve ser investigado com mais profundidade a influência do denominador M em outras bases de texturas. Além disso, outro fator que deverá ser investigado é a utilização de polinômios com termos de mais alto grau para a aproximação do $p(x, y)$ e p_{ij} (Equações 15 e 16, respectivamente) que podem resultar em uma melhora na caracterização de texturas através da Garrafa de Klein, uma vez que, aumentando o grau do polinômio espera-se uma aproximação mais precisa para a função intensidade.

O método proposto para os Padrões Semânticos (Seção 4.2) apresentou uma nova vertente para a caracterização de grãos de pólen, sendo um método que não necessita de alto poder computacional para sua execução. Este método foi aplicado à base Pollen, tendo sido criados conceitos semânticos específicos para a definição do espaço semântico desta base. A representação no nível semântico foi realizada através da caracterização utilizando-se LBP e HOG. Após essa representação, realizou-se o mapeamento do espaço semântico utilizando-se os Coeficientes de Variação (CoV) e os Coeficientes Relativos de Variação (CRoV) obtidos da Correlação Cruzada Normalizada (CCN). Os resultados indicaram que o método proposto apresenta um alto desempenho com um número reduzido de descritores para a base Pollen (Apêndice B). Esses resultados foram comparados com métodos clássicos (estatísticas, wavelet, LBP e Gabor), com os métodos de Chica (2012), Backes, Casanova e Bruno (2013), Ribeiro et al. (2015), Redondo et al. (2015), Chudyk et al. (2015), Daood, Ribeiro e Bush (2016a), Daood, Ribeiro e Bush (2016b) e Daood, Ribeiro e Bush (2016c). Embora os resultados de alguns métodos sejam superiores ao método proposto, o conjunto de imagens utilizado por estes foi composto por poucas classes e um número elevado de imagens por classe. Além disso, a quantidade de atributos utilizadas pela maioria dos métodos é elevada, como é o caso de Redondo et al. (2015), Daood, Ribeiro e Bush (2016a), Daood, Ribeiro e Bush (2016b) e Daood, Ribeiro e Bush (2016c). Conforme exposto, o método proposto utilizando Padrões Semânticos obteve, com vetor de característica de baixa dimensionalidade, altas acurácias, sendo um método promissor para ser aplicado em outras bases de imagens e outros problemas de caracterização de imagens.

Além da caracterização de texturas em imagens estáticas, neste trabalho preocupou-se com a dimensão dos vetores de características, pois estes influenciam diretamente no tamanho para armazenamento e também no tempo de análise e classificação desses dados. Se dois descritores de textura oferecem acurácias semelhantes, um método que apresenta um vetor de características de tamanho menor será preferível, mesmo que alguns possam ser mais demorados para o cálculo.

Com os resultados obtidos pelas duas frentes, pôde-se demonstrar que os métodos propostos conseguiram atingir as expectativas iniciais e puderam contribuir para a área de caracterização de texturas em imagens estáticas. Comparados aos métodos da literatura, conseguem se sobressair pela quantidade reduzida de descritores utilizados e pela acurácia obtida.

5.1 Contribuições em Produção Bibliográfica

Até o momento foram publicados os seguintes trabalhos resultantes dos estudos realizados:

- Ribeiro, T. P., Escarpinati, M. C. and Barcelos, C. A. Z. (2014) Uso da Garrafa de Klein para a Caracterização de Texturas de Deficiências Nutricionais do Café. *In XIV Semana da Matemática (SEMAT) e IV Semana da Estatística (SEMEST)* (pp 133-136), Uberlândia-MG. (RIBEIRO; ESCARPINATI; BARCELOS, 2014).
- Ribeiro T. P., Couto L. N., Backes A. R., Barcelos C. A. Z. (2015) Texture Characterization via Automatic Threshold Selection on Image-Generated Complex Network. In: PARDO, A.; KITTLER, J. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2015)**. Cham: Springer International Publishing, 2015. (Lecture Notes in Computer Science, v. 5197), p. 468-476. ISBN 978-3-319-25751-8. ISSN 0302-9743. (RIBEIRO et al., 2015).
- Couto, L. N., Ribeiro, T. P., Backes, A. R. and Barcelos, C. A. Z. (2015). Texture characterization via improved deterministic walks on image-generated complex network. In: **2015 22th IEEE International Conference on Image Processing (ICIP 2015)**. Quebec, Canada: [s.n.], 2015. p. 4416-4420. (COUTO et al., 2015).
- Ribeiro, T. P., Oliveira, A. L. N. and Barcelos C. A. Z. (2018). Texture Characterization via Projections onto a Klein Bottle Topology. In **25th IEEE International Conference on Image Processing (ICIP 2018)** Athens, Greece, 2018. p. 2097-2101. ISSN 2381-8549. (RIBEIRO; OLIVEIRA; BARCELOS, 2018).

- Ribeiro, T. P., Oliveira, A. L. N. and Barcelos C. A. Z. (2019). Frequency Analysis of Topological Projections onto Klein Bottle for Texture Characterization. In: Vera-Rodriguez R., Fierrez J., Morales A. (eds) **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2018.** Lecture Notes in Computer Science, Cham: Springer International Publishing, 2019. p. 342-350. ISBN 978-3-030-13469-3. (RIBEIRO; OLIVEIRA; BARCELOS, 2019).

5.2 Trabalhos Futuros

Baseado nos experimentos realizados com redes complexas, deve ser explorada a relação entre as medidas estatísticas do histograma dos graus, utilizados para redução de dimensionalidade, e as informações sobre a natureza e a complexidade dos dados. Analisar a junção de métodos estatísticos descritivos para a seleção automática dos limiares. Também analisar os limiares gerados para uma base e aplicá-los em outras bases.

Para as projeções topológicas, investigar com mais profundidade a influência do denominador M em outras bases de texturas. Além disso, outro fator a ser investigado é a utilização de polinômios com termos de mais alto grau para a aproximação do $p(x, y)$ e p_{ij} (Equações 15 e 16, respectivamente) que podem resultar em uma melhora na caracterização de texturas através da Garrafa de Klein. Além disso, investigar os tamanhos e frequências que consigam obter as melhores acurácias.

Na caracterização baseada em semântica, aprofundar nas análises dos algoritmos que podem auxiliar/melhorar a definição do Espaço Semântico tendo a Correlação Cruzada Normalizada (CCN) como mapeador deste Espaço. Explorar o uso de conceitos mantendo o padrão de cores da imagem original. Além disso, pesquisar outros métodos para a determinação e obtenção dos conceitos.

Para todos os métodos propostos, testar em outras bases de texturas e a utilização de outros classificadores, como por exemplo, os baseados em Redes Neurais Artificiais.

Referências

BACKES, A. R.; CASANOVA, D.; BRUNO, O. M. Plant leaf identification based on volumetric fractal dimension. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 23, n. 06, p. 1145–1160, 2009. <https://doi.org/10.1142/S0218001409007508>.

_____. Texture analysis and classification: A complex network-based approach. **Information Sciences**, v. 219, n. 0, p. 168–180, 2013. ISSN 0020-0255. <https://doi.org/10.1016/j.ins.2012.07.003>.

BANICA, D.; SMINCHISESCU, C. Second-order constrained parametric proposals and sequential search-based structured prediction for semantic segmentation in rgb-d images. In: **2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Boston, MA: [s.n.], 2015. p. 3517–3526. ISSN 1063-6919. <https://doi.org/10.1109/CVPR.2015.7298974>.

BOCCALETTI, S. et al. Complex networks: Structure and dynamics. **Physics Reports**, v. 424, n. 4-5, p. 175–308, 2006. ISSN 03701573. <https://doi.org/10.1016/j.physrep.2005.10.009>.

BRODATZ, P. **Textures: a photographic album for artists and designers**. New York: Dover Publications, 1966. (Dover pictorial archives).

BUSH, M. B.; WENG, C. Introducing a new (freeware) tool for palynology. **Journal of Biogeography**, Wiley Online Library, v. 34, n. 3, p. 377–380, 2007. <https://doi.org/10.1111/j.1365-2699.2006.01645.x>.

CAO, L.; WANG, F. Robust latent semantic exploration for image retrieval in social media. **Neurocomputing**, Elsevier, p. 1–5, 2015. ISSN 09252312. <https://doi.org/10.1016/j.neucom.2015.02.082>.

CARLSSON, G. et al. On the local behavior of spaces of natural images. **International Journal of Computer Vision**, v. 76, n. 1, p. 1–12, 2008. ISSN 09205691. <https://doi.org/10.1007/s11263-007-0056-x>.

CARRIÓN, P. et al. Classification of honeybee pollen using a multiscale texture filtering scheme. **Machine Vision and Applications**, v. 15, n. 4, p. 186–193, Oct 2004. ISSN 1432-1769. <https://doi.org/10.1007/s00138-004-0150-9>.

- CHANDRA, P.; WEISSTEIN, E. W. **Fibonacci Number**. From MathWorld—A Wolfram Web Resource, 2018. Acessado em 09/04/2018. Disponível em: <<http://mathworld.wolfram.com/FibonacciNumber.html>>.
- CHEN, X. et al. Wavelet Filter for Signals De-noising in ITER PF Converter I&C System. **Journal of Fusion Energy**, Springer US, p. 2–6, 2015. ISSN 0164-0313. <https://doi.org/10.1007/s10894-015-9946-z>.
- CHICA, M. Authentication of bee pollen grains in bright-field microscopy by combining one-class classification techniques and image processing. **Microscopy Research and Technique**, v. 75, n. 11, p. 1475–1485, 2012. ISSN 1059910X. <https://doi.org/10.1002/jemt.22091>.
- CHUDYK, C. et al. Development of an automatic pollen classification system using shape, texture and aperture features. In: **CEUR Workshop Proceedings**. Germany: [s.n.], 2015. v. 1458, n. 10, p. 65–74. ISSN 16130073.
- COSTA, L. d. F. et al. Characterization of complex networks: A survey of measurements. **Advances in Physics**, Taylor & Francis, v. 56, n. 1, p. 167–242, 2007. ISSN 0001-8732. <https://doi.org/10.1080/00018730601170527>.
- COURANT, R.; ROBBINS, H. **O que é matemática: uma abordagem elementar de métodos e conceitos**. [S.l.]: Ciência Moderna, 2000. ISBN 9788573930214.
- COUTO, L. N. et al. Texture characterization via improved deterministic walks on image-generated complex network. In: **2015 22th IEEE International Conference on Image Processing (ICIP)**. Quebec, Canada: [s.n.], 2015. p. 4416–4420. <https://doi.org/10.1109/ICIP.2015.7351641>.
- CROSIER, M.; GRIFFIN, L. D. Using basic image features for texture classification. **International Journal of Computer Vision**, Springer, v. 88, n. 3, p. 447–460, 2010. <https://doi.org/10.1007/s11263-009-0315-0>.
- CROSS, G. R.; JAIN, a. K. Markov random field texture models. **IEEE transactions on pattern analysis and machine intelligence**, v. 5, n. 1, p. 25–39, 1983. ISSN 0162-8828. <https://doi.org/10.1109/TPAMI.1983.4767341>.
- CUADROS, O. et al. Segmentation of Large Images with Complex Networks. In: **2012 25th SIBGRAPI Conference on Graphics, Patterns and Images**. Ouro Preto: IEEE, 2012. p. 24–31. ISBN 978-0-7695-4829-6. ISSN 15301834. <https://doi.org/10.1109/SIBGRAPI.2012.13>.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: **2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)**. San Diego, CA: [s.n.], 2005. v. 1, p. 886–893 vol. 1. ISSN 1063-6919. <https://doi.org/10.1109/CVPR.2005.177>.
- DANA, K. J. et al. Reflectance and texture of real-world surfaces. **ACM Trans. Graph.**, ACM, New York, NY, USA, v. 18, n. 1, p. 1–34, 1999. ISSN 0730-0301. <https://doi.org/10.1145/300776.300778>.

- DAOOD, A.; RIBEIRO, E.; BUSH, M. Classifying pollen using robust sequence alignment of sparse z-stack volumes. In: BEBIS, G. et al. (Ed.). **Advances in Visual Computing**. Cham: Springer International Publishing, 2016. p. 331–340. ISBN 978-3-319-50835-1. https://doi.org/10.1007/978-3-319-50835-1_31.
- _____. Pollen grain recognition using deep learning. In: BEBIS, G. et al. (Ed.). **Advances in Visual Computing**. Cham: Springer International Publishing, 2016. p. 321–330. ISBN 978-3-319-50835-1. https://doi.org/10.1007/978-3-319-50835-1_30.
- _____. Pollen recognition using a multi-layer hierarchical classifier. In: **2016 23rd International Conference on Pattern Recognition (ICPR)**. Cancun: [s.n.], 2016. p. 3091–3096. <https://doi.org/10.1109/ICPR.2016.7900109>.
- FRIEDMAN, J. H. Regularized discriminant analysis. **Journal of the American Statistical Association**, Taylor & Francis, v. 84, n. 405, p. 165–175, 1989. <https://doi.org/10.1080/01621459.1989.10478752>.
- FUKUNAGA, K. **Introduction to Statistical Pattern Recognition (2Nd Ed.)**. San Diego, CA, USA: Academic Press Professional, Inc., 1990. ISBN 0-12-269851-7.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features for Image Classification. **Systems, Man and Cybernetics, IEEE Transactions on**, SMC-3, n. 6, p. 610–621, 1973. ISSN 0018-9472. <https://doi.org/10.1109/TSMC.1973.4309314>.
- HARALICK, R. M.; STERNBERG, S. R.; ZHUANG, X. Image Analysis Using Mathematical Morphology. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, PAMI-9, n. 4, p. 532–550, 1987. ISSN 0162-8828. <https://doi.org/10.1109/TPAMI.1987.4767941>.
- HAYMAN, E. et al. On the significance of real-world conditions for material classification. In: **Computer Vision - ECCV 2004**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, (Lecture Notes in Computer Science, v. 3024). p. 253–266. ISBN 978-3-540-21981-1. https://doi.org/10.1007/978-3-540-24673-2_21.
- HE, X. et al. Learning a semantic space from user’s relevance feedback for image retrieval. **IEEE Transactions on Circuits and Systems for Video Technology**, v. 13, n. 1, p. 39–48, 2003. ISSN 1051-8215. <https://doi.org/10.1109/TCSVT.2002.808087>.
- HU, J. et al. Topological interactive analysis of power system and its communication module : A complex network approach. **Physica A**, Elsevier B.V., v. 416, p. 99–111, 2014. ISSN 0378-4371. <https://doi.org/10.1016/j.physa.2014.08.015>.
- Jl, L. et al. Median local ternary patterns optimized with rotation-invariant uniform-three mapping for noisy texture classification. **Pattern Recognition**, v. 79, p. 387 – 401, 2018. ISSN 0031-3203. <https://doi.org/10.1016/j.patcog.2018.02.009>.
- Jun Zhang; Jimin Liang; Heng Zhao. Local Energy Pattern for Texture Classification Using Self-Adaptive Quantization Thresholds. **IEEE Transactions on Image Processing**, v. 22, n. 1, p. 31–42, 2013. ISSN 1057-7149. <https://doi.org/10.1109/TIP.2012.2214045>.

- KONSTANTINIDIS, D. et al. Building detection using enhanced hog-lbp features and region refinement processes. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 3, p. 888–905, 2017. ISSN 1939-1404. <https://doi.org/10.1109/JSTARS.2016.2602439>.
- KWITT, R. et al. Endoscopic image analysis in semantic space. **Medical Image Analysis**, Elsevier B.V., v. 16, n. 7, p. 1415–1422, 2012. ISSN 13618415. <https://doi.org/10.1016/j.media.2012.04.010>.
- LAZEBNIK, S.; SCHMID, C.; PONCE, J. A sparse texture representation using local affine regions. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 27, n. 8, p. 1265–1278, 2005. ISSN 0162-8828. <https://doi.org/10.1109/TPAMI.2005.151>.
- LEE, A. B.; PEDERSEN, K. S.; MUMFORD, D. The nonlinear statistics of high-contrast patches in natural images. **International Journal of Computer Vision**, Springer, v. 54, n. 1-3, p. 83–103, 2003. <https://doi.org/10.1023/A:1023705401078>.
- LEWIS, J. P. Fast normalized cross-correlation. **Vision interface**, v. 10, n. 1, p. 120–123, 1995.
- LI, H. et al. Global energy investment structure from the energy stock market perspective based on a Heterogeneous Complex Network Model. **Applied Energy**, v. 194, p. 648–657, 2017. ISSN 03062619. <https://doi.org/10.1016/j.apenergy.2016.05.062>.
- LI, T. et al. Face detection based on depth information using hog-lbp. In: **2016 Sixth International Conference on Instrumentation Measurement, Computer, Communication and Control (IMCCC)**. Harbin: [s.n.], 2016. p. 779–784. <https://doi.org/10.1109/IMCCC.2016.92>.
- LU, P. et al. Finding more relevance: Propagating similarity on Markov random field for object retrieval. **Signal Processing: Image Communication**, Elsevier, v. 32, p. 54–68, 2015. ISSN 09235965. <https://doi.org/10.1016/j.image.2015.01.007>.
- LUM, P. Y. et al. Extracting insights from the shape of complex data using topology. **Scientific reports**, v. 3, p. 1236, 2013. ISSN 2045-2322. <https://doi.org/10.1038/srep01236>.
- MARCOS, J. V. et al. Automated pollen identification using microscopic imaging and texture analysis. **Micron**, v. 68, p. 36–46, 2015. ISSN 09684328. <https://doi.org/10.1016/j.micron.2014.09.002>.
- MARI, S. I.; LEE, Y. H.; MEMON, M. S. Complex network theory-based approach for designing resilient supply chain networks. **International Journal Logistics Systems and Management**, v. 21, p. 365–384, 2015. <http://dx.doi.org/10.1504/IJLSM.2015.069733>.
- MATERKA, A.; STRZELECKI, M. Texture Analysis Methods - A Review. **Methods**, v. 11, p. 1–33, 1998. ISSN 0893-9675.
- NAIR, B. M. et al. Optical flow based Kalman tracker for body joint prediction and tracking using HOG-LBP matching. **Video Surveillance and Transportation Imaging Applications 2014**, v. 9026, 2014. <https://doi.org/10.1117/12.2040392>.

NEWMAN, M. E. J. The Structure and Function of Complex Networks. **SIAM Review**, v. 45, n. 2, p. 167–256, 2003. ISSN 0036-1445. <https://doi.org/10.1137/S003614450342480>.

OJALA, T. et al. Outex-new framework for empirical evaluation of texture analysis algorithms. In: IEEE. **Pattern Recognition, 2002. Proceedings. 16th International Conference on**. Quebec, Canada, 2002. v. 1, p. 701–706. <https://doi.org/10.1109/ICPR.2002.1044854>.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51–59, 1996. ISSN 00313203. [https://doi.org/10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4).

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 24, n. 7, p. 971–987, 2002. ISSN 0162-8828. <https://doi.org/10.1109/TPAMI.2002.1017623>.

OTSU, N. A Threshold Selection Method from Gray-Level Histograms. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 9, n. 1, p. 62–66, 1979. ISSN 0018-9472. <https://doi.org/10.1109/TSMC.1979.4310076>.

PANDEY, S.; KHANNA, P.; YOKOTA, H. An effective use of adaptive combination of visual features to retrieve image semantics from a hierarchical image database. **Journal of Visual Communication and Image Representation**, Elsevier Inc., v. 30, p. 136–152, 2015. ISSN 10473203. <https://doi.org/10.1016/j.jvcir.2015.03.010>.

PEREA, J. A.; CARLSSON, G. A Klein-Bottle-Based Dictionary for Texture Representation. **International Journal of Computer Vision**, Springer US, v. 107, n. 1, p. 75–97, 2014. ISSN 0920-5691. <https://doi.org/10.1007/s11263-013-0676-2>.

QIAO, Y.; CAPPELLE, C.; RUICHEK, Y. Place recognition based visual localization using lbp feature and svm. In: _____. **Advances in Artificial Intelligence and Its Applications: 14th Mexican International Conference on Artificial Intelligence, MICAI 2015, Cuernavaca, Morelos, Mexico, October 25-31, 2015, Proceedings, Part II**. Cham: Springer International Publishing, 2015. p. 393–404. ISBN 978-3-319-27101-9. https://doi.org/10.1007/978-3-319-27101-9_30.

QUAN, Y. et al. Lacunarity analysis on image patterns for texture classification. In: **2014 IEEE Conference on Computer Vision and Pattern Recognition**. Columbus, OH: [s.n.], 2014. p. 160–167. ISSN 1063-6919. <https://doi.org/10.1109/CVPR.2014.28>.

RAGHUWANSHI, G.; TYAGI, V. A Survey on Texture Image Retrieval. In: SATAPATHY, S. C. et al. (Ed.). **Advances in Intelligent Systems and Computing**. New Delhi: Springer India, 2016, (Advances in Intelligent Systems and Computing, v. 380). p. 427–435. ISBN 978-81-322-2522-5. https://doi.org/10.1007/978-81-322-2526-3_44.

REDONDO, R. et al. Pollen segmentation and feature evaluation for automatic classification in bright-field microscopy. **Computers and Electronics in Agriculture**, Elsevier B.V., v. 110, p. 56–69, 2015. ISSN 01681699. <https://doi.org/10.1016/j.compag.2014.09.020>.

REED, M.; SIMON, B. **Methods of Modern Mathematical Physics: Functional analysis**. EUA: Academic Press, 1980. (Methods of Modern Mathematical Physics, v. 1). ISBN 9780125850506.

RIBEIRO, T. P. et al. Texture Characterization via Automatic Threshold Selection on Image-Generated Complex Network. In: PARDO, A.; KITTLER, J. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP)**. Cham: Springer International Publishing, 2015. (Lecture Notes in Computer Science, v. 5197), p. 468–476. ISBN 978-3-319-25751-8. ISSN 0302-9743. https://doi.org/10.1007/978-3-319-25751-8_56.

RIBEIRO, T. P.; ESCARPINATI, M. C.; BARCELOS, C. A. Uso da Garrafa de Klein para a Caracterização de Texturas de Deficiências Nutricionais do Café. In: **XIV Semana da Matemática (SEMAT) e IV Semana da Estatística (SEMEST)**. Uberlândia-MG: [s.n.], 2014. p. 133–136.

RIBEIRO, T. P.; OLIVEIRA, A. L. N. de; BARCELOS, C. A. Z. Texture characterization via projections onto a klein bottle topology. In: IEEE. **2018 25th IEEE International Conference on Image Processing (ICIP)**. Athens, Greece, 2018. p. 2097–2101. ISSN 2381-8549. <https://doi.org/10.1109/ICIP.2018.8451120>.

_____. Frequency analysis of topological projections onto klein bottle for texture characterization. In: VERA-RODRIGUEZ, R.; FIERREZ, J.; MORALES, A. (Ed.). **Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications**. Cham: Springer International Publishing, 2019. p. 342–350. ISBN 978-3-030-13469-3. https://doi.org/10.1007/978-3-030-13469-3_40.

RIJSBERGEN, C. J. V. **Information Retrieval**. 2nd. ed. Newton, MA, USA: Butterworth-Heinemann, 1979. ISBN 0408709294.

RUBINOV, M.; SPORNS, O. Complex network measures of brain connectivity: Uses and interpretations. **NeuroImage**, Elsevier Inc., v. 52, n. 3, p. 1059–1069, 2010. ISSN 10538119. <https://doi.org/10.1016/j.neuroimage.2009.10.003>.

SACHS, L. **Applied Statistics**. New York, NY: Springer New York, 1984. 77-78 p. (Springer Series in Statistics, 1). <https://doi.org/10.1007/978-1-4612-5246-7>. ISSN 13514180. ISBN 978-1-4612-9755-0.

SHAKIBIAN, H.; CHARKARI, N. M. Mutual information model for link prediction in heterogeneous complex networks. **Scientific Reports**, Nature Publishing Group, v. 7, p. 44981, 2017. ISSN 2045-2322. <https://doi.org/10.1038/srep44981>.

SILBERMAN, N. et al. Indoor segmentation and support inference from rgb-d images. In: SPRINGER. **European Conference on Computer Vision**. [S.l.], 2012. p. 746–760. https://doi.org/10.1007/978-3-642-33715-4_54.

SILVERMAN, B. W. **Density Estimation for Statistics and Data Analysis**. London: Chapman & Hall, 1986. (Chapman & Hall/CRC Monographs on Statistics & Applied Probability, v. 26).

SINGH, R. et al. Average Gabor - Wavelet Filter Feature Extraction Technique for Facial Expression Recognition. **International Journal of Computer Science and Network Security**, v. 15, n. 5, p. 102–104, 2015.

- SJÖBERG, M. et al. Retrieval of Multimedia Objects by Combining Semantic Information from Visual and Textual Descriptors. In: **Artificial Neural Networks–ICANN 2006**. Berlin Heidelberg: Springer. p. 75–83. ISBN 3-540-38871-0. https://doi.org/10.1007/11840930_8.
- SONG, L. et al. Using hog-lbp features and mmp learning to recognize imaging signs of lung lesions. In: **2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)**. Rome: [s.n.], 2012. p. 1–4. ISSN 1063-7125. <https://doi.org/10.1109/CBMS.2012.6266313>.
- TIPLER, P. A.; MOSCA, G. **Física para Cientistas e Engenheiros**. 6. ed. Sao Paulo: LTC Editora, 2009. v. 1. 662 p. ISBN 9788521617105.
- VARMA, M.; ZISSERMAN, A. Unifying statistical texture classification frameworks. **Image and Vision Computing**, Elsevier, v. 22, n. 14, p. 1175–1183, 2004. <https://doi.org/10.1016/j.imavis.2004.03.012>.
- VISTEX. **Vision Texture Database**. 2009. Disponível em: <<http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>>.
- WEBER, A. G. The USC-SIPI Image Database. **USC-SIPI Report #315**, Los Angeles, CA 90089-2564 USA, 3740 McClintock Ave, 1997. Disponível em: <<http://sipi.usc.edu/database/database.php?volume=textures>>.
- WEINBERGER, K. Q.; SAUL, L. K. Distance Metric Learning for Large Margin Nearest Neighbor Classification. **Journal of Machine Learning Research**, v. 10, p. 207–244, 2009.
- WELDON, T. P.; HIGGINS, W. E.; DUNN, D. F. Efficient Gabor filter design for texture segmentation. **Pattern Recognition**, v. 29, n. 12, p. 2005–2015, 1996. ISSN 00313203. [https://doi.org/10.1016/S0031-3203\(96\)00047-7](https://doi.org/10.1016/S0031-3203(96)00047-7).
- WIT, T. D. de; FLORIANI, E. Estimating probability densities from short samples: A parametric maximum likelihood approach. **Physical Review E**, v. 58, n. 4, p. 5115–5122, 1998. ISSN 1063-651X. <https://doi.org/10.1103/PhysRevE.58.5115>.
- XIE, X. A review of recent advances in surface defect detection using texture analysis techniques. **ELCVIA Electronic Letters on Computer Vision and Image Analysis**, v. 7, n. 3, p. 1–22, 2008. ISSN 1577-5097. <https://doi.org/10.5565/rev/elcvia.268>. Disponível em: <<http://elcvia.cvc.uab.es/article/download/268/176>>.
- YANG, H.-y. et al. Weibull statistical modeling for textured image retrieval using nonsubsampling contourlet transform. **Soft Computing**, 2018. ISSN 1433-7479. <https://doi.org/10.1007/s00500-018-3127-8>.
- ZHANG, C. et al. Joint image representation and classification in random semantic spaces. **Neurocomputing**, Elsevier, v. 156, p. 79–85, 2015. ISSN 09252312. <https://doi.org/10.1016/j.neucom.2014.12.083>.
- ZHANG, J. et al. Scale invariant texture representation based on frequency decomposition and gradient orientation. **Pattern Recognition Letters**, v. 51, n. 0, p. 57–62, 2015. ISSN 0167-8655. <https://doi.org/10.1016/j.patrec.2014.08.002>.

_____. Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study. In: **Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on**. New York: [s.n.], 2006. p. 13–13. ISSN 2160-7508. <https://doi.org/10.1109/CVPRW.2006.121>.

Fundamentação Teórica

A.1 Projeções Topológicas

Nesta seção serão apresentadas as fundamentações teóricas para complementação da formulação apresentada no Capítulo 3.

A.1.1 Obtenção do *Theta*

Demonstração: Deve-se de encontrar o mínimo da função $\Phi^2(c, d) = \left\| I_P - (cu + d\sqrt{3}u^2) \right\|^2$ sujeita à restrição $c^2 + d^2 = 1$.

Sabe-se que $\Phi^2(c, d) = \iint_{[-1,1]^2} \langle \nabla I_P - c\nabla u - d\sqrt{3}\nabla u^2, \nabla I_P - c\nabla u - d\sqrt{3}\nabla u^2 \rangle dx dy$.
Pela propriedade distributiva do produto interno,

$$\begin{aligned} & \langle \nabla I_P - c\nabla u - d\sqrt{3}\nabla u^2, \nabla I_P - c\nabla u - d\sqrt{3}\nabla u^2 \rangle = \\ & = \left[\langle \nabla I_P, \nabla I_P \rangle - c \langle \nabla I_P, \nabla u \rangle - d\sqrt{3} \langle \nabla I_P, \nabla u^2 \rangle \right. \\ & - c \langle \nabla u, \nabla I_P \rangle + c^2 \langle \nabla u, \nabla u \rangle + cd\sqrt{3} \langle \nabla u, \nabla u^2 \rangle \\ & \left. - d\sqrt{3} \langle \nabla u^2, \nabla I_P \rangle + cd\sqrt{3} \langle \nabla u^2, \nabla u \rangle + 3d^2 \langle \nabla u^2, \nabla u^2 \rangle \right]. \end{aligned} \quad (48)$$

Mas,

$$\iint_{[-1,1]^2} \langle \nabla u, \nabla u \rangle dx dy = 1, \quad (49)$$

$$\iint_{[-1,1]^2} \langle \nabla u^2, \nabla u \rangle dx dy = 0, \quad (50)$$

$$\iint_{[-1,1]^2} \langle \nabla u^2, \nabla u^2 \rangle dx dy = \frac{1}{3}. \quad (51)$$

Assim, $\Phi(c, d)^2 = 1 - 2c \langle \nabla I_P, \nabla u \rangle - 2d\sqrt{3} \langle \nabla I_P, \nabla u^2 \rangle + c^2 + d^2$.

Utilizando o Método dos Multiplicadores de Lagrange encontra-se o mínimo da função $F(x, y) = \Phi(c, d)^2$ sujeita à restrição $G(x, y) = c^2 + d^2 - 1 = 0$. As coordenadas do ponto onde a função assume o mínimo são c^* e d^* (Eq. 29).

Para discretizar as funções $u : [-1, 1] \times [-1, 1] \rightarrow \mathbb{R}$ e $u^2 : [-1, 1] \times [-1, 1] \rightarrow \mathbb{R}$ utiliza-se

$$u_{ij} = \int_{1-\frac{2i}{n}}^{1-\frac{2i-2}{n}} \int_{-1+\frac{2j-2}{n}}^{-1+\frac{2j}{n}} u(x, y) dx dy, \quad i, j = 1, \dots, n$$

onde n é o tamanho do *patch*.

Assim,

$$\nabla u(i, j) = \frac{4}{n^3} (a, b). \quad (52)$$

$$\begin{aligned} \nabla u^2(i, j) = & \left(-8ab \frac{i}{n^4} + 4ab \frac{1}{n^3} + 4ab \frac{1}{n^4} - 4a^2 \frac{1}{n^3} - 4a^2 \frac{1}{n^4} + 8a^2 j \frac{1}{n^4}, \right. \\ & \left. 8ab \frac{j}{n^4} - 4ab \frac{1}{n^3} - 4ab \frac{1}{n^4} + 4b^2 \frac{1}{n^3} + 4b^2 \frac{1}{n^4} - 8b^2 i \frac{1}{n^4} \right). \end{aligned} \quad (53)$$

$$\begin{aligned} \langle I_P, u \rangle_D &= \iint_{[-1, 1]^2} \langle \nabla \mathbf{I}_P, \nabla \mathbf{u} \rangle dx dy \\ &\approx \sum_{i=1}^n \sum_{j=1}^n (P_x(i, j) u_x(i, j) + P_y(i, j) u_y(i, j)). \end{aligned} \quad (54)$$

$$\begin{aligned} \langle I_P, u^2 \rangle_D &= \iint_{[-1, 1]^2} \langle \nabla \mathbf{I}_P, \nabla \mathbf{u}^2 \rangle dx dy \\ &\approx \sum_{i=1}^n \sum_{j=1}^n (P_x(i, j) u_x^2(i, j) + P_y(i, j) u_y^2(i, j)). \end{aligned} \quad (55)$$

$$\begin{aligned} \Phi(c, d) &= \left\| I_P - (c\mathbf{u} + d\sqrt{3}\mathbf{u}^2) \right\| \\ &= \sqrt{\iint_{[-1, 1]^2} \langle \nabla (\mathbf{I}_P - (c\mathbf{u} + d\sqrt{3}\mathbf{u}^2)), \nabla (\mathbf{I}_P - (c\mathbf{u} + d\sqrt{3}\mathbf{u}^2)) \rangle dx dy} \\ &\approx \sqrt{\sum_{i=1}^n \sum_{j=1}^n [\{\mathcal{P}x\}^2 + \{\mathcal{P}y\}^2]}. \end{aligned} \quad (56)$$

onde

$$\begin{aligned} \mathcal{P}x &= P_x(i, j) - cu_x(i, j) - d\sqrt{3}u_x^2(i, j). \\ \mathcal{P}y &= P_y(i, j) - cu_y(i, j) - d\sqrt{3}u_y^2(i, j). \end{aligned}$$

Se $\Phi(c, d)$ é constante, descarta-se P , pois neste caso não tem componentes lineares nem quadráticas. Caso contrário, faz-se $\theta_P \in \left[-\frac{\pi}{2}, \frac{3\pi}{2}\right)$ o minimizador de $\Phi(c, d)$. Utilizando-se (Eq. 29), (Eq. 49), (Eq. 50) e (Eq. 51), obtém-se $\Phi(c^*, d^*) = \sqrt{2 \left(1 - \sqrt{\langle I_P, \nabla u \rangle^2 + 3 \langle I_P, \nabla u^2 \rangle^2}\right)} = \sqrt{2 \left(1 - \sqrt{\varphi(I_P, \alpha_P)}\right)}$.

A.1.2 Obtenção dos Coeficientes de K-Fourier

Sejam $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ e $T = S^1 \times S^1$. Segue-se da parametrização de \mathcal{K} via o conjunto de polinômios

$$p(x, y) = c \frac{(ax + by)}{2} + d\sqrt{3} \frac{(ax + by)^2}{4} \quad (57)$$

com $a^2 + b^2 = c^2 + d^2 = 1$, que cada (z, w) em T determina um único elemento em \mathcal{K} , e que cada (z, w) e (z', w') fornece o mesmo *patch* se e somente se $(z', w') = (-z, -\bar{w})$. \bar{w} denota o complexo conjugado de w . Isto é, \mathcal{K} (e portanto K) pode ser modelado como o espaço obtido de T identificando (z, w) com $(-z, -\bar{w})$ para todo $(z, w) \in T$. Assim, em essência, pode-se interpretar K como

$$\{(z, w), (-z, -\bar{w}) : (z, w) \in T\}$$

e pensar elementos em $L^2(K)$ como o conjunto das funções quadrado integráveis $f : T \rightarrow \mathbb{C}$ satisfazendo a identidade

$$f(z, w) = f(-z, -\bar{w}) \quad (58)$$

para todo $(z, w) \in T$.

Nota-se que com esta caracterização permite-se interpretar $L^2(K)$ como um subespaço linear de $L^2(T)$. Com isto, leva-se a considerar a projeção ortogonal

$$\Pi : L^2(T) \rightarrow L^2(K)$$

de $L^2(T)$ sobre $L^2(K)$, o que fornece uma receita para produção de bases ortonormais para $L^2(K)$ a partir das de $L^2(T)$. Primeiramente, uma fórmula explícita para Π é derivada.

Proposição 1 *A função*

$$\begin{aligned} \Pi : L^2(T) &\longrightarrow L^2(K) \\ g &\longmapsto \frac{g(z, w) + g(-z, -\bar{w})}{2} \end{aligned} \quad (59)$$

é a projeção ortogonal de $L^2(T)$ sobre $L^2(K)$.

Nota-se que se $\{\phi_k | k \in \mathbb{N}\}$ é uma base (ou um conjunto gerador) para $L^2(T)$ então $\{\Pi(\phi_k) | k \in \mathbb{N}\}$ gera $L^2(K)$ e assim contém uma base. Mais ainda, aplicando Gram-Schmidt transforma-se esta base em uma base ortonormal. Em outras palavras, qualquer base ortonormal de $L^2(T)$ fornece uma base ortonormal para $L^2(K)$. Uma observação menos óbvia é a de que qualquer base ortonormal de $L^2(K)$ pode ser obtida de $L^2(T)$ desta forma. A discussão anterior é sintetizada em:

Teorema 1 Se $\{\Pi(\phi_k) | k \in \mathbb{N}\}$ é um conjunto gerador para $L^2(T)$ então

$$\mathcal{L} = \left\{ \frac{\phi_k(z, w) + \phi_k(-z, -\bar{w})}{2} \mid k \in \mathbb{N} \right\}$$

é um conjunto gerador para $L^2(K)$.

Se $\mathcal{B} \subset \mathcal{L}$ é uma base para $L^2(K)$, então a aplicação de Gram-Schmidt a \mathcal{B} fornece uma base ortonormal para $L^2(K)$. Mais ainda, qualquer base ortonormal para $L^2(K)$ pode ser obtida de uma base ortonormal para $L^2(T)$ desta forma.

Em particular, tem-se

Corolário 1 Seja $\{\phi_n\}_{n \in \mathbb{N}}$ uma base para $L^2(S^1)$. Então em coordenadas polares

$$\{\phi_n(\alpha)\phi_m(\theta) \mid n, m \in \mathbb{N}\}$$

é uma base para $L^2(T)$ e portanto

$$\left\{ \frac{\phi_n(\alpha)\phi_m(\theta) + \phi_n(\alpha + \pi)\phi_m(\pi - \theta)}{2} \mid n, m \in \mathbb{N} \right\}$$

é um conjunto gerador para $L^2(K)$.

Os próximos dois resultados fornecem bases do tipo Fourier para $L^2(K)$ e $L^2(T)$ aplicando o corolário 1 à $\{\phi_k(\alpha) = e^{ik\alpha}\}_{k \in \mathbb{Z}}$.

Corolário 2 Sejam $n \in \mathbb{Z}$ e $m \in \mathbb{N} \cup \{0\}$. Então o conjunto de funções

$$\frac{e^{in\alpha + im\theta} + (-1)^{n+m} e^{in\alpha - im\theta}}{2}, \quad m = 0 \quad \text{implica } n \text{ par}$$

é uma base ortonormal para $L^2(K)$. Refere-se a esta base como uma base trigonométrica para $L^2(K)$.

Quando se restringe a funções a valores reais $f : K \rightarrow \mathbb{R}$, pode-se começar com bases \mathcal{B} para $L^2(T, \mathbb{R})$ consistindo de funções da forma $\phi(n\alpha)\psi(m\theta)$, onde ϕ e ψ são senos ou cossenos, e considerar $\Pi(\mathcal{B})$, o que nos dá o seguinte resultado:

Corolário 3 Sejam $n, m \in \mathbb{N}$ e $\pi_{n,m} = \frac{(1 - (-1)^{n+m})\pi}{4}$. Então o conjunto de funções

$$\begin{aligned} &1, \quad \sqrt{2} \cos(m\theta - \pi_{0,m}), \quad \sqrt{2} \cos(2n\alpha), \quad \sqrt{2} \sin(2n\alpha), \\ &2 \cos(n\alpha) \cdot \cos(m\theta - \pi_{n,m}), \quad 2 \sin(n\alpha) \cdot \cos(m\theta - \pi_{n,m}) \end{aligned}$$

é uma base ortonormal para $L^2(K, \mathbb{R})$. Refere-se à esse conjunto como uma base trigonométrica para $L^2(K)$.

As simetrias fornecidas pela Equação 58 provêm $L^2(K)$ como um subespaço de $L^2(T)$. Isto simplifica o produto interno $\langle \cdot, \cdot \rangle_T$ quando restrito à $L^2(K)$. Se $f, g \in L^2(K)$ então pode-se mostrar que

$$\begin{aligned} \langle f, g \rangle_T &= \frac{1}{(2\pi)^2} \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \int_{\frac{\pi}{4}}^{\frac{9\pi}{4}} f(\alpha, \theta) \bar{g}(\alpha, \theta) d\alpha d\theta \\ &= \frac{1}{2\pi^2} \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} f(\alpha, \theta) \bar{g}(\alpha, \theta) d\alpha d\theta \end{aligned} \quad (60)$$

e assim pode-se munir $L^2(K)$ com seu próprio produto interno:

$$\langle f, g \rangle_K = \frac{1}{(2\pi)^2} \int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}} \int_{\frac{\pi}{4}}^{\frac{5\pi}{4}} f(\alpha, \theta) \bar{g}(\alpha, \theta) d\alpha d\theta. \quad (61)$$

Definição 1 *Sejam $f \in L^2(K, \mathbb{R}), \langle \cdot, \cdot \rangle_K$ conforme a Equação 61 e sejam*

$$\begin{aligned} a_m &= \langle f, \sqrt{2} \cos(m\theta - \pi_{0,m}) \rangle_K \\ b_n &= \langle f, \sqrt{2} \cos(2n\alpha) \rangle_K \\ c_n &= \langle f, \sqrt{2} \sin(2n\alpha) \rangle_K \\ d_{n,m} &= \langle f, 2 \cos(n\alpha) \cos(m\theta - \pi_{n,m}) \rangle_K \\ e_{n,m} &= \langle f, 2 \sin(n\alpha) \cos(m\theta - \pi_{n,m}) \rangle_K \end{aligned}$$

os coeficientes de f na base trigonométrica para $L^2(K, \mathbb{R})$.

$K\mathcal{F}(f)$ serão os Coeficientes de K-Fourier de f .

A.1.3 Detalhes de Implementação

Para a implementação correta das Equações 20, 21, 22 e 23, devido a inversão computacional da orientação dos eixos x e y , deve-se utilizar as Equações 62, 63, 64 e 64.

$$\nabla \mathbf{P}(i, j) = \frac{1}{2} \begin{bmatrix} p_{i,j+1} - p_{i,j-1} \\ p_{i-1,j} - p_{i+1,j} \end{bmatrix} \quad (62)$$

$$HP(i, j) = \begin{bmatrix} p_{i,j+1} - 2p_{i,j} + p_{i,j-1} & H_{xy}P(i, j) \\ H_{xy}P(i, j) & p_{i+1,j} - 2p_{i,j} + p_{i-1,j} \end{bmatrix} \quad (63)$$

$$H_{xy}P(i, j) = \frac{p_{i-1, j+1} - p_{i-1, j-1} + p_{i+1, j-1} - p_{i+1, j+1}}{4}.$$

$$\nabla P(r, t) = \nabla P(i, j) + HP(i, j) \begin{bmatrix} t - j \\ i - r \end{bmatrix} \quad (64)$$

A.2 Padrões Semânticos

Nesta seção serão apresentadas as fundamentações teóricas para complementação da formulação apresentada no Capítulo 4.

A.2.1 Correlação Cruzada Normalizada

A Correlação Cruzada Normalizada (CCN) (LEWIS, 1995) é definida por:

$$\gamma(u, v) = \frac{\sum_{x,y} [I(x, y) - \bar{I}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\left\{ \sum_{x,y} [I(x, y) - \bar{I}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2 \right\}^{0.5}} \quad (65)$$

sendo I a imagem, a soma sobre (x, y) sob a janela que contém o atributo t posicionado em (u, v) , \bar{t} é a média do conceito e $\bar{I}_{u,v}$ é a média de $I(x, y)$ na região sob o conceito.

A.2.2 Coeficiente de Variação

O Coeficiente de Variação (CoV) (SACHS, 1984) é definido por:

$$CoV = \frac{S}{\bar{x}} \text{ para todo } x > 0 \quad (66)$$

sendo \bar{x} a média dos valores utilizados e S é o desvio-padrão.

A.2.3 Coeficiente Relativo de Variação

O Coeficiente Relativo de Variação (CRoV) (SACHS, 1984) é definido por:

$$CRoV = \frac{S}{\bar{x}} \frac{100}{\sqrt{n}} \text{ para todo } x > 0 \quad (67)$$

sendo $\frac{S}{\bar{x}}$ o Coeficiente de Variação e n é a quantidade de elementos.

Bases de Imagens de Texturas

A validação dos experimentos realizados foi feita através da aplicação das propostas de caracterização às bases de imagens de textura. As bases utilizadas para os testes e validações dos métodos propostos foram:

- ❑ **Brodatz** é uma base de texturas proveniente do livro (BRODATZ, 1966) com 1.776 amostras de texturas agrupadas em 111 classes. Cada imagem tem 128×128 *pixels* com 256 tons de cinza.
- ❑ **CUReT** (*Columbia-Utrecht Reflectance and Texture*) das Universidades Columbia e Utrecht contendo imagens de 61 materiais diferentes, cada um fotografado em 92 combinações distintas de condições de iluminação e visualização. As imagens são em escala de cinza, com dimensões 200×200 *pixels* com 8 *bits/pixel* em formato PNG (DANA et al., 1999).
- ❑ **Outex_TC_00013** ou **Outex** inclui uma coleção de cenas naturais tendo 1.360 imagens em tons de cinza textura (128×128), sendo construído através da divisão de cada uma das 68 imagens de textura originais (746×538) em 20 sub-imagens não sobrepostas (OJALA et al., 2002).
- ❑ **KTH-TIPS** é uma base de texturas sob variações de iluminação, posição e escala, consistindo de 10 materiais, cada um capturado em 9 escalas diferentes e 9 posições e iluminações diferentes, totalizando 810 imagens com dimensões 200×200 *pixels* com 8 *bits/pixel* em formato PNG (HAYMAN et al., 2004).
- ❑ **UIUCTex** da Universidade de Illinois em Urbana-Champaign. Nesta base são 25 classes de texturas com 40 amostras de cada. Todas as imagens estão em escala de cinza com dimensões 640×480 *pixels* com 8 *bits/pixel* em formato JPG. A base inclui superfícies com variações de albedo (por exemplo, madeira e mármore), formas 3D (por exemplo, cascalho e pele), bem como uma mistura de ambos (por exemplo, carpete e tijolo). Também ocorrem mudanças significativas no ponto de

vista, diferenças de escala nas classes e condições de iluminação não controladas (LAZEBNIK; SCHMID; PONCE, 2005).

- ❑ **VisTex** (*The Vision Texture*) é mantida pelo grupo de Visão e Modelagem do MIT Media LAB. A base de imagens completa contém imagens representantes de texturas do mundo real em condições práticas (iluminação, perspectiva, etc.). As imagens são em tons de cinza com resolução de 512×512 *pixels*, divididos em 16 sub-imagens não sobrepostas de 128×128 *pixels* (VISTEX, 2009).
- ❑ **USC_SIPi** do Instituto de Processamento de Sinal e Imagem da Universidade do Sul da Califórnia. A base de imagem é dividida em 4 volumes e, para este trabalho foram utilizadas as texturas rotacionadas que são baseadas nas imagens da Brodatz. Essas texturas são de 13 imagens em sete diferentes ângulos: 0, 30, 60, 90, 120, 150 e 200 graus. As imagens tem dimensões de 512×512 *pixels* com 8 *bits/pixel* em formato TIFF (WEBER, 1997).
- ❑ **Pollen** da *Florida Institute of Technology* - Melbourne - USA, conta com 30 classes com uma média de 36 imagens (coloridas e tons de cinza) por classe com dimensões variadas, obtidas utilizando câmera digital acoplada a um microscópio Zeiss Axioskopmicroscope com lentes objetivas Plan-apochromatic ($\times 100$, 1.4) e Plan Neofluar ($\times 40$, 0.85) (BUSH; WENG, 2007).

A Figura 21 ilustra alguns dos padrões de texturas presentes nas base de imagens anteriormente descritas.

A base de imagens de grãos de Pólen (denominada no projeto como *Pollen*) foi disponibilizada pelo prof. Dr. Eraldo Ribeiro da *Florida Institute of Technology* - Melbourne - USA. As classes disponibilizadas e suas respectivas quantidades de imagens foram:

- ❑ *Alternanthera filifolia* - 40 imagens
- ❑ *Ambrosia arborens* - 39 imagens
- ❑ *Apium leptophyllum* - 40 imagens
- ❑ *Bidens triplinervia* - 33 imagens
- ❑ *Bursera simaruba* - 40 imagens
- ❑ *Chaenactis steroides* - 22 imagens
- ❑ *Clibadium surinamensis* - 45 imagens
- ❑ *Dictyocarium sp* - 29 imagens

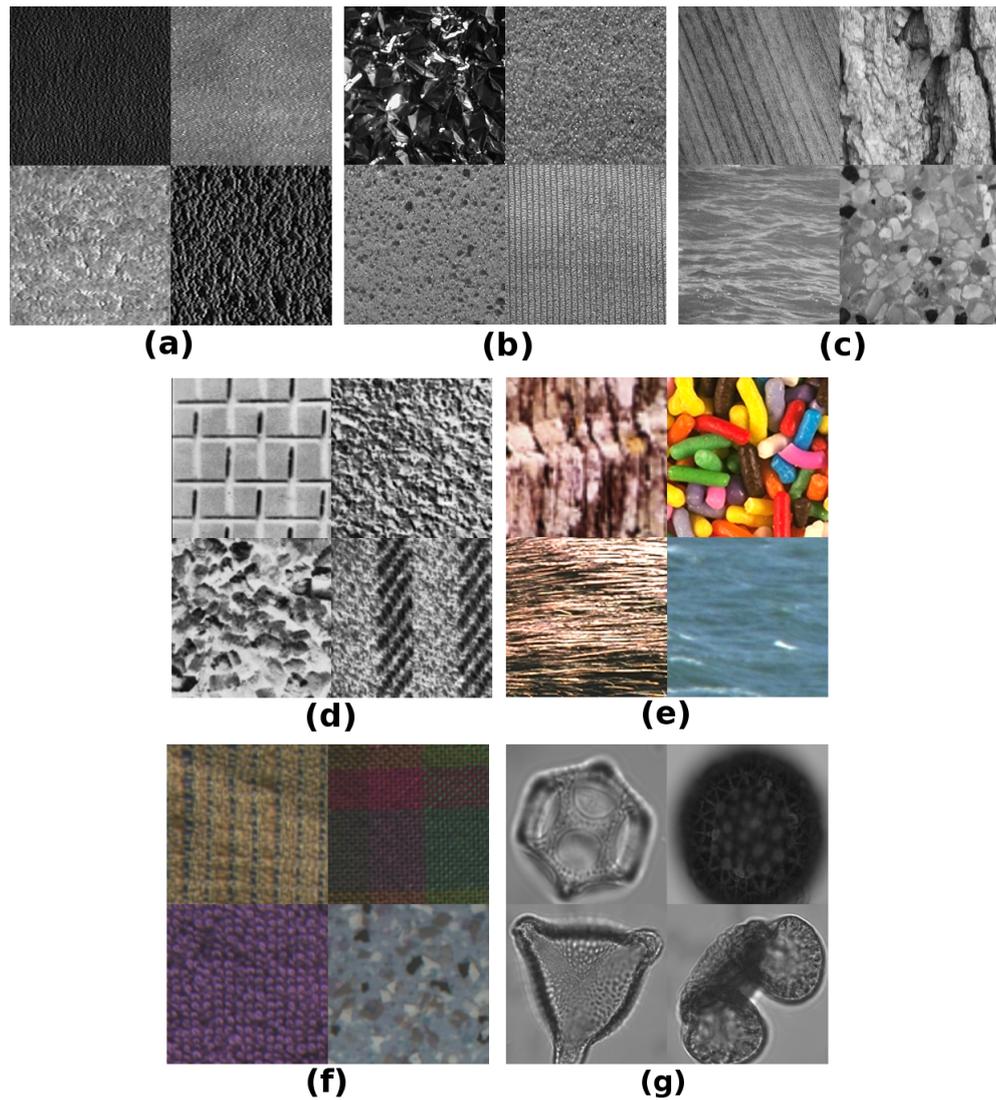


Figura 21 – Exemplos de imagens das bases: (a) CURET; (b) KTH-TIPS; (c) UIUCTex; (d) Brodatz; (e) VisTex; (f) Outex_TC_00013; (g) Pollen.

- ❑ *Dilenia indica* - 40 imagens
- ❑ *Euphorbiaceae Manihot leptophylla* - 40 imagens
- ❑ *Euterpe edulis* - 29 imagens
- ❑ *Exothea paniculata* - 40 imagens
- ❑ *Galapageium cardiospermum* - 40 imagens
- ❑ *Iriarteia deltoidea* - 29 imagens
- ❑ *Kangenekia lanceolata* - 40 imagens
- ❑ *Marila laxiflora* - 38 imagens
- ❑ *Mauritia flexuosa* - 28 imagens

- ❑ *Mauritiella armata* - 32 imagens
- ❑ *Pinus palustris* - 40 imagens
- ❑ *Psychotria nervosa* - 40 imagens
- ❑ *Puya roldanii* - 17 imagens
- ❑ *Quercus virginiana* - 42 imagens
- ❑ *Sabal palmetto* - 13 imagens
- ❑ *Stylogyne ambigua* - 40 imagens
- ❑ *Syagrus botryophora* - 35 imagens
- ❑ *Trifolium repens* - 40 imagens
- ❑ *Unonopsis floribunda* - 38 imagens
- ❑ *Urera elata* - 33 imagens
- ❑ *Vibrinium oboratum* - 40 imagens
- ❑ *Walteria ovata* - 41 imagens